



HAL
open science

New procedures controlling the false discovery proportion via Romano-Wolf's heuristic

Sylvain Delattre, Etienne Roquain

► **To cite this version:**

Sylvain Delattre, Etienne Roquain. New procedures controlling the false discovery proportion via Romano-Wolf's heuristic. 2014. hal-00905060v2

HAL Id: hal-00905060

<https://hal.science/hal-00905060v2>

Preprint submitted on 16 Jun 2014 (v2), last revised 3 Jun 2015 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NEW PROCEDURES CONTROLLING THE FALSE DISCOVERY PROPORTION VIA ROMANO-WOLF'S HEURISTIC

BY SYLVAIN DELATTRE AND ETIENNE ROQUAIN

The false discovery proportion (FDP) is a convenient way to account for false positives when a large number of tests are performed simultaneously. Romano and Wolf (2007) have proposed a general principle that builds FDP controlling procedures from k -family-wise error rate controlling procedures while incorporating dependencies in an appropriate manner, see Korn et al. (2004); Romano and Wolf (2007). However, the theoretical validity of the latter is still largely unknown. This paper provides a careful study of this heuristic: first, we extend this approach by using a notion of “bounding device” that allows to cover a wide range of critical values, including those that adapt to m_0 , the number of true null hypotheses. Second, the theoretical validity of the latter is investigated both non-asymptotically and asymptotically. Third, we introduce suitable modifications of this heuristic that provide new methods overcoming the existing procedures with a proven FDP control.

1. Introduction.

1.1. *Motivation.* Assessing significance in massive data is an important challenge of contemporary statistics, which becomes especially difficult when the underlying errors are correlated. Pertaining to this class of high-dimensional problems, a common issue is to make simultaneously a huge number of 0/1 decisions with a valid control of the overall amount of false discoveries (items declared to be wrongly significant). In this context, a convenient way to account for false discoveries is the false discovery proportion (FDP) that corresponds to the proportion of errors among the items declared as significant (i.e. “1”) by the procedure.

The Benjamini and Hochberg (BH) procedure has been widely popularized after the celebrated paper Benjamini and Hochberg (1995) and is shown to control the *expectation* of the FDP, called the false discovery rate (FDR), either theoretically under constrained dependency structures (see Benjamini and Yekutieli (2001)), or with simulations (see Kim and van de Wiel (2008)). However, many authors have noticed that the distribution of the FDP of BH procedure can be affected by the dependencies, see, e.g., Korn et al. (2004); Delattre and Roquain (2011); Guo et al. (2014), which makes the use of the BH procedure questionable.

To illustrate further this phenomenon, Figure 1 displays the distribution of the FDP of the BH procedure in the classical one-sided Gaussian multiple testing framework, when the test statistics are all ρ -equicorrelated. As ρ increases, the distribution of the FDP gets less concentrated and turns out to be drastically skewed for $\rho = 0.1$ (in particular it falls outside the Gaussian regime). Clearly, in that case, the mean fails to describe accurately the overall behavior of the FDP distribution. In particular, although the mean of the FDP is below $\alpha = 0.2$, the true value of FDP is not ensured to be small in that case.

An alternative proposed in Genovese and Wasserman (2004); Perone Pacifico et al. (2004);

AMS 2000 subject classifications: Primary 62H15; secondary 60F17

Keywords and phrases: multiple testing, false discovery rate, equi-correlation, Gaussian multivariate distribution, positive dependence, Simes' inequality

Lehmann and Romano (2005) is to control the $(1 - \zeta)$ -quantile of the FDP distribution at level α , that is, to assert

$$(1) \quad \mathbb{P}(\text{FDP} > \alpha) \leq \zeta.$$

While taking $\zeta = 1/2$ into (1) provides a control of the median of the FDP, taking $\zeta = 0.05$ ensures that the FDP does not exceed α with probability at least 95%. Markedly, Figure 1 shows that the $(1 - \zeta)$ -quantiles of the FDP distribution are substantially affected by the dependencies, but not equally for all the ζ 's: due to the increasingly heavy upper-tail, while the 95%-quantile gets substantially larger, the median gets slightly smaller. This suggests that the BH procedure is much too optimistic for a 95%-quantile control, but is actually too conservative for a FDP median control. Overall, this reinforces the fact that controlling the $(1 - \zeta)$ -quantile of the FDP is essential in the presence of dependence.

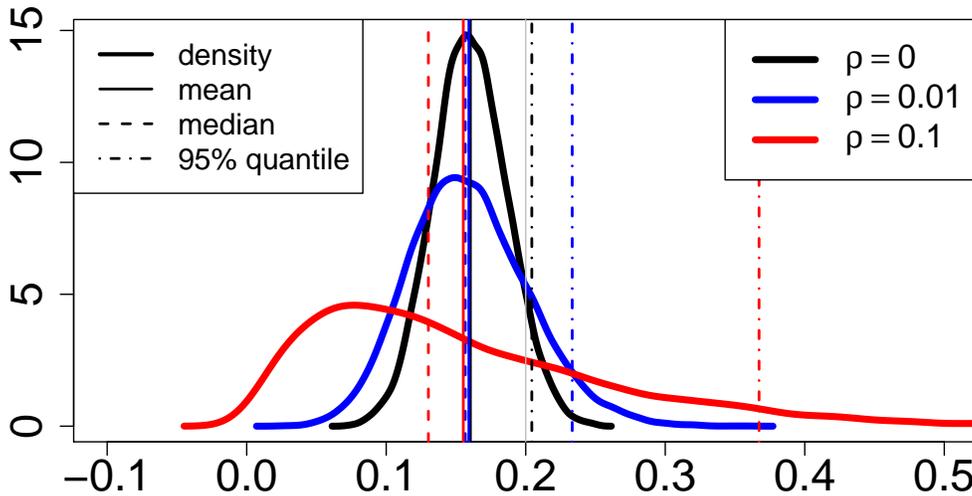


FIG 1. Fitted density of the false discovery proportion of the BH procedure when increasing the dependence. $m = 1000$, $m_0 = 800$, 10^4 simulations, Gaussian one-sided equicorrelated model.

1.2. *RW's heuristic and main contributions of this paper.* The problem of finding multiple testing procedures ensuring the control (1) has received a growing attention in the last decades, see for instance Lehmann and Romano (2005); Romano and Shaikh (2006a,b); Romano and Wolf (2007); Guo and Romano (2007); Dudoit and van der Laan (2008); Chi and Tan (2008a); Roquain (2011); Roquain and Villers (2011); Guo et al. (2014). However, existing procedures with a proven FDP control are in general too conservative. This increases the interest of simple and general heuristics that work “fairly”. Romano and Wolf (2007) Romano and Wolf (2007), themselves referring to Korn et al. (2004) Korn et al. (2004), have proposed such an heuristic. It is called RW's heuristic in the sequel and can be formulated as follows: *start from a family \mathcal{R}_k , $k \in \{1, \dots, m\}$, of procedures such that for all k , with probability at least $1 - \zeta$, the procedure \mathcal{R}_k makes less than $k - 1$ false discoveries. Then, choose some \hat{k} such that $(\hat{k} - 1)/R_{\hat{k}} \leq \alpha$, where R_k denotes the number of rejections of \mathcal{R}_k . Finally use $\mathcal{R}_{\hat{k}}$.* Note that,

in the original formulation, \widehat{k} was constrained to be chosen such that any k' with $k' < \widehat{k}$ should also satisfy $(k' - 1)/R_{k'} \leq \alpha$ (“step-down” approach). The rationale behind this principle is that, for each k , the FDP of \mathcal{R}_k is bounded by $(k - 1)/R_k$ with probability $1 - \zeta$, so that the FDP of $\mathcal{R}_{\widehat{k}}$ should be smaller than $(\widehat{k} - 1)/R_{\widehat{k}} \leq \alpha$ with probability $1 - \zeta$, which entails (1). However, as it is, this argument is not rigorous, because it does not take into account the fluctuations of \widehat{k} .

This heuristic has been theoretically justified (in the step-down form) in settings where the p -values under the null are independent of the p -values under the alternative (full independence in Guo and Romano (2007); alternative p -values all equal to 0 in Romano and Wolf (2007)). Since these situations rely on an independence assumption, and since the FDP is particularly interesting under dependence, it seems appropriate to study the precise behavior of this method in a “simple dependent case”. Thus, our study is guided by the case where the dependencies are *known* and Gaussian multivariate, from which an important particular case is the equi-correlated case mentioned above with a known ρ (the estimation of ρ will be also discussed). In a nutshell, this paper makes the following main contributions.

- It provides a general framework in which RW’s heuristic can be investigated, by building the initial k -FWE critical values with “bounding devices”: a strong interest is the possibility to build critical values that “adapt” to m_0 , the number of true nulls. This allows to encompass many procedures, either new or previously known.
- We show that RW’s heuristic may fail to control the FDP non-asymptotically (even under its step-down form). Two corrections that provably control the FDP are introduced. By using simulations, we show that the resulting procedures are more powerful than those previously existing.
- In an ρ -equicorrelated one-sided Gaussian asymptotical framework we show that one of our bounding device suggests to use the following new critical values:

$$(2) \quad \tau_\ell = \overline{\Phi} \left(\rho^{1/2} \overline{\Phi}^{-1}(\zeta) + (1 - \rho)^{1/2} \overline{\Phi}^{-1}(\alpha \ell / m) \right),$$

where $\overline{\Phi}$ is the upper-tail of the standard normal distribution. Interestingly, we show that the step-up procedure associated to (2) asymptotically controls the FDP. Furthermore, we prove that this control is maintained when we plug an estimator $\widehat{\rho}_m$ of ρ into (2) (provided that the estimator has an estimation rate faster than logarithmic).

Finally, let us emphasize that the novel critical values (2) allow to describe how the quantities α , ζ and ρ come into play when controlling (asymptotically) the FDP. Taking $\rho = 0$ just gives Simes’ critical values and thus the BH procedure, whatever ζ is. The asymptotic FDP control can be explained in that case by the fast concentration of the FDP of BH around its expectation as m grows to infinity under independence (see, e.g., Neuvial (2008)). Now, for $\rho > 0$, the new critical values are markedly different from the BH critical values: taking $\zeta = 1/2$ leads to less conservative critical values (if $\alpha \leq 1/2$), while taking ζ smaller can lead to more conservativeness (as expected), see Figure 2 (a) for an illustration. Finally, we plot in Figure 2 (b) the density of the FDP of the step-up procedure using the new critical values (2) for $\zeta = 0.05$. As one might expect, compared to the BH procedure, the density has been shifted to the left so that the 95%-quantile of the FDP of the novel procedure is below α .

1.3. *Multiple testing framework.* We observe a random variable X , whose distribution belongs to some set \mathcal{P} . For $m \geq 2$, we define a setting for performing m tests simultaneously

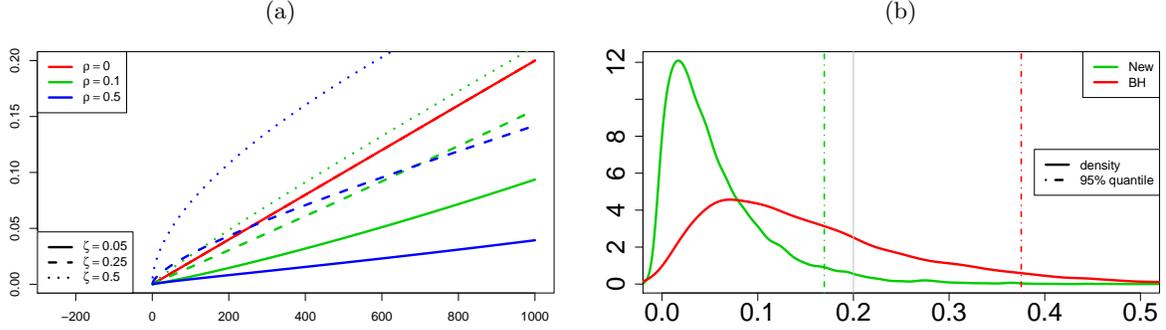


FIG 2. Left: plot of the critical values (2) in function of l . Right: same as Figure 1 but only for $\rho = 0.1$ and by adding the new step-up procedure using (2). $m = 1000$, $\alpha = 0.2$, Gaussian one-sided equicorrelated model.

by introducing a true/false null parameter $H \in \{0, 1\}^m$ and a set of associated distributions $\mathcal{P}_H \subset \mathcal{P}$ (assumed to be non-empty) which are candidates to be the distribution of X under the configuration H . We denote $\mathcal{H}_0(H) = \{i : H_i = 0\}$, $m_0(H) = \sum_{i=1}^m (1 - H_i)$ and $\mathcal{H}_1(H) = \{i : H_i = 1\}$, $m_1(H) = \sum_{i=1}^m H_i$ the set/number of true and false nulls, respectively. The basic assumption is the following: for all $i \in \{1, \dots, m\}$, there is a p -value $p_i(X)$ satisfying the following assumption

$$\forall H \in \{0, 1\}^m \text{ with } H_i = 0, \text{ we have } p_i(X) \sim U(0, 1) \text{ when } X \sim P \text{ with } P \in \mathcal{P}_H.$$

In this paper, a leading example is the one-sided *Gaussian multivariate framework*, for which $X = H\mu + Y$, where $H\mu = (H_i\mu_i)_{1 \leq i \leq m}$ for $H \in \{0, 1\}^m$, $\mu \in (\mathbb{R}_+ \setminus \{0\})^m$ and Y is a m -dimensional centered Gaussian vector with covariance matrix Γ such that $\Gamma_{i,i} = 1$. We focus on the (one-sided) multiple testing problem of “ $\mathbb{E}X_i = 0$ ” versus “ $\mathbb{E}X_i > 0$ ”, which is equivalent to test “ $H_i = 0$ ” against “ $H_i = 1$ ” here. The p -values are given by $p_i(X) = \bar{\Phi}(X_i)$. A special case of interest is the equi-correlated case:

$$(\rho\text{-equi}) \quad \Gamma_{i,j} = \rho, \text{ for all } i \neq j, \text{ where } \rho \in [-(m-1)^{-1}, 1],$$

which is extensively used throughout the paper. When (ρ -equi) does not hold, the joint distribution of the p -values under the null ($p_i, i \in \mathcal{H}_0(H)$) depends in general on the subset $\mathcal{H}_0(H)$. Obviously, in that case, we do not want to explore the $\binom{m}{m_0(H)}$ possible subsets of $\{1, \dots, m\}$ in our inference, which inevitably should arise when our procedure fits to such a dependence structure. To circumvent this technical difficulty, we can add random effects to our model. This makes H becoming random. More formally, we distinguish between the two following models:

- Fixed mixture model: the parameter H is fixed by advance and unknown. Overall, the parameters of the model are given by $\theta = (H, P)$ to be chosen in the set

$$\Theta^F = \{(H, P) : H \in \{0, 1\}^m, P \in \mathcal{P}_H\}.$$

- Uniform mixture model: the number of true null $m_0 \in \{0, 1, \dots, m\}$ is unknown and fixed by advance, while H is a random vector distributed in such a way that $\mathcal{H}_0(H)$ is randomly generated (independently and previously of the other variables), uniformly in the subsets of $\{1, \dots, m\}$ of cardinal m_0 . The parameters of the model are given by $\theta = (m_0, (P_H)_H)$, to be chosen in the set

$$\Theta^U = \{(m_0, (P_H)_H) : m_0 \in \{0, 1, \dots, m\}, P_H \in \mathcal{P}_H \text{ for all } H\}.$$

In this model, the distribution of X conditionally on H is P_H .

While the fixed mixture model is the most commonly used model for multiple testing, the uniform mixture model is new to our knowledge and follows the general philosophy of models with random effects, see [Efron et al. \(2001\)](#). It is convenient for the adaptation issue w.r.t. m_0 , as we will see later on. With some abuse, we denote $m_0(\theta)$ (or m_0 when not ambiguous) the number of true nulls in the fixed/uniform mixture models. In the sequel, Θ denotes either Θ^F or Θ^U (when unspecified, Θ is Θ^F).

1.4. *Type I error rates.* First, for $t \in [0, 1]$, denote $V_m(t) = \sum_{i=1}^m (1 - H_i) \mathbf{1}\{p_i(X) \leq t\}$ and $R_m(t) = \sum_{i=1}^m \mathbf{1}\{p_i(X) \leq t\}$ the number of false discoveries and the number of discoveries (at threshold t), respectively. For some pre-specified $k \in \{1, \dots, m\}$ and some thresholding method $\hat{t}_m \in [0, 1]$ (potentially depending on the data), the k -family-wise error rate (k -FWER) is defined as the probability that more than k true nulls have a p -value smaller than \hat{t}_m , see, e.g., [Hommel and Hoffman \(1988\)](#); [Lehmann and Romano \(2005\)](#). Formally, for $\theta \in \Theta$ (in one of the models defined in Section 1.3 and Θ being the corresponding parameter space),

$$(3) \quad k\text{-FWER}(\hat{t}_m) = \mathbb{P}_\theta(V_m(\hat{t}_m) \geq k).$$

Note that $k = 1$ corresponds to the traditional family-wise error rate (FWER). From (3), providing $k\text{-FWER}(\hat{t}_m) \leq \zeta$ (for all $\theta \in \Theta$ and some $\zeta \in (0, 1)$), ensures that, with probability at least $1 - \zeta$, less than $k - 1$ false discoveries are made by the thresholding procedure \hat{t}_m .

Next, for some threshold $t \in [0, 1]$, define the false discovery proportion at threshold t as follows:

$$(4) \quad \text{FDP}_m(t) = \frac{V_m(t)}{R_m(t) \vee 1}.$$

The quantity $\text{FDP}_m(t)$ is not observable because it depends on the unknown process $V_m(t)$. Furthermore, $\text{FDP}_m(t)$ is random, so that $\text{FDP}_m(t) \leq \alpha$ can hold only with high probability. Controlling the FDP via a threshold $t = \hat{t}_m$ (potentially depending on the data) thus corresponds to the following probabilistic bound:

$$(5) \quad \forall \theta \in \Theta, \quad \mathbb{P}_\theta(\text{FDP}_m(\hat{t}_m) \leq \alpha) \geq 1 - \zeta,$$

for some pre-specified values $\alpha, \zeta \in (0, 1)$. As mentioned before, (5) corresponds to upper-bound the $(1 - \zeta)$ -quantile of the distribution of $\text{FDP}_m(\hat{t}_m)$ by α . Since $\text{FDP}_m(t) > \alpha$ is equivalent to $V_m(t) \geq \lfloor \alpha R_m(t) \rfloor + 1$, the FDP control and the k -FWER control are intrinsically linked.

From an historical point of view, the introduction of the FDP goes back to Eklund in the 1960's (as reported in [Seeger \(1968\)](#)), that has presented the FDP as a solution to the ‘‘mass-significance problem’’. Much later, the seminal paper of Benjamini and Hochberg [Benjamini and Hochberg \(1995\)](#) has widely popularized the use of the FDP in practical problems by introducing and studying the false discovery rate (FDR), which corresponds to the expectation of the FDP.

1.5. *Step-up and step-down procedures.* Consider a nondecreasing sequence $(\tau_\ell)_{1 \leq \ell \leq m}$ of nonnegative values, with the convention $\tau_0 = 0$. Classically, the latter is referred to as *critical*

values. The corresponding step-up (resp. step-down) procedure is defined as rejecting the p -values smaller than $\tau_{\hat{\ell}}$, where $\hat{\ell}$ is defined by either of the two following quantities:

$$\begin{aligned} \text{(SU)} \quad & \max\{\ell \in \{0, 1, \dots, m\} \text{ such that } p_{(\ell)} \leq \tau_{\ell}\}; \\ \text{(SD)} \quad & \max\{\ell \in \{0, 1, \dots, m\} \text{ such that } \forall \ell' \in \{0, 1, \dots, \ell\}, p_{(\ell')} \leq \tau_{\ell'}\}. \end{aligned}$$

Let us also recall the so-called ‘‘switching relation’’: $p_{(\ell)} \leq \tau_{\ell}$ is equivalent to $R_m(\tau_{\ell}) \geq \ell$. This entails $R_m(\tau_{\hat{\ell}}) = \hat{\ell}$ both in the step-up and step-down cases.

2. Building k -FWE-based critical values.

2.1. *Revisiting RW’s heuristic.* Starting from arbitrary critical values $(\tau_{\ell})_{1 \leq \ell \leq m}$, and by taking an integer $\hat{\ell}$ such that $R_m(\tau_{\hat{\ell}}) = \hat{\ell}$, we have

$$\begin{aligned} \mathbb{P}_{\theta}(\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha) &= \mathbb{P}_{\theta}(V_m(\tau_{\hat{\ell}}) > \alpha R_m(\tau_{\hat{\ell}})) \\ \text{(6)} \quad &= \mathbb{P}_{\theta}(V_m(\tau_{\hat{\ell}}) \geq \lfloor \alpha \hat{\ell} \rfloor + 1). \end{aligned}$$

Hence, by taking τ_{ℓ} such that $(\lfloor \alpha \ell \rfloor + 1)\text{-FWER}(\tau_{\ell}) \leq \zeta$ for all ℓ , we should get that (6) is below ζ . However, as already mentioned, the above reasoning does not rigorously establish (5) (with $\hat{t}_m = \tau_{\hat{\ell}}$), because it implicitly assumes that $\hat{\ell}$ is determinist. Nevertheless, this heuristic is a good starting point for building critical values related to the FDP control.

2.2. *Bounding device.* Let us consider either the fixed model $\Theta = \Theta^F$ or the uniform model $\Theta = \Theta^U$. First, let us define a bounding device as any function $B_m^0 : (t, k, u) \mapsto B_m^0(t, k, u) \in [0, 1]$, defined for $t \in [0, 1]$, $k \in \{1, \dots, m\}$ and $u \in \{0, \dots, m\}$, which is non-increasing in k , with $B_m^0(0, k, u) = 0$ for all u, k , $B_m^0(t, k, u) = 0$ for all $t \in [0, 1]$ whenever $u < k$, and such that for all $t \in [0, 1]$, $k \in \{1, \dots, m\}$ and $u \in \{k, \dots, m\}$, we have

$$\text{(Bound)} \quad B_m^0(t, k, u) \geq \sup_{\substack{\theta \in \Theta \\ m_0(\theta) = u}} \{\mathbb{P}_{\theta}(V_m(t) \geq k)\}.$$

Now, define for $t \in [0, 1]$, $k \in \{1, \dots, m\}$ and $\ell \in \{k, \dots, m\}$, the quantities

$$\begin{aligned} \text{(Bound-nonadapt)} \quad & \bar{B}_m(t, k) = \sup_{0 \leq u \leq m} \{B_m^0(t, k, u)\}; \\ \text{(Bound-adapt)} \quad & \tilde{B}_m(t, k, \ell) = \sup_{k \leq k' \leq \ell} \left\{ \sup_{0 \leq u \leq m - \ell + k'} B_m^0(t, k', u) \right\}; \end{aligned}$$

which are additionally assumed to be non-decreasing and left-continuous in t . Note that $\bar{B}_m(t, k)$ and $\tilde{B}_m(t, k, \ell)$ are both non-increasing in k .

DEFINITION 2.1. *Let us consider a bounding device $B_m^0(t, k, u)$ and the above associated quantities $\bar{B}_m(t, k)$ and $\tilde{B}_m(t, k, \ell)$. Then the non adaptive (resp. adaptive, oracle) k -FWE-based critical values associated to the bounding function B_m^0 are defined as follows (respectively):*

$$\begin{aligned} \text{(7)} \quad & \bar{\tau}_{\ell} = \max \{t \in [0, 1] : \bar{B}_m(t, \lfloor \alpha \ell \rfloor + 1) \leq \zeta\}, \quad 1 \leq \ell \leq m; \\ \text{(8)} \quad & \tilde{\tau}_{\ell} = \max \left\{ t \in [0, 1] : \tilde{B}_m(t, \lfloor \alpha \ell \rfloor + 1, \ell) \leq \zeta \right\}, \quad 1 \leq \ell \leq m; \\ \text{(9)} \quad & \tau_{\ell}^0 = \max \{t \in [0, 1] : B_m^0(t, \lfloor \alpha \ell \rfloor + 1, m_0) \leq \zeta\}, \quad 1 \leq \ell \leq m. \end{aligned}$$

The above definition implies that $(\bar{\tau}_\ell)_{1 \leq \ell \leq m}$, $(\tilde{\tau}_\ell)_{1 \leq \ell \leq m}$ and $(\tau_\ell^0)_{1 \leq \ell \leq m}$ are nondecreasing sequences, so that they can be used as critical values. The critical values $\tilde{\tau}_\ell$, $\ell = 1, \dots, m$, are said *adaptive* because they implicitly (over-)estimate m_0 by

$$(10) \quad m(\ell) = m - \ell + \lfloor \alpha \ell \rfloor + 1.$$

In the literature, this way to adapt to π_0 is often referred to as “one-stage” (by contrast with “two-stage”, see [Benjamini et al. \(2006\)](#); [Sarkar \(2008\)](#); [Blanchard and Roquain \(2009\)](#)). It has been used in [Gavrilov et al. \(2009\)](#); [Finner et al. \(2009\)](#) and has been proved to be asymptotically optimal in a specific sense (AORC), see [Finner et al. \(2009\)](#). Also, $\bar{\tau}_\ell \leq \tilde{\tau}_\ell$ for all ℓ , that is, adaptation always leads to less conservative critical values. Finally, it is worth to check that $\bar{\tau}_m \leq \tilde{\tau}_m < 1$ (this comes from $B_m^0(1, k, u) = 1$ for all $u \geq k$) so that the output $\hat{\ell}$ of the step-up algorithm is not identically equal to m .

2.3. Examples. We provide below three examples of bounding devices: Markov, K -Markov and Exact. Instances of resulting critical values are displayed in [Figure 3](#). As we will see, while the exact bounding device leads to the largest critical values the Markov-type devices are still useful because they can offer finite sample controls.

Markov. By Markov’s inequality, we have

$$(11) \quad \mathbb{P}_\theta(V_m(t) \geq k) \leq \frac{\mathbb{E}_\theta(V_m(t))}{k} = \frac{m_0 t}{k} =: B_m^0(t, k, m_0).$$

Since $\bar{B}_m(t, k) = mt/k$ and $\tilde{B}_m(t, k, \ell) = \sup_{k \leq k' \leq \ell} \sup_{0 \leq u \leq m - \ell + k'} \{ut/k'\} = (m - \ell + k)t/k$, this gives rise to the critical values

$$(12) \quad \bar{\tau}_\ell = \frac{\zeta(\lfloor \alpha \ell \rfloor + 1)}{m}; \quad \tilde{\tau}_\ell = \frac{\zeta(\lfloor \alpha \ell \rfloor + 1)}{m(\ell)}, \quad 1 \leq \ell \leq m,$$

where $m(\ell)$ is defined by (10). The adaptive critical values $(\tilde{\tau}_\ell)_{1 \leq \ell \leq m}$ are those proposed by Lehmann and Romano in [Lehmann and Romano \(2005\)](#). Note that these critical values do not adapt to the underlying dependence structure of the p -values.

K -Markov. When ζ is small, the Markov device can be too conservative and we might want to use a sharper tool. Let $K \geq 1$ be an integer. As suggested in [Guo et al. \(2014\)](#) (for $K = 2$), we can use the following bound: for $k \geq K$,

$$\mathbb{P}_\theta(V_m(t) \geq k) \leq \frac{\mathbb{E}_\theta[V_m(t)(V_m(t) - 1) \cdots (V_m(t) - K + 1)]}{k(k-1) \cdots (k-K+1)}.$$

By essence, this device is useless whenever $k < K$, so we replace it by the simple Markov device. This yields

$$(13) \quad B_m^0(t, k, u) = \begin{cases} \frac{A_m(t, K, u)}{k(k-1) \cdots (k-K+1)} & \text{if } K \leq k \\ \frac{ut}{k} \vee \frac{A_m(t, K, m)}{K!} & \text{if } 1 \leq k \leq K-1, \end{cases}$$

where $A_m(t, K, u) = \sup\{\mathbb{E}_\theta[V_m(t)(V_m(t) - 1) \cdots (V_m(t) - K + 1)], \theta \in \Theta, m_0(\theta) = u\}$. Note that the operator “ \vee ” in the last display is added to keep the non-increasing property w.r.t. k . As a first illustration, assuming

(Indep) $p_i, i \in \mathcal{H}_0$, are mutually independent (cond. on H in model Θ^U),

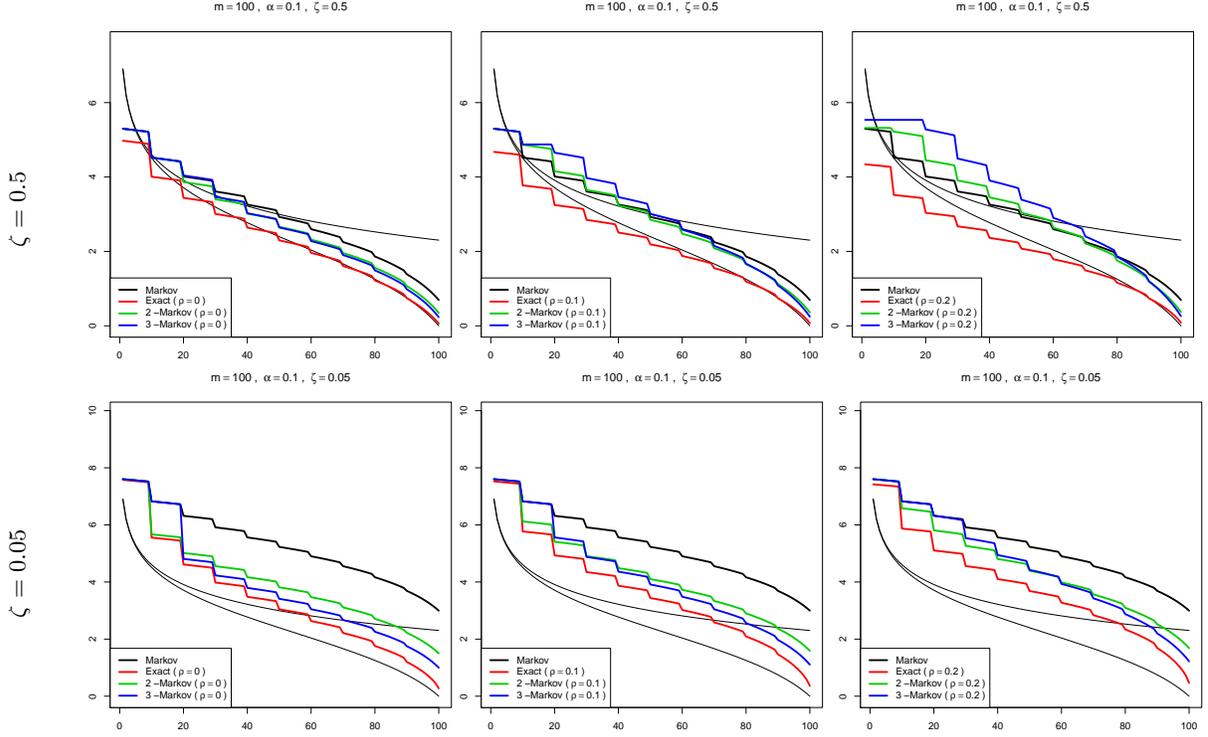


FIG 3. Plot of $-\log(\tilde{\tau}_\ell)$ in function of $\ell \in \{1, \dots, m\}$, for k -FWE-based critical values obtained with several types of bounding devices and assuming $(\rho$ -equi). For comparison, the solid thin black line corresponds either to the BH critical values $-\log(\alpha\ell/m)$ (non adaptive) or the AORC critical values $-\log(\alpha\ell/(m - (1 - \alpha)\ell))$ defined in [Finner et al. \(2009\)](#) (adaptive).

$B_m^0(t, k, u)$ equals $\frac{u(u-1)\dots(u-K+1)}{k(k-1)\dots(k-K+1)}t^K$ for $k \geq K$, which entails for $\ell \in \{1, \dots, m\}$, $\tilde{\tau}_\ell = \zeta^{1/K} \left(\frac{(\lfloor \alpha\ell \rfloor + 1)(\lfloor \alpha\ell \rfloor + 1 - 1)\dots(\lfloor \alpha\ell \rfloor + 1 - K + 1)}{m(\ell)(m(\ell) - 1)\dots(m(\ell) - K + 1)} \right)^{1/K}$ for $\ell \geq \ell_K$, where ℓ_K denotes $\lceil (K-1)/\alpha \rceil$ with $\lceil n \rceil$ being the smallest integer larger than or equal to n . A second illustration is the equicorrelated case $(\rho$ -equi) with $\rho \in [0, 1)$, for which

$$(14) \quad B_m^0(t, k, u) = \frac{u(u-1)\dots(u-K+1)}{k(k-1)\dots(k-K+1)} \int_0^1 (F_{0,\rho}(t, x))^K dx$$

for $k \geq K$, where we let

$$(15) \quad F_{0,\rho}(t, x) = \bar{\Phi} \left((\bar{\Phi}^{-1}(t) - \sqrt{\rho}\bar{\Phi}^{-1}(x))/\sqrt{1-\rho} \right).$$

The latter comes from the well known decomposition

$$(16) \quad X_i - \mathbb{E}X_i \sim \sqrt{\rho}\xi_0 + \sqrt{1-\rho}\xi_i,$$

where the ξ_i 's are all i.i.d. $\mathcal{N}(0, 1)$. Inverting $B_m^0(t, \lfloor \alpha\ell \rfloor + 1, m(\ell)) = \zeta$ gives rise to critical values $\bar{\tau}_\ell$ and $\tilde{\tau}_\ell$, both taking into account the value of ρ .

Exact. In some cases, closed-formulas can be derived for the RHS of (Bound). First, assuming (Indep), the distribution of $V(t)$ is a binomial with parameters (u, t) . Hence, $B_m^0(t, k, u) = \sum_{j=k}^u \binom{u}{j} t^j (1-t)^{u-j}$. It is easy to check that $\tilde{B}_m(t, k, \ell) = B_m^0(t, k, m - \ell + k)$, so that the corresponding adaptive critical values can be obtained by a numerical inversion (these critical values were already proposed in Guo and Romano (2007)). Second, the following exact formula can be used under (ρ -equi) when $\rho \in [0, 1)$:

$$(17) \quad \begin{aligned} B_m^0(t, k, u) &= \int_0^1 \sum_{j=k}^u \binom{u}{j} (F_{0,\rho}(t, x))^j (1 - F_{0,\rho}(t, x))^{u-j} dx \\ \tilde{B}_m(t, k, \ell) &= B_m^0(t, k, m - \ell + k), \end{aligned}$$

where $F_{0,\rho}(t, x)$ is defined in (15). Third, in the Gaussian case where Γ is known, the quantities in the RHS of (Bound) can be computed by a Monte-Carlo method as follows: on the one hand, in the non adaptive case, we can upper-bound $V_m(t)$ by the full null process

$$(18) \quad V'_m(t) = m^{-1} \sum_{i=1}^m \mathbf{1} \{ \bar{\Phi}(Y_i) \leq t \},$$

whose distribution is known when Γ is known, because $Y \sim \mathcal{N}(0, \Gamma)$. On the other hand, in the adaptive case, the original null process $V_m(t)$ can be easily generated in the uniform model Θ^U in the case $m_0(\theta) = u$, for an arbitrary u . This leads to non-adaptive and adaptive critical values that incorporate any pre-specified covariance matrix Γ .

3. Finite sample results.

3.1. *Preliminary results.* The following theorem gathers the only existing cases, quite restrictive, where RW's heuristic has been proved to provide FDP control (to our knowledge). For the sake of generality, we reformulate these results by using our setting.

PROPOSITION 3.1 (Romano and Wolf (2007); Guo and Romano (2007); Guo et al. (2014)). *Consider some bounding device B_m^0 and the associated k -FWE-based critical values $(\tau_\ell)_{1 \leq \ell \leq m}$, being either adaptive or not and computed either in the fixed mixture model ($\Theta = \Theta^F$) or in the uniform mixture model ($\Theta = \Theta^U$). Let us consider the corresponding number of rejections $\hat{\ell}$ of the associated step-down (SD) or step-up (SU) procedure. Then the FDP control (5) holds (with $\hat{t}_m = \tau_{\hat{\ell}}$) in the following cases:*

- (i) *step-down algorithm and the null p -values $(p_i, i : H_i = 0)$ are independent of the alternative p -values $(p_i, i : H_i = 1)$ Romano and Wolf (2007); Guo and Romano (2007);*
- (ii) *step-down or step-up algorithm with the Lehmann-Romano critical values, that is, with $(\tau_\ell)_{1 \leq \ell \leq m}$ given by (12), and when Simes' inequality is valid:*

$$(19) \quad \mathbb{P}_\theta \left(\bigcup_{k=1}^{m_0} \{q_{(k)} \leq \zeta k / m_0\} \right) \leq \zeta,$$

where $q_{(1)} \leq q_{(2)} \leq \dots \leq q_{(m_0)}$ denote the ordered p -values under the null Guo et al. (2014).

The case (i) comes from inequalities established in Lehmann and Romano (2005); Romano and Wolf (2007), that we recall in Section 7.1 under an unified form (see also Theorem 5.2 in

Roquain (2011)). Note that it contains the case where all the p -values under the alternative are equal to zero (Dirac-Uniform configuration). Case (ii) has been solved more recently in Guo et al. (2014). Here, it can be seen as a consequence of the following general inequality, that will be useful throughout the paper (see Section 7.3 for a proof).

PROPOSITION 3.2. *Consider the setting of Proposition 3.1 in the step-down or step-up case. Then we have for all $\theta \in \Theta$,*

$$(20) \quad \mathbb{P}_\theta (\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha) \leq \mathbb{P}_\theta \left(V_m(\nu_k^0) \geq \hat{k} \geq 1 \right) = \mathbb{P}_\theta \left(q_{(\hat{k})} \leq \nu_k^0, \hat{k} \geq 1 \right),$$

for $\hat{k} = V_m(\tau_{\hat{\ell}})$, where $q_{(1)} \leq \dots \leq q_{(m_0)}$ denotes the ordered p -values under the null and where

$$(21) \quad \nu_k^0 = \max\{t \in [0, 1] : B_m^0(t, k, m_0) \leq \zeta\}.$$

Proposition 3.1 (ii) thus follows from Proposition 3.2, used with the adaptive Markov bounding device coming from (11). Markedly, Proposition 3.2 establishes that the FDP control for adaptive k -FWE based critical values is linked to a specific inequality between the null p -values and the bounding device using the true value of m_0 .

Further note that (20) in Proposition 3.2 is sharp whenever $m_0(\theta) = m$: in that case, the LHS and RHS are both equal to the probability that $\hat{k}(=\hat{\ell})$ is not zero, i.e., that at least one $\ell \in \{1, \dots, m\}$ is such that $p_{(\ell)} \leq \tau_\ell$. For instance, in the independent case and using the exact device (17), when $m = m_0 = 2$ and $\alpha = 0.5$, we have $\bar{\tau}_1 = \tilde{\tau}_1 = 1 - (1 - \zeta)^{1/2}$ and $\bar{\tau}_2 = \tilde{\tau}_2 = \zeta^{1/2}$ and

$$\mathbb{P}_\theta (\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha) = 2\zeta - (1 - (1 - \zeta)^{1/2})(2\zeta^{1/2} - 1 + (1 - \zeta)^{1/2}).$$

We merely check that the latter is larger than ζ for all $\zeta \in (0, 1)$. This establishes the following:

FACT 3.3. *RW's heuristic does not always provide a valid FDP control for finite m in its step-up form.*

Also, non-reported simulations show that the FDP control can be violated for a larger value of m .

Now, an important question is to know whether RW's heuristic always provides a valid FDP control for finite m in its *step-down* form. First, we can merely check that the following cases can be added in Proposition 3.1 in the step-down case:

- (iii) for all $\theta \in \Theta$, $\lfloor \alpha b_\alpha(m_0(\theta)) \rfloor = 0$ (e.g., $m_0(\theta) \in \{1, m\}$ or $\lfloor \alpha m \rfloor = 0$);
- (iv) under (ρ -equi) when $\rho = 1$.

Note that (iii) contains the case $m_0 = m$ which was problematic in the step-up case. A consequence is that any configuration for which the FDP control fails should be searched outside cases (i), (ii), (iii) and (iv). As a matter of fact, we found a numerical example under equi-correlation when using the critical values $(\tau_\ell^0)_\ell$ defined by (9), with the exact device. To this end, we have evaluated the probability exceedance of the FDP by the *exact calculations* proposed in Roquain and Villers (2011); Blanchard et al. (2014). This method is time consuming for large m but avoids the undesirable fluctuations due to the Monte-Carlo approximation while performing simulations. Precisely, under Assumption (ρ -equi), when $m = 30$, $\alpha = 0.2$,

$\zeta = 0.05$, $\rho = 0.3$, $m_0 = 15$, $\mu_i = 1.5$, $1 \leq i \leq m$, we obtained $\mathbb{P}_\theta \left(\text{FDP}_m(\tau_\ell^0) > \alpha \right) > \zeta + 10^{-3}$. Admittedly, the FDP control is just slightly violated. Nevertheless, this gives a numerical evidence of the following fact.

FACT 3.4. *RW's heuristic does not always provide a valid FDP control for finite m in its (oracle) step-down form.*

Fact 3.4 is interesting from a theoretical point of view: it annihilates any hope of finding a general finite sample proof of FDP control in the step-down case (the case of the *non-oracle* adaptive version is studied with extensive simulations, see Section 5.1).

3.2. Existing modifications. Fact 3.4 indicates that, to obtain a provable finite sample control, it is appropriate to slightly decrease the initial k -FWE-based critical values $(\tau_\ell)_{1 \leq \ell \leq m}$. Interestingly, several existing procedures that provably control the FDP can be reinterpreted as modifications of the τ_ℓ 's. In the literature, we have identified the following principles:

- The “diminution” principle [Romano and Shaikh \(2006a,b\)](#); [Guo et al. \(2014\)](#): first, establish a rigorous upper-bound for $\mathbb{P}(\text{FDP} > \alpha)$ for a step-down or step-up procedure with arbitrary critical values $(c_\ell(x))_{1 \leq \ell \leq m}$ depending on a single parameter x . Second, adjust x to make the bound smaller than ζ . As an illustration, [Romano and Shaikh \(2006a,b\)](#) have proposed a bound that can be rewritten as follows (see Section 7.2 for a proof):

$$(22) \quad C^{RS}(x) = \max_{1 \leq u \leq m} \left\{ u \sum_{\ell=1}^{b_\alpha(u)} \frac{c_\ell(x) - c_{\ell-1}(x)}{d(\ell, m, u)} \right\};$$

where for all $u, \ell \in \{0, \dots, m\}$, we let

$$(23) \quad b_\alpha(u) = \begin{cases} (\lfloor (m-u)/(1-\alpha) \rfloor + 1) \wedge (\lceil u/\alpha \rceil - 1) \wedge m & \text{(step-down)} \\ \lceil u/\alpha \rceil - 1 \wedge m & \text{(step-up)} \end{cases};$$

$$(24) \quad d(\ell, m, u) = \begin{cases} \lfloor \alpha \ell \rfloor + 1 & \text{(step-down)} \\ (\lfloor \alpha \ell \rfloor + 1) \vee (\ell - m + u) & \text{(step-up)} \end{cases}$$

This bound holds under general dependence, but does not incorporate it. Finally, the diminution principle has been recently followed by using much more sophisticated bounds that incorporate the pairwise dependence, see Theorems 3.7 and 3.8 in [Guo et al. \(2014\)](#).

- The “augmentation” principle [van der Laan et al. \(2004\)](#); [Farcomeni \(2009\)](#): consider the 1-FWE controlling procedure at level ζ rejecting the null hypotheses corresponding to the set $\mathcal{R}^{(1)} = \{1 \leq i \leq m : p_i \leq \tau_1(\zeta)\}$, denote $\ell^{(1)}$ the number of rejections of $\mathcal{R}^{(1)}$ and

$$\tilde{\ell}^{Aug} = \lfloor \ell^{(1)} / (1 - \alpha) \rfloor \wedge m.$$

Then the “augmented” procedure rejects the nulls associated to the $\tilde{\ell}^{Aug}$ smallest p -values. This procedure can incorporate the dependence if $\mathcal{R}^{(1)}$ is appropriately chosen.

- The “simultaneous” k -FWE control proposed in [Genovese and Wasserman \(2006\)](#): consider critical values $(\tau_\ell(\zeta/m))_{1 \leq \ell \leq m}$ (with ζ divided by m) and let

$$\tilde{\ell}^{sim} = \left\lfloor \frac{\max \{R(\tau_\ell(\zeta/m)) - \lfloor \alpha \ell \rfloor : \ell \leq R(\tau_\ell(\zeta/m)), \ell \geq 0\}}{1 - \alpha} \right\rfloor \wedge m.$$

Then the “simultaneous” procedure rejects the nulls corresponding to the $\tilde{\ell}^{sim}$ smallest p -values. Again, this procedure is able to incorporate the dependence if the τ_ℓ 's are suitably built.

3.3. Two new modifications. This section presents new results that can be seen as modifications of k -FWE based procedures that ensure finite sample FDP control. Both modifications incorporate the dependence between the p -values. Furthermore, the numerical experiments of Section 5 show that they are more powerful than the the state-of-the-art procedures described above in Section 3.2.

The first result follows the “diminution” principle by using a new bound on the probability exceedance. For any arbitrary critical values $(c_\ell(x))_{1 \leq \ell \leq m}$ (depending on a variable x), we let

$$C^{ex}(x) = \max_{1 \leq u \leq m} \left\{ \sum_{\ell=1}^{b_\alpha(u)} \max_{\substack{\theta \in \Theta \\ m_0(\theta)=u}} \{(\mathbb{P}_\theta(V_m(c_\ell(x)) \geq d(\ell-1, m, u)) - \mathbb{P}_\theta(V_m(c_{\ell-1}(x)) \geq d(\ell-1, m, u))) \right. \\ \left. \wedge (\mathbb{P}_\theta(V_m(c_\ell(x)) \geq d(\ell, m, u)) - \mathbb{P}_\theta(V_m(c_{\ell-1}(x)) \geq d(\ell, m, u))) \right\}, \quad (25)$$

where $b_\alpha(u)$ and $d(\ell, m, u)$ are given by (23) and (24), respectively. The following result is established in Section 7.2.

THEOREM 3.5. *Let us consider either the fixed model ($\Theta = \Theta^F$) or the uniform model ($\Theta = \Theta^U$) and any family of critical values $(c_\ell(x))_{1 \leq \ell \leq m}$, $x \geq 0$, such that $c_m(0) = 0$. Consider some $x^* \geq 0$ satisfying*

$$C^{ex}(x^*) \wedge C^{RS}(x^*) \leq \zeta,$$

where $C^{ex}(\cdot)$ is defined by (25) and $C^{RS}(\cdot)$ by (22). Let $\hat{\ell}$ be the number of rejections of the step-down (SD) (resp. step-up (SU)) algorithm associated to the critical values $(c_\ell(x^*))_{1 \leq \ell \leq m}$. Then the FDP control (5) holds (with $\hat{t}_m = \tau_{\hat{\ell}}$).

Theorem 3.5 can be applied with any starting critical values $(c_\ell(x))_{1 \leq \ell \leq m}$. A choice in accordance with RW's heuristic is

$$c_\ell(x) = x \tilde{\tau}_\ell, \quad \ell \in \{1, \dots, m\}, \quad x \geq 0, \quad (26)$$

where $(\tilde{\tau}_\ell)_{1 \leq \ell \leq m}$ are the adaptive k -FWE based critical values (8) for some appropriate bounding device. Next, while Theorem 3.5 does not require any assumption on the dependence, it implicitly assumes that the function $C^{ex}(\cdot)$ is known or easily computable. This is the case for instance in the equicorrelated case (ρ -equi) with $\rho \in [0, 1]$, because we have

$$C^{ex}(x) = \max_{1 \leq u \leq m} \left\{ \sum_{\ell=1}^{b_\alpha(u)} (B_m^0(c_\ell(x), d(\ell-1, m, u), u) - B_m^0(c_{\ell-1}(x), d(\ell-1, m, u), u)) \right. \\ \left. \wedge (B_m^0(c_\ell(x), d(\ell, m, u), u) - B_m^0(c_{\ell-1}(x), d(\ell, m, u), u)) \right\}, \quad (27)$$

where $B_m^0(t, k, u)$ is the exact bounding device defined by (17). A second illustration is the case where Γ is known but arbitrary and where the model is $\Theta = \Theta^U$. In that situation, $C^{ex}(x)$ in (25) can be approximated by Monte-Carlo calculations.

The second result presented in this section relies on the K -Markov device $B_m^0(t, k, u)$ given by (14) (for some integer $K \geq 1$) and specifically assumes equicorrelation (ρ -equi) with $\rho \in [0, 1)$. We consider the critical values defined as follows: for $\ell \in \{1, \dots, m\}$,

$$(28) \quad \tau_\ell^{new} = \begin{cases} \tilde{\tau}_\ell(\lambda\zeta, m(\ell)) & \text{if } \ell \geq \ell_K \\ \left(\frac{(1-\lambda)\zeta(\lfloor \alpha\ell \rfloor + 1)}{m(\ell)} \right) \wedge \tilde{\tau}_{\ell_K}(\lambda\zeta, m) & \text{if } \ell < \ell_K, \end{cases}$$

where $\ell_K = \lceil (K-1)/\alpha \rceil$, $\lambda \in [0, 1]$ is some tuning parameter and where $\tilde{\tau}_\ell(\zeta, u)$ denotes the value of t obtained by solving the equation

$$\int_0^1 (F_{0,\rho}(t, x))^K dx = \zeta \frac{(\lfloor \alpha\ell \rfloor + 1)(\lfloor \alpha\ell \rfloor + 1 - 1) \cdots (\lfloor \alpha\ell \rfloor + 1 - K + 1)}{u(u-1) \cdots (u-K+1)}$$

with $F_{0,\rho}(t, x)$ given by (15). The following result holds (see Section 7.4 for a proof).

THEOREM 3.6. *Consider the step-up procedure associated to the critical values $(\tau_\ell^{new})_{1 \leq \ell \leq m}$ given by (28) with rejection number $\hat{\ell}$. Then the finite sample FDP control (5) holds for $\hat{t}_m = \tau_{\hat{\ell}}^{new}$ in the equicorrelated case (ρ -equi) with $\rho \in [0, 1)$.*

When $K = 1$ and $\lambda = 1$, the critical values $(\tau_\ell^{new})_{1 \leq \ell \leq m}$ are the Lehmann-Romano critical values (12) and thus Theorem 3.6 is in accordance with Proposition 3.1 (ii) and Theorem 3.1 of Guo et al. (2014), because Simes' inequality is valid in that case. The originality of Theorem 3.6 lies in the case $K > 1$, that allows to incorporate ρ in a FDP controlling procedure. The experiments of Section 5 are made with the choice $K = 2$ and $\lambda = 0.5$ or 0.95 .

Finally, let us underline that the proof of Theorem 3.6 uses the MTP2 property (see, e.g., Karlin and Rinott (1980)) in a novel manner via the K -Markov device. It is thus of independent interest, see Section 7.4.

4. Asymptotic results. The goal of this section is to study RW's heuristic from an asymptotic point of view.

4.1. *Setting and assumptions.* In this section, the FDP control under study is asymptotic:

$$(29) \quad \limsup_m \{ \mathbb{P}_{\theta^{(m)}} (\text{FDP}_m(\hat{t}_m) > \alpha) \} \leq \zeta,$$

for some \hat{t}_m . This requires to consider a *sequence* of models $(\Theta^{(m)}, m \geq 1)$ and a sequence of parameters $(\theta^{(m)}, m \geq 1)$ with $\theta^{(m)} \in \Theta^{(m)}$ for all $m \geq 1$. The latter sequence is assumed to be fixed once for all throughout this section. Also, the family of models is composed of *fixed mixture* models ($H^{(m)}$ is not assumed to be random). Moreover, we will assume throughout this study the following common assumption:

$$(30) \quad m_0(\theta^{(m)})/m \rightarrow \pi_0, \text{ where } \pi_0 \in (0, 1).$$

Also, under (30), we let $\pi_1 = 1 - \pi_0$.

Useful assumptions on $(\theta^{(m)}, m \geq 1)$ are the following weak dependence assumptions on the processes $\widehat{\mathbb{G}}_m(t) = R_m(t)/m$ and $\widehat{\mathbb{G}}_{0,m}(t) = V_m(t)/m_0$, $t \in [0, 1]$:

$$\text{(weakdep)} \quad \|\widehat{\mathbb{G}}_m - G\|_\infty = o_P(1), \text{ for some } G : [0, 1] \rightarrow [0, 1];$$

$$\text{(weakdep0)} \quad \|\widehat{\mathbb{G}}_{0,m} - I\|_\infty = o_P(1), \text{ for } I(t) = t, t \in [0, 1].$$

In the Gaussian multivariate setting, [\(weakdep\)](#) and [\(weakdep0\)](#) are satisfied under [\(30\)](#) if the following conditions hold (see [Lemma A.1](#)):

$$\text{(Conv-alt)} \quad (m_1)^{-1} \sum_{i=1}^m H_i \delta_{\mu_i} \xrightarrow{\text{weak}} \nu, \text{ for a distribution } \nu \text{ on } \mathbb{R}^+ \text{ with } \nu(\{0\}) = 0,$$

$$\text{(weakdepGauss)} \quad m^{-2} \sum_{i,j=1}^m (\Gamma_{i,j})^2 \rightarrow 0.$$

Here, $G(t) = \pi_0 t + \pi_1 F_1(t)$ and $F_1(t) = \int_0^\infty \overline{\Phi}(\overline{\Phi}^{-1}(t) - \beta) d\nu(\beta)$. Condition [\(Conv-alt\)](#) extends the usual situation where the p -value distribution is fixed with m under the alternative (as assumed in [Genovese and Wasserman \(2004\)](#); [Neuivial \(2008\)](#); [Delattre and Roquain \(2011\)](#); [Delattre and Roquain \(2012\)](#)). Assumption [\(weakdepGauss\)](#) seems to have been introduced in [Fan et al. \(2012\)](#), see also [Schwartzman and Lin \(2011\)](#); [Delattre and Roquain \(2012\)](#). Finally, let us underline that under [\(\$\rho\$ -equi\)](#), Assumption [\(weakdepGauss\)](#) is satisfied whenever $\rho = \rho_m \rightarrow 0$.

4.2. The BH procedure and FDP control under weak dependence. Let us go back to [Figure 1](#). Clearly, the BH procedure is inappropriate to control the FDP for $\zeta = 0.05$ in the equi-correlated case $\rho = 0.1$. However, when $\rho = 0$, even if the BH procedure is only intended to control the expectation of the FDP at level α , the 95% quantile of the FDP is still close to α . This comes from the concentration of the FDP of the BH procedure around $\pi_0 \alpha < \alpha = 0.2$ as m grows to infinity. It is well known that this quantile converges to $\pi_0 \alpha$ as m grows to infinity, so that the limit in [\(29\)](#) is equal to zero, see, e.g., [Neuivial \(2008\)](#). In other words, the FDP concentration combined with the slight amount of conservativeness due to $\pi_0 < 1$ “prevents” the FDP to exceed α . The consequence is simple: the BH procedure controls the FDP asymptotically in the sense of [\(29\)](#) under independence. As a matter of fact, the latter also holds under weak dependence (see [Section 8.1](#) for a proof):

LEMMA 4.1. *Consider the BH procedure, that is, the step-up procedure associated to the linear critical values $\tau_\ell = \alpha \ell / m$, $\ell \in \{1, \dots, m\}$, and denote its rejection number by $\widehat{\ell}$. Assume that $(\theta^{(m)}, m \geq 1)$ satisfies [\(30\)](#), [\(weakdep\)](#), [\(weakdep0\)](#) and further assume that G satisfies the following property:*

$$\text{(Exists)} \quad \text{there exists } t \in (0, 1), \text{ such that } G(t) > t/\alpha.$$

Then we have $\mathbb{P}_{\theta^{(m)}}(\text{FDP}_m(\tau_{\widehat{\ell}}) > \alpha) \rightarrow 0$. In particular, [\(29\)](#) always holds.

Note that all the assumptions of [Theorem 4.1](#) are always satisfied under [\(30\)](#), [\(Conv-alt\)](#) and [\(weakdepGauss\)](#).

In the literature, even under independence, it is common to exclude the BH procedure while studying [\(29\)](#). For instance, [Proposition 4.1](#) in [Chi and Tan \(2008b\)](#) shows that the

“oracle” version of the BH procedure, that is, the step-up procedure with critical values $\alpha\ell/m_0$, $\ell \in \{1, \dots, m\}$, has a FDP exceeding α with a probability tending to $1/2$. Since the oracle BH procedure is often considered as “better” than the original BH procedure, it is thus tempting to exclude the BH procedure when studying an FDP control of type (29). Lemma 4.1 shows that, perhaps surprisingly, this is a mistake: BH procedure is interesting for providing (29) and does not suffer from the same drawback than the oracle BH procedure.

However, as Figure 1 suggests, the BH procedure can fail to control the FDP as m is tending to infinity if the dependence is not weak. Under (ρ -equi) and when the p -values under the alternative are zero (Dirac uniform configuration), this fact has been established formally in Finner et al. (2007), by showing that the limit of the FDP of the BH procedure is not deterministic anymore and hence can exceed α with a positive probability, which is obviously not related to ζ (because BH critical values does not depend on ζ).

4.3. *Asymptotic modification of RW’s heuristic.* First, let us mention that RW’s heuristic controls the FDP under weak dependence: we have proved in a former version of this manuscript that a step-up procedure with k -FWE based critical values also satisfies (29) in a weak dependence framework slightly more constrained than the one of Section 4.1. Roughly speaking, the essence of the argumentation is as follows: when $\hat{\ell}/m$ converges in probability to some deterministic quantities, then the fluctuations of $\hat{\ell}$ asymptotically disappear in (6), and thus this probability is bounded by ζ from (Bound). Let us note that a previous occurrence of this reasoning can be found at the end of Section 7 in Genovese and Wasserman (2006).

However, according to Lemma 4.1, to guarantee an asymptotic control of the FDP under weak dependence, one could simply apply the BH procedure, which is much simpler and more powerful (at least when ζ is small). Hence, to this respect, the “weak dependent” case is not so crucial and we do not report here the corresponding study; we refer the reader to the first version of this manuscript for more details on this issue.

Instead, we consider the asymptotical framework under assumption (ρ -equi). By using the decomposition (16), we easily check that the exact device (17) is such that, for $k = k_m$ with $k_m/m \rightarrow c$,

$$\bar{B}_m(t, k) = \mathbb{P}_U(N_m/m \geq k/m) \rightarrow \mathbb{P}(F_{0,\rho}(t, U) \geq c),$$

where $F_{0,\rho}(t, x)$ is defined by (15), where $U \sim U(0, 1)$ and N_m follows a binomial distribution of parameters m and $F_{0,\rho}(t, U)$ (conditionally on U). Making the latter probability equal to c entails $c = F_{0,\rho}(t, \zeta)$, which, by replacing respectively c by $\alpha\ell/m$ and t by τ_ℓ leads to the new critical values (2) mentioned in the introduction of the paper. We now state the main result of this section, to be proved in Section 8.2.

THEOREM 4.2. *Assume that $\theta^{(m)}$ satisfies (30), assume (ρ -equi) with an arbitrary $\rho = \rho_m \in [0, 1]$, possibly depending on m , but bounded away from 1 asymptotically. Assume (Conv-alt) with a constant alternative β . Consider the step-up procedure associated to the critical values τ_ℓ , $\ell = 1, \dots, m$ defined by (2) with rejection number $\hat{\ell}$. Then we have the asymptotic FDP control (29) with $\hat{t} = \tau_{\hat{\ell}}$.*

In practice, ρ is often unknown and we can legitimately ask whether it is possible to incorporate an estimate $\hat{\rho}_m$ of ρ in the critical values (2). The next theorem gives a positive answer, provided that the estimation rate is faster than logarithmic.

THEOREM 4.3. *In the setting of Theorem 4.2, consider a random variable $\hat{\rho}_m$ such that*

$$(31) \quad (\log m)(\hat{\rho}_m - \rho_m)^2 = o_P(1).$$

Consider the critical values $\hat{\tau}_\ell$, $\ell = 1, \dots, m$ defined by (2) where ρ has been replaced by $\hat{\rho}_m$. Then the corresponding step-up procedure (with rejection number $\hat{\ell}$) satisfies (29) with $\hat{t} = \hat{\tau}_{\hat{\ell}}$.

Finally, let us now provide an example of estimator satisfying (31) in the equicorrelated setting of Theorem 4.2. For this, we make the light assumption that X is obtained as an average between two i.i.d. variables: $X = (X^{(1)} + X^{(2)})/\sqrt{2}$, with $X_i^{(j)} = \beta H_i/\sqrt{2} + \rho_m^{1/2} \xi_0^{(j)} + (1 - \rho_m)^{1/2} \xi_i^{(j)}$, where $\xi_i^{(j)}$, $0 \leq i \leq m$, $1 \leq j \leq 2$, are all i.i.d. $\mathcal{N}(0, 1)$. Letting $Z = (X^{(2)} - X^{(1)})/\sqrt{2}$ and considering the empirical variance $\hat{\sigma}_{m,Z}^2 = \frac{1}{m} \sum_{i=1}^m (Z_i - \bar{Z})^2$ of Z . We have classically that $m\hat{\sigma}_{m,Z}^2/(1 - \rho_m) \sim \chi^2(m - 1)$. Thus $\hat{\rho}_m = 1 - \hat{\sigma}_{m,Z}^2$ satisfies (31).

5. Numerical experiments.

5.1. *Studying finite sample FDP control for the exact bounding device.* This section complements Fact 3.4 by considering the FDP control for the step-down procedure associated to the *adaptive* (non-oracle) k -FWE based critical values coming from the exact bounding device (17). It took us some efforts to identify a parameter configuration for which the FDP control is violated. As a matter of fact, m should be quite large to make the probability exceeds ζ , hence, to that respect, the exact calculations of Section 3 are not usable anymore. We thus evaluate the probability that the FDP exceeds α with extensive Monte-Carlo simulations (10^6 replicates), performed in the one-sided equi-correlated Gaussian multivariate framework (ρ -*equi*) defined in Section 1.3, for alternative means μ_i all equal to some β . Figure 4 reports a situation where the FDP control is violated (admittedly not by much). This is a numerical support to the fact that, in general, modifications should be used to provide a finite sample FDP control via RW's heuristic.

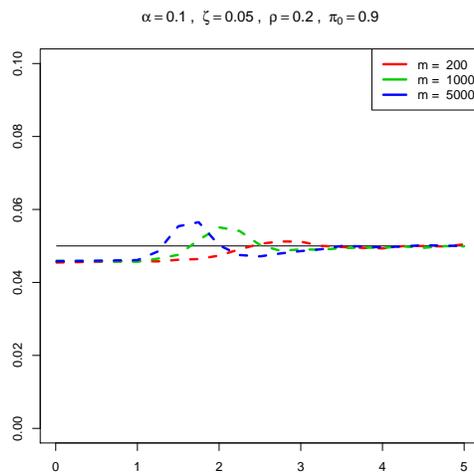


FIG 4. $\mathbb{P}(\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha)$ in function of β for the adaptive step-down procedure using the exact bounding device in the equi-correlated case.

Procedures not using the value of ρ	Procedures incorporating the value of ρ
[LR]: Lehmann Romano procedure (12)	[AugEx] augmentation with exact bounding device
[AugBonf]: augmentation with $\tau_1(\zeta) = \zeta/m$	[SimEx] simultaneous k -FWE with exact bounding device
[SimLR]: simultaneous k -FWE with (12)	[DimExEx] new diminution with (25)
[DimMarkovLR]: diminution with (22)	[DimGuoLR] diminution proposed in Theorem 3.8 of Guo et al. (2014)
[Bonf]: the raw Bonferroni procedure	[Split1/2] new procedure (28) with $\lambda = 1/2$ and $K = 2$
	[Split0.95] new procedure (28) with $\lambda = 0.95$ and $K = 2$
	[RWExact] non-modified k -FWE based with exact bounding device

TABLE 1

Procedures used in Figures 5 and 6, see Sections 3.2 and 3.3 for more details. All the procedures are step-up and adaptive (non oracle).

5.2. *Power comparison of the finite sample modifications.* This section evaluates the power of the procedures considered in Section 3 with a proven FDP control. The power is evaluated by using the standard false non-discovery rate (FNR), defined as the expected ratio of errors among the accepted null hypotheses. Table 1 summarizes the procedures that have been considered. The simulation setting is the same as in the previous section.

Figure 5 displays the power of procedures that do not incorporate the value of ρ (left column in Table 1). Note that [LR] controls the FDP because Simes' inequality is valid here, see Proposition 3.1 (ii). Hence, it does not use the true value of ρ but uses nevertheless an assumption on the dependence structure. This is not the case of [AugBonf], [SimLR], [DimMarkovLR] and [Bonf] that control the FDP for any dependence structure. As one can expect, [LR] essentially dominates the other procedures. Also, while [AugBonf] comes in second position, [SimLR] is even worst than [Bonf] so should be avoided here.

Now, while incorporating the value of ρ , we will loosely say that a procedure is “admissible” if it performs better than [LR] at least for “a reasonable amount” of parameter configurations. Figure 6 displays the power of procedures incorporating the value of ρ (except [LR] and [Bonf] that we have added only for comparison), see the right column of Table 1. Note that, except [RWExact], all the procedures have a proven FDP control, so that the power comparison is fair. First, [DimGuoLR] is “not admissible”, which indicates that the interest of the bounds found in Guo et al. (2014) are mainly theoretical in our setting. Second, [AugEx] only improves [LR] in a very small region, which shows that, as one can expect, providing 1-FWE control for controlling the FDP is too conservative in general. As for [SimEx], things are more balanced: when $\rho = 0$, it improves [LR] when many rejections are possible (π_0 not large or β large) but does worst otherwise. We think that this is due to the nature of the [LR] critical values, which are design to perform well when only few nulls are expected to be rejected. When ρ is larger, however, [SimEx] quickly deteriorates. An explanation is that the simultaneity in [SimEx] is obtained via an union bound, which is conservative when the dependence is strong. Finally, our new procedures [Split1/2], [Split0.95] and [DimExEx] seem to be all “admissible” in these simulations and substantially outperform the other procedures. Also, none of the three procedures uniformly dominates the others. For instance, taking $\lambda = 1/2$ rather than $\lambda = 0.95$ is better when less rejections are expected, but worst otherwise, while [DimExEx] seems often better than [Split1/2]. In conclusion, the three new proposed procedures have interesting power properties. Finally note that among these three procedures, [DimExEx] has the additional interest of offering FDP control outside the equi-correlated regime, see Theorem 3.5.

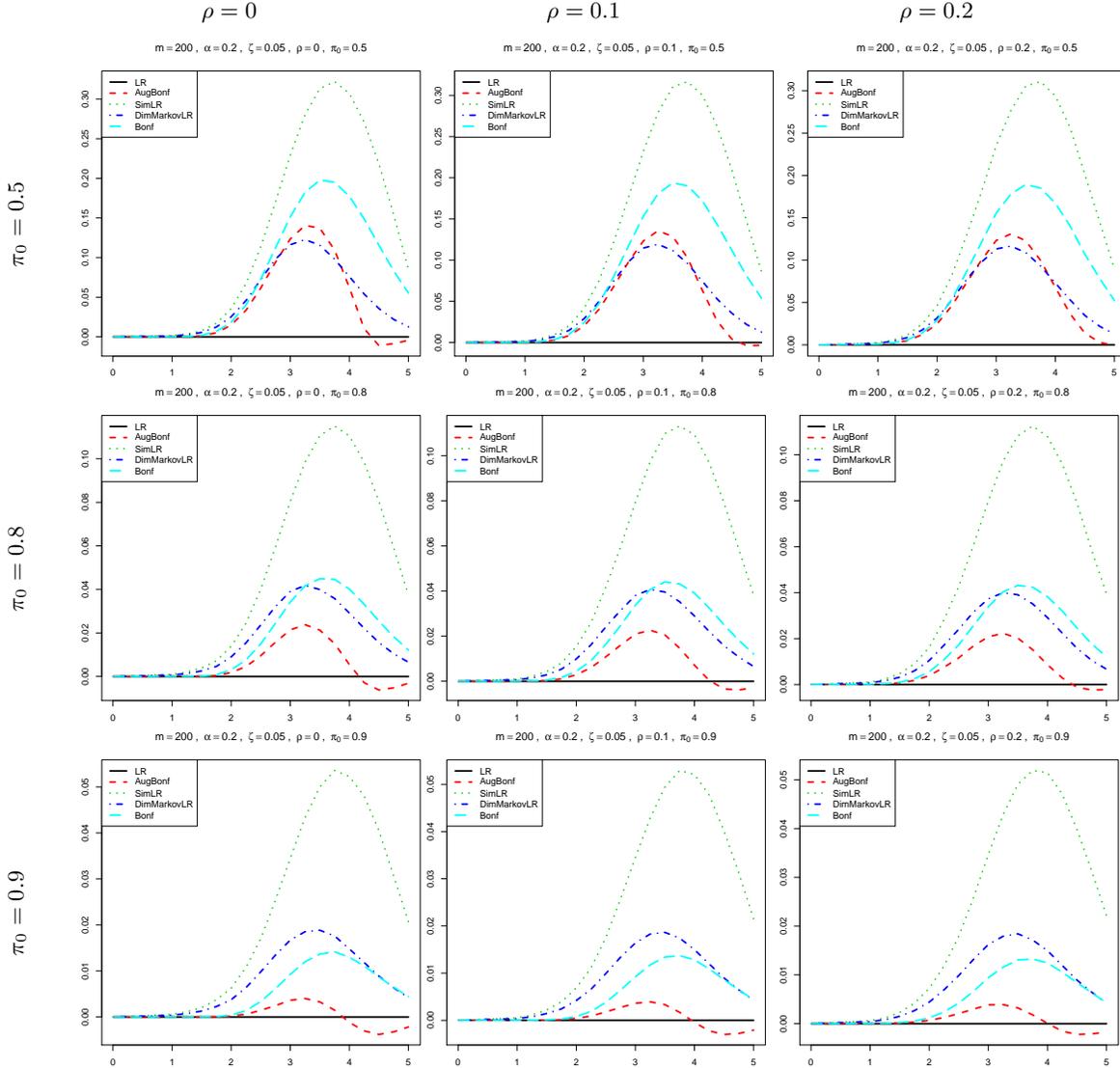


FIG 5. Relative FNR to Lehmann Romano procedure in function of β , see text and Table 1. Procedures not using the value of ρ .

6. Conclusion and discussion. This paper investigated the FDP control in the case where the dependence is partly/fully incorporated, by using an extension of RW's heuristic. We provided two new approaches that offer finite sample control: the first one (Theorem 3.5) followed the diminution principle and can be used as soon as the joint distribution of the null p -values can be computed. The second one (Theorem 3.6) offered a finite sample control under equicorrelation and interestingly relied on the K -Markov device. Next, an important part of our work concerned the asymptotic FDP control: we first noticed that the weak dependent case seems to be of marginal interest, because the simple BH procedure can be used in that case. Then, still based on RW's heuristic, we proposed new critical values that provide asymptotic control under equicorrelation (Theorems 4.2 and 4.3). A strong point of the latter is the capacity to plug an estimator of ρ in the critical values. Hence, it can undertake the difficult task of incorporating unknown dependence while controlling the FDP (asymptotically).

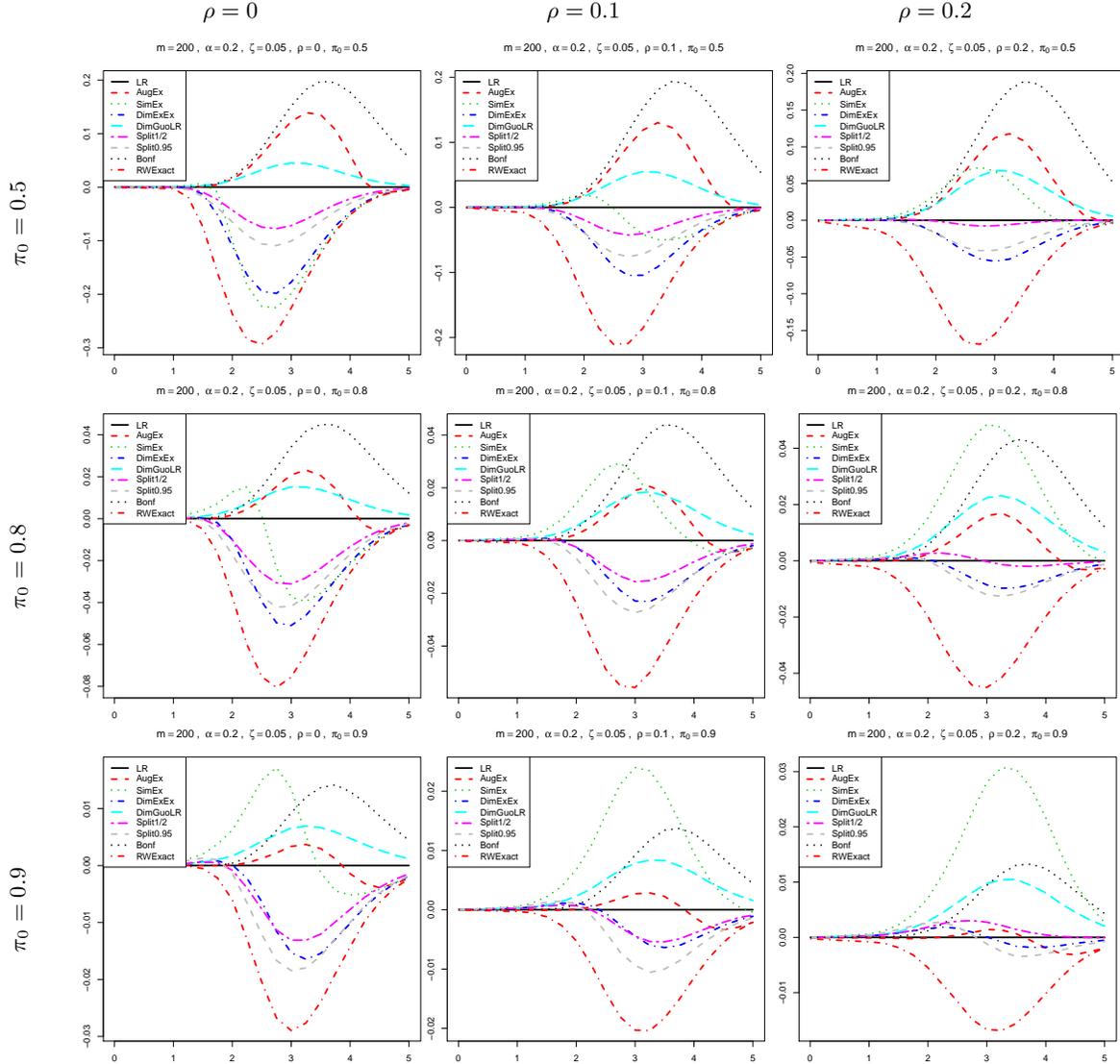


FIG 6. Relative FNR to Lehmann Romano procedure in function of β , see text and Table 1. Procedures using the value of ρ (except [LR] and [Bonf]).

Our leading example is related to one-sided testing, so we can legitimately ask whether our results can be extended to two-sided testing, that is, when $p_i = 2\bar{\Phi}(|X_i|)$ (by using the notation of Section 1.3). The bounding device calculations done in Section 2.3 can be clearly generalized to that case by replacing $F_{0,\rho}(t, x)$ (see (15)) by

$$F_{0,\rho}^{(2)}(t, x) = F_{0,\rho}(t/2, x) + F_{0,\rho}(t/2, 1 - x).$$

Hence, we can define new critical values coming from the corresponding exact bounding device and combine it with the diminution principle presented in Theorem 3.5. However, the other results of the paper cannot be directly generalized to the two-sided case: the proof of Theorem 3.6 uses a positive dependence assumption that is not valid in the two-sided case (see, e.g., Benjamini and Yekutieli (2001)). As for Theorems 4.2 and 4.3, the proofs rely on the fact that for all $t, x \mapsto F_{0,\rho}(t, x)$ is non-increasing w.r.t. x , which is a property not maintained

for $x \mapsto F_{0,\rho}^{(2)}(t, x)$.

While this paper solved some issues, it opened several directions of research. For instance, is the asymptotic FDP control of Theorem 4.2 still true when using the original critical values of RW's method rather than their asymptotic counterpart (2)? We believe that this issue intrinsically relies on the ‘‘Poisson’’ asymptotic regime, which was (essentially) not considered here in our asymptotic FDP controlling results. Finally, a crucial (but probably very challenging) issue is the validity of RW's approach in the case of permutation tests with an arbitrary and unknown dependence structure.

7. Proofs for finite sample results.

7.1. An unifying bound.

PROPOSITION 7.1. *For any critical values $(\tau_\ell)_{1 \leq \ell \leq m}$, consider either the corresponding step-down procedure (SD) or the corresponding step-up procedure (SU), with rejection number $\widehat{\ell}$. Then the following holds, both in the fixed model $(\Theta = \Theta^F)$ and the uniform model $(\Theta = \Theta^U)$: for all $\theta \in \Theta$,*

$$(32) \quad \mathbb{P}_\theta (\text{FDP}_m(\tau_{\widehat{\ell}}) > \alpha) \leq \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{P}_\theta \left(V_m(\tau_\ell) \geq d(\ell, m, m_0), \widetilde{\ell} = \ell \right),$$

where $b_\alpha(m_0)$, $d(\ell, m, m_0)$ are defined in (23) and (24) respectively, and $\widetilde{\ell}$ is taken as follows:

- (i) *Step-down case: $\widetilde{\ell} = \widehat{\ell}^{(1)}$, where $\widehat{\ell}^{(1)} = \min\{\ell \in \{1, \dots, m\} : S_m(\tau_\ell) < (1 - \alpha)\ell\}$ (with the convention $\min \emptyset = m + 1$) and by denoting $S_m(t) = R_m(t) - V_m(t)$ the number of true discoveries at threshold t .*
- (ii) *Step-up case: $\widetilde{\ell} = \widehat{\ell}$.*

Moreover, in the step-up case, (32) is an equality.

Proposition 7.1 (i) is a reformulation of Theorem 5.2 in Roquain (2011) in our framework and is based on ideas presented in the proofs of Lehmann and Romano (2005); Romano and Wolf (2007). Proposition 7.1 (ii) is essentially based on Romano and Shaikh (2006b) and we provide a short proof below.

PROOF. Since $\text{FDP}_m(\tau_{\widehat{\ell}}) > \alpha$ implies $\lfloor \alpha \widehat{\ell} \rfloor + 1 \leq m_0$, we have $\widehat{\ell} \leq b_\alpha(m_0)$. Also, $\widehat{\ell} = R_m(\tau_{\widehat{\ell}}) \leq m_1 + V_m(\tau_{\widehat{\ell}})$, which implies $V_m(\tau_{\widehat{\ell}}) \geq \widehat{\ell} - m_1$. This implies (32) in case (ii). \square

7.2. A new bound.

PROPOSITION 7.2. *In the setting of Proposition 7.1, assume moreover that there exists a family of random variables $(Z_{\ell,\ell'})_{1 \leq \ell, \ell' \leq m}$ satisfying: for all ℓ, ℓ' ,*

$$(33) \quad \mathbf{1} \{V_m(\tau_\ell) \geq d(\ell', m, m_0)\} \leq Z_{\ell,\ell'} \text{ a.s.}$$

and, a.s., $Z_{\ell,\ell'}$ is nondecreasing in ℓ and nonincreasing in ℓ' . Then for all $\theta \in \Theta$,

$$(34) \quad \mathbb{P}_\theta (\text{FDP}_m(\tau_{\widehat{\ell}}) > \alpha) \leq \sum_{\ell=1}^{b_\alpha(m_0)} (\mathbb{E}_\theta (Z_{\ell,\ell-1}) - \mathbb{E}_\theta (Z_{\ell-1,\ell-1})) \wedge (\mathbb{E}_\theta (Z_{\ell,\ell}) - \mathbb{E}_\theta (Z_{\ell-1,\ell})),$$

by letting $Z_{0,\ell'} = 0$ and $Z_{\ell,0} = 1$ for $\ell' \geq 0, \ell \geq 1$.

Applied with $Z_{\ell,\ell'} = V_m(\tau_\ell)/d(\ell', m, m_0)$, Proposition 7.2 establishes the Romano-Shaikh bound (22). Applied with $Z_{\ell,\ell'} = \mathbf{1}\{V_m(\tau_\ell) \geq d(\ell', m, m_0)\}$, Proposition 7.2 entails Theorem 3.5.

PROOF. From (32), we derive

$$\begin{aligned} \mathbb{P}_\theta(\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha) &\leq \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{P}_\theta\left(V_m(\tau_\ell) \geq d(\ell, m, m_0), \tilde{\ell} = \ell\right) \\ &\leq \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{E}_\theta\left(Z_{\ell,\ell} \mathbf{1}\{\tilde{\ell} = \ell\}\right). \end{aligned}$$

Now, the RHS of the previous display is equal to

$$\begin{aligned} &\sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{E}_\theta\left(Z_{\ell,\ell} \mathbf{1}\{\tilde{\ell} \geq \ell\}\right) - \sum_{\ell=1}^{b_\alpha(m_0)-1} \mathbb{E}_\theta\left(Z_{\ell,\ell} \mathbf{1}\{\tilde{\ell} \geq \ell+1\}\right) \\ &= \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{E}_\theta\left(Z_{\ell,\ell} \mathbf{1}\{\tilde{\ell} \geq \ell\}\right) - \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{E}_\theta\left(Z_{\ell-1,\ell-1} \mathbf{1}\{\tilde{\ell} \geq \ell\}\right) \\ &= \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{E}_\theta\left((Z_{\ell,\ell} - Z_{\ell-1,\ell-1}) \mathbf{1}\{\tilde{\ell} \geq \ell\}\right) \\ &\leq \sum_{\ell=1}^{b_\alpha(m_0)} \mathbb{E}_\theta(Z_{\ell,\ell-1} - Z_{\ell-1,\ell-1}) \wedge \mathbb{E}_\theta(Z_{\ell,\ell} - Z_{\ell-1,\ell}), \end{aligned}$$

which proves the result. \square

7.3. *Proof of Proposition 3.2.* Let $\hat{k} = V_m(\tau_{\hat{\ell}})$ and note that $\hat{k} \leq m_0$ and $\{\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha\} = \{\hat{k} \geq \lfloor \alpha \hat{\ell} \rfloor + 1\}$. First, in the non-adaptive case, we have by definition of \bar{B}_m , for all t and $k \leq m_0$,

$$\bar{B}_m(t, k) = \sup_{0 \leq u \leq m} \{B_m^0(t, k, u)\} \geq B_m^0(t, k, m_0).$$

Hence, we have by definition of the (non-adaptive) critical values,

$$\zeta \geq \bar{B}_m(\tau_{\hat{\ell}}, \lfloor \alpha \hat{\ell} \rfloor + 1) \geq B_m^0(\tau_{\hat{\ell}}, \lfloor \alpha \hat{\ell} \rfloor + 1, m_0),$$

which is larger than or equal to $B_m^0(\tau_{\hat{\ell}}, \hat{k}, m_0)$ whenever $\hat{k} \geq \lfloor \alpha \hat{\ell} \rfloor + 1$. Hence, we obtain

$$\{\text{FDP}_m(\tau_{\hat{\ell}}) > \alpha\} \subset \left\{B_m^0(\tau_{\hat{\ell}}, \hat{k}, m_0) \leq \zeta, \hat{k} \geq 1\right\} \subset \left\{\tau_{\hat{\ell}} \leq \nu_k^0, \hat{k} \geq 1\right\},$$

and thus (20) holds. Second, in the adaptive case, we use that $m_0 \leq m - R_m(\tau_{\hat{\ell}}) + V_m(\tau_{\hat{\ell}}) = m - \hat{\ell} + \hat{k}$. Thus, whenever $\hat{k} \geq \lfloor \alpha \hat{\ell} \rfloor + 1$, we have for all t ,

$$\begin{aligned} \zeta &\geq \bar{B}_m(\tau_{\hat{\ell}}, \lfloor \alpha \hat{\ell} \rfloor + 1, \hat{\ell}) = \sup_{\lfloor \alpha \hat{\ell} \rfloor + 1 \leq k' \leq \hat{\ell}} \left\{ \sup_{0 \leq u \leq m - \hat{\ell} + k'} B_m^0(\tau_{\hat{\ell}}, k', u) \right\} \\ (35) \quad &\geq \sup_{0 \leq u \leq m - \hat{\ell} + \hat{k}} B_m^0(\tau_{\hat{\ell}}, \hat{k}, u) \geq B_m^0(\tau_{\hat{\ell}}, \hat{k}, m_0). \end{aligned}$$

Hence, this implies $\tau_{\hat{\ell}} \leq \nu_k^0$ and $\hat{k} \geq 1$ and the proof is finished.

7.4. *Proof of Theorem 3.6.* First observe that the critical values (28) can be obtained by modifying the K -Markov Bounding device $B_m^0(t, k, u)$ defined by (14) as follows :

$$(B_m^0)'(t, k, u) = \begin{cases} B_m^0(t, k, u)/\lambda & \text{if } k \geq K \\ \frac{ut}{(1-\lambda)^k} \vee (B_m^0(t, K, m)/\lambda) & \text{if } k < K, \end{cases}$$

(the second bounding value being infinite when $\lambda = 1$). Note that the associated adaptive bounding device (**Bound-adapt**) is equal to $(B_m^0)'(t, k, m - \ell + k)$ and thus gives rise to the adaptive critical values (28). By using Proposition 3.2 and by letting $\widehat{k} = V_m(\tau_{\widehat{\ell}}^{new})$, we get

$$\mathbb{P}_{\theta} \left(\text{FDP}_m(\tau_{\widehat{\ell}}^{new}) > \alpha \right) = \sum_{k=1}^{m_0} \mathbb{P}_{\theta} \left(V_m(\nu_k^0) \geq k, \widehat{k} = k \right),$$

where ν_k^0 is such that $(B_m^0)'(\nu_k^0, k, m_0) \leq \zeta$ for all k . It follows that $\mathbb{P}_{\theta} \left(\text{FDP}_m(\tau_{\widehat{\ell}}^{new}) > \alpha \right)$ is smaller than or equal to $T_1 + T_2$, where we let

$$\begin{aligned} T_1 &= \sum_{k=K}^{m_0} \mathbb{P}_{\theta} \left(V_m(\nu_k^0) \geq k, \widehat{k} = k \right); \\ T_2 &= \sum_{k=1}^{(K-1) \wedge m_0} \mathbb{P}_{\theta} \left(V_m(\nu_k^0) \geq k, \widehat{k} = k \right), \end{aligned}$$

with by convention $T_1 = 0$ when $K > m_0$. For the first term, we note that for $k \geq K$,

$$\mathbf{1} \{ V_m(\nu_k^0) \geq k \} \leq \frac{1}{\binom{K}{k}} \sum_{X \subset \mathcal{H}_0: |X|=K} \mathbf{1} \left\{ \max_{i \in X} \{q_i\} \leq \nu_k^0 \right\}.$$

Then, since \widehat{k} is permutation invariant (as a function of the p -values) and since (ρ -**equi**) holds, we obtain

$$\begin{aligned} T_1 &\leq \sum_{k=K}^{m_0} \frac{m_0(m_0-1) \cdots (m_0-K+1)}{k(k-1) \cdots (k-K+1)} \mathbb{P}_{\theta} \left(\widehat{k} = k, \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) \\ &\leq \lambda \zeta \sum_{k=K}^{m_0} \mathbb{P}_{\theta} \left(\widehat{k} = k \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) \\ &\leq \lambda \zeta \sum_{k=K}^{m_0} \mathbb{P}_{\theta} \left(\widehat{k} \leq k \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) - \mathbb{P}_{\theta} \left(\widehat{k} \leq k-1 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right), \end{aligned}$$

where q_1, \dots, q_{m_0} denotes the p -values under the null, i.e., the p -values of the set $\{p_i, i \in \mathcal{H}_0\}$. Now, under assumption (ρ -**equi**) with $\rho \geq 0$, since the inverse of the covariance matrix has only nonpositive off-diagonal entries, the p -value family $(p_i)_{1 \leq i \leq m}$ is MTP2, see [Karlin and Rinott \(1980\)](#). Then, since the set $\{p \in [0, 1]^m : \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0\}$ is stable w.r.t. the lattice operations \wedge and \vee , the p -value family $(p_i)_{1 \leq i \leq m}$ is also MTP2 conditionally on the event $\{\max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0\}$. Therefore, the p -value family $(p_i)_{1 \leq i \leq m}$ is associated conditionally on $\{\max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0\}$. Since the function $p \in [0, 1]^m \mapsto \mathbf{1} \{ \widehat{k} \leq k-1 \}$ is nondecreasing and

the function $p \in [0, 1]^m \mapsto \mathbf{1} \{ \max_{1 \leq i \leq K} \{q_i\} \leq \nu_{k-1}^0 \}$ is nonincreasing, we thus have

$$\begin{aligned} & \mathbb{P}_\theta \left(\widehat{k} \leq k-1, \max_{1 \leq i \leq K} \{q_i\} \leq \nu_{k-1}^0 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) \\ & \leq \mathbb{P}_\theta \left(\widehat{k} \leq k-1 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) \times \mathbb{P}_\theta \left(\max_{1 \leq i \leq K} \{q_i\} \leq \nu_{k-1}^0 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right), \end{aligned}$$

which entails

$$\mathbb{P}_\theta \left(\widehat{k} \leq k-1 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) \geq \mathbb{P}_\theta \left(\widehat{k} \leq k-1 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_{k-1}^0 \right).$$

This shows

$$T_1 \leq \lambda \zeta \sum_{k=K}^{m_0} \mathbb{P}_\theta \left(\widehat{k} \leq k \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_k^0 \right) - \mathbb{P}_\theta \left(\widehat{k} \leq k-1 \mid \max_{1 \leq i \leq K} \{q_i\} \leq \nu_{k-1}^0 \right) \leq \lambda \zeta,$$

by using a telescopic argument.

Now, for T_2 , we use $\mathbf{1} \{V_m(\nu_k^0) \geq k\} \leq \frac{1}{k} \sum_{i=1}^{m_0} \mathbf{1} \{q_i \leq \nu_k^0\}$ and $m_0 \nu_k^0 \leq (1-\lambda)\zeta k$ for $k < K$,

$$\begin{aligned} T_2 & \leq \sum_{i=1}^{m_0} \sum_{k=1}^{(K-1) \wedge m_0} \frac{1}{k} \mathbb{P}_\theta \left(\widehat{k} = k, q_i \leq \nu_k^0 \right) \\ & \leq (1-\lambda)\zeta m_0^{-1} \sum_{i=1}^{m_0} \sum_{k=1}^{(K-1) \wedge m_0} \mathbb{P}_\theta \left(\widehat{k} = k \mid q_i \leq \nu_k^0 \right) \\ & \leq (1-\lambda)\zeta m_0^{-1} \sum_{i=1}^{m_0} \sum_{k=1}^{(K-1) \wedge m_0} \mathbb{P}_\theta \left(\widehat{k} \leq k \mid q_i \leq \nu_k^0 \right) - \mathbb{P}_\theta \left(\widehat{k} \leq k-1 \mid q_i \leq \nu_k^0 \right) \\ & \leq (1-\lambda)\zeta m_0^{-1} \sum_{i=1}^{m_0} \sum_{k=1}^{(K-1) \wedge m_0} \mathbb{P}_\theta \left(\widehat{k} \leq k \mid q_i \leq \nu_k^0 \right) - \mathbb{P}_\theta \left(\widehat{k} \leq k-1 \mid q_i \leq \nu_{k-1}^0 \right) \\ & \leq (1-\lambda)\zeta, \end{aligned}$$

by using again the MTP2 property. This finishes the proof

8. Proofs for asymptotic results.

8.1. *Proof of Lemma 4.1.* Actually, we prove the result for a more general class of procedures, where $\widehat{t} = \tau_{\widehat{\ell}}$ is obtained by (43) for a sequence of functions $f_m = \widehat{f}_m$ (possibly random) which is uniformly close to $f_\infty(t) = t/\alpha$ on every compact of $(0, \alpha]$, that is,

$$(36) \quad \sup_{b \leq t \leq \alpha} |\widehat{f}_m(t) - t/\alpha| \rightarrow 0, \quad \text{a.s.} \quad \text{for all } b \in (0, \alpha).$$

Note that $\widehat{f}_m = f_\infty$ gives the BH procedure by Lemma A.2. Next, since $R_m(\widehat{t}) \geq m\widehat{f}_m(\widehat{t})$,

$$\begin{aligned} \mathbb{P}(\text{FDP}_m(\widehat{t}) > \alpha) & \leq \mathbb{P} \left(V_m(\widehat{t})/m > \alpha \widehat{f}_m(\widehat{t}) \right) \\ & = \mathbb{P} \left((m_0/m)(\widehat{\mathbb{G}}_{0,m}(\widehat{t}) - \widehat{t}) - \alpha(\widehat{f}_m(\widehat{t}) - \widehat{t}/\alpha) > (1 - m_0/m)\widehat{t} \right) \\ & \leq \mathbb{P} \left((m_0/m) \|\widehat{\mathbb{G}}_{0,m} - I\|_\infty + \alpha \sup_{t^* \leq t \leq \alpha} |\widehat{f}_m(t) - t/\alpha| > (1 - m_0/m)t^* \right) \\ & \quad + \mathbb{P}(\widehat{t} \leq t^*), \end{aligned}$$

for some $t^* > 0$ satisfying $G(t^*) > t^*/\alpha$ (which exists by (Exists)). By (30) and (weakdep0), it is sufficient to check that $\mathbb{P}(\widehat{t} \leq t^*)$ tends to zero. For this, we use (weakdep) that ensures

$$\begin{aligned} \mathbb{P}(\widehat{t} > t^*) &\geq \mathbb{P}(\widehat{\mathbb{G}}_m(t^*) > \widehat{f}_m(t^*)) \\ &\geq \mathbb{P}(G(t^*) > f_\infty(t^*) + |G(t^*) - \widehat{\mathbb{G}}_m(t^*)| + |\widehat{f}_m(t^*) - f_\infty(t^*)|) \rightarrow 1, \end{aligned}$$

which finishes the proof.

8.2. *Proof of Theorems 4.2 and 4.3.* First observe that Theorem 4.3 is more general than Theorem 4.2, simply by taking $\widehat{\rho}_m = \rho_m$ (which of course satisfies (31)). Hence, we focus in the sequel only on the proof of Theorem 4.3.

By Lemma A.2, $\widehat{t} = \widehat{\tau}_{\widehat{\rho}_m}$ is given by (43) with \widehat{f}_m defined by (42). We merely check that

$$(37) \quad \widehat{f}_m(t) = F_{0,\widehat{\rho}_m}(t, \zeta)/\alpha,$$

where $F_{0,\widehat{\rho}_m}$ is given by (15). Next, up to consider a subsequence, we can assume that the sequence ρ_m converge to some limit $\rho^* \in [0, 1)$ and that (31) holds almost surely.

First, when $\rho^* = 0$, since $F_{0,\widehat{\rho}_m}(t, \zeta)$ converges uniformly to t on any compact of $(0, 1)$ and by using the result of Section 8.1, we obtain that $\mathbb{P}(\text{FDP}_m(\widehat{t}) > \alpha)$ converges to zero. Now assume $\rho^* \in (0, 1)$. Still using a subsequence argument in combination with the Skorokhod representation theorem, we can assume that $(\widehat{t}, U) \in [0, 1]^2$ is almost surely converging to some (T, U) (on appropriate subspaces). We aim now at identifying the values of U for which $T > 0$ a.s. From (37), observe that

$$(38) \quad F_{0,\widehat{\rho}_m}(\widehat{t}, \zeta) = \max\{t' \in [0, 1] \mid \widehat{\mathbb{G}}'_m(t') \geq t'/\alpha\},$$

by writing $X_i = \beta H_i + \rho_m^{1/2} \overline{\Phi}^{-1}(U) + (1 - \rho_m)^{1/2} \xi_i$, for $U \sim U(0, 1)$ independent of ξ_i , $1 \leq i \leq m$, i.i.d. $\mathcal{N}(0, 1)$ and where

$$(39) \quad \widehat{\mathbb{G}}'_m(t') = m^{-1} \sum_{i=1}^m \mathbf{1}\{F_{0,\widehat{\rho}_m}(p_i, \zeta) \leq t'\} \geq m^{-1} \sum_{i=1}^m H_i \mathbf{1}\{\overline{\Phi}(\xi_i) \leq \widehat{H}_m(t')\},$$

by letting

$$\widehat{H}_m(t') = \overline{\Phi} \left(\frac{\overline{\Phi}^{-1}(t')(1 - \widehat{\rho}_m)^{1/2} + (\widehat{\rho}_m)^{1/2} \overline{\Phi}^{-1}(\zeta) - \rho_m^{1/2} \overline{\Phi}^{-1}(U) - \beta}{(1 - \rho_m)^{1/2}} \right).$$

Since the latter converges a.s. to

$$H(t') = \overline{\Phi} \left(\overline{\Phi}^{-1}(t') - \mu(U) \right), \quad \mu(U) = (1 - \rho^*)^{-1/2} \left(\beta + (\rho^*)^{1/2} \left(\overline{\Phi}^{-1}(U) - \overline{\Phi}^{-1}(\zeta) \right) \right),$$

the RHS of (39) converges a.s. to $\pi_1 H(t')$. Now, we should consider two cases:

- Case $\mu(U) > 0$: the slope of H is infinite in 0. From (38), for a large m , we have $F_{0,\widehat{\rho}_m}(\widehat{t}, \zeta) > t'_0$ where t'_0 is equal to any $t' \in (0, 1)$ satisfying $H(t') > t'/\alpha$. This entails that \widehat{t} is almost surely asymptotically bounded away from 0, and $T > 0$ a.s. Moreover, when $H_i = 0$, the assertion $p_i \leq t$ is equivalent to $\overline{\Phi}(\xi_i) \leq F_{0,\rho_m}(t, U)$. This entails that for any compact $K \subset (0, 1)$,

$$\sup_{t \in K} |\widehat{\mathbb{G}}_{0,m}(t) - F_{0,\rho^*}(t, U)| \leq \sup_{t \in K} |\widehat{\mathbb{G}}_{0,m}(t) - F_{0,\rho_m}(t, U)| + \sup_{t \in K} |F_{0,\rho_m}(t, U) - F_{0,\rho^*}(t, U)|,$$

which tends a.s. to zero by the Glivenko-Cantelli theorem and because ρ_m tends to ρ^* . As a consequence, when $\mu(U) > 0$, a.s., as m grows to infinity,

$$(40) \quad \text{FDP}(\hat{t}) = \frac{m_0}{m} \frac{\widehat{\mathbb{G}}_{0,m}(\hat{t})}{\widehat{\mathbb{G}}_m(\hat{t})} \leq \frac{m_0}{m} \frac{\widehat{\mathbb{G}}_{0,m}(\hat{t})}{\widehat{f}_m(\hat{t})} \rightarrow \pi_0 \alpha \frac{F_{0,\rho^*}(T, U)}{F_{0,\rho^*}(T, \zeta)}.$$

- Case $\mu(U) < 0$: in that case, we show that the probability that the procedure makes at least one rejection is tending to zero. For this, let $\varepsilon > 0$, and consider $t_\varepsilon \in (0, 1)$ such that for all $t' \in (0, t_\varepsilon]$, we have $\overline{\Phi}(\overline{\Phi}^{-1}(t') - \mu(U)/2) \leq \varepsilon t'$. Let

$$\widehat{\mu}_m = \overline{\Phi}^{-1}(\alpha/m) \left| \left(\frac{1 - \widehat{\rho}_m}{1 - \rho_m} \right)^{1/2} - 1 \right| + \frac{\beta + \rho_m^{1/2} \overline{\Phi}^{-1}(U) - (\widehat{\rho}_m)^{1/2} \overline{\Phi}^{-1}(\zeta)}{(1 - \rho_m)^{1/2}}.$$

By (31), $\widehat{\mu}_m$ converges a.s. to $\mu(U)$ and thus is smaller than $\mu(U)/2$ for m large enough (a.s.). As a consequence, a.s., for all $t' \geq \alpha/m$, for a large m , $\widehat{H}_m(t') \leq \overline{\Phi}(\overline{\Phi}^{-1}(t') - \widehat{\mu}_m) \leq \overline{\Phi}(\overline{\Phi}^{-1}(t) - \mu(U)/2)$. Then, by denoting $q_i = \overline{\Phi}(\xi_i)$, so that the q_i 's are all i.i.d. uniform, we have

$$\begin{aligned} \mathbb{P}(\widehat{\ell} \geq 1) &\leq \mathbb{P}\left(\exists \ell \in \{1, \dots, m\} : q_{(\ell)} \leq \widehat{H}_m(\alpha \ell / m)\right) \\ &\leq \mathbb{P}\left(\exists \ell \in \{1, \dots, m\} : q_{(\ell)} \leq \overline{\Phi}(\overline{\Phi}^{-1}(\alpha \ell / m) - \mu(U)/2)\right) + o(1) \\ &\leq \mathbb{P}\left(\exists \ell \in \{1, \dots, m\}, \alpha \ell / m \leq t_\varepsilon : q_{(\ell)} \leq \varepsilon \alpha \ell / m\right) \\ &\quad + \mathbb{P}\left(\exists \ell \in \{1, \dots, m\}, \alpha \ell / m > t_\varepsilon : q_{(\ell)} \leq \alpha \ell / m\right) + o(1) \\ &\leq \varepsilon \alpha + \mathbb{P}\left(\exists t \in (t_\varepsilon, \alpha] : \widehat{\mathbb{F}}(t) \geq t/\alpha\right) + o(1) \\ &\leq \varepsilon \alpha + \mathbb{P}\left(\sup_{t \in [0, 1]} |\widehat{\mathbb{F}}(t) - t| \geq (1/\alpha - 1)t_\varepsilon\right) + o(1) \end{aligned}$$

by using Simes' inequality (see, e.g., (19)) and by denoting $\widehat{\mathbb{F}}$ the e.c.d.f. of the q_i 's. By taking the lim sup in m and then by making ε tends to zero we get that $\mathbb{P}(\widehat{\ell} \geq 1)$ tends to zero. As a consequence, when $\mu(U) < 0$, we have a.s. (up to consider a subsequence),

$$(41) \quad \text{FDP}(\hat{t}) \leq \mathbf{1}\{\widehat{\ell} \geq 1\} \rightarrow 0.$$

Finally, combining (40) and (41) yields

$$\begin{aligned} \limsup_m \mathbb{P}(\text{FDP}_m(\hat{t}) > \alpha) &\leq \limsup_m \mathbb{P}(\text{FDP}_m(\hat{t}) > \alpha, \mu(U) > 0) \\ &\quad + \limsup_m \mathbb{P}(\text{FDP}_m(\hat{t}) > \alpha, \mu(U) < 0) \\ &= \mathbb{P}\left(\pi_0 \alpha \frac{F_{0,\rho^*}(T, U)}{F_{0,\rho^*}(T, \zeta)} > \alpha, \mu(U) > 0\right) \\ &\leq \mathbb{P}(F_{0,\rho^*}(T, U) > F_{0,\rho^*}(T, \zeta)) = \mathbb{P}(U < \zeta) = \zeta, \end{aligned}$$

because $U \sim U(0, 1)$.

APPENDIX A: AUXILIARY RESULTS

LEMMA A.1. *In the Gaussian setting, conditions (30), (Conv-alt) and (weakdepGauss) imply (weakdep).*

PROOF. From (weakdepGauss) and (30), Proposition 2.1 in Delattre and Roquain (2012) (for instance) ensures that, for all t , $V_m(t)/m$ converges in probability to $\pi_0 t$. Similarly, let $p'_i = \bar{\Phi}(\bar{\Phi}^{-1}(p_i) - \mu_i) = \bar{\Phi}(Y_i)$ for i such that $H_i = 1$. We have that for all t , $m^{-1} \sum_{i=1}^m H_i \mathbf{1}\{p'_i \leq t\}$ converges to $\pi_1 t$. Let us now consider the case of $S_m(t)/m = m^{-1} \sum_{i=1}^m H_i \mathbf{1}\{p'_i \leq \bar{\Phi}(\bar{\Phi}^{-1}(t) - \mu_i)\}$. For this, we define the uniform random index $I \sim m_1^{-1} \sum_{i=1}^m H_i \delta_i$. By (Conv-alt), we have $\mu_I \rightsquigarrow \nu$. Moreover, we merely check that, up to consider a subsequence, a.s., the joint convergence $(p'_I, \mu_I) \rightsquigarrow U(0, 1) \otimes \nu$ holds. Hence, $S_m(t)/m$ is converging to $\pi_1 F_1(t)$. \square

LEMMA A.2. *Let us consider a step-up procedure with critical values $(\tau_\ell)_{1 \leq \ell \leq m}$ and a rejection threshold $\tau_{\hat{\ell}}$, where $\hat{\ell}$ is defined by (SU). Consider the function f_m defined by*

$$(42) \quad f_m(t) = m^{-1} \times \min\{\ell \in \{0, \dots, m+1\} : \tau_\ell \geq t\},$$

with the conventions $\tau_0 = 0$, $\tau_{m+1} = 1$. Let \hat{t} be defined by

$$(43) \quad \hat{t} = \sup\{t \in [0, 1] : \hat{\mathbb{G}}_m(t) \geq f_m(t)\}.$$

Then the supremum into (43) is a maximum, that is, $\hat{\mathbb{G}}_m(\hat{t}) \geq f_m(\hat{t})$. Furthermore, $\hat{t} = \tau_{\hat{\ell}}$.

PROOF. The first point holds because f_m is left-continuous. Next, since $f_m(t) \leq \ell/m$ is equivalent to $\tau_\ell \geq t$, we have

$$\tau_\ell = \max\{t \in [0, 1] : f_m(t) \leq \ell/m\}.$$

Hence, $\hat{\mathbb{G}}_m(\tau_{\hat{\ell}}) \geq \hat{\ell}/m \geq f_m(\tau_{\hat{\ell}})$ and $\tau_{\hat{\ell}} \leq \hat{t}$. Conversely, let us prove $\hat{t} \leq \tau_{\hat{\ell}}$. Since $f_m(\hat{t}) \leq 1$, we have $\hat{t} \leq \tau_m$. We can thus define

$$\tilde{k} = \min\{k \in \{0, \dots, m\} : \hat{t} \leq \tau_k\}.$$

By definition of $\hat{\ell}$, the result is shown if we prove $\hat{\mathbb{G}}_m(\tau_{\tilde{k}}) \geq \tilde{k}/m$. Assume $\tilde{k} \geq 1$ (otherwise the result is trivial). By definition of \tilde{k} , $\tau_{\tilde{k}-1} < \hat{t}$, which in turn implies $f_m(\hat{t}) > (\tilde{k}-1)/m$ (by definition of f_m). This gives $\hat{\mathbb{G}}_m(\hat{t}) \geq f_m(\hat{t}) > (\tilde{k}-1)/m$. Hence, $\hat{\mathbb{G}}_m(\tau_{\tilde{k}}) \geq \hat{\mathbb{G}}_m(\hat{t}) \geq \tilde{k}/m$, which concludes the proof. \square

ACKNOWLEDGMENTS

We acknowledge an associate editor and three referees for their relevant suggestions that have substantially improved the quality of the manuscript. We are very grateful to Yoav Benjamini, Gilles Blanchard and Pierre Neuvial for helpful discussions. We also thank Bertrand Michel and Jacques Portes for their support during the numerical experiments. The second author was supported by the French Agence Nationale de la Recherche (ANR grant references: ANR-09-JCJC-0027-01, ANR-PARCIMONIE, ANR-09-JCJC-0101-01).

REFERENCES

- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57(1):289–300.
- Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.
- Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.*, 29(4):1165–1188.
- Blanchard, G., Dickhaus, T., Roquain, E., and Villers, F. (2014). On least favorable configurations for step-up-down tests. *Statist. Sinica*, 24(1):1–23.
- Blanchard, G. and Roquain, E. (2009). Adaptive false discovery rate control under independence and dependence. *J. Mach. Learn. Res.*, 10:2837–2871.
- Chi, Z. and Tan, Z. (2008a). Positive false discovery proportions: intrinsic bounds and adaptive control. *Statist. Sinica*, 18(3):837–860.
- Chi, Z. and Tan, Z. (2008b). Positive false discovery proportions: intrinsic bounds and adaptive control. *Statist. Sinica*, 18(3):837–860.
- Delattre, S. and Roquain, E. (2011). On the false discovery proportion convergence under Gaussian equi-correlation. *Statist. Probab. Lett.*, 81(1):111–115.
- Delattre, S. and Roquain, E. (2012). On empirical distribution function of high-dimensional Gaussian vector components with an application to multiple testing. *ArXiv e-prints*.
- Dudoit, S. and van der Laan, M. J. (2008). *Multiple testing procedures with applications to genomics*. Springer Series in Statistics. Springer, New York.
- Efron, B., Tibshirani, R., Storey, J. D., and Tusher, V. (2001). Empirical Bayes analysis of a microarray experiment. *J. Amer. Statist. Assoc.*, 96(456):1151–1160.
- Fan, J., Han, X., and Gu, W. (2012). Estimating false discovery proportion under arbitrary covariance dependence. *Journal of the American Statistical Association*, 107(499):1019–1035.
- Farcomeni, A. (2009). Generalized augmentation to control the false discovery exceedance in multiple testing. *Scand. J. Stat.*, 36(3):501–517.
- Finner, H., Dickhaus, T., and Roters, M. (2007). Dependency and false discovery rate: asymptotics. *Ann. Statist.*, 35(4):1432–1455.
- Finner, H., Dickhaus, T., and Roters, M. (2009). On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Statist.*, 37(2):596–618.
- Gavrilov, Y., Benjamini, Y., and Sarkar, S. K. (2009). An adaptive step-down procedure with proven FDR control under independence. *Ann. Statist.*, 37(2):619–629.
- Genovese, C. and Wasserman, L. (2004). A stochastic process approach to false discovery control. *Ann. Statist.*, 32(3):1035–1061.
- Genovese, C. R. and Wasserman, L. (2006). Exceedance control of the false discovery proportion. *J. Amer. Statist. Assoc.*, 101(476):1408–1417.
- Guo, W., He, L., and Sarkar, S. K. (2014). Further results on controlling the false discovery proportion. *The Annals of Statistics*, 42(3):1070–1101.
- Guo, W. and Romano, J. (2007). A generalized Sidak-Holm procedure and control of generalized error rates under independence. *Stat. Appl. Genet. Mol. Biol.*, 6:Art. 3, 35 pp. (electronic).
- Hommel, G. and Hoffman, T. (1988). Controlled uncertainty. In *Multiple Hypotheses Testing*, pages 154–161. Springer, Heidelberg.
- Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.*, 10(4):467–498.
- Kim, K. I. and van de Wiel, M. (2008). Effects of dependence in high-dimensional multiple testing problems. *BMC Bioinformatics*, 9(1):114.
- Korn, E. L., Troendle, J. F., McShane, L. M., and Simon, R. (2004). Controlling the number of false discoveries: application to high-dimensional genomic data. *J. Statist. Plann. Inference*, 124(2):379–398.
- Lehmann, E. L. and Romano, J. P. (2005). Generalizations of the familywise error rate. *Ann. Statist.*, 33:1138–1154.
- Neuival, P. (2008). Asymptotic properties of false discovery rate controlling procedures under independence. *Electron. J. Stat.*, 2:1065–1110.
- Perone Pacifico, M., Genovese, C., Verdinelli, I., and Wasserman, L. (2004). False discovery control for random fields. *J. Amer. Statist. Assoc.*, 99(468):1002–1014.
- Romano, J. P. and Shaikh, A. M. (2006a). On stepdown control of the false discovery proportion. volume 49 of *Lecture Notes-Monograph Series*, pages 33–50.

- Romano, J. P. and Shaikh, A. M. (2006b). Stepup procedures for control of generalizations of the familywise error rate. *Ann. Statist.*, 34(4):1850–1873.
- Romano, J. P. and Wolf, M. (2007). Control of generalized error rates in multiple testing. *Ann. Statist.*, 35(4):1378–1408.
- Roquain, E. (2011). Type I error rate control for testing many hypotheses: a survey with proofs. *J. Soc. Fr. Stat.*, 152(2):3–38.
- Roquain, E. and Villers, F. (2011). Exact calculations for false discovery proportion with application to least favorable configurations. *Ann. Statist.*, 39(1):584–612.
- Sarkar, S. K. (2008). On methods controlling the false discovery rate. *Sankhya, Ser. A*, 70:135–168.
- Schwartzman, A. and Lin, X. (2011). The effect of correlation in false discovery rate estimation. *Biometrika*, 98(1):199–214.
- Seeger, P. (1968). A note on a method for the analysis of significances en masse. *Technometrics*, 10(3):586–593.
- van der Laan, M. J., Dudoit, S., and Pollard, K. S. (2004). Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives. *Stat. Appl. Genet. Mol. Biol.*, 3:Art. 15, 27 pp. (electronic).

UNIVERSITÉ PARIS DIDEROT, LPMA,
E-MAIL: sylvain.delattre@univ-paris-diderot.fr

UPMC UNIVERSITÉ PARIS 6, LPMA,
E-MAIL: etienne.roquain@upmc.fr