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

By Sylvain Delattre and Etienne Roquain

This supplement presents additional materials for the paper Delattre and Roquain (2013).

S-1. *p*-value correction under Gaussian equi-correlation. Consider the one-sided location model satisfying the particular condition (Gauss-ρ-equi) with $\rho \geq 0$. Remember that X_i can be realized via

$$X_i = \mu_i + \rho^{1/2} W + (1-\rho)^{1/2} \xi_i,$$

where W, $\xi_1, \dots, \xi_m$ are i.i.d. $\mathcal{N}(0,1)$. Hence, a simple transformation that removes W is to consider

$$X_i^\star = (X_i - \overline{X}) \left\{m/(m-1)\right\}^{1/2} (1-\rho)^{-1/2} = \mu_i' + \xi_i'.$$

where $\mu_i' = (\mu_i - \overline{\mu})\{m/(m-1)\}^{1/2}(1-\rho)^{-1/2}$ and $\xi_i' = (\xi_i - \overline{\xi})\{m/(m-1)\}^{1/2}$. Although this operation adds a bias in the signal μ_i', it considerably "improves" the noise, because the multivariate vector ξ' is closed to be independent (it satisfies (Gauss-ρ-equi) with $\rho = -(m-1)^{-1}$). While using (almost) i.i.d. test statistics $X_i^\star$ seems overall more desirable than the original strongly test statistics X_i, we can compare the performance of the test statistics in the context of our work via the power of the resulting FDP controlling procedure.

Assume that the aim is the asymptotic FDP control:

$$\limsup_m \mathbb{P}(\mathrm{FDP} > \alpha) \leq \zeta, \tag{S-1}$$

for some pre-specified (small) ζ. First, we can easily check that the BH procedure using the $X_i^\star$ (say, pro1) controls the FDP asymptotically, that is, satisfies (S-1) for any $\zeta > 0$ (a reasoning similar to that of Lemma 4.1 can be applied). Second, the new procedure derived from RW's heuristic and using the original p-values (say, pro2) also provides (S-1) (see Theorem 4.5). In Figure S-1, we display the difference of power between pro2 and pro1. Here, the power is simply defined here as the number of correct rejections. As we can see, the conclusions strongly depend on the parameter configurations π_0, ζ, ρ and β: for a large ρ, the factor W has a major influence, thus pro1, which removes W, does better. By contrast, when ρ is close to 0, no correction is needed at all and pro1, which has a bias, does worst. Furthermore, when π_0 close to 1 (sparse situation), the bias disappears and pro1 should be preferred. Of course, decreasing ζ makes pro2 more conservative while it let pro1 unchanged.

Overall, this experiment re-inforces the interest in correcting the p-values under Gaussian equi-correlation, at least in some situations. However, to prevent this procedure from a possible loss, it seems desirable to correct the signal bias into $X_i^\star$ (as, e.g., in Fan et al. (2012)). However, the corresponding FDP control is to our knowledge not theoretically established. Finally, while this type of corrections seems impossible to apply in great generality (e.g., in (facmod) when $\mu_i H_i$ has a distribution close to the one of $c_i W$), identifying assumptions under which corrected p-values can be used to theoretically control the FDP is a research direction of primary interest.

imsart-aos ver. 2014/07/30 file: DR2012a2-revised3_supp.tex date: December 19, 2014

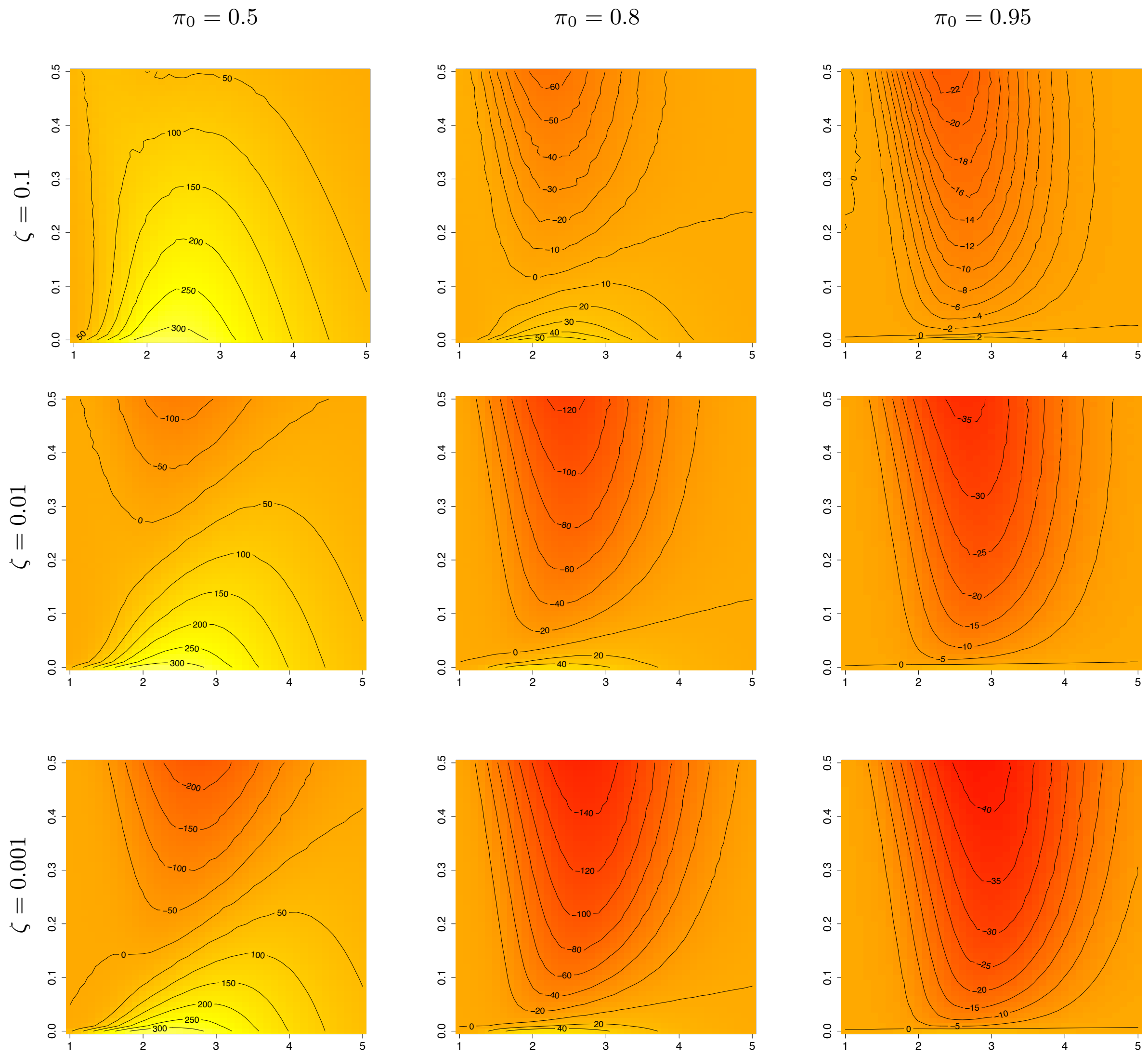

FIG S-1. *Difference of power between pro2 and pro1, see text. Heatmap in function of β (X-axis) and ρ (Y-axis). $m = 1000$; $\alpha = 0.2$; 10^4 simulations.*

S-2. Incorporating an estimator of ρ in (asymptotic) RW's method.

THEOREM S-2.1. *Consider the one-sided testing problem* (3) *with all alternative means equal to some $\beta > 0$ and assume that $\theta^{(m)}$ satisfies* (29). *Assume* (Gauss-ρ-equi) *with an arbitrary $\rho = \rho_m \in [0, 1]$, possibly depending on m, but bounded away from 1 asymptotically. Consider the step-up procedure with rejection number $\widehat{\ell}$ associated to the critical values τ_ℓ, $\ell = 1, \dots, m$ defined by* (2) *where $\rho = \rho_m$ is replaced by any random variable $\widehat{\rho}_m$ such that* (34) *holds. Then we have the asymptotic FDP control* (28) *with $\widehat{t} = \tau_{\widehat{\ell}}$.*

The interest of Theorem S-2.1 is that, when ρ is unknown, it can be suitably replaced by any estimator, provided that the estimation rate is faster than logarithmic. For instance, if we make the (fair) assumption that X is obtained as an average between two i.i.d. variables: $X = (X^{(1)} + X^{(2)})/\sqrt{2}$, with $X_i^{(j)} = \mu_i H_i/\sqrt{2} + {\rho_m}^{1/2} W^{(j)} + (1 - \rho_m)^{1/2} \xi_i^{(j)}$, where $W^{(j)}, \xi_i^{(j)}$,

$1 \leq i \leq m$, $1 \leq j \leq 2$, are all i.i.d. $\mathcal{N}(0,1)$. Letting $Z = (X^{(2)} - X^{(1)})/\sqrt{2}$ and considering the empirical variance $\widehat{\sigma}^2_{m,Z} = \frac{1}{m}\sum_{i=1}^m (Z_i - \overline{Z})^2$ of Z. Since $m\widehat{\sigma}^2_{m,Z}/(1-\rho_m) \sim \chi^2(m-1)$, we have that $\widehat{\rho}_m = 1 - \widehat{\sigma}^2_{m,Z}$ satisfies (34).

Proof. We follow here the proof scheme of Theorem 4.5. First, let us denote by $F_{0,\rho}$ the function F_0 defined by (F_0-Gauss-ρ-equi) and denote $q_\zeta = \Phi^{-1}(\zeta)$. Also, we write $X_i = \beta H_i + {\rho_m}^{1/2}\, W + (1-\rho_m)^{1/2}\, \xi_i$, for W, ξ_i, $1 \leq i \leq m$, i.i.d. $\mathcal{N}(0,1)$.

By Lemma 8.1, $\widehat{t} = \widehat{\tau}_{\widehat{\ell}}$ is given by (40) with $\widehat{f}_m$ defined by (39). We merely check that

$$\widehat{f}_m(t) = F_{0,\hat{\rho}_m}(t, q_\zeta)/\alpha. \tag{S-2}$$

Next, up to consider a subsequence, we can assume that the sequence ρ_m converge to some limit $\rho^\star \in [0,1)$ and that (34) holds almost surely.

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

$$F_{0,\hat{\rho}_m}(\widehat{t}, q_\zeta) = \max\{t' \in [0,1] \mid \widehat{\mathbb{G}}'_m(t') \geq t'/\alpha\}, \tag{S-3}$$

where

$$\widehat{\mathbb{G}}'_m(t') = m^{-1}\sum_{i=1}^m \mathbf{1}\left\{F_{0,\hat{\rho}_m}(p_i, q_\zeta) \leq t'\right\} \geq m^{-1}\sum_{i=1}^m H_i\, \mathbf{1}\left\{\Phi(\xi_i) \leq \widehat{H}_m(t')\right\}, \tag{S-4}$$

by letting

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

Since the latter converges a.s. to

$$H(t') = \Phi\left(\Phi^{-1}(t') - \mu(W)\right),\ \ \mu(W) = (1-\rho^\star)^{-1/2}\left(\beta + (\rho^\star)^{1/2}\,(W - q_\zeta)\right),$$

the RHS of (S-4) converges a.s. to $\pi_1 H(t')$. Now, we consider two cases:

- Case $\mu(W) > 0$: the slope of H is infinite in 0. From (S-3), for a large m, we have $F_{0,\hat{\rho}_m}(\widehat{t}, q_\zeta) > t'_0$ where t'_0 denotes 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 $\Phi(\xi_i) \leq F_{0,\rho_m}(t, W)$. This entails that for any compact $K \subset (0,1)$,

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

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

  $$\mathrm{FDP}(\widehat{t}) = \frac{m_0}{m}\frac{\widehat{\mathbb{G}}_{0,m}(\widehat{t})}{\widehat{\mathbb{G}}_m(\widehat{t})} \leq \frac{m_0}{m}\frac{\widehat{\mathbb{G}}_{0,m}(\widehat{t})}{\widehat{f}_m(\widehat{t})} \to \pi_0\alpha\frac{F_{0,\rho^\star}(T,W)}{F_{0,\rho^\star}(T,q_\zeta)}. \tag{S-5}$$

- Case $\mu(W) < 0$: 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 $\Phi(\Phi^{-1}(t') - \mu(W)/2) \leq \varepsilon\, t'$. Let

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

  By (34), $\widehat{\mu}_m$ converges a.s. to $\mu(W)$ and thus is smaller than $\mu(W)/2$ for m large enough (a.s.). As a consequence, a.s., for all $t' \geq \alpha/m$, for a large m, $\widehat{H}_m(t') \leq \Phi(\Phi^{-1}(t') - \widehat{\mu}_m) \leq \Phi(\Phi^{-1}(t) - \mu(W)/2)$. Then, by denoting $q_i = \Phi(\xi_i)$, so that the q_i's are all i.i.d. uniform, we have

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

  by using Simes' inequality and by denoting $\widehat{\mathbb{F}}$ the e.c.d.f. of the q_i's. By taking the $\limsup$ 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 $\mu(W) < 0$, we have a.s. (up to consider a subsequence),

$$\text{FDP}(\widehat{t}) \leq \mathbf{1}\left\{ \widehat{\ell} \geq 1 \right\} \to 0. \tag{S-6}$$

Finally, combining (S-5) and (S-6) yields

$$\begin{aligned}
\limsup_m \mathbb{P}\left(\text{FDP}_m(\widehat{t}) > \alpha \right) &\leq \limsup_m \mathbb{P}\left(\text{FDP}_m(\widehat{t}) > \alpha, \mu(W) > 0 \right) \\
&\quad + \limsup_m \mathbb{P}\left(\text{FDP}_m(\widehat{t}) > \alpha, \mu(W) < 0 \right) \\
&= \mathbb{P}\left(\pi_0 \alpha \frac{F_{0,\rho^\star}(T, W)}{F_{0,\rho^\star}(T, q_\zeta)} > \alpha, \mu(W) > 0, T > 0 \right) \\
&\leq \mathbb{P}\left(F_{0,\rho^\star}(T, W) > F_{0,\rho^\star}(T, q_\zeta), T \in (0,1) \right) \\
&\leq \mathbb{P}\left(W \geq q_\zeta \right) = \zeta.
\end{aligned}$$

□

S-3. Auxiliary results.

Lemma S-3.1. *In the Gaussian setting, conditions (29), (Conv-alt) and (weakdepGauss) imply (weakdep).*

PROOF. From (weakdepGauss) and (29), Proposition 2.1 in Delattre and Roquain (2014) (for instance) ensures that, for all t, $V_m(t)/m$ converges in probability to $\pi_0 t$. Similarly, let $p'_i = \Phi(\Phi^{-1}(p_i) - \mu_i) = \Phi(Y_i)$ for i such that $H_i = 1$. We have that for all t, $m^{-1}\sum_{i=1}^m H_i \mathbf{1}\{p'_i \le t\}$ converges to $\pi_1 t$. Let us now consider the case of $S_m(t)/m = m^{-1}\sum_{i=1}^m H_i \mathbf{1}\{p'_i \le \Phi(\Phi^{-1}(t) - \mu_i)\}$. 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 $\pi_1 F_1(t)$. □

LEMMA S-3.2. *Let $\gamma \in (0,1)$ and V_m be a process valued in $D([0,1])$. Assume that for some compact $K \subset [0,1]$, the following convergence holds*

$$(V_m(t))_{t\in K} \rightsquigarrow (V(t))_{t\in K},$$

where $V \in C(K)$, a.s., and where for all $t \in K$, $V(t)$ as a continuous increasing c.d.f. Let

$$q_m(t) = \min\{x \; : \; \mathbb{P}(V_m(t) \le x) \ge \gamma\}.$$

Then the function sequence q_m converge uniformly on K to the function

$$q(t) = \min\{x \; : \; \mathbb{P}(V(t) \le x) \ge \gamma\}.$$

PROOF. Consider an arbitrary sequence $t_m \in K$ tending to $t \in K$ and show that $q_m(t_m)$ tends to $q(t)$. Up to consider a subsequence assume that $q_m(t_m)$ is tending to some $q^\star$. Then we have

$$\gamma \le \lim_m \mathbb{P}(V_m(t_m) \le q_m(t_m)) = \mathbb{P}(V(t) \le q^\star),$$

because $V_m(t_m) = \{V_m(t_m) - V_m(t)\} + V_m(t) \xrightarrow{P} V(t)$ and the c.d.f. of $V(t)$ is continuous. Similarly, for all $\varepsilon > 0$,

$$\gamma \ge \lim_m \mathbb{P}(V_m(t_m) \le q_m(t_m) - \varepsilon) = \mathbb{P}(V(t) \le q^\star - \varepsilon).$$

By making ε decreases to zero, we get $\mathbb{P}(V(t) \le q^\star) = \gamma$ which entails $q^\star = q(t)$ because the e.d.f. of $V(t)$ is one to one.

□

S-4. Critical values under non-positive dependence. We provide in Figure S-2 an analogue of Figure 3 but in the alternate equi-correlated model (alt-ρ-equi); this means that the bounding devices of Section 2 are computed with F_0 given by (F_0-alt-ρ-equi). We have chosen $a \in \{0.5, 0.75\}$, $\rho \in \{0.1, 0.2, 0.5\}$ ($\rho = 0$ is already treated in Figure 3).

S-5. Additional numerical experiments. Figures S-3, S-4, S-5 and S-6 display the probability exceedance/the power of the procedures of Table 1 that incorporate the known dependence (except [LR] and [Bonf]). The models considered are (Gauss-ρ-equi) and (alt-ρ-equi) for $a = 0.5$. The alternative means μ_i are all equal to some β. The parameters are $m = 200$; $\alpha = 0.2$; $\zeta = 0.05$; and the quantities are evaluated with 3×10^5 simulations.

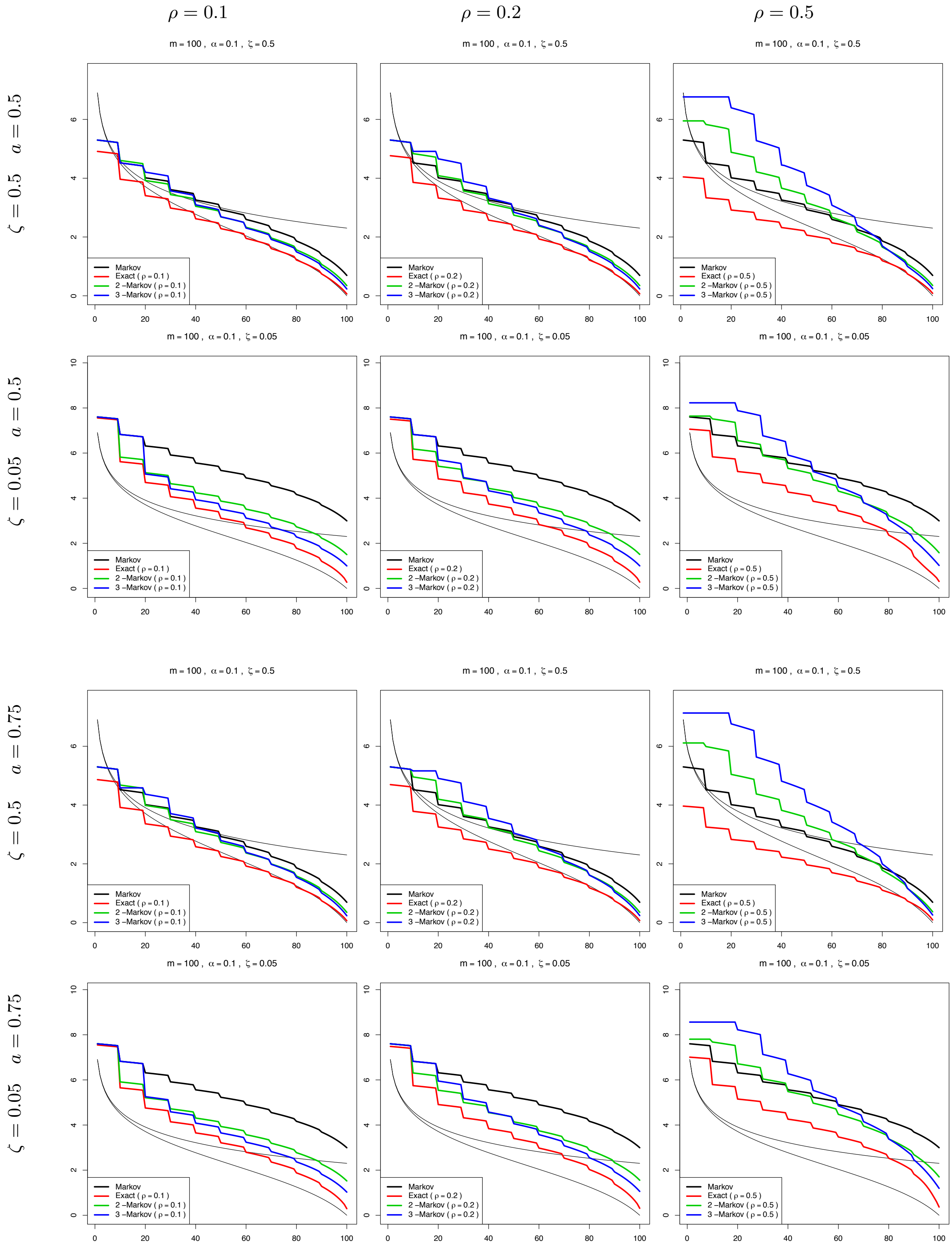

Fig S-2. *Same as Figure 3, but in the alternate equi-correlated case, see text.*

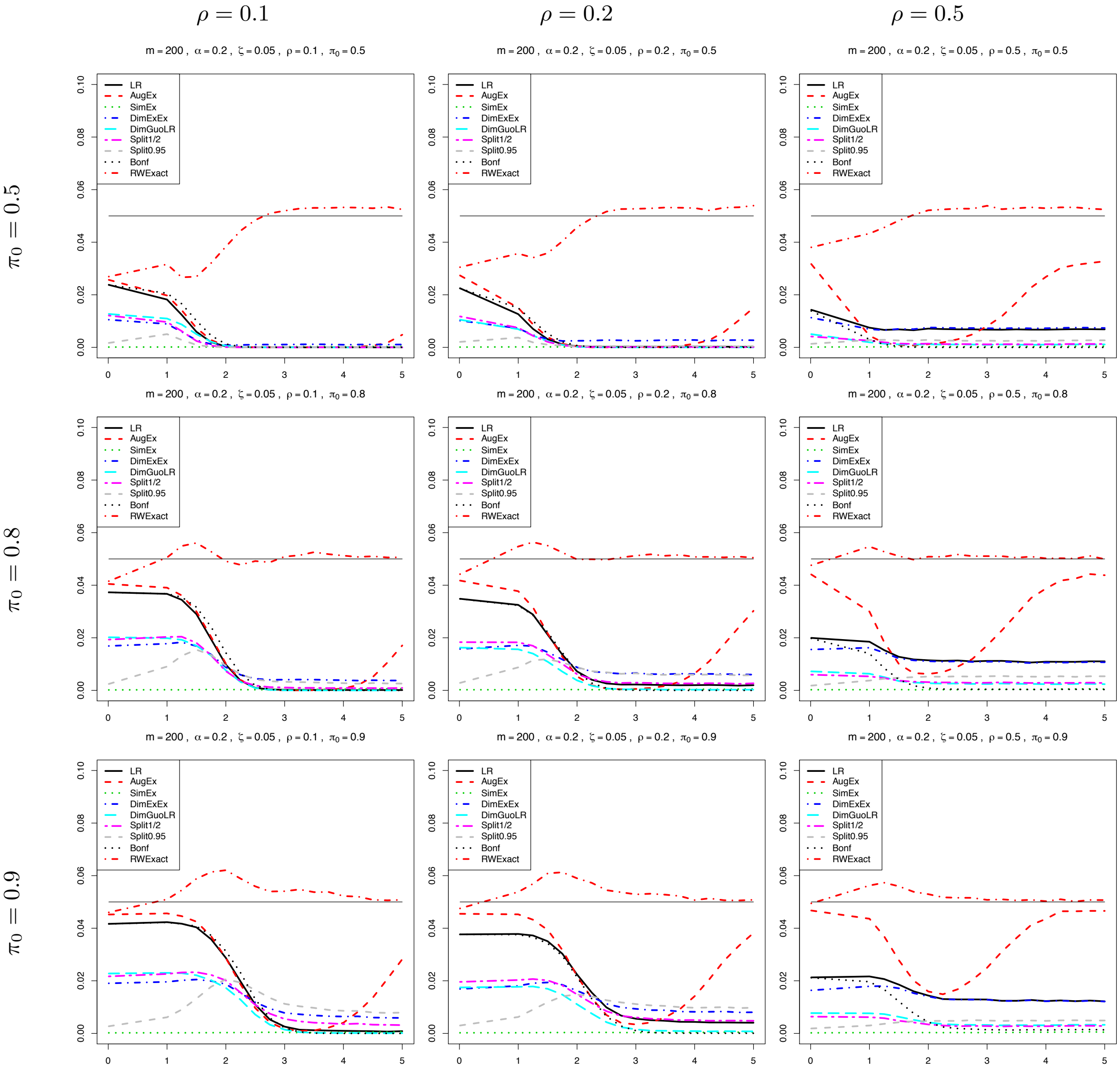

FIG S-3. *Equicorrelated case. Probability that the FDP exceeds α in function of β, see text and Table 1.*

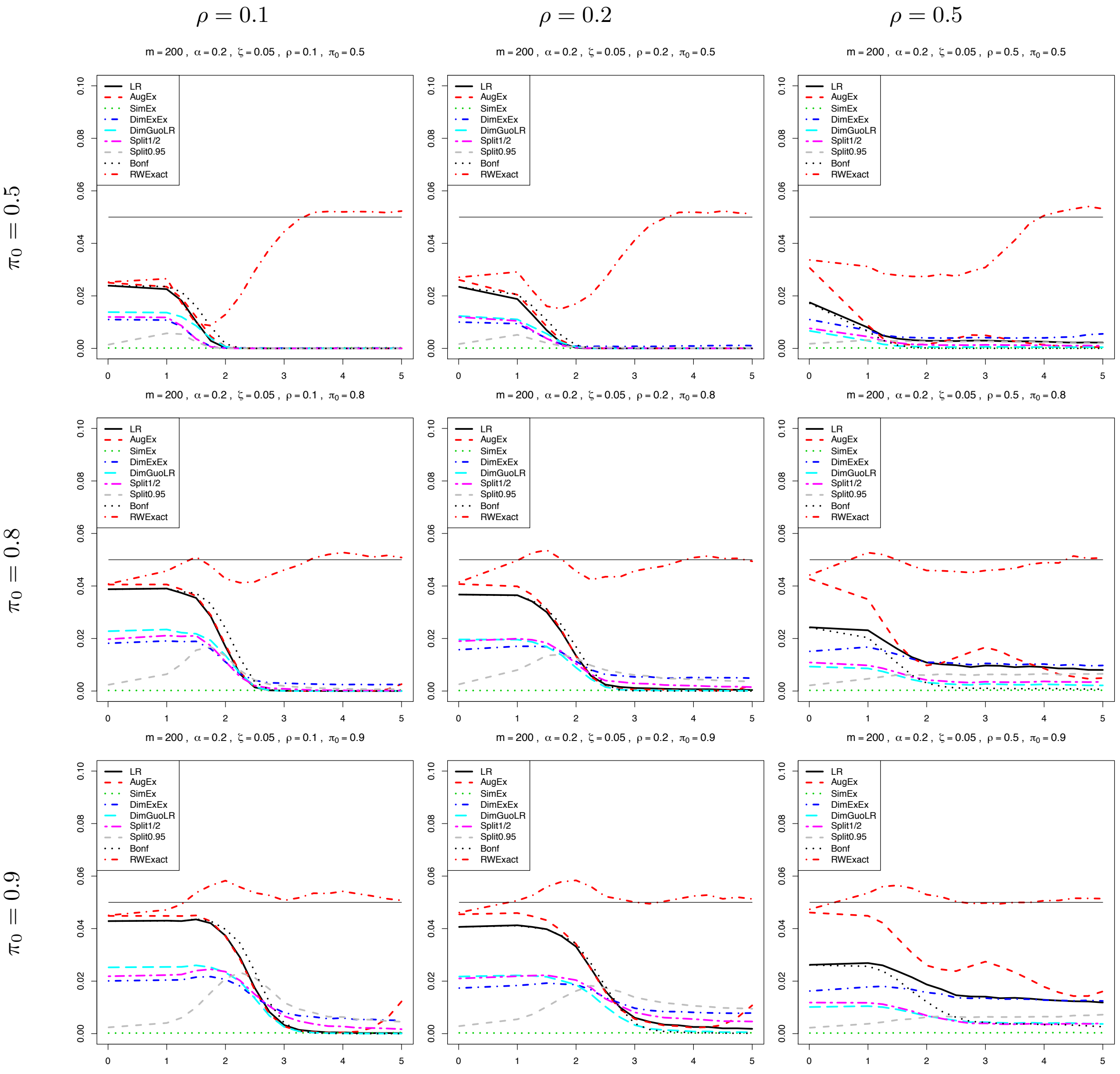

Fig S-4. *Alternate equicorrelated case. Probability that the FDP exceeds* α *in function of* β*, see text and Table* 1.

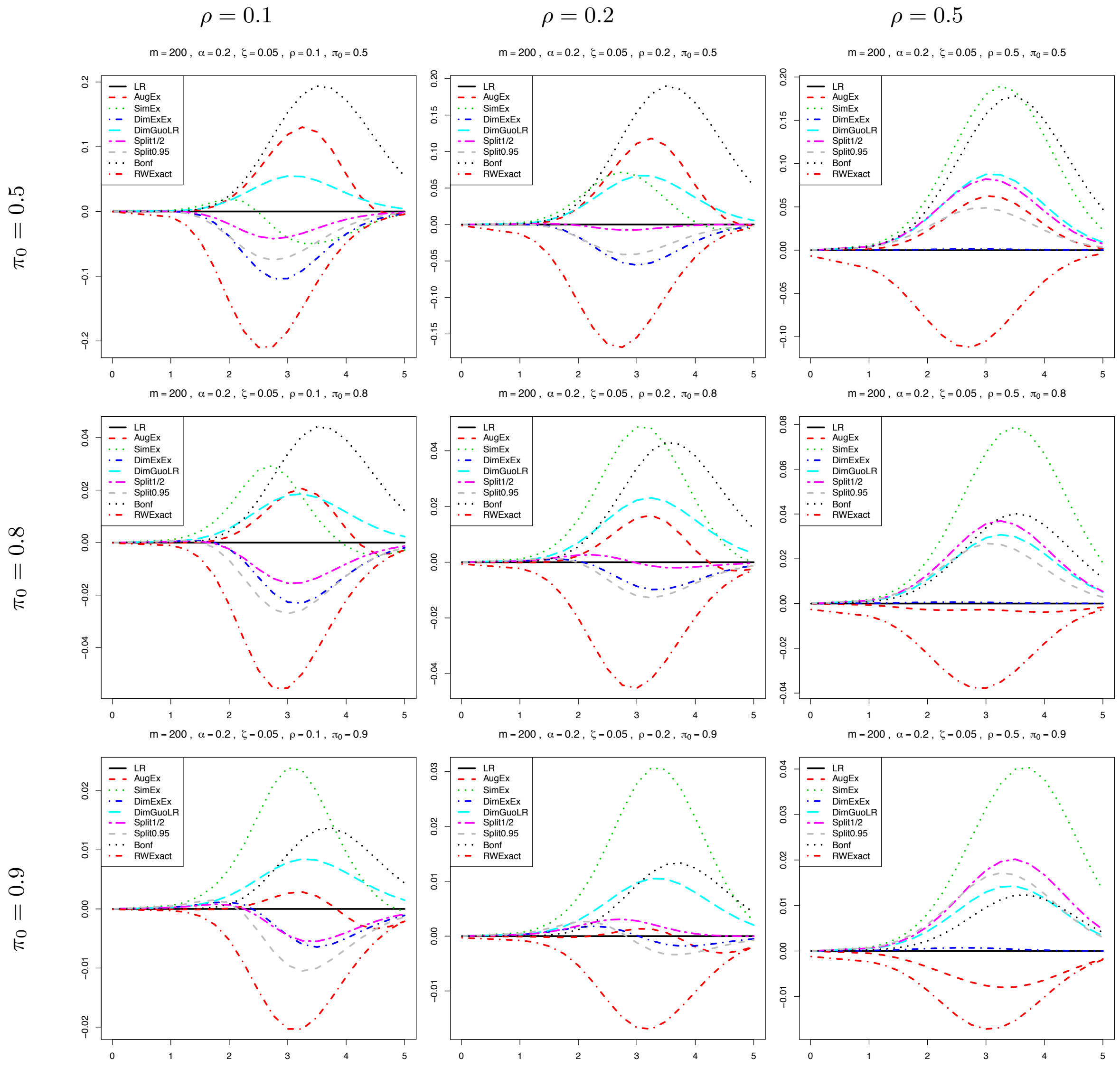

FIG S-5. *Equicorrelated case. Relative FNR to Lehmann Romano procedure in function of β, see text and Table 1.*

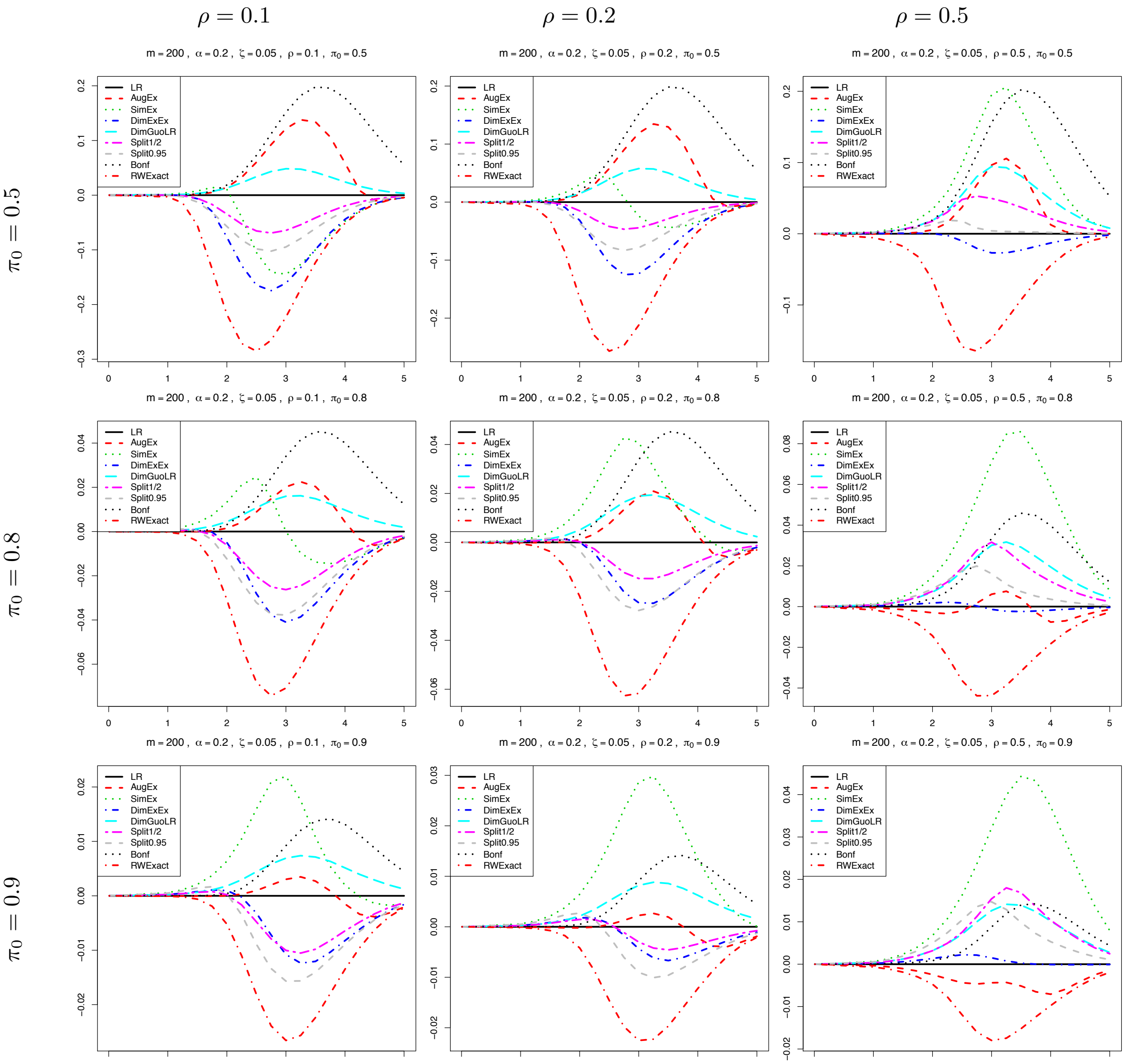

FIG S-6. *Alternate equicorrelated case. Relative FNR to Lehmann Romano procedure in function of β, see text and Table 1.*

S-6. 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 (16). 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 model (Gauss-ρ-equi) (the alternative means μ_i are all equal to some β). Figure S-7 reports a situation where the FDP control is violated (admittedly not by much).

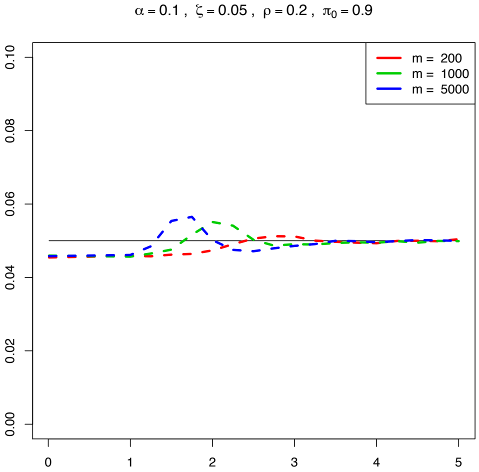

Fig S-7. $\mathbb{P}(\mathrm{FDP}_m(\tau_{\hat{\ell}}) > \alpha)$ *in function of* β *for the adaptive step-down procedure using the exact bounding device in the equi-correlated case.*

References.

Delattre, S. and Roquain, E. (2013). New procedures controlling the false discovery proportion via Romano-Wolf's heuristic. *ArXiv e-prints.*

Delattre, S. and Roquain, E. (2014). On empirical distribution function of high-dimensional Gaussian vector components with an application to multiple testing. *Bernoulli.* To appear.

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.

Université Paris Diderot, LPMA,
E-mail: sylvain.delattre@univ-paris-diderot.fr

UPMC Université Paris 6, LPMA,
E-mail: etienne.roquain@upmc.fr