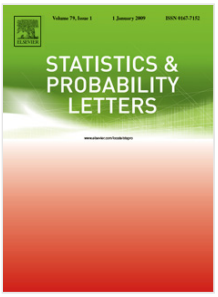
# On the false discovery proportion convergence under Gaussian equi-correlation

S. Delattre, E. Roquain




On the false discovery proportion convergence under Gaussian equi-correlation


S. Delattre, E. Roquain



PII: S0167-7152(10)00275-0
DOI: 10.1016/j.spl.2010.09.025
Reference: STAPRO 5811

To appear in: *Statistics and Probability Letters*

Received date: 3 July 2010
Accepted date: 30 September 2010

Please cite this article as: Delattre, S., Roquain, E., On the false discovery proportion convergence under Gaussian equi-correlation. *Statistics and Probability Letters* (2010), doi:10.1016/j.spl.2010.09.025

# On the false discovery proportion convergence under Gaussian equi-correlation

Delattre, S.[a], Roquain, E.[b]

[a] *University Paris Diderot, LPMA*
[b] *UPMC University of Paris 6, LPMA*

**Abstract**

We study the convergence of the false discovery proportion (FDP) of the Benjamini-Hochberg procedure in the Gaussian equi-correlated model, when the correlation $\rho_m$ converges to zero as the hypothesis number $m$ grows to infinity. By contrast with the standard convergence rate $m^{1/2}$ holding under independence, this study shows that the FDP converges to the false discovery rate (FDR) at rate $\{\min(m, 1/\rho_m)\}^{1/2}$ in this equi-correlated model.

*Keywords:* False discovery rate, Donsker theorem, equi-correlation, functional Delta method, $p$-value.
*2000 MSC:* 62G10, 62J15, 60F05

## 1. Introduction

When testing simultaneously a large number $m$ of null hypotheses, a popular global type I error, that can be traced back to Seeger (1968), is the false discovery proportion (FDP), defined as the ratio of the number of erroneous rejections to the number of rejections. The average of this random variable, called the false discovery rate (FDR, introduced by Benjamini and Hochberg (1995)), has been studied by a considerable number of works, see for instance Sarkar (2008) and Blanchard and Roquain (2009) for a review. However, studying the FDR is not sufficient to catch the full behavior of the FDP, for instance a FDR control does not prevent the FDP from having large variations. Therefore, some other studies aim to directly control the upper-quantile of the FDP distribution, see e.g. Genovese and Wasserman (2006); Lehmann and Romano (2005), or to directly compute the distribution of the FDP, either non-asymptotically Chi and Tan (2008); Roquain and Villers (2010), or asymptotically Genovese and Wasserman (2004). Recently, Neuvial (2008, 2009) computed the asymptotic distribution of the FDP *actually* achieved by the Benjamini-Hochberg (BH) procedure (and some other adaptive procedures) under independence of the $p$-values. It is proved that the FDP converges to the FDR at the parametric rate $\sqrt{m}$. Furthermore, Farcomeni (2007) showed that this convergence is unchanged under a specific short-range dependency between the $p$-values.

In this paper, we are interested in studying the convergence of the FDP of the BH procedure in the model where the test statistics have exchangeable Gaussian errors, with equi-correlation $\rho$ (allowing for instance long-range dependencies). This model has become quite standard in multiple testing (see e.g. Benjamini et al. (2006); Finner et al. (2007)), as it is a very simple instance of dependent $p$-value model. From an intuitive point of view, the test statistics can be seen as independent test statistics plus a disturbance variable whose importance depends on the

value of $\rho$. When $\rho \in (0,1)$ is fixed with $m$ (and in the "ideal" setting where the $p$-values under the alternative are all equal to zero), Finner et al. (2007) proved that the FDP of the BH procedure converges to a non-deterministic random variable that still depends on the disturbance variable. When $\rho_m \to 0$, we show here that this disturbance variable has no effect on the limit of the FDP anymore, which equals $\pi_0\alpha$ (where $\pi_0$ is the proportion of true nulls), but can still have an effect on the asymptotic variance of the FDP or even on the convergence rate. More precisely, when $\rho_m \to 0$ our main result states that $\{\min(m, 1/\rho_m)\}^{1/2}(\mathrm{FDP}_m - \pi_0\alpha) \rightsquigarrow \mathcal{N}(0, V)$ holds for a given $V > 0$; in comparison with the independent case, we may distinguish the two following cases, recovering all the possible convergence regimes of $\rho_m$ to zero:

- when $\lim_m m\rho_m = \theta \in [-1, +\infty)$, the limit of the FDP and the convergence rate are the same as in the independent case. The asymptotic variance $V$ is larger if $\theta > 0$, smaller if $\theta < 0$ and is the same whenever $\theta = 0$ (i.e. $m\rho_m \to 0$).
- when $\lim_m m\rho_m = +\infty$ and $\lim_m \rho_m = 0$ the convergence rate is $\rho_m^{-1/2}$ instead of $m^{1/2}$.

On the one hand, this shows that the FDP of the BH procedure is still well concentrated around $\pi_0\alpha$ under weak equi-correlation such that $\rho_m = O(1/m)$. On the other hand, this puts forward that the concentration of the FDP of the BH procedure around the FDR may be arbitrarily slow when $\rho_m \to 0$, which is a striking result that has not been reported before to the best of our knowledge. Therefore, our recommendation is that the BH procedure can be used under Gaussian equi-correlation when $\rho_m = O(1/m)$ (including the case of a negative equi-correlation) but should be used carefully as soon as $m\rho_m \to \infty$, as the actual convergence rate of the FDP to the FDR might be much slower.

The paper is organized as follows: Section 2 presents the model, the notation and the main result. The latter is proved in Section 3, including a generalization to any "regular" thresholding procedure, recovering the so-called $\pi_0$-adaptive procedures studied in Neuvial (2008). Finally, some further points in connexion with our methodology are discussed in Section 4.

## 2. Setting and main result

We observe $X_i = \tau_i + Y_i$, $1 \le i \le m$, where the parameter of interest is $(\tau_i)_i \in \{0, \mu\}^m$ (for a given $\mu > 0$) and the (unobservable) error vector $(Y_1, ..., Y_m)$ is an exchangeable Gaussian vector with $\mathbb{E}Y_1 = 0$ and $\mathrm{Var}\, Y_1 = 1$. We let $\rho_m = \mathrm{Cov}(Y_1, Y_2) \in [-(m-1)^{-1}, 1]$. We consider the problem of the one-sided testing of the null "$\tau_i = 0$" against the alternative "$\tau_i = \mu$", simultaneously for any $1 \le i \le m$. To test each null, we define the $p$-value $p_i = \Phi(X_i)$, where $\Phi(z) = \mathbb{P}(Z \ge z)$ is the standard Gaussian upper-tail function. The c.d.f. of each $p$-value is denoted by $G_0(t) = t$ under the null and by $G_1(t) = \Phi(\Phi^{-1}(t) - \mu)$ under the alternative. The number of true nulls is denoted by $m_0(m) = |\{i \mid \tau_i = 0\}|$ and is assumed to be of the form $\lfloor m\pi_0 \rfloor$ for a given proportion of true null $\pi_0 \in (0,1)$ independent of $m$. The "mixture" c.d.f. of the $p$-values is denoted by $G(t) = \pi_0 G_0(t) + (1-\pi_0)G_1(t)$. Next, we define the e.c.d.f.'s $\widehat{\mathbb{G}}_{0,m}(t) = (m_0(m))^{-1} \sum_{i=1}^m \mathbf{1}\{\tau_i = 0\}\mathbf{1}\{p_i \le t\}$, $\widehat{\mathbb{G}}_{1,m}(t) = (m - m_0(m))^{-1} \sum_{i=1}^m \mathbf{1}\{\tau_i > 0\}\mathbf{1}\{p_i \le t\}$ and $\widehat{\mathbb{G}}_m(t) = m^{-1} \sum_{i=1}^m \mathbf{1}\{p_i \le t\}$.

Given a pre-specified level $\alpha \in (0,1)$, the procedure of Benjamini and Hochberg (1995) can be defined as the procedure rejecting the nulls corresponding to $p_i \le \mathcal{T}^{\mathrm{BH}}(\widehat{\mathbb{G}}_m)$ where the (data-driven) threshold $\mathcal{T}^{\mathrm{BH}}(\widehat{\mathbb{G}}_m)$ is $\max\{t \in [0,1] \mid \widehat{\mathbb{G}}_m(t) \ge t/\alpha\}$. Next, the false discovery proportion

at a given threshold $t \in [0,1]$ is defined as the proportion of true nulls among the hypotheses having a $p$-value smaller than or equal to $t$:

$$\mathrm{FDP}_m(t) = \frac{|\{1 \le i \le m \mid \tau_i = 0, p_i \le t\}|}{|\{1 \le i \le m \mid p_i \le t\}| \vee 1} = \frac{m_0(m)}{m} \frac{\widehat{\mathbb{G}}_{0,m}(t)}{\widehat{\mathbb{G}}_m(t) \vee m^{-1}},$$

where $|\cdot|$ denotes the cardinality function.

We now state our main result.

**Theorem 2.1.** *There is a unique point $t^\star \in (0,1)$ such that $G(t^\star) = t^\star/\alpha$ and we have*

*(i) if $\lim_m m\rho_m = \theta \in [-1, +\infty)$, then*

$$\sqrt{m}(FDP_m(\mathcal{T}^{BH}(\widehat{\mathbb{G}}_m)) - \pi_0\alpha) \rightsquigarrow \mathcal{N}\left(0, \pi_0\alpha^2 \frac{1-t^\star}{t^\star} + \theta \frac{\pi_0^2\alpha^2}{2\pi(t^\star)^2} e^{-(\Phi^{-1}(t^\star))^2}\right); \tag{1}$$

*(ii) if $\lim_m m\rho_m = +\infty$ and $\lim_m \rho_m = 0$, then*

$$\rho_m^{-1/2}(FDP_m(\mathcal{T}^{BH}(\widehat{\mathbb{G}}_m)) - \pi_0\alpha) \rightsquigarrow \mathcal{N}\left(0, \frac{\pi_0^2\alpha^2}{2\pi(t^\star)^2} e^{-(\Phi^{-1}(t^\star))^2}\right). \tag{2}$$

## 3. Proof of Theorem 2.1

### *3.1. A more general result*

In what follows, we denote the space of functions from $[0,1]$ to $\mathbb{R}$ which are right-continuous and with left-hand limits (Skorokhod's space) by $D(0,1)$ and the space of continuous functions from $[0,1]$ to $\mathbb{R}$ by $C(0,1)$. The method for proving our result relies on the methodology let down by Neuvial (2008) which consider the case of a general threshold function $\mathcal{T} : D(0,1) \to [0,1]$ which is Hadamard differentiable at $G$, tangentially to $C(0,1)$ (see van der Vaart (1998) for a formal definition). The proof of Theorem 2.1 is presented here as a consequence of a more general theorem, true for any such threshold. The derivative of the threshold $\mathcal{T}$ at $G$, which is a continuous linear form on $C(0,1)$, is denoted by $\dot{\mathcal{T}}_G$. According to the Riesz representation theorem, the continuous linear form $\dot{\mathcal{T}}_G$ can be written as $\dot{\mathcal{T}}_G(F) = \int_0^1 F(t)\dot{\mathcal{T}}_G(dt)$, where we identified the linear form $\dot{\mathcal{T}}_G$ and the corresponding signed measure.

**Theorem 3.1.** *Let $\mathcal{T} : D(0,1) \to [0,1]$ be Hadamard differentiable at $G$, tangentially to $C(0,1)$, with derivative $\dot{\mathcal{T}}_G$. Let $q(t) = \pi_0 t/G(t)$ for $t > 0$, let $t^\star = \mathcal{T}(G)$ and assume $t^\star > 0$. We set $\zeta_0 = \frac{q(t^\star)(1-q(t^\star))}{t^\star}\delta_{t^\star} + \dot{q}(t^\star)\pi_0\dot{\mathcal{T}}_G$, $\zeta_1 = -\frac{q(t^\star)(1-q(t^\star))}{G_1(t^\star)}\delta_{t^\star} + \dot{q}(t^\star)(1-\pi_0)\dot{\mathcal{T}}_G$ and*

$$c(\mathcal{T}) = (2\pi)^{-1/2}\int_0^1 e^{-\frac{1}{2}(\Phi^{-1}(t))^2}\zeta_0(dt) + (2\pi)^{-1/2}\int_0^1 e^{-\frac{1}{2}(\Phi^{-1}(t)-\mu)^2}\zeta_1(dt);$$

$$\sigma^2(\mathcal{T}) = \pi_0^{-1}\int_{[0,1]^2}(s\vee t - st)\zeta_0(ds)\zeta_0(dt) + (1-\pi_0)^{-1}\int_{[0,1]^2}(G_1(s\vee t) - G_1(s)G_1(t))\zeta_1(ds)\zeta_1(dt).$$

*Then the following holds:*

*(i) if $\lim_m m\rho_m = \theta \in [-1, +\infty)$,*

$$\sqrt{m}(FDP_m(\mathcal{T}(\widehat{\mathbb{G}}_m)) - q(t^\star)) \rightsquigarrow \mathcal{N}\left(0, \sigma^2(\mathcal{T}) + \theta c(\mathcal{T})^2\right); \tag{3}$$

*(ii) if* $\lim_m m\rho_m = +\infty$ *and* $\lim_m \rho_m = 0$,

$$\rho_m^{-1/2}(FDP_m(\mathcal{T}(\widehat{\mathbb{G}}_m)) - q(t^\star)) \rightsquigarrow \mathcal{N}\left(0, c(\mathcal{T})^2\right). \tag{4}$$

Let us now check that Theorem 2.1 follows from Theorem 3.1. From Neuvial (2008) Corollary 7.12, $\mathcal{T}^{\text{BH}} : F \mapsto \max\{t \in [0,1] \mid F(t) \geq t/\alpha\}$ is Hadamard differentiable at $G$, tangentially to $C(0,1)$, with derivative $\dot{\mathcal{T}}_G^{\text{BH}} = (1/\alpha - \dot{G}(t^\star))^{-1}\delta_{t^\star}$. Moreover, $t^\star = \max\{t \in [0,1] \mid G(t) \geq t/\alpha\}$ is positive, because $\lim_{t\to 0^+} t/G(t) = 0$. Also, since $G(t^\star) = t^\star/\alpha$ and $\dot{q}(t^\star) = (1/\alpha - \dot{G}(t^\star))\pi_0\alpha^2/t^\star$, we may check that $\zeta_1 = 0$ and $\zeta_0 = (\pi_0\alpha/t^\star)\delta_{t^\star}$ in the above theorem, which leads to Theorem 2.1.

### *3.2. Proof of Theorem 3.1*

Let us now prove Theorem 3.1. First write $\text{FDP}_m(\mathcal{T}(\widehat{\mathbb{G}}_m)) = (m_0(m)/m)\pi_0^{-1}\Psi(\widehat{\mathbb{G}}_{0,m}, \widehat{\mathbb{G}}_{1,m})$, where for any $F_0, F_1$ in $D(0,1)$ with $F(\mathcal{T}(F)) > 0$ (letting $F = \pi_0 F_0 + (1-\pi_0)F_1$), we put $\Psi(F_0, F_1) = \pi_0 \frac{F_0(\mathcal{T}(F))}{F(\mathcal{T}(F))}$. From standard computations, $\Psi$ is Hadamard differentiable at $(G_0, G_1)$, tangentially to $C(0,1)^2$ and the derivative takes the form, for $(H_0, H_1) \in C(0,1)^2$, $\dot{\Psi}_{G_0,G_1}(H_0, H_1) = q(t^\star)(1-q(t^\star))\left(\frac{H_0(t^\star)}{t^\star} - \frac{H_1(t^\star)}{G_1(t^\star)}\right) + \dot{q}(t^\star)\dot{\mathcal{T}}_G(H)$, where $H = \pi_0 H_0 + (1-\pi_0)H_1$. Applying the functional Delta method, this leads to the following useful result, which was essentially stated in Neuvial (2008).

**Proposition 3.2.** *Let* $\mathcal{T} : D(0,1) \to [0,1]$ *be Hadamard differentiable at* $G$*, tangentially to* $C(0,1)$*, with derivative* $\dot{\mathcal{T}}_G$*. Let* $q(t) = \pi_0 t/G(t)$ *for* $t > 0$*, let* $t^\star = \mathcal{T}(G)$ *and assume* $t^\star > 0$*. If for a given sequence* $a_m \to \infty$ *with* $a_m = o(m)$*,*

$$a_m \begin{pmatrix} \widehat{\mathbb{G}}_{0,m} - G_0 \\ \widehat{\mathbb{G}}_{1,m} - G_1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{W}_0 \\ \mathbb{W}_1 \end{pmatrix}, \tag{5}$$

*where the convergence in distribution is relative to the Skorokhod topology and where* $\mathbb{W}_0$ *and* $\mathbb{W}_1$ *are processes with continuous paths, then we have*

$$a_m(FDP_m(\mathcal{T}(\widehat{\mathbb{G}}_m)) - q(t^\star)) \rightsquigarrow X, \tag{6}$$

*where* $X = \zeta_0(\mathbb{W}_0) + \zeta_1(\mathbb{W}_1)$ *and* $\zeta_0, \zeta_1$ *are defined as in Theorem 3.1.*

A convergence of the type (5) in the particular Gaussian equi-correlated model is stated in Lemma 3.3. Using Proposition 3.2, this proves that (6) holds both in the cases (i) and (ii) with $a_m = \sqrt{m}$ and $a_m = \rho_m^{-1/2}$, which respectively leads to (3) and (4) (the variance computations are straightforward).

### *3.3. Convergence of the e.c.d.f.'s in the Gaussian equi-correlated model*

**Lemma 3.3.** *Let* $(\mathbb{Z}_0, \mathbb{Z}_1, Z)$ *be a random variable such that* $\mathbb{Z}_0 \overset{(d)}{=} \pi_0^{-1/2}\mathbb{B}$, $\mathbb{Z}_1 \overset{(d)}{=} (1-\pi_0)^{-1/2}\mathbb{B} \circ G_1$, $\mathbb{B}$ *being a standard Brownian bridge on* $[0,1]$, $\mathbb{Z}_0$ *is independent from* $\mathbb{Z}_1$, $Z \sim \mathcal{N}(0,1)$, $\text{Cov}(Z, \mathbb{Z}_0(t)) = (2\pi)^{-1/2}\exp(-\{\Phi^{-1}(t)\}^2/2)$ *and* $\text{Cov}(Z, \mathbb{Z}_1(t)) = (2\pi)^{-1/2}\exp(-\{\Phi^{-1}(t) - \mu\}^2/2)$. *Let also* $U \sim \mathcal{N}(0,1)$ *be independent of the vector* $(\mathbb{Z}_0, \mathbb{Z}_1, Z)$*. Then we have the following convergences in law for the Skorokhod topology:*

*(i) if* $\lim_m m\rho_m = \theta \in [-1, +\infty)$,

$$\sqrt{m}\begin{pmatrix} \widehat{\mathbb{G}}_{0,m} - G_0 \\ \widehat{\mathbb{G}}_{1,m} - G_1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} \mathbb{Z}_0 + (Z - \sqrt{1+\theta}U)\,\dot{\Phi} \circ \Phi^{-1} \\ \mathbb{Z}_1 + (Z - \sqrt{1+\theta}U)\,\dot{\Phi} \circ (\Phi^{-1} - \mu) \end{pmatrix}; \tag{7}$$

*(ii) if* $\lim_m m\rho_m = +\infty$ *and* $\lim_m \rho_m = 0$,

$$\rho_m^{-1/2}\begin{pmatrix} \widehat{\mathbb{G}}_{0,m} - G_0 \\ \widehat{\mathbb{G}}_{1,m} - G_1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} U\,\dot\Phi \circ \Phi^{-1} \\ U\,\dot\Phi \circ (\Phi^{-1} - \mu) \end{pmatrix}. \tag{8}$$

To prove Lemma 3.3, first remark that the distribution of the $X_i$'s may be realized as $X_i = \sqrt{1-\rho_m}(\xi_i - \overline{\xi}) + \sqrt{(1+(m-1)\rho_m)/m}\,U + \mu\mathbf{1}\{\tau_i > 0\}$, where $(\xi_1, ..., \xi_m, U)$ are all i.i.d. $\mathcal{N}(0,1)$ variables and $\overline{\xi}$ denotes the empirical mean of the $\xi_i$'s. Let $\widehat{\mathbb{G}}'_{0,m}(t) = (m_0(m))^{-1}\sum_{i:\tau_i=0}\mathbf{1}\{\Phi(\xi_i) \le t\}$, $\widehat{\mathbb{G}}'_{1,m}(t) = (m - m_0(m))^{-1}\sum_{i:\tau_i>0}\mathbf{1}\{\Phi(\xi_i+\mu) \le t\}$ and

$$f_m(t, U, \rho_m) = (1-\rho_m)^{-1/2}\left(\Phi^{-1}(t) - \sqrt{(1+(m-1)\rho_m)/m}\,U\right).$$

The process $(\widehat{\mathbb{G}}_{0,m} - G_0, \widehat{\mathbb{G}}_{1,m} - G_1)$ is then equal to $V_m + W_m$ where

$$V_m(t) = \begin{pmatrix} (\widehat{\mathbb{G}}'_{0,m} - G_0)(\Phi(f_m(t,U,\rho_m) + \overline{\xi})) \\ (\widehat{\mathbb{G}}'_{1,m} - G_1)(\Phi(f_m(t,U,\rho_m) + \overline{\xi} - \mu(1-\rho_m)^{-1/2} + \mu)) \end{pmatrix}$$
$$W_m(t) = \begin{pmatrix} \Phi(f_m(t,U,\rho_m) + \overline{\xi}) - t \\ \Phi(f_m(t,U,\rho_m) + \overline{\xi} - \mu(1-\rho_m)^{-1/2}) - \Phi(\Phi^{-1}(t) - \mu) \end{pmatrix}.$$

Next, applying Donsker's theorem, we derive $\sqrt{m}(\widehat{\mathbb{G}}'_{0,m} - G_0, \widehat{\mathbb{G}}'_{1,m} - G_1, \overline{\xi}) \rightsquigarrow (\mathbb{Z}_0, \mathbb{Z}_1, Z)$, where $(\mathbb{Z}_0, \mathbb{Z}_1, Z)$ is defined as in Lemma 3.3. Since $\rho_m \to 0$, the inverse functions of $t \mapsto \Phi(f_m(t,U,\rho_m) + \overline{\xi})$ and $t \mapsto \Phi(f_m(t,U,\rho_m) + \overline{\xi} - \mu(1-\rho_m)^{-1/2} + \mu)$ converge uniformly on $[0,1]$ to the identity a.s. Therefore, applying the Skorokhod's representation theorem, we get

$$\sqrt{m}(V_m, \overline{\xi}) \rightsquigarrow (\mathbb{Z}_0, \mathbb{Z}_1, Z). \tag{9}$$

Let us now consider the case (i), in which $\lim_m m\rho_m = \theta \in [-1, +\infty)$. In that case, a standard reasoning involving Taylor expansions of $\Phi$ and $y \mapsto \Phi(y\Phi^{-1}(t))$ leads to

$$\sqrt{m}W_m(t) = \begin{pmatrix} \dot\Phi(\Phi^{-1}(t))(\sqrt{m}\,\overline{\xi} - \sqrt{1+\theta}\,U) \\ \dot\Phi(\Phi^{-1}(t) - \mu)(\sqrt{m}\,\overline{\xi} - \sqrt{1+\theta}\,U) \end{pmatrix} + \begin{pmatrix} R_{0,m}(t) \\ R_{1,m}(t) \end{pmatrix},$$

with remainder terms satisfying $\|R_{0,m}\|_\infty \vee \|R_{1,m}\|_\infty \to 0$ in probability. Since $U$ is independent of all the other variables, we derive from (9) that $\sqrt{m}(V_m, W_m) \rightsquigarrow (\mathbb{Z}_0, \mathbb{Z}_1, (Z - \sqrt{1+\theta}U))\,\dot\Phi \circ \Phi^{-1}, (Z - \sqrt{1+\theta}U))\,\dot\Phi \circ (\Phi^{-1} - \mu))$. This implies (7). Consider now the case (ii), in which $\lim_m m\rho_m = +\infty$ and $\lim_m \rho_m = 0$. In that situation, we deduce from (9) that $\rho_m^{-1/2}V_m$ converges in probability to 0. Furthermore, using that $\rho_m^{-1/2}\overline{\xi}$ tends to zero in probability, we obtain that

$$\rho_m^{-1/2}W_m(t) = \begin{pmatrix} \dot\Phi(\Phi^{-1}(t))(-U) \\ \dot\Phi(\Phi^{-1}(t) - \mu)(-U) \end{pmatrix} + \begin{pmatrix} T_{0,m}(t) \\ T_{1,m}(t) \end{pmatrix},$$

with remainder terms satisfying $\|T_{0,m}\|_\infty \vee \|T_{1,m}\|_\infty \to 0$ in probability. This implies (8).

## 4. Discussion: FDP convergence in the case $\rho_m = \rho \in (0,1)$

When $\rho_m = \rho \in (0,1)$, we cannot expect that the FDP concentrates around the FDR as in Theorem 2.1 (see e.g. Finner et al. (2007) Theorem 2.1). As a consequence, even if the FDP has

a mean below $\pi_0\alpha$ (because the false discovery rate of the BH procedure is below $\pi_0\alpha$ for each $m$ for PRDS statistics, see Theorem 1.2 in Benjamini and Yekutieli (2001)), the FDP can exceed $\pi_0\alpha + \varepsilon$ ($\varepsilon > 0$) with a probability that does not vanish when $m$ grows to infinity.

We claim here that in the ideal situation where the parameters of the model $\pi_0$, $\mu$, $\rho$ are perfectly known, it is possible to modify the $p$-values so that the FDP convergence to the FDR keeps the parametric convergence rate $\sqrt{m}$. For this, we replace each test statistic $X_i$ by $\widetilde{X}_i = \sqrt{m/((m-1)(1-\rho))}(X_i - \overline{X} + (1-\pi_0)\mu)$, so that $(\widetilde{X}_1, ..., \widetilde{X}_m)$ is a Gaussian vector with variances equal to 1, equi-correlation $\widetilde{\rho}_m = -(m-1)^{-1}$ and means $\mathbb{E}\widetilde{X}_i = \sqrt{m/((m-1)(1-\rho))}\tau_i$. We build the corresponding $p$-values by letting $\widetilde{p}_i = \Phi(\widetilde{X}_i)$, which are uniform under the null and have the c.d.f. $\widetilde{G}_{1,m}(t) = \Phi(\Phi^{-1}(t) - \widetilde{\mu}_m)$ for $\widetilde{\mu}_m = (m/(m-1))^{1/2}\mu(1-\rho)^{-1/2}$ under the alternative. Although the latter depends (slightly) on $m$, we easily check that our methodology applies using $\widetilde{G}_1(t) = \Phi(\Phi^{-1}(t) - \widetilde{\mu})$ for $\widetilde{\mu} = \mu(1-\rho)^{-1/2}$ and that the following convergence holds:

$$\sqrt{m}(\widetilde{\mathrm{FDP}}_m - \pi_0\alpha) \rightsquigarrow \mathcal{N}\left(0\,,\ \pi_0\alpha^2\frac{1-t^\star_\rho}{t^\star_\rho} - \frac{\pi_0^2\alpha^2}{2\pi(t^\star_\rho)^2}e^{-(\Phi^{-1}(t^\star_\rho))^2}\right),$$

where $\widetilde{\mathrm{FDP}}_m$ denotes the FDP of the BH threshold $\mathcal{T}^{\mathrm{BH}}$ used with the $p$-values $\widetilde{p}_i$'s and where $t^\star_\rho \in (0,1)$ is the unique point $t \in (0,1)$ satisfying $\pi_0 t + (1-\pi_0)\Phi(\Phi^{-1}(t) - \mu(1-\rho)^{-1/2}) = t/\alpha$ (which depends on $\rho$). Of course, while this $p$-value modification greatly improves the concentration of the FDP, this approach is oracle because $\pi_0, \mu, \rho$ are generally unknown. A correct estimation of the model parameters within such a procedure stays an open issue.

## References


Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Statist. Soc. Ser. B 57, 289–300.

Benjamini, Y., Krieger, A.M., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93, 491–507.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Statist. 29, 1165–1188.

Blanchard, G., Roquain, E., 2009. Adaptive FDR control under independence and dependence. J. Mach. Learn. Res. 10, 2837–2871.

Chi, Z., Tan, Z., 2008. Positive false discovery proportions: intrinsic bounds and adaptive control. Statist. Sinica 18, 837–860.

Farcomeni, A., 2007. Some results on the control of the false discovery rate under dependence. Scand. J. Statist. 34, 275–297.

Finner, H., Dickhaus, T., Roters, M., 2007. Dependency and false discovery rate: asymptotics. Ann. Statist. 35, 1432–1455.

Genovese, C., Wasserman, L., 2004. A stochastic process approach to false discovery control. Ann. Statist. 32, 1035–1061.

Genovese, C.R., Wasserman, L., 2006. Exceedance control of the false discovery proportion. J. Amer. Statist. Assoc. 101, 1408–1417.

Lehmann, E.L., Romano, J.P., 2005. Generalizations of the familywise error rate. Ann. Statist. 33, 1138–1154.

Neuvial, P., 2008. Asymptotic properties of false discovery rate controlling procedures under independence. Electron. J. Stat. 2, 1065–1110.

Neuvial, P., 2009. Corrigendum to "Asymptotic properties of false discovery rate controlling procedures under independence" [mr2460858]. Electron. J. Stat. 3, 1083.

Roquain, E., Villers, F., 2010. Exact calculations for false discovery proportion with application to least favorable configurations. ArXiv.org:1002.2845.

Sarkar, S.K., 2008. On methods controlling the false discovery rate. Sankhya, Ser. A 70, 135–168.

Seeger, P., 1968. A note on a method for the analysis of significances en masse. Technometrics 10, 586–593.

van der Vaart, A.W., 1998. Asymptotic statistics. volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.