Web-based Supplementary Materials for *Screening for Partial Conjunction Hypotheses*

Yoav Benjamini*

Department of Statistics and Operations Research

Tel Aviv University, Tel Aviv 69978, Israel

and

Ruth Heller†

Department of Statistics, University of Pennsylvania

Philadelphia, USA

November 6, 2007

Appendix A

Proof of Theorem 1

We have to show that $P_0(P^{u/n}(s) \leq q) \leq q$ where the subscript 0 indicates that the probability is calculated under the partial conjunction null hypothesis.

Since $p^{u/n}(s)$ in equation (2) is an increasing function of the p-values, Lemma 1 tells us that the stochastically smallest distribution of $p^{u/n}(s)$ oc-

* *email:* ybenja@post.tau.ac.il
† *email:* ruheller@wharton.upenn.edu

curs when $u-1$ p-values are zero (i.e. $h_i(P_i)=0$ for $i=1,\ldots,u-1$) and the remaining $n-u+1$ p-values are $U(0,1)$ random variables. So it is enough to show for this case that $P^{u/n}(s) \succeq_{\mathrm{st}} U(0,1)$.

Let $U_{(1)} \leq \ldots \leq U_{(n-u+1)}$ be the order statistics of $n-u+1$ $U(0,1)$ random variables.

$$P_0(P^{u/n}(s) \leq q) \leq P(\min_{i=1,\ldots,n-u+1}\{\frac{(n-u+1)}{i}U_{(i)}\} \leq q) \leq q$$

where the last inequality follows since the $n-u+1$ null p-values satisfy one of the conditions (D1)-(D3) for which the Simes test is valid.

Appendix B

Proof of Theorem 3

The individual p-value maps are PRDS, so for every map $j=1,\ldots,n$ with vector of p-values $\vec{p}_j=(p_j(1),\ldots,p_j(S))$ we can say the following: $P(\vec{P}_j \in A|P_j(s)=x)$ for any increasing set A is non-decreasing in x for all $s \in I_0^j$, where I_0^j is the subset of null locations in map j.

Let the combined p-value in location s be $p^{u/n}(s)=f(p_1(s),\ldots,p_n(s))$ and $\vec{p}^{u/n}=(p^{u/n}(1),\ldots,p^{u/n}(S))$ be the vector of all combined p-values.

Let $I_0^{u/n}$ be the subset of locations where the partial conjunction null is true, $I_0^{u/n}=\{s: H_0^{u/n}(s) \text{ is true}\}$. Benjamini and Yekutieli (2001) showed in equation (10) that the FDR can be expressed as follows:

$$FDR=\sum_{k=1}^{S}\sum_{s\in I_0^{u/n}}\frac{1}{k}Pr(P^{u/n}(s) \leq \frac{kq}{S} \cap C_k^{(s)}) \tag{B.1}$$

where $C_k^{(s)}$ is the event that if s is rejected, $k-1$ hypotheses are rejected alongside with it.

The combined p-value under the null will have $n-u+1$ p-values distributed uniformly (or from a distribution larger than the uniform), as well as $u-1$ p-values that can have any distribution since they can come either from the null or from the alternative hypothesis. Assume without loss of generality that for a given pooled p-value from the partial conjunction null, $P^{u/n}(s) = f(P_1(s), \ldots, P_n(s))$, the $n-u+1$ first p-value maps contain uniformly distributed p-values in location s, i.e. $P_i(s) \sim U(0,1), i = 1, \ldots, n-u+1$. Since the combining function is increasing, $P^{u/n}(s) \geq f(P_1(s), \ldots, P_{n-u+1}(s), 0, \ldots, 0)$ and moreover since it is a null location p-value, $Pr(f(P_1(s), \ldots, P_{n-u+1}(s), 0, \ldots, 0) \leq \frac{kq}{S}) \leq \frac{kq}{S}$. For convenience, let $P_0^{u/n}(s) = f(P_1(s), \ldots, P_{n-u+1}(s), 0, \ldots, 0)$. An upper bound on the FDR is therefore

$$
\begin{aligned}
FDR &= \sum_{k=1}^{S} \sum_{s \in I_0^{u/n}} \frac{1}{k} \frac{Pr(P^{u/n}(s) \leq \frac{kq}{S} \cap C_k^{(s)}) Pr(P_0^{u/n}(s) \leq \frac{kq}{S})}{Pr(P_0^{u/n}(s) \leq \frac{kq}{S})} \\
&\leq \sum_{k=1}^{S} \sum_{s \in I_0^{u/n}} \frac{q}{S} \frac{Pr(P_0^{u/n}(s) \leq \frac{kq}{S} \cap C_k^{(s)})}{Pr(P_0^{u/n}(s) \leq \frac{kq}{S})} \qquad \text{(B.2)}
\end{aligned}
$$

In order to show that this upper bound is below the nominal q level, we will first show that $P(\vec{P}^{u/n} \in A | P_0^{u/n}(s) = x)$ for any increasing set A is non-decreasing in x. Let

$$
h(u_1, \ldots, u_{n-u+1}) = Pr(\vec{P}^{u/n} \in A | P_1(s) = u_1, \ldots, P_{n-u+1}(s) = u_{n-u+1})
$$

$h(u_1, \ldots, u_{n-u+1})$ is a non-decreasing function of u_j for $j \in \{1, \ldots, n-u+1\}$. This follows immediately from the fact that $Pr(\vec{P}^{u/n} \in A | P_j(s) = u_j)$ is increasing in u_j for $j = 1, \ldots, n-u+1$. To see this fact, note that if A is non-decreasing in its elements $\vec{P}^{u/n}$, then A is non-decreasing also in the original p-values $\{P_j(s) : j = 1, \ldots, n, s = 1, \ldots, S\}$ because the combining function is an increasing function. For a fixed map j, the p-values within map j satisfy the PRDS property, and moreover since the p-values in other maps are independent from these in map j (therefore trivially satisfy the PRDS property as well on the subset of null p-values from map j), we have that $Pr(\vec{P}^{u/n} \in A | P_j(s) = u_j)$ is increasing in u_j.

We will use the following theorem due to Efron (1965):

THEOREM 1. *Let* $X_1, \ldots, X_n$ *be* n *independent random variables with* PF_2 *densities* $r_1(x), \ldots, r_n(x)$ *respectively, and let* $H(x_1, \ldots, x_n)$ *be a real measurable function on Euclidean n-space which is non-decreasing in each of its arguments. Then* $E(H(x_1, \ldots, x_n) | \sum_{i=1}^{n} X_i = y)$ *is a non-decreasing function of* y.

Since $P_0^{u/n}(s) = G(\sum_{i=1}^{n-u+1} g(P_i(s)))$ and both $G(\cdot)$ and $g(\cdot)$ are increasing, for every x there exists a constant c such that $\{p^{u/n}(s) = x\} = \{\sum_{i=1}^{n-u+1} g(P_i(s)) = c\}$, and $c(x)$ is increasing in x.

$$Pr(\vec{P}^{u/n} \in A | \{P_0^{u/n}(s) = x\}) = Pr(\vec{P}^{u/n} \in A | \sum_{i=1}^{n-u+1} g(P_i(s)) = c)$$

$$= Eh(U_1, \ldots, U_{n-u+1}) | \sum_{i=1}^{n-u+1} g(U_i) = c)$$

We can apply Theorem 4 to conclude that $Pr(\vec{P}^{u/n} \in A | \{P_0^{u/n}(s) = x\})$ increases in x.

The result will follow as in the proof of Theorem 1.2 in Benjamini and Yekutieli (2001). Let $p_{(1)}^{u/n(s)} \leq \ldots \leq p_{(S-1)}^{u/n(s)}$ be the ordered set of $S - 1$ p-values excluding s, let $D_k^{(s)} = \{\frac{(k+1)q}{S} < p_{(k)}^{u/n(s)}, \frac{(k+2)q}{S} < p_{(k+1)}^{u/n(s)}, \ldots, q <$

$p^{u/n(s)}_{(S-1)}\}$ for $k = 1, \ldots, S-1$, and $D^{(s)}_S$ be the entire space. Clearly for each k, $D^{(s)}_k$ is a non-decreasing set. Therefore

$$Pr(P^{u/n}_0(s) \leq \frac{kq}{S} \cap D^{(s)}_k | \{P^{u/n}_0(s) = \frac{kq}{S}\}) \leq Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S} \cap D^{(s)}_k | \{P^{u/n}_0(s) = \frac{(k+1)q}{S}\})$$

and so

$$Pr(P^{u/n}_0(s) \leq \frac{kq}{S} \cap D^{(s)}_k | \{P^{u/n}_0(s) \leq \frac{kq}{S}\}) \leq Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S} \cap D^{(s)}_k | \{P^{u/n}_0(s) \leq \frac{(k+1)q}{S}\})$$

Since $D^{(s)}_{k+1} = D^{(s)}_k \cup C^{(s)}_{k+1}$, we get that

$$\begin{aligned}
&\frac{Pr(P^{u/n}_0(s) \leq \frac{kq}{S} \cap D^{(s)}_k)}{Pr(P^{u/n}_0(s) \leq \frac{kq}{S})} + \frac{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S} \cap C^{(s)}_{k+1})}{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S})} \\
&\leq \frac{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S} \cap D^{(s)}_k)}{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S})} + \frac{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S} \cap C^{(s)}_{k+1})}{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S})} \\
&\leq \frac{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S} \cap D^{(s)}_{k+1})}{Pr(P^{u/n}_0(s) \leq \frac{(k+1)q}{S})}
\end{aligned}$$

Since $C_1 = D_1$ repeatedly using the above inequality for $s = 1, \ldots, S-1$ leads to

$$\sum_{k=1}^{S} \frac{Pr(P^{u/n}_0(s) \leq \frac{kq}{S} \cap C^{(s)}_k)}{Pr(P^{u/n}_0(s) \leq \frac{kq}{S})} \leq 1$$

Going back to (B.1),

$$FDR \leq \sum_{s \in I^{u/n}_0} \frac{q}{S} \sum_{k=1}^{S} \frac{Pr(P^{u/n}_0(s) \leq \frac{kq}{S} \cap C^{(s)}_k)}{Pr(P^{u/n}_0(s) \leq \frac{kq}{S})} \leq \frac{S_0}{S} q$$

Appendix C

Simulations

We conducted a simulations study in order to 1) validate that the screening procedure controls the FDR in finite settings that are common in fMRI, and 2) examine the robustness of the method in more extreme settings of dependency.

We simulated a typical fMRI setting. Maps of size 32×32 locations were simulated, each containing a region of 100 locations that contain signal. An additional control map of size 32×32 contained no signal. White noise convolved with a Gaussian filter created spatially correlated noise. The filter width varied in the simulations, ranging from a positive correlation between all locations in the map to a positive correlation with the nearest 8 neighbors only. The number of maps n was either 3 or 10. The number of locations that satisfy the partial conjunction alternative was 100. For screening for conjunction hypotheses that at least u out of n p-values are non-null, the worst configuration (i.e. highest FDR) is when the null locations has $u-1$ p-values with value zero. Therefore, the number of maps that contain signal in a location was k for 100 locations (where k ranged from u to n) and $u-1$ for all remaining locations. The signal size for each of the 100 locations that satisfy the partial conjunction alternative was sampled from a $N(\mu, 2)$ distribution, where μ ranged from 2 to 6 (if the signal size sampled was smaller than zero it was set to 0.1). The signal size for the locations that satisfy the partial conjunction null was set at 10. We applied the method for all values of $u = 1, \ldots, n$.

In order to examine the robustness of the method, we also simulated maps within which we had negative correlation. We considered maps of size 100 and 300. The region of activity varied between 5 and 60 locations. The negative correlation was either equal across all locations (in which case the correlation of the noise was -1/100 or -1/300), or equal only within blocks of 5, 10, 20 and 30 with correlations of size -1/5, -1/10, -1/20 and -1/30 respectively.

A control map that contained only noise was also simulated for each dependency structure. The p-value in each location, for each map, was the tail Gaussian probability of the difference in values in the map and in the control map normalized by the factor $\sqrt{2}$. The pooled p-value motivated by Simes (equation 2) was used since the dependency per location is PRDS (the case of all treatments being compared to the same control treatment). We next applied the BH procedure at level 0.05.

The simulation results show that the FDR is controlled in all settings. The power of the method depends both on the configuration of signal μ and on the proportion k/n of false hypotheses. The representative Figure 1 for a typical fMRI setting shows the FDR level and power as a function of average signal size μ where the Gaussian spatial filter had a standard deviation of 4, the number of maps was 3 and the number of p-values per location that come from the alternative is either null or 2.

The simulations with negative correlations show that the FDR is controlled in all settings. The FDR level is higher when the fraction of locations containing signal is small, k/n is large and u approaches k (or is greater than k, i.e. all partial conjunction hypotheses are null). Figure 2 shows the FDR and the power as a function of μ (with standard error ≤ 0.0008) in the setting in which we received the highest FDR level: a map of size 300 with 5 locations containing signal and a correlation of $-1/20$ for blocks of size 20 locations. The number of maps containing signal is $k = 7$. The highest FDR was achieved when $u = 8$, where it was exactly 0.0500 (Standard error

0.0007).

[Figure 1 about here.]

[Figure 2 about here.]

Appendix D

Asymptotic Results

If the p-values within the individual maps have local dependencies, then the dependencies between the p-values within the combined map remain local.

In this case applying the BH procedure after combining using equations (2)-(5) as appropriate will control the FDR asymptotically as the number of locations goes to infinity if the following asymptotic conditions on every map i, $i = 1, \ldots, n$ are satisfied:

$$\lim_{S\to\infty} \frac{S_{0i}}{S} = A_{0i} \text{ Exists and } A_{0i} < 1 \tag{D.3}$$

$$F_{Si} = \frac{1}{S}\sum_{s=1}^{S} 1[p_i(s) < t|H_{0i}(s)] \underset{S\to\infty}{\overset{a.s.}{\to}} A_{0i}F_i(t), F_i(t) \leq t \forall t \in (0,1] \tag{D.4}$$

$$G_{Si} = \frac{1}{S}\sum_{s=1}^{S} 1[p_i(s) < t|H_{1i}(s)] \underset{S\to\infty}{\overset{a.s.}{\to}} (1 - A_{0i})G_i(t) \quad \forall t \in (0,1] \tag{D.5}$$

where S_{0i} is the number of null locations in map i.

The threshold in the BH procedure is

$$t_S^* = \sup\{t : \frac{t}{F_S(t) + G_S(t)} \leq q\},$$

where $F_S = \frac{1}{S}\sum_{s=1}^{S} 1[p^{u/n}(s) < t|H_{0s}^{u/n}]$ and $G_S = \frac{1}{S}\sum_{s=1}^{S} 1[p^{u/n}(s) < t|H_{1s}^{u/n}]$. It controls the FDR asymptotically at level q (as $S \to \infty$) for any valid pooled p-value (i.e. not only using equation (4) or (5), but also using (2) when valid or (3) if conditions (D.3)-(D.5) hold, and $\delta \equiv \sup\{t : t/\lim(F_S(t) + G_S(t)) \leq q\} \in (0,1]$:

$$\begin{aligned} FDR &= E(\frac{F_S(t_S^*)}{(F_S(t_S^*) + G_S(t_S^*)) \vee \frac{1}{S}}) = E(\frac{t_S^*}{(F_S(t_S^*) + G_S(t_S^*)) \vee \frac{1}{S}} + \frac{(F_S(t_S^*) - t_S^*)}{(F_S(t_S^*) + G_S(t_S^*)) \vee \frac{1}{S}}) \\ &\leq q + \sup_{t \geq \delta}\{\frac{(F_S(t) - t)}{(F_S(t) + G_S(t)) \vee \frac{1}{S}}\} + I\{t_S^* < \delta\} \end{aligned}$$

From equations (D.4)-(D.5) the second term is asymptotically negative (because these conditions guarantee that the variance of $F_S(t)$ is asymptotically zero, so $\lim F_S(t) \leq t$) , and from the definition of δ the third term is asymptotically zero. It follows that the asymptotic upper bound for the FDR is q.

References

Benjamini, Y. and Yekutieli, Y. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29** **(4)**, 1165–1188.

Efron, B. (1965). Increasing properties of polya frequency functions. *The Annals of Mathematical Statistics* **36**, 272–279.

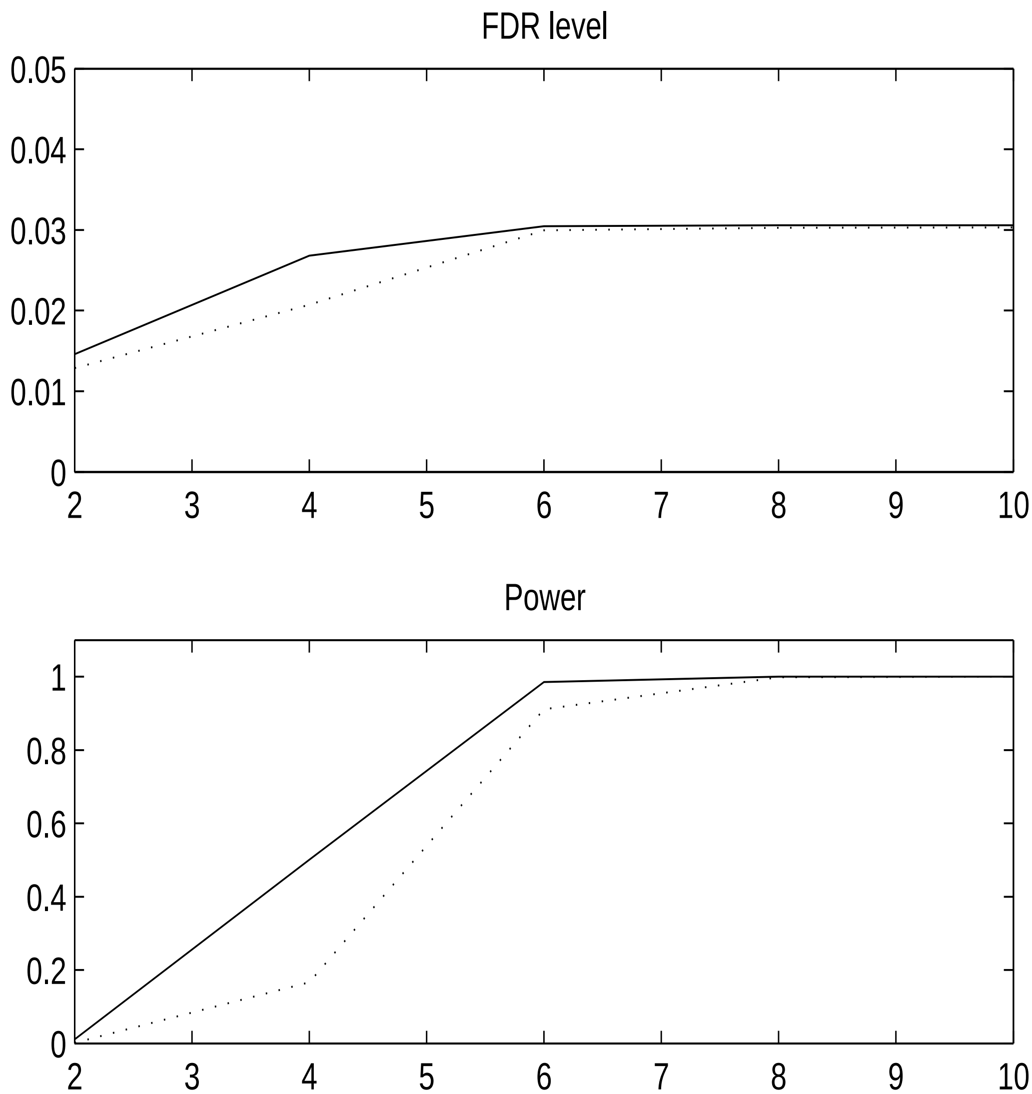

Figure 1. FDR (top) and Power (bottom) as a function of μ for a BH procedure at level 0.05. The simulated setting is that in which the Gaussian spatial filter had a standard deviation of 4, the number of maps was 3 and the number of p-values per location that come from the alternative is either null or 2. We combined the maps using equation (2). The partial conjunction parameter is $u = 1$ (solid line) or $u = 2$ (dotted line). The FDR does not exceed the nominal 0.05 level.

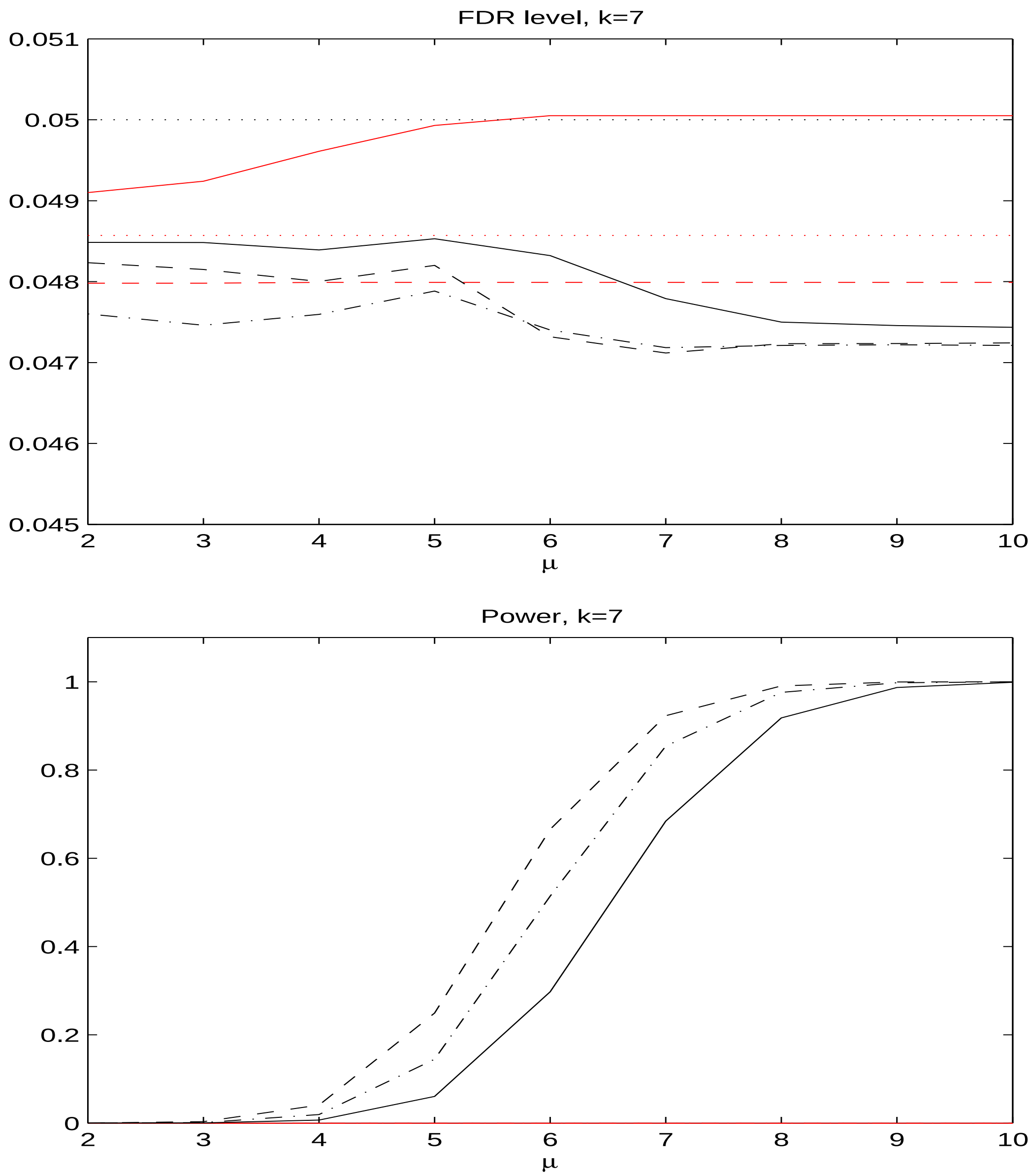

Figure 2. FDR (top) and Power (bottom) as a function of μ for a BH procedure at level 0.05. The simulated setting is that in which the correlation is negative in blocks of 20, the number of maps was 10 and the number of p-values per location that come from the alternative is either null or 7. We combined the maps using equation (2). The partial conjunction parameter is $u = 10$ (red dotted line), $u = 9$ (red dashed line), $u = 8$ (red solid line), $u = 7$ (black solid line), $u = 6$ (dash-dot line) or $u = 5$ (dashed line). The FDR does not exceed the nominal 0.05 level.