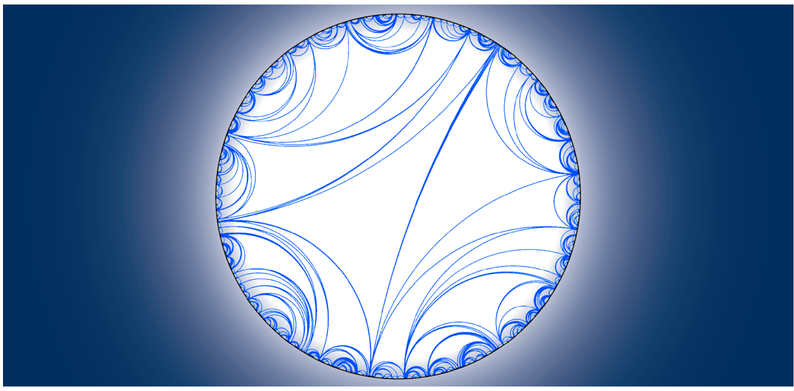
Universität Potsdam



Gilles Blanchard | Sylvain Delattre | Étienne Roquain

Testing over a continuum of null hypotheses

Preprints des Instituts für Mathematik der Universität Potsdam
1 (2012) 1

Preprints des Instituts für Mathematik der Universität Potsdam

Preprints des Instituts für Mathematik der Universität Potsdam
1 (2012) 1

Gilles Blanchard | Sylvain Delattre | Étienne Roquain

Testing over a continuum of null hypotheses

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.de abrufbar.

Universitätsverlag Potsdam 2012
http://info.ub.uni-potsdam.de/verlag.htm

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Preprints des Instituts für Mathematik der Universität Potsdam** wird herausgegeben vom Institut für Mathematik der Universität Potsdam.

ISSN (online) 2193-6943

Kontakt:
Institut für Mathematik
Am Neuen Palais 10
14469 Potsdam
Tel.: +49 (0)331 977 1028
WWW: http://www.math.uni-potsdam.de

Titelabbildungen:
1. Karla Fritze | Institutsgebäude auf dem Campus Neues Palais
2. Nicolas Curien, Wendelin Werner | Random hyperbolic triangulation
Published at: http://arxiv.org/abs/1105.5089

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2012/5687/
URN urn:nbn:de:kobv:517-opus-56877
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-56877

Testing over a continuum of null hypotheses

Gilles Blanchard
Institut für Mathematik, Universität Potsdam

Sylvain Delattre
Laboratoire de Probabilités et Modèles Aléatoires
Université Paris Diderot (Paris 7)

Étienne Roquain
Laboratoire de Probabilités et Modèles Aléatoires
Université Pierre et Marie Curie (Paris 6)

Abstract

We introduce a theoretical framework for performing statistical hypothesis testing simultaneously over a fairly general, possibly uncountably infinite, set of null hypotheses. This extends the standard statistical setting for multiple hypotheses testing, which is restricted to a finite set. This work is motivated by numerous modern applications where the observed signal is modeled by a stochastic process over a continuum.

As a measure of type I error, we extend the concept of false discovery rate (FDR) to this setting. The FDR is defined as the average ratio of the measure of two random sets, so that its study presents some challenge and is of some intrinsic mathematical interest. Our main result shows how to use the p-value process to control the FDR at a nominal level, either under arbitrary dependence of p-values, or under the assumption that the finite dimensional distributions of the p-value process have positive correlations of a specific type (weak PRDS). Both cases generalize existing results established in the finite setting, the latter one leading to a less conservative procedure. The interest of this approach is demonstrated in several non-parametric examples: testing the mean/signal in a Gaussian white noise model, testing the intensity of a Poisson process and testing the c.d.f. of i.i.d. random variables. Conceptually, an interesting feature of the setting advocated here is that it focuses directly on the intrinsic hypothesis space associated with a testing model on a random process, without referring to an arbitrary discretization.

1 Introduction

Multiple testing is a long-established topic of statistics which has known a surge of interest in the past two decades. This renewed popularity is due to a growing range of applications (such as bioinformatics and medical imaging), enabled in particular by modern computational possibilities, through which collecting, manipulating and processing massive amounts of data in very high dimension has become commonplace. Multiple testing is in essence a multiple decision problem: each individual test output is a yes/no (or accept/reject) decision about a particular question (or null hypothesis) concerning the generating distribution of some random observed data, the decision being taken only on the basis on this data.

The standard framework for these problems is to consider a finite family of hypotheses and of associated tests. However, in many cases of interest, it is natural to interpret the observed data as the discretization of an underlying continuous random process; each decision (test) is then associated to one of the discretization points. A first example is that of detecting unusually frequent words in DNA sequences: a classical model is to consider a Poisson model for the (non-overlapping) word occurrence process (Robin, 2002), the observed data being interpreted as a discretized version of this process. A second example is given in the context of medical imaging, where the observed pixelized image can be interpreted as a sampled random process, and the decision to take is, for each pixel, whether the observed value is due to pure noise or reveals some relevant activity (pertaining to this setting, see in particular the work of Perone Pacifico et al., 2004; see also Schwartzman et al., 2010).

As the discretization becomes finer with ever increasing computing power, it is a natural theoretical question to consider whether results concerning the testing of finite families of hypotheses have a suitable continuous counterpart when considered directly for the underlying continuous process. It is this issue that the present paper tries to tackle. More specifically, we focus on the extension to continuous observation (and decision) processes of so-called *step-up* multiple testing procedures, and the control of the (continuous analogue of) their *false discovery rate (FDR)*. The motivation for concentrating on the latter specific criterion is as follows. Thinking of the output of the multiple decision process as a random subset of rejected hypotheses out of the initial set of candidate hypotheses, the FDR is a collective type I error measure defined (in the finite case) as the average proportion of errors in the rejected set. From this definition follows that FDR-controlling procedures scale without becoming trivial as the size of the set of null hypotheses grows; and that the FDR has a natural counterpart in a continuous setting, namely when the ratio of set cardinalities defining the proportion is replaced by a *volume* ratio (for a certain reference finite measure). This natural scaling property of the FDR contrasts with another standard collective type I error measure criterion, the *family-wise error rate* (FWER), defined as the probability of making one error or more in the rejection set. Procedures that control the FWER generally strongly depend on the number of hypotheses to be tested and thus generally become trivial (rejecting nothing) when the discretization become very fine. For this reason, it is more natural to focus on the FDR criterion when the space of hypotheses is infinite (and possibly continuous).

To illustrate the above concepts, let us consider the following example, which will be more developed in Section 4.2.4: observe $(N_t)_{t\in[0,1]}$ a Poisson process with intensity $\lambda : [0,1] \to \mathbb{R}^+$. Assume that we are interested at detecting the times $t \in [0,1]$ such that $\lambda(t)$ exceeds some benchmark intensity λ_0. In statistical terms, we want to test the null hypothesis "$\lambda(t) \le \lambda_0$" against the alternative"$\lambda(t) > \lambda_0$" — and this for all elements t in the inteval $[0,1]$. For each individual t, a statistical test can be easily built (by assuming $\lambda(\cdot)$ smooth enough), that can take the following form: reject "$\lambda(t) \le \lambda_0$" at level α if $p_t \le \alpha$, where p_t is the p-value of the test. The resulting p-value process $(p_t)_{t\in[0,1]}$ is displayed in the bottom of Figure 2, for a particular realization of the Poisson process and by choosing some triangular signal $\lambda(\cdot)$. The gray area indicates the regions where the null hypothesis is true. The goal is to find a threshold (dashed line) such that the corresponding FDR is smaller than a nominal level $\alpha \in (0,1)$. In this continuous setting, the FDR is the mean of the ratio of the measure of the false rejection set (red area) to the measure of the total rejection set (red plus green area).

At this point, it is legitimate to ask whether it is justified at all to make the extra effort of defining and studying explicitly a mathematical model for testing over a continuum. One could argue that in any practical application, the hypotheses space will be effectively discretized, after which step existing theory for finite testing can be applied. To this regard, the motivations of the present work are the following.

- Since in many applications the underlying signal is modeled as a random process on a continuum, it is more coherent to study this object intrinsically; in particular, the FDR error criterion is more naturally defined in its continuous version in that setting, rather than depending on a possibly arbitrary choice of discretization.

- In some cases, such as the Poisson model described above, it turns out that the studied testing procedures take a very simple form in the continuous setting, a property that can be obfuscated by an a priori discretization.

- Finally, a mathematical motivation is to study in a rigorous way some specific points arising from the continuous setting. In particular, since the FDR involves ratios of measures of random subsets of $[0, 1]$, some measurability issues arise, and its control is not straightforward.

The principal contributions of the present work are the following. We define a precise mathematical setting for multiple testing over a (possibly) continuous set of hypotheses, taking particular attention to specific measurability issues. We extend to this setting existing results on the step-up procedure, using for this the tools and analysis developed by Blanchard and Roquain (2008) (a programmatic sketch of the present work can be found in Section 4.4 of the latter paper). In particular, we extend suitably to the continuous setting the notion of positive regressively dependent on a subset (PRDS) condition, which plays a crucial role in the analysis. The latter is a general type of dependency condition on the individual tests' p-values allowing to ensure FDR control. An important difference between the continuous and finite setting is that the continuous case precludes the possibility of independent p-values, which is the simplest reference setting considered in the finite case, so that a more general assumption on dependency structure is necessary (on this point, see the discussion at the end of Section 2.2).

We have tried as much as possible to make this work self-contained, and accessible to readers having little background knowledge in multiple testing. Section 2 of the paper correspondingly introduces the necessary notions with a an angle towards stochastic processes, and some specific examples for the introduced setting. The main result is exposed in Section 3, followed by its applications to the examples introduced in Section 2. The proof for the main theorem is found in Section 4. Extensions and discussions come in Section 5, while some technical results are deferred to the Appendix.

2 Setting

2.1 Multiple testing: mathematical framework

Let X be a random variable defined from a measurable space $(\Omega, \mathfrak{F})$ to some observation space $(\mathcal{X}, \mathfrak{X})$. We assume that there is a family of probability distributions on $(\Omega, \mathfrak{F})$ that induces a subset $\mathcal{P}$ of probability distributions on $(\mathcal{X}, \mathfrak{X})$, which is called the model. The distribution of X on $(\mathcal{X}, \mathfrak{X})$ is denoted by P; for each $P \in \mathcal{P}$ there exists a distribution on $(\Omega, \mathfrak{F})$ for which $X \sim P$; it is referred to as $\mathbb{P}_{X \sim P}$ or simply by $\mathbb{P}$ when unambiguous. The corresponding expectation operator is denoted $\mathbb{E}_{X \sim P}$ or $\mathbb{E}$ for short.

We consider a general multiple testing problem for P, defined as follows. Let $\mathcal{H}$ denote an index space for (null) hypotheses. To each $h \in \mathcal{H}$ is associated a known subset $H_h \subset \mathcal{P}$ of probability measures on $(\mathcal{X}, \mathfrak{X})$. Multiple hypothesis testing consists in taking a decision, based on a single realization of the variable X, of whether for each $h \in \mathcal{H}$ it holds or not that $P \in H_h$ (which is read "P satisfies H_h", or "H_h is true"). We denote by $\mathcal{H}_0(P) := \{h \in \mathcal{H} : P \text{ satisfies } H_h\}$ the set of true null hypotheses, and by its complementary $\mathcal{H}_1(P) := \mathcal{H} \backslash \mathcal{H}_0(P)$ the set of false nulls. These sets are of course unknown because they depend on the unknown distribution P. For short, we will write sometimes $\mathcal{H}_0$ and $\mathcal{H}_1$ instead of $\mathcal{H}_0(P)$ and $\mathcal{H}_1(P)$, respectively.

As an illustration, if we observe a continuous Gaussian process $X = (X_h)_{h\in[0,1]^d}$ with a continuous mean function $\mu : h \in [0,1]^d \mapsto \mu(t) := \mathbb{E}X_t$, then P is the distribution of this process, $(\mathcal{X}, \mathfrak{X})$ is the Wiener space and $\mathcal{P}$ is the set of distributions generated by continuous Gaussian processes having a continuous mean function. Typically, $\mathcal{H} = [0,1]^d$ and, for any h, we choose H_h equal to the set of distributions in $\mathcal{P}$ for which the mean function μ satisfies $\mu(h) \leq 0$. This is usually denoted H_h: "$\mu(h) \leq 0$". Finally, the set $\mathcal{H}_0(P) = \{h \in [0,1]^d : \mu(h) \leq 0\}$ corresponds to the true null hypotheses. Several other examples are provided below in Section 2.4.

Next, for a more formal definition of a multiple testing procedure, we first assume the following:

(A1) The index space $\mathcal{H}$ is endowed with a σ-algebra $\mathfrak{H}$ and for all $P \in \mathcal{P}$, the set $\mathcal{H}_0(P)$ of true nulls is assumed to be measurable, that is, $\mathcal{H}_0(P) \in \mathfrak{H}$.

Definition 2.1 (Multiple testing procedure)**.** Assume **(A1)** holds. A multiple testing procedure on $\mathcal{H}$ is a function $R : X(\Omega) \subset \mathcal{X} \to \mathfrak{H}$ such that the set

$$\{(\omega, h) \in \Omega \times \mathcal{H} : h \in R(X(\omega))\}$$

is a $\mathfrak{F} \otimes \mathfrak{H}$-measurable set; or in other terms, that the process $(\mathbf{1}\{h \in R(X)\})_{h\in\mathcal{H}}$ is a *measurable process.*

The fact that R need only be defined on the image $X(\Omega)$, rather than on the full space $\mathcal{X}$, is a technical detail necessary for later coherence; this introduces no restriction since R will only be ever applied to possible observed values of X.

A multiple testing procedure R is interpreted as follows: based on the observation $x = X(\omega)$, $R(x)$ is the set of null hypotheses that are deemed to be false, also called set of *rejected* hypotheses. The set $\mathcal{H}_0(P) \cap R(x)$ formed of true null hypotheses that are rejected in error is called the set of *type I errors*. Similarly, the set $\mathcal{H}_1(P) \cap R^c(x)$ is that of *type II errors*.

2.2 The p-value functional and process

We will consider a very common framework for multiple testing, where the decision for each null hypothesis $H_h, h \in \mathcal{H}$, is taken based on a scalar statistic $p_h(x) \in [0,1]$ called a p-value. The characteristic property of a p-value statistic is that if the generating distribution P is such that the corresponding null hypothesis is true (i.e. $h \in \mathcal{H}_0(P)$), then the random variable $p_h(X)$ should be stochastically lower bounded by a uniform random variable. Conversely, this statistic is generally constructed in such a way that if the null hypothesis H_h is false, its distribution will be more concentrated towards the value 0. Therefore, a p-value close to 0 is interpreted as evidence from the data against the validity of the null hypothesis, and one will want to reject hypotheses having lower p-values. Informally speaking, based on observation x, the construction of a multiple testing procedure generally proceeds as follows:

(i) compute the p-value $p_h(x)$ for each individual null index $h \in \mathcal{H}$.

(ii) determine a threshold $t_h(x)$ for each $h \in \mathcal{H}$, depending on the whole family $(p_h(x))_{h\in\mathcal{H}}$.

(iii) put $R(x) = \{h \in \mathcal{H} : p_h(x) \le t_h(x)\}$.

To summarize, the rejection set consists of hypotheses whose p-values are lower than a certain threshold, this threshold being itself random, depending on the observation x and possibly depending on h. This will be elaborated in more detail in Section 3.2, in particular how the threshold function $t_h(x)$ is chosen. For now, we focus on properly defining the p-value functional itself, the associated process, and the assumptions we make on them.

Formally, we define the *p-value functional* as a mapping $\mathbf{p} : \mathcal{X} \to [0,1]^{\mathcal{H}}$, or equivalently as a collection of functions $\mathbf{p} = (p_h(x))_{h\in\mathcal{H}}$, each of the functions $p_h : \mathcal{X} \to [0,1]$, $h \in \mathcal{H}$, being considered as a scalar statistic that can be computed from the observed data $x \in \mathcal{X}$.

We will consider correspondingly the random p-values $\omega \in \Omega \mapsto p_h(X(\omega))$, and p-value process $\omega \in \Omega \mapsto \mathbf{p}(X(\omega))$. With some notation overload, we will sometimes drop the dependency on X and use the notation p_h and $\mathbf{p}$ to also denote the *random variables* $p_h(X)$ and $\mathbf{p}(X)$ (the meaning – function of x, or random variable on Ω – should be clear from the context).

We shall make the following assumptions on the p-value process:

- Joint measurability over $\Omega \times \mathcal{H}$: we assume that the random process $(p_h(X))_{h\in\mathcal{H}}$ is a measurable process, that is:

$$(\omega, h) \in \big(\Omega \times \mathcal{H}, \mathfrak{F} \otimes \mathfrak{H}\big) \mapsto p_h(X(\omega)) \in [0,1] \text{ is (jointly) measurable.} \qquad \textbf{(A2)}$$

- For any $P \in \mathcal{P}$, the marginal distributions of the p-values corresponding to true nulls are stochastically lower bounded by a uniform random variable on $[0,1]$:

$$\forall h \in \mathcal{H}_0(P), \quad \forall u \in [0,1], \quad \mathbb{P}_{X\sim P}(p_h(X) \le u) \le u. \qquad \textbf{(A3)}$$

  (The distribution of p_h if h lies in $\mathcal{H}_1(P)$ can be arbitrary).

Condition **(A2)** is specific to the continuous setting considered here and will be discussed in more detail in the next section. Condition **(A3)** is the standard characterization of a single p-value statistic in classical (single or multiple) hypothesis testing. In general, an arbitrary scalar statistic used to take the rejection decision on hypothesis H_h can be monotonically normalized into a p-value as follows: assume $S_h(x)$ is a scalar test statistic, then

$$p_h(x) = \sup_{P\in H_h} F_{h,P}(S_h(x))$$

is a p-value in the sense of (**A3**), where $F_{h,P}(t) = \mathbb{P}_{X\sim P}(S_h(X) \ge t)$ (and where the supremum is assumed to maintain the measurability in x, for any fixed h). If the scalar statistic $S_h(x)$ is constructed so that it tends to be stochastically larger when hypothesis H_h is false, the corresponding p-value indeed has the desirable property that it is more concentrated towards 0 in this case. Such test statistics abound in the (single) testing literature, and a few examples will be given below.

2.3 Discussion on measurability assumptions

Since the focus of the present work is to be able to deal with uncountable spaces of hypotheses $\mathcal{H}$, we have to be somewhat careful with corresponding measurability assumptions over $\mathcal{H}$ (a problem that does not arise when $\mathcal{H}$ is finite or countable). The main assumption needed to this regard in order to state properly the results to come is the joint measurability assumption appearing in either Definition 2.1 (for the multiple testing procedure) or in (**A2**) (for the p-value process), both of which are specific to the uncountable setting. Essentially, joint measurability will be necessary in order to use Fubini's theorem on the space $\big(\Omega \times \mathcal{H}, \mathfrak{F} \otimes \mathfrak{H}\big)$, and have the expectation operator w.r.t. ω and the integral operator over $\mathcal{H}$ commute.

If $\mathcal{H}$ has at most countable cardinality, and is endowed with the trivial σ-field comprising all subsets of $\mathcal{H}$, then **(A2)** is automatically satisfied whenever all individual p-value functions $p_h : \mathcal{X} \to [0,1]$, $h \in \mathcal{H}$, are separately measurable, which is the standard setting in multiple testing.

If $\mathcal{H}$ is uncountable, a sufficient condition ensuring (**A2**) is the joint measurability of the p-value *functional*,

$$(x,h) \in \big(\mathcal{X} \times \mathcal{H}, \mathfrak{X} \otimes \mathfrak{H}\big) \mapsto p_h(x) \in [0,1] \text{ is (jointly) measurable,} \qquad \textbf{(A2')}$$

which implies (**A2**) by composition. Unfortunately, (**A2'**) might not always hold. To see this, consider the following canonical example. Assume the observation takes the form of a stochastic process indexed by the hypothesis space itself, $X = \{X_h, h \in \mathcal{H}\}$. In this case, the observation space $\mathcal{X}$ is included in $\mathbb{R}^{\mathcal{H}}$. Furthermore, assume the p-value function $p_h(x)$ is given by a fixed measurable mapping ψ of the value of x at point h, i.e. $p_h(x) = \psi(x_h)$, $\forall h \in \mathcal{H}$. In this case assumption (**A2'**) boils down to the joint measurability of the evaluation mapping $(x,h) \in \mathcal{X} \times \mathcal{H} \mapsto x_h$. Whether this holds depends on the nature of the space $\mathcal{X}$. We give some classical examples in the next section where the assumption holds; for example, it is true if $\mathcal{X}$ is the Wiener space.

However, the joint measurability of the evaluation mapping does not hold if $\mathcal{X}$ is taken to be the product space $\mathbb{R}^{\mathcal{H}}$ endowed with the canonical product σ-field (indeed, this would imply that any $x \in \mathbb{R}^{\mathcal{H}}$, i.e., any function from $\mathcal{H}$ into $\mathbb{R}$, is measurable). The more general assumption (**A2**) may still hold, though, but it generally requires some additional regularity or structural assumptions on the paths of the process X. In particular, in the above example if $X = \{X_h, h \in \mathcal{H}\}$ is a stochastic process having a (jointly) measurable modification (and more generally for other examples, if there exists a modification of X such that (**A2**) is satisfied), we will always assume that we observe such a modification, so that assumption (**A2**) holds.

We have gathered in Appendix A some auxiliary (mostly classical) results related to the existence and properties of such modifications. Lemma A.2 shows that such a (jointly) measurable modification exists as soon as the process is continuous in probability. The latter is not an iff condition, but is certainly much weaker than having continuous paths.

On the other hand, it is important to observe here that a jointly measurable modification of X, or, for that matter, of the p value process, might not exist. Lemma A.1 reproduces a classical argument showing that for $\mathcal{H} = [0,1]$, assumption (**A2**) is violated for any modification of a mutually independent p-value process. Therefore, for an uncountable space of hypotheses $\mathcal{H}$, assumption (**A2**) precludes the possibility that the p-values $\{p_h, h \in \mathcal{H}\}$ are mutually independent. This contrasts strongly with the situation of a finite hypothesis set $\mathcal{H}$, where mutual independence of the p-values is generally considered the reference case.

A final issue is to which extent the results exposed in the remainder of this work depend on the (jointly) measurable modification chosen for the underlying stochastic process. Lemma A.4 elucidates this issue by showing that this is not the case, because the FDR (the main measure of type I error, which will be formally defined in Section 3.1) is identical for two such modifications.

2.4 Examples

To illustrate the above generic setting, let us consider the following examples.

Example 2.2 (Testing the mean of a process). Assume that we observe the realization of a real-valued process $X = (X_t)_{t\in[0,1]^d}$ with an unknown (measurable) mean function $\mu : t \in [0,1]^d \mapsto \mu(t) := \mathbb{E}X_t$. We take $\mathcal{H} = [0,1]^d$ and want to test simultaneously for each $t \in [0,1]^d$ the null hypothesis H_t : "$\mu(t) \leq 0$". Assume that for each t the marginal distribution of $(X_t - \mu(t))$ is known, does not depend on t and has upper-tail function G (for instance, X is a Gaussian process with marginals $X_t \sim \mathcal{N}(\mu(t), 1)$). We correspondingly define the p-value process $\forall t \in [0,1]^d$, $p_t(X) = G(X_t)$, which satisfies (**A3**). Next, the measurability assumption (**A2**) follows from a regularity assumption on X:

- if we assume that the process X has continuous paths, $X : \omega \mapsto (X_t(\omega))_t$ can be seen as taking values in the Wiener space $\mathcal{X} = \mathcal{C}_{[0,1]^d} = C([0,1]^d, \mathbb{R})$ of continuous functions from $[0,1]^d$ to $\mathbb{R}$. (In this case, the Borel σ-field corresponding to the supremum norm topology on $\mathcal{C}_{[0,1]^d}$ is the trace of the product σ-field on $\mathcal{C}_{[0,1]^d}$, and X is measurable iff all its coordinate projections are.) Furthermore, the p-value function can be written as

  $$(x, t) \in \mathcal{C}_{[0,1]^d} \times [0,1]^d \mapsto p_t(x) = G(x(t)) \in [0,1].$$

  The evaluation functional $(x, t) \in \mathcal{C}_{[0,1]^d} \times [0,1]^d \mapsto x(t)$ is jointly measurable because it is continuous, thus $p_t(x)$ is jointly measurable by composition and (**A2'**) holds, hence also (**A2**).

- if $d = 1$ and the process X is càdlàg , the random variable X can be seen as taking values in the Skorohod space $\mathcal{X} = \mathcal{D} := D([0,1], \mathbb{R})$ of càdlàg functions from $[0,1]$ to $\mathbb{R}$. In this case, the Borel σ-field generated by the Skorohod topology is also the trace of the product σ-field on $\mathcal{D}$ (see, e.g., Theorem 14.5 p.121 of Billingsley, 1999). Moreover, the evaluation functional $(x, t) \mapsto x(t)$ is jointly measurable, as for any càdlàg funtion x, it is the pointwise limit of the jointly measurable functions ζ_n: $(x, t) \mapsto \zeta_n(x, t) := \sum_{k=1}^{2^n} x(k2^{-n})\mathbf{1}\{(k-1)2^{-n} \leq t < k2^{-n}\} + x(1)\mathbf{1}\{t = 1\}$, therefore (**A2'**) is fulfilled by composition, hence also (**A2**).

- assume that X is a Gaussian process defined on the space $\mathcal{X} = \mathbb{R}^{[0,1]^d}$ endowed with the canonical product σ-field, and with a covariance function $\Sigma(t, t')$ such that Σ is continuous on all points (t, t) of the diagonal and takes a contant (known) value σ^2 on those points.

  This assumption is not sufficient to ensure that X has a continuous version, but it ensures that (X_t) is continuous in L^2 and hence in probability; Lemma A.2 then states that X has a modification such that the evaluation functional is jointly measurable. Assuming that such a jointly measurable modification is observed, we deduce that (**A2**) holds for the associated p-value process.

Example 2.3 (Testing the signal in a Gaussian white noise model). Let us consider the Gaussian white noise model $dZ_t = f(t)dt + \sigma dB_t$, $t \in [0,1]$, where B is a Wiener process on $[0,1]$ and $f \in C([0,1])$ is a continuous signal function. For simplicity, the standard deviation σ is assumed to be equal to 1. Equivalently, we assume that we can observe the stochastic integral of Z_t against any test function in $L^2([0,1])$, that is, that we observe the Gaussian process $(X_g)_{g\in L^2([0,1])}$ defined by

$$X_g := \int_0^1 g(t)f(t)dt + \int_0^1 g(t)dB_t, \;\; g \in L^2([0,1]).$$

Formally, the observation space is the whole space $\mathcal{X} = \mathbb{R}^{L^2([0,1])}$, endowed with the product σ-field. However, in the sequel, we will use the observation of the process X only against a "small" subspace of functions of $L^2([0,1])$.

Let us consider $\mathcal{H} = [0,1]$ and the problem of testing for each $t \in \mathcal{H}$, the null H_t : "$f(t) \leq 0$" (signal nonpositive). We can build p-values based upon a kernel estimator in the following way. Consider a kernel function $K \in L^2(\mathbb{R})$, assumed positive on $[-1,1]$ and zero elsewhere, and denote by $K_t \in L^2([0,1])$ the function $K_t(s) := K((t-s)/\eta)$, where $0 < \eta \leq 1$ is a bandwidth to be chosen. Let us consider the process $\widetilde{X}_t := X_{K_t}, t \in [0,1]$. From Lemma A.3, $\widetilde{X}$ has a modification which is jointly measurable in (ω, t). Clearly, this implies that there exists a modification of the original process X such that $\widetilde{X}$ is jointly measurable in (ω,t), and we assume that we observe such a modification. For any $t \in [0,1]$, letting $c_{K,t} := \int_0^1 K((t-s)/\eta)ds > 0$ and $v_{K,t} := \int_0^1 K^2((t-s)/\eta)ds \geq c_{K,t}^2 > 0$, we can consider the following standard estimate of $f(t)$:

$$\begin{aligned}\widehat{f}_\eta(t) &:= c_{K,t}^{-1} X_{K_t} \\ &= c_{K,t}^{-1}\int_0^1 K\left(\frac{t-s}{\eta}\right) f(s)ds + c_{K,t}^{-1}\int_0^1 K\left(\frac{t-s}{\eta}\right) dB_s. \end{aligned} \tag{1}$$

Assume that there is a known $\delta_{t,\eta} > 0$ such that for any t with $f(t) \leq 0$, we have the upper-bound

$$\mathbb{E}\widehat{f}_\eta(t) = c_{K,t}^{-1}\int_0^1 K\left(\frac{t-s}{\eta}\right) f(s)ds \leq \delta_{t,\eta}. \tag{2}$$

For instance, this holds if we can assume a priori knowledge on the regularity of f, of the form $\sup_{s:|s-t|\leq\eta} |f(s) - f(t)| \leq \delta_{t,\eta}$. Then, the statistics $(\widehat{f}_\eta(t))_t$ can be transformed into a p-value process in the following way:

$$p_t(X) = \overline{\Phi}\left(\frac{\widehat{f}_\eta(t) - \delta_{t,\eta}}{v_{K,t}^{1/2}/c_{K,t}}\right), \tag{3}$$

where $\overline{\Phi}(w) := \mathbb{P}(W \geq w)$, $W \sim \mathcal{N}(0,1)$, is the upper tail distribution of a standard Gaussian distribution. The p-value process (3) satisfies (**A3**), because for any t with $f(t) \leq 0$ and any $u \in [0,1]$,

$$\begin{aligned}\mathbb{P}(p_t(X) \leq u) &= \mathbb{P}(\widehat{f}_\eta(t) - \delta_{t,\eta} \geq v_{K,t}^{1/2}/c_{K,t}\overline{\Phi}^{-1}(u)) \\ &\leq \mathbb{P}(c_{K,t}(\widehat{f}_\eta(t) - \mathbb{E}\widehat{f}_\eta(t))/v_{K,t}^{1/2} \geq \overline{\Phi}^{-1}(u)) \\ &= u,\end{aligned}$$

because $\int_0^1 K_t(s)dB_s \sim \mathcal{N}(0, v_{K,t})$. Moreover, the p-value process (3) satisfies (**A2**), since we assumed $(X_{K_t})_t \in [0,1]$ to be jointly measurable in (ω, t).

Example 2.4 (Testing the c.d.f.). Let $X = (X_1, \dots, X_m) \in \mathcal{X} = \mathbb{R}^m$ be a m-uple of i.i.d. real random variables of common continuous c.d.f. F. For $\mathcal{H} = I$ an interval of $\mathbb{R}$ and a given benchmark c.d.f. F_0, we aim to test simultaneously for all $t \in I$ the null H_t : "$F(t) \leq F_0(t)$". The individual hypothesis H_t may be tested using the p-value

$$p_t(X) = G_t(m\mathbb{F}_m(X,t)), \tag{4}$$

where $\mathbb{F}_m(X,t) = m^{-1}\sum_{i=1}^m \mathbf{1}\{X_i \leq t\}$ is the empirical c.d.f. of $X_1, \dots, X_m$ and where $G_t(k) = \mathbb{P}[Z_t \geq k]$, $Z_t \sim \mathcal{B}(m, F_0(t))$, is the upper-tail function of a binomial distribution of parameter $(m, F_0(t))$. The conditions (**A2**) and (**A3**) are both clearly satisfied.

Figure 1 provides a realization of the p-value process (4) when testing for all $t \in [0,1]$ the null H_t : "$F(t) \leq t$" when F comes from a mixture of beta distributions. The correct/erroneous rejections are also pictured for the simple procedure $R(X) = \{t \in [0,1] : p_t(X) \leq 0.4\}$.

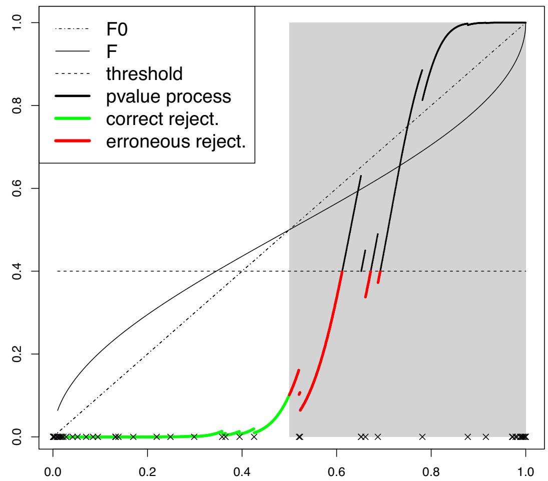


Figure 1: Plot of a realization of the p-value process as defined in (4) for the c.d.f. testing, together with F_0 and F, for $F_0(t) = t$ and $F(t) = 0.5F_1(t) + 0.5F_2(t)$, where F_1 (resp. F_2) is the c.d.f. of a beta distribution of parameter $(0.5, 1.5)$ (resp. $(1.5, 0.5)$). The region where the null hypothesis "$F(t) \leq F_0(t)$" is true is depicted in gray color. The crosses correspond to the elements of $\{X_i, 1 \leq i \leq m\}$; $m = 50$. The correct/erroneous rejections refer to the procedure $R(X) = \{t \in [0,1] : p_t(X) \leq 0.4\}$ using the threshold 0.4.

Example 2.5 (Testing the intensity of a Poisson process). Assume we observe $(N_t)_{t \in [0,1]} \in \mathcal{X} = D([0,1], \mathbb{R})$ a Poisson process with intensity $\lambda : [0,1] \to \mathbb{R}^+ \in L^1(d\Lambda)$, where Λ denotes the Lebesgue measure on $[0,1]$. For each $t \in [0,1]$, we aim to test H_t : "$\lambda(t) \leq \lambda_0(t)$" where $\lambda_0(\cdot) > 0$ is a given benchmark intensity. Assume that for a given bandwidth $\eta \in (0,1]$, there is a known upper bound $\delta_{t,\eta}$ for $\int_{(t-\eta)\vee 0}^{(t+\eta)\wedge 1} \lambda(s)ds$ that holds true for any t such that $\lambda(t) \le \lambda_0(t)$. For instance, we can choose $\delta_{t,\eta} = ((t+\eta)\wedge 1 - (t-\eta)\vee 0)(\lambda_0(t) + \sup_{s:|t-s|\le\eta} |\lambda(t) - \lambda(s)|)$ (assuming knowledge on the regularity of λ is available a priori). For any $t \in [0,1]$, the variable $N_{(t+\eta)\wedge 1} - N_{(t-\eta)\vee 0}$ follows a Poisson variable of parameter $\int_{(t-\eta)\vee 0}^{(t+\eta)\wedge 1} \lambda(s)ds$. Since the latter parameter is smaller than $\delta_{t,\eta}$ as soon as $\lambda(t) \le \lambda_0(t)$, the following p-value process satisfies (**A3**):

$$p_t(X) = G_t\big(N_{(t+\eta)\wedge 1} - N_{(t-\eta)\vee 0}\big), \tag{5}$$

where for any $k \in \mathbb{N}$, $G_t(k)$ denotes $\mathbb{P}[Z \ge k]$ for Z a Poisson distribution of parameter $\delta_{t,\eta}$. Moreover, the p-value process fulfills condition (**A2'**), because (N_t) is a càdlàg process, so that arguments similar to those of Example 2.2 apply. Thus (**A2**) also holds.

3 Main concepts and tools

3.1 False discovery rate

Following the usual philosophy of hypothesis testing, the first thing one wants to ensure is some form of control over type I errors commited by the procedure. For multiple testing, there are different scalar criteria available to assess the amount of type I errors over the whole family of hypotheses. In the present work we focus on a generalization to a continuum of hypotheses of the false discovery rate (FDR). We choose this criterion for two reasons: first, it has now become a widely used standard after its introduction by Benjamini and Hochberg (1995) (see also Seeger, 1968); secondly, it has a natural extension to an uncountable spaces of hypotheses. For a finite number of null hypotheses, the FDR is defined as the average proportion of type I errors in the set of all rejected hypotheses. To extend this definition to a possibly uncountable space, we quantify this proportion by a volume ratio, defined with respect to a finite measure Λ on $(\mathcal{H}, \mathfrak{H})$ (the usual definition over a finite space is recovered by taking Λ equal to the counting measure).

Definition 3.1 (False discovery proportion, false discovery rate)**.** Let Λ be a finite positive measure on $(\mathcal{H}, \mathfrak{H})$. Let R be a multiple testing procedure on $\mathcal{H}$. The false discovery rate (FDR) of R is defined as the average of the the false discovery proportion (FDP):

$$\forall P \in \mathcal{P}\,,\;\; \forall x \in X(\Omega)\,,\;\; \mathrm{FDP}(R(x), P) := \frac{\Lambda\big(R(x) \cap \mathcal{H}_0(P)\big)}{\Lambda\big(R(x)\big)} \mathbf{1}\{\Lambda\big(R(x)\big) > 0\}, \tag{6}$$

and

$$\forall P \in \mathcal{P}\,,\;\; \mathrm{FDR}(R, P) := \mathbb{E}_{X\sim P}\left[\mathrm{FDP}(R(X), P)\right]. \tag{7}$$

The indicator function in (6) means that the ratio is taken equal to zero whenever the denominator is zero. Observe that, due to the joint measurability assumption in definition 2.1 of a multiple testing procedure, both of the above quantities are well-defined (the FDP is only formally defined over the image of Ω through X since only on this set is the measurability of $R(x)$ guaranteed by the definition. In particular, it is defined for P-almost all $x \in \mathcal{X}$).

As illustration, in the particular realization of the p-value process pictured in Figure 1, if we denote by "Red" (resp. "Green") the length of the interval corresponding to the projection of the red (resp. green) part of the p-value process on the X-axis, the FDP of the procedure

$R(X) = \{t \in [0,1] \,:\, p_t(X) \le 0.4\}$ is Red/(Red + Green). A similar interpretation for the FDP holds in Figure 2.

Finding a procedure R with a FDR smaller than or equal to α has the following interpretation: on average, the volume proportion of type I errors among the rejected hypotheses is smaller than α. This means that the procedure is allowed to reject in error some true nulls but in a small (average) proportion among the rejections. For a pre-specified level α, the goal is then to determine multiple testing procedures R such that for any $P \in \mathcal{P}$, it holds that $\mathrm{FDR}(R,P) \le \alpha$. (In fact, the statement need only hold for $P \in \mathcal{P} \cap \bigcup_{h\in\mathcal{H}} H_h$, since outside of this set $\mathcal{H}_0(P) = \emptyset$ and the FDR is 0.) The rest of the paper will concentrate on establishing sufficient conditions under which the FDR is controlled at a fixed level α. Under this constraint, in order to get a procedure with good power properties (that is, low type II error), it is, generally speaking, desirable that R rejects as many nulls as possible, that is, has volume $\Lambda(R)$ as large as possible.

3.2 Step-up procedures

In what follows, we will focus on a particular form of multiple testing procedures which can be written as function of the p-value family $\mathbf{p}(x) = (p_h(x))_{h\in\mathcal{H}}$.

First, we define a parametrized family of possible rejection sets having the following form: for a given *threshold function* $\Delta : (h,r) \in \mathcal{H} \times \mathbb{R}^+ \mapsto \mathbb{R}^+$, we define for any $r \ge 0$ the sub-level set

$$\forall x \in \mathcal{X}\,, \quad L_\Delta(x,r) := \{h \in \mathcal{H} \,:\, p_h(x) \le \Delta(h,r)\} \subset \mathcal{H}. \tag{8}$$

For short, we sometimes write $L_\Delta(r)$ instead of $L_\Delta(x,r)$ when unambiguous. We will more particularly focus on threshold functions Δ of the product form $\Delta(h,r) = \alpha\pi(h)\beta(r)$, where $\alpha \in (0,1)$ is a positive scalar (*level*), $\pi : \mathcal{H} \to \mathbb{R}^+$ is measurable (*weight function*), and $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ is nondecreasing and right-continuous (*shape function*). Clearly this decomposition is not unique, but will be practical for the formulation of the main result.

Given a threshold function Δ of the above form, we will be interested in a particular, data-dependent choice of the parameter r determining the rejection set, called *step-up procedure*.

Definition 3.2 (Step-up procedure)**.** Let $\Delta(h,r) = \alpha\pi(h)\beta(r)$ a threshold function with $\alpha \in (0,1)$; $\pi : \mathcal{H} \to \mathbb{R}^+$ measurable and $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ nondecreasing and right-continuous. Then the *step-up multiple testing procedure* R on $(\mathcal{H},\Lambda)$ associated to Δ, is defined by

$$\forall x \in X(\Omega)\,, \quad R(x) = L_\Delta(x, \widehat{r}(x))\,, \text{ where } \widehat{r}(x) := \max\{r \ge 0 \,:\, \Lambda(L_\Delta(x,r)) \ge r\}\,. \tag{9}$$

Note that $\widehat{r}$ above is well-defined: first, since $x \in X(\Omega)$ and from assumption (**A2**), the function $h \mapsto p_h(x) - \alpha\pi(h)\beta(r)$ is measurable; thus $L_\Delta(x,r)$ is a measurable set of $\mathcal{H}$, which in turn implies that $\Lambda(L_\Delta(x,r))$ is well-defined. Secondly, the supremum of $\{r \ge 0 \,:\, \Lambda(L_\Delta(x,r)) \ge r\}$ exists because $r = 0$ belongs to this set and $M = \Lambda(\mathcal{H})$ is an upper bound. Third, this supremum is a maximum because the function $r \mapsto \Lambda(L_\Delta(x,r))$ is nondecreasing (right-continuity is not needed for this).

We should ensure in Definition 3.2 that a step-up procedure satisfies the measurability requirements of Definition 2.1. This is proved separately in Section 5.2. In that section, we also check that the equality $\Lambda(L_\Delta(x,\widehat{r}(x))) = \widehat{r}(x)$ always holds. Hence, $\widehat{r}(x)$ is the largest intersection point between the function $r \mapsto \Lambda(L_\Delta(x,r))$ giving the volume of the candidate rejection sets as a function of r, and the identity line $r \mapsto r$.

To give some basic intuition behind the principle of a step-up procedure, consider for simplicity that π is a constant function, so that the family defined by (8) are ordinary sub-level sets of the p-value family. The goal is to find a suitable common rejection threshold t giving rise to rejection set R_t. Assume also without loss of generality that $\Lambda(\mathcal{H}) = 1$. Now consider the following heuristic. If the threshold t is deterministic, any p-value associated to a true null hypothesis, being stochastically lower bounded by a uniform variable, has probability less than t of being rejected in error. Thus, we expect on average a volume $t\Lambda(\mathcal{H}_0) \le t$ of erroneously rejected null hypotheses. If we therefore use t as a rough upper bound of the numerator in the definition (6) of the FDP or FDR, and we want the latter to be less than α, we obtain the constraint $t/\Lambda(R_t) \le \alpha$, or equivalently $\Lambda(R_t) \ge \alpha^{-1}t$. Choosing the largest t satisfying this heuristic constraint is equivalent to the step-up procedure wherein $\beta(u) = u$. The choice of a different shape function with $\beta(u) \le u$ can be interpreted roughly as a pessimistic discount to compensate for various inaccuracies in the above heuristic argument (in particular the fact that the obtained threshold is really a random quantity).

In the case where $\mathcal{H}$ is finite and Λ is the counting measure, it can be seen that the above definition recovers the usual notion of step-up procedures (see, e.g., Blanchard and Roquain, 2008); in particular, the linear shape function $\beta(u) = u$ gives rise to the celebrated linear step-up procedure of Benjamini and Hochberg (1995).

3.3 PRDS conditions

To ensure control of the FDR criterion, an important role is played by structural assumptions on the dependency of the p-values. While the case of independent p-values is considered as the reference setting in the case where $\mathcal{H}$ is finite, we recall that for an uncountable set $\mathcal{H}$, we cannot assume mutual independency of the p-values since this would contradict our measurability assumptions (see concluding discussion of Section 2.3).

We will consider two different situations in our main result: first, if the dependency of the p-values can be totally arbitrary, and secondly, if a form of positive dependency is assumed. This is the latter condition which we define more precisely now. We consider a generalization to the case of infinite, possibly uncountable space $\mathcal{H}$, of the notion of positive regression dependency on each one from a subset (PRDS) introduced by Benjamini and Yekutieli (2001) in the case of a finite set of hypotheses. For any finite set $\mathcal{I}$, a subset $D \subset [0,1]^{\mathcal{I}}$ is called *nondecreasing* if for all $\mathbf{z}, \mathbf{z}' \in [0,1]^{\mathcal{I}}$ such that $\mathbf{z} \le \mathbf{z}'$ (i.e. $\forall h \in \mathcal{I}, z_h \le z'_h$), we have $\mathbf{z} \in D \Rightarrow \mathbf{z}' \in D$.

Definition 3.3. (PRDS conditions for a finite p-value family) Assume $\mathcal{H}$ to be finite. For $\mathcal{H}'$ a subset of $\mathcal{H}$, the p-value family $\mathbf{p}(X) = (p_h(X))_{h\in\mathcal{H}}$ is said to be *weak PRDS on* $\mathcal{H}'$ for the distribution P, if for any $h \in \mathcal{H}'$, for any measurable nondecreasing set D in $[0,1]^{\mathcal{H}}$, the function $u \in [0,1] \mapsto \mathbb{P}(\mathbf{p}(X) \in D \mid p_h(X) \le u)$ is nondecreasing on $\{u \in [0,1] : \mathbb{P}(p_h(X) \le u) > 0\}$; it is said to be *strong PRDS* if the function $u \mapsto \mathbb{P}(\mathbf{p}(X) \in D \mid p_h(X) = u)$ is nondecreasing.

To be completely rigorous, observe that the conditional probability with respect to the event $\{p_h(X) \le u\}$ is defined pointwise unequivocally whenever this event has positive probability, using a ratio of probabilities; while the conditional probability with respect to $p_h(X) = u$ can only be defined via conditional expectation, and is therefore only defined up to a $p_h(X)$-negligible set. Hence, in the definition of strong PRDS, strictly speaking, we only require that the conditional probability coincides $p_h(X)$-a.s. with a nondecreasing function.

Definition 3.4. (Finite dimensional PRDS conditions for a p-value process) For $\mathcal{H}'$ a subset of $\mathcal{H}$, the p-value process $\mathbf{p}(X) = (p_h(X))_{h\in\mathcal{H}}$ is said to be *finite dimensional weak PRDS on* $\mathcal{H}'$ (resp. *finite dimensional strong PRDS on* $\mathcal{H}'$) for the distribution P, if for any finite subset $\mathcal{S}$ of $\mathcal{H}$, the finite p-value family $\mathbf{p}_{\mathcal{S}}(X) = (p_h(X))_{h\in\mathcal{H}\cap\mathcal{S}}$ is weak PRDS on $\mathcal{H}'\cap\mathcal{S}$ (resp. strong PRDS on $\mathcal{H}'\cap\mathcal{S}$) for the distribution P.

To state our main result, the finite dimensional weak PRDS property will be sufficient. However, on practical examples it is sometimes easier to establish first the finite dimensional strong PRDS property and use the following lemma.

Lemma 3.5. *The finite dimensional strong PRDS property implies the finite dimensional weak PRDS property.*

Proof. We just have to replace "=" by "≤" in the conditional probability. This can be done using the following standard argument (also used by (Benjamini and Yekutieli, 2001) with a reference to (Lehmann, 1966)). Put $f(u) := \mathbb{P}\left[\mathbf{p}\in D \mid p_h = u\right]$ and let $u \geq 0$ be such that $\mathbb{P}(p_h(X)\leq u) > 0$. For all $u'\geq u$, putting $\gamma = \mathbb{P}\left[p_h \leq u \mid p_h \leq u'\right]$ (which is well-defined by the probability quotient),

$$\begin{aligned}\mathbb{P}\left[\mathbf{p}\in D \mid p_h \leq u'\right] &= \mathbb{E}\left[f(p_h) \mid p_h \leq u'\right]\\ &= \gamma\mathbb{E}\left[f(p_h) \mid p_h \leq u\right] + (1-\gamma)\mathbb{E}\left[f(p_h) \mid u < p_h \leq u'\right]\\ &\geq \mathbb{E}\left[f(p_h) \mid p_h \leq u\right] = \mathbb{P}\left[\mathbf{p}\in D \mid p_h \leq u\right],\end{aligned}$$

where we used in the inequality that f is nondecreasing. □

Finally, Benjamini and Yekutieli (2001) (Section 3.1 therein) proved that the p-value family corresponding to a finite Gaussian random vector are (finite) strong PRDS as soon as all the coefficient of the covariance matrix are non-negative. This equivalently proves the following result:

Lemma 3.6. *Let* $\mathbf{p}(X) = (p_h(X))_{h\in\mathcal{H}}$ *be a p-value process of the form* $p_h(X) = G(X_h)$, $h\in\mathcal{H}$, *where* $X = (X_h)_{h\in\mathcal{H}}$ *is a Gaussian process and where* G *is continuous decreasing from* $\mathbb{R}$ *to* $[0,1]$. *Assume that the covariance function* Σ *of* X *satisfies*

$$\forall h, h'\in\mathcal{H},\ \ \Sigma(h,h')\geq 0. \tag{10}$$

Then the p-value process is finite dimensional strong PRDS (on any subset).

4 Control of the FDR

In this section, our main result is stated and then illustrated with several examples.

4.1 Main result

The following theorem establishes our main result on sufficient conditions to ensure FDR control at a specified level for step-up procedures. It is proved in Section 5.

Theorem 4.1. *Assume that the hypothesis space $\mathcal{H}$ satisfies* (**A1**) *and is endowed with a finite measure Λ. Let $\mathbf{p}(X) = (p_h(X))_{h\in\mathcal{H}}$ be a p-value process satisfying the conditions* **(A2)** *and* **(A3)**. *Denote R the step-up procedure on $(\mathcal{H}, \Lambda)$ associated to a threshold function of the product form $\Delta(h, r) = \alpha\pi(h)\beta(r)$, with $\alpha \in (0, 1)$, β a non-decreasing right-continuous shape function and π a probability density function on $\mathcal{H}$ with respect to Λ. Then for any $P \in \mathcal{P}$, letting $\Pi(\mathcal{H}_0(P)) := \int_{h\in\mathcal{H}_0(P)} \pi(h)d\Lambda(h)$, the inequality*

$$FDR(R, P) \leq \alpha\, \Pi(\mathcal{H}_0(P)) \quad (\leq \alpha) \tag{11}$$

holds in either of the two following cases:

1. *$\beta(x) = x$ and the p-value process $\mathbf{p}$ is finite dimensional weak PRDS on $\mathcal{H}_0(P)$ for the distribution P;*

2. *the function β is of the form*

$$\beta_\nu(x) = \int_0^x u d\nu(u)\,, \tag{12}$$

   where ν is an arbitrary probability distribution on $(0, \infty)$.

Since π is taken as a probability density function on $\mathcal{H}$ with respect to Λ, the FDR in (11) is upper bounded by $\alpha\Pi(\mathcal{H}_0) \leq \alpha\Pi(\mathcal{H}) = \alpha$, so that the corresponding step-up procedure provides FDR control at level α. As an illustration, a typical choice for π is the constant probability density function $\forall h \in \mathcal{H}, \pi(h) = 1/\Lambda(\mathcal{H}) = M^{-1}$.

According to the standard philosophy of (multiple) testing, while the FDR is controlled at level α as in (11), we aim to have a procedure that rejects a volume of hypotheses as large as possible. In that sense, choosing a step-up procedure with $\beta(x) = x$ always leads to a better step-up procedure than choosing $\beta(x)$ of the form (12), because $\int_0^x u d\nu(u) \leq x$. Hence, in Theorem 4.1, the PRDS assumption allows us to get a result which is less conservative (i.e. rejecting more) than under arbitrary dependencies. Therefore, when we want to apply Theorem 4.1, an important issue is to obtain, if possible, the finite dimensional PRDS condition, see the examples of Section 4.2. When the PRDS assumption does not hold, we refer to Blanchard and Roquain (2008) for an extended discussion on choices of the shape function β of the form (12) (which can be suitably adapted to the uncountable case).

4.2 Applications

4.2.1 FDR control for testing the mean of a Gaussian process

Consider the multiple testing setting of Example 2.2. More specifically, we consider here the particular case where we observe $\{X_t, t \in [0, 1]^d\}$ a Gaussian process with measurable mean μ, with unit variance and covariance function Σ. Recall that the problem is to test for all $t \in [0, 1]^d$ the hypothesis H_t : "$\mu(t) \leq 0$". Taking for Λ the d-dimensional Lebesgue measure, the FDR control at level α of a step-up procedure of shape function β and weight function $\pi(h) = 1$ can be rewritten as

$$\mathbb{E}\left[\frac{\Lambda\Big(\big\{t \in [0, 1]^d \,:\, \mu(t) \leq 0,\, \overline{\Phi}(X_t) \leq \alpha\beta(\hat{r}(X))\big\}\Big)}{\Lambda\Big(\big\{t \in [0, 1]^d \,:\, \overline{\Phi}(X_t) \leq \alpha\beta(\hat{r}(X))\big\}\Big)}\right] \leq \alpha\,, \tag{13}$$

where $\overline{\Phi}$ is the upper-tail distribution function of a standard Gaussian variable and

$$\hat{r}(X) = \max\left\{r \in [0,1] \,:\, \Lambda\Big(\left\{t \in [0,1]^d \,:\, \overline{\Phi}(X_t) \leq \alpha\beta(r)\right\}\Big) \geq r\right\}.$$

Thus, Theorem 4.1 and Lemma 3.6 entail the following result.

Corollary 4.2. *For any jointly measurable Gaussian process $\{X_t\}_{t\in[0,1]^d}$ over $[0,1]^d$ with a measurable mean μ and unit variances, the FDR control (13) holds in either of the two following cases:*

- *$\beta(x) = x$ and the covariance function of the process is coordinates-wise non-negative, that is, satisfies (10);*
- *β is of the form (12), under no assumption on the covariance function.*

For instance, any Gaussian process with continuous paths is measurable and thus can be used in Corollary 4.2. More generally, Lemma A.2 states that any Gaussian process with a covariance function $\Sigma(t,t')$ such that

$$\forall t \in [0,1]^d,\ \lim_{t'\to t} \Sigma(t',t) = \Sigma(t,t) \text{ and } \lim_{t'\to t} \Sigma(t',t') = \Sigma(t,t) = 1$$

has a measurable modification and hence can be used in Corollary 4.2.

4.2.2 FDR control for testing the signal in a Gaussian white noise model

We continue Example 2.3, in which we observe the Gaussian process X defined by $X_g = \int_0^1 g(t)f(t)dt + \int_0^1 g(t)dB_t$, $g \in L^2([0,1])$, where B is a Wiener process on $[0,1]$ and $f \in C([0,1])$ is a continuous signal function. Remember that we aim at testing H_t : "$f(t) \leq 0$" for any $t \in [0,1]$, using the integration of the process against a smoothing kernel K_t. Assuming condition (2) holds, the p-value process is obtained via (3) as $p_t(X) = \overline{\Phi}^{-1}(Y_t)$, where $Y_t = v_{K,t}^{-1/2}(X_{K_t} - \delta_{t,\eta}c_{K,t})$ is a Gaussian process. Applying Lemma 3.6, we can prove that the p-value process defined by (3) is finite dimensional strong PRDS (on any subset) by checking that the covariance function of $(Y_t)_t$ has nonnegative values: the latter holds because the kernel K has been taken nonnegative and $\forall t,s$, $\mathrm{Cov}(Y_t,Y_s) = c\int_0^1 K((t-u)/\eta)K((s-u)/\eta)du$, for a nonnegative constant c. As a consequence, Theorem 4.1 shows that a step-up procedure using $\beta(x) = x$ controls the FDR.

To illustrate this result, let us consider a simple particular case where the kernel K is rectangular, i.e., $K(s) = \mathbf{1}\{|s| \leq 1\}/2$ and f is L-Lipschitz. Also, to avoid the boundary effects due to the kernel smoothing, we assume that the observation X is made against functions of $L^2([-1,2])$ while the test of H_t : "$f(t) \leq 0$" has only to be performed for $t \in [0,1]$ only. In that case, for $t \in [0,1]$, $\delta_{t,\eta} = L\eta$, $c_{K,t} = \eta$, $v_{K,t} = \eta/2$, so that $Y_t = (2\eta)^{-1/2}(Z_{t+\eta} - Z_{t-\eta} - L\eta^2)$. Therefore, the following statement holds.

Corollary 4.3. *Let us consider the Gaussian process $Z_t = \int_{-1}^t f(s)ds + B_t$, $t \in [-1,2]$, where B is a Wiener process on $[-1,2]$ and f is a L-Lipschitz function on $[-1,2]$ ($L > 0$). Let $\eta \in (0,1]$ and $Y_t = (2\eta)^{-1/2}(Z_{t+\eta} - Z_{t-\eta} - L\eta^2)$. Denote the Lebesgue measure on $[0,1]$ by Λ. Consider the volume rejection of the step-up procedure using $\pi(t) = 1$ and $\beta(x) = x$, that is,*

$$\hat{r}(X) = \max\left\{r \in [0,1] \,:\, \Lambda\Big(\left\{t \in [0,1] \,:\, \overline{\Phi}(Y_t) \leq \alpha r\right\}\Big) \geq r\right\},$$

where $\overline{\Phi}$ denotes the upper-tail distribution function of a standard Gaussian variable. Then the following FDR control holds:

$$\mathbb{E}\left[\frac{\Lambda\Big(\ \{t\in[0,1]\,:\,f(t)\le 0,\ \overline{\Phi}(Y_t)\le \alpha\hat{r}(X)\}\ \Big)}{\Lambda\Big(\ \{t\in[0,1]\,:\,\overline{\Phi}(Y_t)\le \alpha\hat{r}(X)\}\ \Big)}\right]\le \alpha\,. \tag{14}$$

4.2.3 FDR control for testing the c.d.f.

Consider the testing setting of Example 2.4 where we aim at testing whether "$F(t)\le F_0(t)$" for any t in an interval $I\subset\mathbb{R}$. Lemma C.1 states that the p-value process defined by (4) is finite dimensional weak PRDS (on any subset). As a consequence, Theorem 4.1 applies and leads to a control of the FDR.

For instance, let us consider the simple case where $I=[0,1]$, $F_0(t)=t$ and Λ is the Lebesgue measure on $[0,1]$. In this case, for any $k\in\{1,\dots,m\}$, the function $G_t(k)=\mathbb{P}(Z_t\ge k)$, with $Z_t\sim\mathcal{B}(m,t)$, is continuous increasing in the variable $t\in[0,1]$. Moreover, for any $t\in(0,1)$, the function $G_t(k)$ is decreasing in $k=0,\dots,m$. Therefore, denoting $0=X_{(0)}\le X_{(1)}\le\ldots\le X_{(m)}\le X_{(m+1)}=1$ the order statistics of $X_1,\dots,X_m$, the p-value process $t\mapsto p_t(X)=G_t(|\{1\le i\le m\,:\,X_i\le t\}|)$ is equal to 1 on $[0,X_{(1)})$, is increasing on each interval $(X_{(j)},X_{(j+1)}]$, $j=1,\dots,m$, and is left-discontinuous and right-continuous in each $X_{(j)}$, $1\le j\le m$, with a left limit larger than $p_{X_{(j)}}(X)=G_{X_{(j)}}(j)$ (see Figure 1).

As a consequence, for any threshold $u\in(0,1)$, we obtain the following relation for the Lebesgue measure $\gamma(u)$ of the level set $\{t\in[0,1]\,:\,p_t(X)\le u\}$:

$$\begin{aligned}\gamma(u)&=\sum_{j=0}^{m}\mathbf{1}\{G_{X_{(j)}}(j)\le u\}\,\Lambda\big(\{t\ge X_{(j)}\,:\,G_t(j)\le u\text{ and }t<X_{(j+1)}\}\big)\\&=\sum_{j=0}^{m}(X_{(j+1)}\wedge t_j(u)-X_{(j)})_+\end{aligned} \tag{15}$$

where $t_j(u), j=0,\dots,m$ is the unique solution of the equation $G_t(j)=u$, which can be easily computed numerically. Choosing for simplicity a uniform weighting $\pi(x)\equiv 1$, the choice of the rejection threshold given by the linear step-up procedure is then $\widehat{u}=\alpha\widehat{r}$, where $\widehat{r}$ is the largest solution of the equation $\gamma(\alpha r)=r$. To sum up, we have shown the following result:

Corollary 4.4. *Let $X=(X_1,\dots,X_m)$ be a vector of m i.i.d. real random variables of common continuous c.d.f. F. Consider $(p_t(X))_{t\in[0,1]}$ the p-value process $p_t(X)=G_t(|\{1\le i\le m\,:\,X_i\le t\}|)$ for $G_t(k)=\mathbb{P}(Z_t\ge k)$, where Z_t is a binomial variable of parameters (m,t). Assume that the hypothesis space $[0,1]$ is endowed with the Lebesgue measure Λ. Consider the volume rejection of the step-up procedure given by*

$$\widehat{r}(X)=\max\{r\in[0,1]\,:\,\gamma(\alpha r)\ge r\}\,, \tag{16}$$

where $\gamma(\cdot)$ is defined by (15). Then the following FDR control holds:

$$\mathbb{E}\left[\frac{\Lambda\big(\{t\in[0,1]\,:\,F(t)\le t,\ p_t(X)\le\alpha\hat{r}(X)\}\big)}{\Lambda\big(\{t\in[0,1]\,:\,p_t(X)\le\alpha\hat{r}(X)\}\big)}\right]\le\alpha. \tag{17}$$

4.2.4 FDR control for testing the intensity of a Poisson process

Let us consider the testing setting of Example 2.5. Lemma C.2 states that the p-values process is finite dimensional strong PRDS (on any subset). Thus, it is also finite dimensional weak PRDS (on any subset) by Lemma 3.5, and Theorem 4.1 leads to a control of the FDR.

Now, we aim at finding a closed formula for the linear step-up procedure ($\beta(x) = x$) using the p-value process $(p_t(X))_t$. Let us consider the particular case where the benchmark intensity $\lambda_0(\cdot)$ is constantly equal to some $\lambda_0 > 0$ while $\lambda(\cdot)$ is L-Lipschitz. Also, to avoid the boundary effects, assume that the process $(N_t)_t$ is observed for $t \in [-1, 2]$ while H_t: "$\lambda(t) \leq \lambda_0$" is tested only for $t \in [0, 1]$. In this case, the p-value process is simply given by

$$p_t(X) = G\big(N_{t+\eta} - N_{t-\eta}\big), \tag{18}$$

where for any $k \in \mathbb{N}$, $G(k)$ denotes $\mathbb{P}\left[Z \geq k\right]$ for Z a Poisson distribution of parameter $2\eta\lambda_0 + L\eta^2$ (note that $G(\cdot)$ is independent of t). Consider the jumps $\{T_j\}_j$ of the process $(N_t)_{t\in[-1,2]}$ and the set $\mathcal{S} = \{s_i\}_{2\leq i\leq m}$ of the the distinct and ordered values of the set $\cup_j\{T_j - \eta, T_j + \eta\} \cap (0, 1)$. Moreover, we let $s_1 = 0$ and $s_{m+1} = 1$. Next, since the p-value process is constant on each interval $[s_i, s_{i+1})$, $1 \leq i \leq m$, we have for any $u \geq 0$,

$$\begin{aligned}\Lambda\left(\{t \in [0, 1] \,:\, p_t(X) \leq u\}\right) &= \sum_{i=1}^{m}(s_{i+1} - s_i)\mathbf{1}\{p_{s_i}(X) \leq u\}\\ &= \sum_{k=1}^{m} w_k\mathbf{1}\{q_{\sigma(k)}(X) \leq u\},\end{aligned}$$

where we let $q_i(X) = p_{s_i}(X)$, where σ is a permutation of $\{1, \ldots, m\}$ such that $q_{\sigma(1)} \leq \ldots \leq q_{\sigma(m)}$ and where $w_k = s_{\sigma(k)+1} - s_{\sigma(k)} > 0$ can be interpreted as a "weighting" associated to $q_{\sigma(k)}$. As a consequence, we get

$$\begin{aligned}\widehat{r}(X) &= \max\left\{r \in [0, 1] \,:\, \sum_{\ell=0}^{m} w_\ell\mathbf{1}\{q_{\sigma(\ell)}(X) \leq \alpha r\} \geq r\right\}\\ &= \max\left\{\sum_{\ell=0}^{k} w_\ell,\ \text{for } k \in \{0, \ldots, m\} \text{ s.t. } q_{\sigma(k)}(X) \leq \alpha\sum_{\ell=0}^{k} w_\ell\right\},\end{aligned} \tag{19}$$

because since $\widehat{r}(X)$ is a maximum, it is of the form $\sum_{\ell=0}^{k} w_\ell$, $k \in \{0, \ldots, m\}$. Note that we should let $q_{\sigma(0)} = 0$ and $w_0 = 0$ to cover the case $\widehat{r}(X) = 0$. Relation (19) only involves a finite number of variables. Thus, $\widehat{r}(X)$ can be easily computed in practice. This is illustrated in Figure 2.

We have proved the following result:

Corollary 4.5. *Let $X = (N_t)_{t\in[-1,2]}$ be a Poisson process with an intensity $\lambda : [-1, 2] \to \mathbb{R}^+$ L-Lipschitz ($L > 0$) and let $\lambda_0 > 0$. For $\eta \in (0, 1]$, consider the p-value process $\{p_t(X)\}_{t\in[0,1]}$ given by (18). Assume that the hypothesis space $[0, 1]$ is endowed with the Lebesgue measure Λ. Then $\widehat{r}(X)$ defined by (19) satisfies the following:*

$$\mathbb{E}\left[\frac{\Lambda\big(\{t \in [0, 1] \,:\, \lambda(t) \leq \lambda_0,\, p_t(X) \leq \alpha\,\hat{r}(X)\}\big)}{\Lambda\big(\{t \in [0, 1] \,:\, p_t(X) \leq \alpha\,\hat{r}(X)\}\big)}\right] \leq \alpha. \tag{20}$$

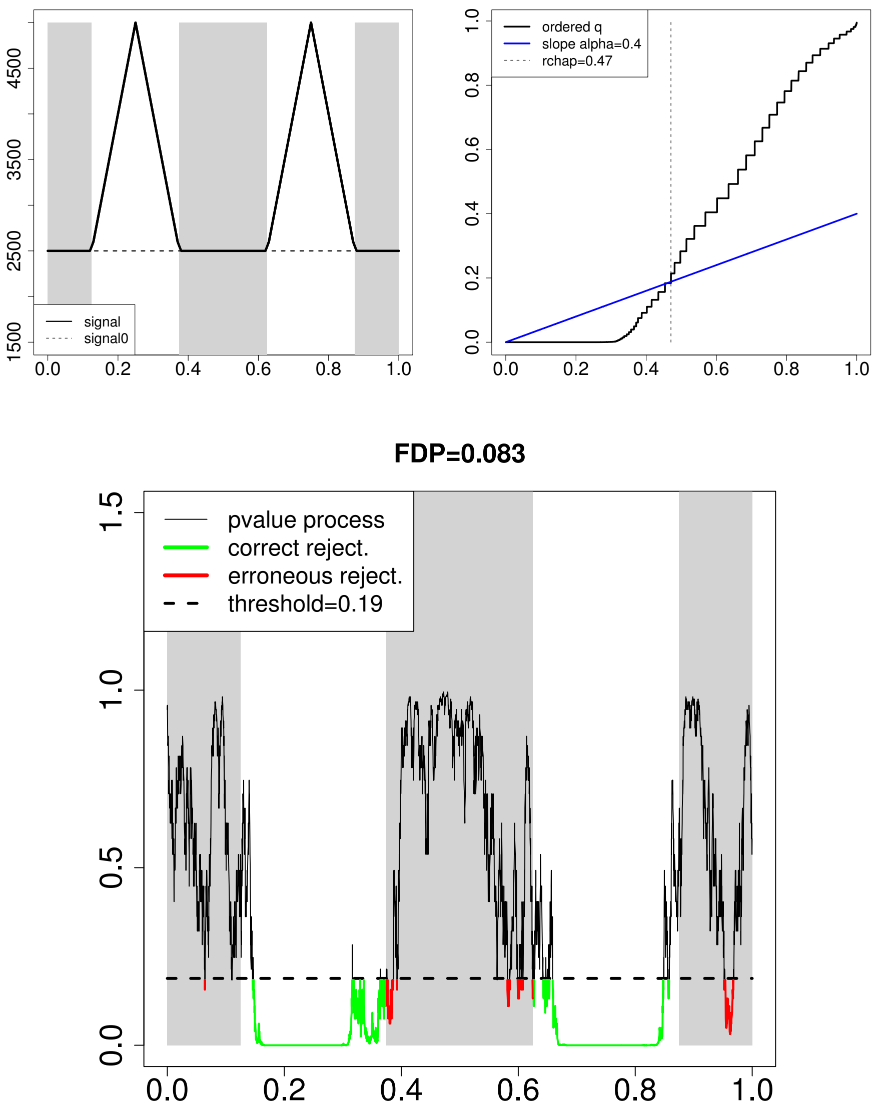


Figure 2: Several plots versus $t \in [0,1]$. Top left: $\lambda(t)$ (solid) and λ_0 (dashed). Top right: $q_{\sigma(k)}(X)$ and $\alpha \sum_{\ell=0}^{k} w_\ell$ in function of $\sum_{\ell=0}^{k} w_\ell$, for $k = 1, \ldots, m$. Bottom: p-value process $p_t(X)$ defined by (18). $\eta = 0.015$, $\alpha = 0.4$. The gray areas indicate regions where the null hypotheses are true.

To illustrate Corollary 4.5, Figure 2 displays the case where $\lambda(t)$ is a truncated triangular signal. The choice of the bandwidth η has been made manually, see Section 6.2 for a discussion on this point.

Remark 4.6. Up to increase the set $\mathcal{S} = \{s_i\}_i$ so that $t \mapsto \mathbf{1}\{\lambda(t) \leq \lambda_0\}$ is constant over each $[s_i, s_{i+1})$, the FDR control (20) can be rewritten as

$$\mathbb{E}\left[\frac{\sum_{i=1}^{m}(s_{i+1}-s_i)\mathbf{1}\{\lambda(s_i) \leq \lambda_0\}\mathbf{1}\{p_{s_i}(X) \leq \alpha\,\hat{r}(X)\}}{\sum_{i=1}^{m}(s_{i+1}-s_i)\mathbf{1}\{p_{s_i}(X) \leq \alpha\,\hat{r}(X)\}}\right] \leq \alpha. \tag{21}$$

Hence, the procedure (19) appears as controlling the discrete FDR-weighting on $\{1,\ldots,m\}$ where the weight for rejecting "$\lambda(s_i) \leq \lambda_0$" is $(s_{i+1}-s_i)$ and where the initial p-values are $q_i(X) = p_{s_i}(X)$. The rationale behind this is that if $q_i(X) = p_{s_i}(X)$ is below $\hat{r}(X)$, then so are all $p_t(X)$, $t \in [s_i, s_{i+1})$. Hence a rejection for a p-value $q_i(X) = p_{s_i}(X)$ accounts for the length of the entire interval in the FDR. From an intuitive point of view, this means that the type I error importance in the FDR is larger for "isolated" points of the process.

Finally, let us underline that (discrete) weighted FDR control results of the type (21) have been found in Benjamini and Hochberg (1997) and Blanchard and Roquain (2008), but only for a weighting that does not dependent on the data. Here, since $\{s_i\}_i$ depends on X, these former results do not apply and Corollary 4.5 is a novel finding.

5 Proof of Theorem 4.1

5.1 Two conditions for controlling the FDR

Similarly to Proposition 2.7 of Blanchard and Roquain (2008) (which we refer to as BR08 for short from now on), we can prove that the FDR control $\mathrm{FDR}(R,P) \leq \alpha\Pi(\mathcal{H}_0(P))$ holds true for any $P \in \mathcal{P}$ as soon as the two following sufficient conditions hold:

- the multiple testing procedure R satisfies the "self-consistency condition"

$$R(x) \subset \{h \in \mathcal{H} : p_h(x) \leq \alpha\pi(h)\beta(\Lambda(R(x)))\} \qquad \text{for } P\text{-almost all } x \in \mathcal{X} \tag{$\mathbf{SC}(\alpha,\pi,\beta)$}$$

- for any $P \in \mathcal{P}$, for any $h \in \mathcal{H}_0(P)$ the couple of real random variables $(U_h, V) := (p_h(X), \Lambda(R(X)))$ satisfies the "dependency control condition"

$$\forall c > 0, \quad \mathbb{E}\left[\frac{\mathbf{1}\{U_h \leq c\beta(V)\}}{V}\mathbf{1}\{V > 0\}\right] \leq c\,. \tag{$\mathbf{DC}(\beta)$}$$

The proof is as follows: from the definition (7) of the FDR and using Fubini's theorem, we have

$$\begin{aligned}
\mathrm{FDR}(R,P) &= \mathbb{E}\left[\frac{\Lambda(R\cap\mathcal{H}_0)}{\Lambda(R)}\mathbf{1}\{\Lambda(R)>0\}\right] \\
&= \mathbb{E}\left[\int_{h\in\mathcal{H}_0}\frac{\mathbf{1}\{h\in R\}}{\Lambda(R)}\mathbf{1}\{\Lambda(R)>0\}d\Lambda(h)\right] \\
&= \int_{h\in\mathcal{H}_0}\mathbb{E}\left[\frac{\mathbf{1}\{h\in R\}}{\Lambda(R)}\mathbf{1}\{\Lambda(R)>0\}\right]d\Lambda(h) \\
&\leq \int_{h\in\mathcal{H}_0}\mathbb{E}\left[\frac{\mathbf{1}\{p_h\leq\alpha\pi(h)\beta(\Lambda(R))\}}{\Lambda(R)}\mathbf{1}\{\Lambda(R)>0\}\right]d\Lambda(h) \\
&\leq \alpha\int_{h\in\mathcal{H}_0}\pi(h)d\Lambda(h),
\end{aligned}$$

where we have used the shortened notation R for $R(X)$ and p_h for $p_h(X)$, and used successively conditions (**SC**(α,π,β)) and(**DC**(β)) for the two above inequalities. Observe that the use of Fubini's theorem is granted by the measurability assumption of Definition 2.1.

Therefore, to obtain the FDR bound of Theorem 4.1 in each case, we simply have to check conditions (**SC**(α,π,β)) and (**DC**(β)) in the different settings.

5.2 Any step-up procedure satisfies (SC(α,π,β))

From the definition of a step-up procedure, for all $\varepsilon>0$, we have $\Lambda(L_\Delta(\widehat{r}))\leq\Lambda(L_\Delta(\widehat{r}+\varepsilon))<\widehat{r}+\varepsilon$. This entails that $\widehat{r}$ satisfies $\Lambda(L_\Delta(\widehat{r}))=\widehat{r}$. Hence the step-up procedure R satisfies **SC**(α,π,β) with equality.

We now check that any step-up procedure is a multiple testing procedure, that is, that $(\omega,h)\mapsto\mathbf{1}\{h\in R(X(\omega))\}=\mathbf{1}\{p_h(X(\omega))\leq\alpha\pi(h)\beta(\widehat{r}(X(\omega)))\}$ is (jointly) measurable. From (**A2**) and since β and π are measurable, it is enough to check that $\omega\mapsto\widehat{r}(X(\omega))$ is measurable. For any $x\in X(\Omega)$, let us consider the function

$$f:r\in\mathbb{R}^+\mapsto\Lambda(L_\Delta(x,r))=\int_{\mathcal{H}}\mathbf{1}\{p_h(x)\leq\alpha\pi(h)\beta(r)\}d\Lambda(h).$$

We observe that f is right-continuous and nondecreasing (because β is) and bounded, and that $\widehat{r}=\max\{r\geq 0\,:\,f(r)\geq r\}$. Applying Lemma D.2, we deduce that for any $x\in X(\Omega)$,

$$\begin{aligned}
\widehat{r}(x) &= \inf_{\varepsilon>0,\varepsilon\in\mathbb{Q}}\sup\left\{r\in\mathbb{Q}^+\,:\,\Lambda(L_\Delta(x,r))\geq r-\varepsilon\right\} \\
&= \inf_{\varepsilon>0,\varepsilon\in\mathbb{Q}}\sup_{r\in\mathbb{Q}^+}\Big(r\mathbf{1}\{\Lambda(L_\Delta(x,r))\geq r-\varepsilon\}\Big). \qquad (22)
\end{aligned}$$

Since from (**A2**), for all $\varepsilon>0$, $\varepsilon\in\mathbb{Q}$ and $r\in\mathbb{Q}^+$, the function

$$\omega\mapsto r\mathbf{1}\{\Lambda(L_\Delta(X(\omega),r))\geq r-\varepsilon\}=r\mathbf{1}\{\Lambda\big(\{h\in\mathcal{H}\,:\,p_h(X(\omega))\leq\alpha\pi(h)\beta(r)\}\big)\geq r-\varepsilon\}$$

is measurable, expression (22) implies that $\omega\mapsto\widehat{r}(X(\omega))$ is measurable. Hence, a step-up procedure satisfies the measurability requirements of Definition 2.1.

5.3 Conditions implying (DC(β))

We use the following lemma which was proved in (BR08) (see Lemma 3.2, items (ii,iii) therein):

Lemma 5.1. *Let (U,V) be a couple of nonnegative random variables such that U is stochastically lower bounded by a uniform variable on $[0,1]$, i.e. $\forall t \in [0,1], \mathbb{P}(U \leq t) \leq t$. Then the dependency control condition* **DC**(β) *is satisfied by (U,V) in either one of the following situations:*

(i) $\beta(x) = x$ and

$$\forall r \in \mathbb{R}^+,\ u \mapsto \mathbb{P}(V < r \,|\, U \leq u) \text{ is nondecreasing on } \{u : P(U \leq u) > 0\}. \tag{23}$$

(ii) The shape function β is of the form (12).

The point (ii) above, together with the results of the two previous sections, establishes point 2 of Theorem 4.1. To establish point 1 and finish the proof, we have to prove that (23) holds in the finite dimensional weak PRDS dependence context, which is done in the following proposition:

Proposition 5.2. *Assume that the p-values process $\mathbf{p} = (p_h, h \in \mathcal{H})$ is finite dimensional weak PRDS on $\mathcal{H}_0(P)$ for any $P \in \mathcal{P}$. Consider R the step-up procedure defined by Definition 3.2 with $\beta(x) = x$. Then for any $P \in \mathcal{P}$, for any $h \in \mathcal{H}_0(P)$, the couple of variables $(U_h, V) = (p_h, \Lambda(R))$ satisfies (23) and thus* **DC**(β) *holds for $\beta(x) = x$.*

Proof. In the above statement and the present proof, we use the shortened notation R, p_h, and $L_\Delta(r)$ for the random quantities $R(X), p_h(X)$, and $L_\Delta(X,r)$, respectively. The goal of the proof is to establish (23), that is for any $h_0 \in \mathcal{H}_0$ (h_0 is assumed to be fixed in $\mathcal{H}_0$ in the rest of the proof), for any t, and $0 \leq u \leq u'$ with $\mathbb{P}(p_{h_0} \leq u) > 0$:

$$\mathbb{P}\left[\Lambda(R) < t \,|\, p_{h_0} \leq u\right] \leq \mathbb{P}\left[\Lambda(R) < t \,|\, p_{h_0} \leq u'\right] ;$$

From Definition 3.2, the real random variable $\Lambda(R)$ can be rewritten as $\Lambda(R) = \widehat{r} = \max\{r : f(r) \geq r\}$ with $f : r \mapsto \Lambda(L_\Delta(r))$. Furthermore, denoting $G_u = \frac{\mathbf{1}\{p_{h_0} \leq u\}}{\mathbb{P}[p_{h_0} \leq u]}$, we are equivalently aiming at proving that for any t and $0 \leq u \leq u'$ with $\mathbb{P}(p_{h_0} \leq u) > 0$:

$$\mathbb{E}\left[\mathbf{1}\{\widehat{r} < t\} G_u\right] \leq \mathbb{E}\left[\mathbf{1}\{\widehat{r} < t\} G_{u'}\right] . \tag{24}$$

By using Lemma B.1 (and the notation therein) there exists a fixed sequence of finitely supported measures Λ_n on $\mathcal{H}$ such that, denoting $\widehat{r}_{n,k} = \max\{r \geq 0 : \Lambda_n(L_\Delta(r)) \geq r - k^{-1}\}$, it holds that

$$\widehat{r} = \lim_{k \to \infty} \widehat{r}_k^+ = \lim_{k \to \infty} \widehat{r}_k^- \quad \text{almost surely,} \tag{25}$$

where we let $\widehat{r}_k^+ = \limsup_{n \to \infty} \widehat{r}_{n,k}$ and $\widehat{r}_k^- = \liminf_{n \to \infty} \widehat{r}_{n,k}$.

Let $\mathcal{S}_n$ be the (finite) support of Λ_n and $\mathcal{S}'_n = \mathcal{S}_n \cup \{h_0\}$. Writing $\widehat{r}_{n,k}$ as a function of the finite p-value set $\{p_h, h \in \mathcal{S}'_n\}$, the function $\widehat{r}_{n,k} : \mathbf{z} = (z_h)_{h \in \mathcal{S}'_n} \in [0,1]^{\mathcal{S}'_n} \mapsto \widehat{r}_{n,k}(\mathbf{z})$ is measurable (where the space $[0,1]^{\mathcal{S}'_n}$ is endowed with the standard product Borel σ-field), and is additionally non-increasing in each p-value. Hence the set $\{\mathbf{z} = (z_h)_{h \in \mathcal{S}'_n} : \widehat{r}_{n,k}(\mathbf{z}) < t + k^{-1}\}$ is a non-decreasing measurable subset of $[0,1]^{\mathcal{S}'_n}$. Using that the p-value process

$\mathbf{p} = (p_h, h \in \mathcal{H})$ is finite dimensional weak PRDS on $\mathcal{H}_0$, the p-values $(p_h, h \in \mathcal{S}'_n)$ are PRDS on $\mathcal{H}_0 \cap \mathcal{S}'_n$, which implies that for any $t \geq 0$ and $u \leq u'$ with $\mathbb{P}(p_{h_0} \leq u) > 0$,

$$\mathbb{E}\left[\mathbf{1}\{\widehat{r}_{n,k} - k^{-1} < t\} G_u\right] \leq \mathbb{E}\left[\mathbf{1}\{\widehat{r}_{n,k} - k^{-1} < t\} G_{u'}\right] . \tag{26}$$

Now, to prove (24), it suffices to carefully make n and k tend to infinity. By Fatou's lemma and by (26), we have for all $k \geq 1$:

$$\begin{aligned} \mathbb{E}\left[\liminf_n \mathbf{1}\{\widehat{r}_{n,k} - k^{-1} < t\} G_u\right] &\leq \liminf_n \mathbb{E}\left[\mathbf{1}\{\widehat{r}_{n,k} - k^{-1} < t\} G_u\right] \\ &\leq \limsup_n \mathbb{E}\left[\mathbf{1}\{\widehat{r}_{n,k} - k^{-1} < t\} G_{u'}\right] \\ &\leq \mathbb{E}\left[\limsup_n \mathbf{1}\{\widehat{r}_{n,k} - k^{-1} < t\} G_{u'}\right] . \end{aligned}$$

Notice that the following inclusions of events hold: $\{\widehat{r}_k^+ < t + k^{-1}\} \subset \liminf_n \{\widehat{r}_{n,k} < t + k^{-1}\}$, $\limsup_n \{\widehat{r}_{n,k} < t + k^{-1}\} \subset \{\widehat{r}_k^- \leq t + k^{-1}\}$. Hence, we obtain for all k:

$$\mathbb{E}\left[\mathbf{1}\{\widehat{r}_k^+ - k^{-1} < t\} G_u\right] \leq \mathbb{E}\left[\mathbf{1}\{\widehat{r}_k^- - k^{-1} \leq t\} G_{u'}\right] .$$

Then, if t is such that $\mathbb{P}[\widehat{r} = t] = 0$, the above expression can be rewritten as

$$\mathbb{E}\left[\mathbf{1}\{\widehat{r}_k^+ - k^{-1} < t\} G_u \mathbf{1}\{\widehat{r} \neq t\}\right] \leq \mathbb{E}\left[\mathbf{1}\{\widehat{r}_k^- - k^{-1} \leq t\} G_{u'} \mathbf{1}\{\widehat{r} \neq t\}\right] .$$

We now let $k \to \infty$ in the above expression by using (25) and the dominated convergence theorem: for any $u \leq u'$ with $\mathbb{P}(p_{h_0} \leq u) > 0$, and any $t \notin D := \{s \geq 0 : \mathbb{P}[\widehat{r} = s] > 0\}$, we have

$$\mathbb{E}\left[\mathbf{1}\{\widehat{r} < t\} G_u\right] \leq \mathbb{E}\left[\mathbf{1}\{\widehat{r} < t\} G_{u'}\right] . \tag{27}$$

Since the above expectations may be interpreted as (conditional) probabilities, the LHS and RHS in (27) are left-continuous functions of t. Using that $\mathbb{R}^+ \cap D^c$ is dense in $\mathbb{R}^+$ (because D is at most countable), we obtain that (27) holds for any t. Finally, the condition (**DC**(β)) comes from Lemma 5.1. □

6 Discussion

6.1 FDR control for self-consistent, non step-up procedures

In some cases, for instance, after a discretization in r or under a global constraint over the admissible geometry of sets of rejected hypotheses, the procedure of interest may not be of the step-up form, while still satisfying the more general condition (**SC**(α, π, β)) (called self-consistency, see Section 5.1). In that situation, Theorem 4.1 does not apply, because the procedure is not step-up. We proved an extension of Theorem 4.1 holding more generally for (nonincreasing) self-consistent procedures, but point 1 of the theorem is established only under a stronger PRDS condition called general PRDS. (On the other hand, the fact that point 2 of Theorem 4.1 remains valid under the more general condition (**SC**(α, π, β)) is quite immediate.) The general PRDS condition is defined in terms of the entire process X and not only its finite dimensional projections. Therefore, it is substantially more technical than finite dimensional PRDS. In particular, it is an open question to characterize when does finite dimensional PRDS imply general PRDS (we provide some sufficient conditions). For simplicity, we deferred the corresponding study to a supplementary material (Blanchard et al., 2011).

6.2 Adaptive procedures

This work has focused on suitable control of the type I error. Under this constraint, the quality of a testing procedure depends on its power, that is, its ability to reject as many false hypotheses as possible. Ideally, power should be improved by making procedures adaptive with respect to various types of underlying regularity structure:

- adaptivity of single tests to various alternatives
- adaptivity to the proportion of true nulls
- adaptivity to the dependence structure of the p-values

These issues have received significant attention in recent literature on testing over discrete space of hypotheses; we discuss them briefly in the light of the framework developed in the present paper.

Adaptivity of single tests

The power of a multiple testing procedure depends primarily on the power of the underlying single tests and p-values it is built upon. It is of course desirable to have powerful individual tests in the first place. While this issue actually pertains to the domain of single hypothesis testing, and is to this extent quite independent of the methodology studied here, we briefly discuss this issue in the light of the specific examples studied here.

In the Gaussian white noise or Poisson model, the p-value process depends on the bandwidth η and on the regularity of the signal (Lipschitz constant L). The parameter L is often unknown and η should also be chosen based on the (unknown) regularity of the signal in order to produce a powerful test. One can therefore ask whether we can build individual tests that are adaptive to the regularity of the signal. Unfortunately, as argued in Low (1997) for adaptive confidence intervals, this is a delicate issue because the null hypothesis "$f(t_0) \leq 0$" (for a $t_0 \in [0,1]$) is a "huge" non-parametric class. This contrasts with the situation where the null is "$\forall t \in [0,1], f(t) \leq 0$", for which adaptive non parametric testing can be applied, see, e.g., Baraud et al. (2005); Durot and Rozenholc (2006); Horowitz and Spokoiny (2001); Butucea and Tribouley (2006). However, recent work Picard and Tribouley (2000); Giné and Nickl (2010); Nickl and Hoffmann (2011) showed that adaptive confidence intervals can be designed under some additional restrictions on the signal. Correspondingly adaptive p-values can be derived, though whether they satisfy or not the PRDS condition would have to be checked.

Adaptivity to $\Pi(\mathcal{H}_0)$

The main result of the present paper indicates that we can increase the power of the step-up procedure by considering $\alpha_0 = \alpha/\Pi(\mathcal{H}_0(P)) > \alpha$ instead of α as pre-specified level in the procedure. The latter still leads to a FDR control at level α because this yields $\text{FDR}(R,P) \leq \alpha_0 \Pi(\mathcal{H}_0(P)) = \alpha$. However, since α_0 depends on P, this parameter is unknown: this is only an "oracle" value which is not accessible in practice. In the case of finite hypotheses set, various approaches have been proposed to build multiple-testing procedures that are adaptive to $\pi_0 := \Pi(\mathcal{H}_0(P))$, see, e.g., Benjamini et al. (2006); Sarkar (2008); Blanchard and Roquain (2009); Finner et al. (2009). It is naturally an interesting perspective to extend this kind of methodology to the present continuous context. Note however, that the mentioned works in the discrete case almost always posit joint independence of the p-values, which we recall is not an acceptable assumption when the hypotheses space is uncountable. Therefore, this extension is not straightforward.

Adaptivity to the dependence structure

Another source of possible power improvement for multiple testing procedures is to take into account the dependence structure of the concerned statistics more finely than through the fairly broad PRDS assumption. This point is particularly relevant when testing over a continuum, as hypotheses which are "close" to each other will have very correlated statistics. The extent to which the upper bound on the main theorem can be improved depends in a complex way on this structure. In the discrete case, this issue has been studied for elementary settings, such as equi-correlated Gaussian statistics, see, e.g., Finner et al. (2007); Roquain and Villers (2011); Delattre and Roquain (2011). In the present continuous setting, studying precisely how dependence structures coming from stochastic process models affect the value of the FDR is an issue that remains to be explored.

A Auxiliary results pertaining to measurability issues

Lemma A.1 (Revuz and Yor (1991) p. 36)**.** *Let $(Z_t)_{t\in[0,1]}$ a real stochastic process on $(\Omega, \mathfrak{F}, \mathbb{P})$ where $[0,1]$ is endowed with its Borel σ-field and the Lebesgue measure Λ. Suppose that for all t, Z_t is square-integrable and* $\operatorname{Var} Z_t > 0$*. Then, if the variables of $(Z_t)_{t\in[0,1]}$ are mutually independent, the application $(\omega, t) \mapsto Z_t(\omega)$ is not jointly measurable in its variables.*

Proof. We essentially reproduce here an argument given p. 36 of Revuz and Yor (1991). Without loss of generality, let us assume that $\forall u \in [0,1]$, $\mathbb{E}Z_u = 0$ and $Z_u \in [0,1]$. If the joint measurability assumption holds, we can use Fubini's theorem: for all $t \in [0,1]$,

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^t Z_u d\Lambda(u)\right)^2\right] &= \mathbb{E}\left[\int_{[0,t]^2} Z_u Z_v d\Lambda^{\otimes 2}(u,v)\right] \\ &= \int_{[0,t]^2} \mathbb{E}\left[Z_u Z_v\right] d\Lambda^{\otimes 2}(u,v) \\ &= \int_{[0,t]^2} \mathbf{1}\{u=v\} d\Lambda^{\otimes 2}(u,v) = 0.\end{aligned}$$

Therefore, for all t, a.s. in ω, we have $\int_0^t Z_u(\omega)d\Lambda(u) = 0$. Which implies (by separability of $[0,1]$ and applying the Lebesgue differentiation theorem) that a.s. in (t,ω), $Z_t(\omega) = 0$. If follows that

$$0 = \mathbb{E}\left[\int_0^1 Z_t^2 d\Lambda(t)\right] = \int_0^1 \operatorname{Var}(Z_t) d\Lambda(t),$$

which contradicts that for all t, $\operatorname{Var} Z_t > 0$. □

The next lemma is a variation of Theorem 30 in Dellacherie and Meyer (1975).

Lemma A.2. *Let $\mathcal{H}$ be metric σ-compact space, endowed with the Borel σ-field $\mathfrak{H}$ and take a real stochastic process $Z = (Z_h, h \in \mathcal{H})$ defined on $(\Omega, \mathfrak{F}, \mathbb{P})$ and satisfying*

$$\forall h_0 \in \mathcal{H}, Z_h \to Z_{h_0} \text{ in probability when } h \to h_0. \tag{28}$$

Then there exists a process $Z' = (Z'_h, h \in \mathcal{H})$ which is jointly measurable in (ω, h) and which is a modification of $Z = (Z_h, h \in \mathcal{H})$, that is, such that for any $h \in \mathcal{H}$, for $\mathbb{P}$-almost every ω, $Z_h(\omega) = Z'_h(\omega)$.

Proof. Let assume first that the space $\mathcal{H}$ is compact. First, considering a metric of probability convergence, the convergence (28) is uniform and thus $\forall \delta > 0$, $\sup_{d(h,h')\leq\varepsilon} \mathbb{P}(|Z_h - Z_{h'}| > \delta) \xrightarrow{\varepsilon\to 0} 0$. Thus there exists $\varepsilon_n \to 0$ such that

$$\sup_{d(h,h')\leq\varepsilon_n} \mathbb{P}(|Z_h - Z_{h'}| > n^{-1}) \leq n^{-2}.$$

Next, taking a finite partition $\{A_i^n\}_{1\leq i\leq N_n}$ such that A_i^n is measurable and the diameter of each A_i^n is smaller than ε_n and fixing $h_i^n \in A_i^n$ for each i, we may define for each $h \in \mathcal{H}$ and $\omega \in \Omega$,

$$Z_h^n(\omega) = \sum_{i=1}^{N_n} \mathbf{1}\{h \in A_i^n\} Z_{h_i^n}(\omega).$$

Clearly, the function $(\omega, h) \mapsto Z_h^n(\omega)$ is jointly measurable in (ω, h) for each n and we have for each $h \in \mathcal{H}$,

$$\sum_{n\geq 1} \mathbb{P}(|Z_h - Z_h^n| > n^{-1}) \leq \sum_{n\geq 1} n^{-2} < \infty.$$

Applying the Borel-Cantelli theorem, for all $h \in \mathcal{H}$, for $\mathbb{P}$-almost every $\omega \in \Omega$, $Z_h^n(\omega)$ converges to $Z_h(\omega)$. Hence, $Z'_h(\omega) = \limsup_n Z_h^n(\omega)$ defines a jointly measurable modification of $(Z_h)_h$. The extension to a σ-compact space $\mathcal{H}$ is straightforward, by considering $\mathcal{H} = \cup_k \mathcal{H}_k$ with $\mathcal{H}_k$ compact and $(\mathcal{H}_k)_k$ nondecreasing. □

Lemma A.3. *Let $(W_g)_{g\in L^2([0,1])}$ be the Gaussian white noise process. Consider $K \in L^2(\mathbb{R})$ positive on $[-1, 1]$ and zero elsewhere. Denote by $K_t \in L^2([0, 1])$ the function $K_t(s) = K((t - s)/\eta)$, where $0 < \eta \leq 1$. Then there exists a modification of $(W_{K_t})_t$ that is jointly measurable in (ω, t).*

Proof. To prove this, we apply Lemma A.2 and check that the process $(W_{K_t})_t$ is continuous in probability i.e. that for any $t_0 \in [0, 1]$, W_{K_t} converges to $W_{K_{t_0}}$ in probability when t converges to t_0. We establish this by simply noting that $\mathbb{E}(W_{K_t} W_{K_{t_0}}) = \int_0^1 K_t(s)K_{t_0}(s)ds$ and $\mathbb{E}(W_{K_t}^2) = \int_0^1 K_t^2(s)ds$ are continuous functions w.r.t. the variable t, because the map $t \in [0, 1] \mapsto K_t \in L^2([0, 1])$ is continuous (this is classical and can be proved by using that the continuous functions are dense in $L^2([0, 1])$). □

The following lemma establishes that the FDR of a step-up procedure does not change if we consider a (measurable) modification of the p-value process.

Lemma A.4. *Let us consider a p-value functional $\mathbf{p} : \mathcal{X} \to [0, 1]^{\mathcal{H}}$ and two observations X' and X'' such that $(p_h(X'(\omega)))_{h,\omega}$ and $(p_h(X''(\omega)))_{h,\omega}$ are (jointly) measurable and are modification of each other, that is, for all $h \in \mathcal{H}$, for $\mathbb{P}$-almost every $\omega \in \Omega$, we have $p_h(X'(\omega)) = p_h(X''(\omega))$. Consider the two corresponding step-up procedures $R(X')$ and $R(X'')$ defined by Definition 3.2, using the observations X' and X'', respectively. Then the following holds:*

- *for $\mathbb{P}$-almost every $\omega \in \Omega$, for Λ-almost every h, $\mathbf{1}\{h \in R(X'(\omega))\} = \mathbf{1}\{h \in R(X''(\omega))\}$;*

- *for $\mathbb{P}$-almost every $\omega \in \Omega$, FDP$(R(X'(\omega)), P) =$ FDP$(R(X''(\omega)), P)$ and therefore we have FDR$(R(X'), P) =$ FDR$(R(X''), P)$.*

Proof. Let us first observe that by the joint measurability assumption, we may use Fubini's theorem to get

$$(\Lambda \otimes \mathbb{P})(\{(h, \omega) \,:\, p_h(X'(\omega)) \neq p_h(X''(\omega))\}) = 0,$$

which implies that, for $\mathbb{P}$-almost every $\omega \in \Omega$ and for Λ-almost every h, for any $r \geq 0$, we have $\mathbf{1}\{p_h(X'(\omega)) \leq \Delta(h, r)\} = \mathbf{1}\{p_h(X''(\omega)) \leq \Delta(h, r)\}$ and thus $\widehat{r}(X'(\omega)) = \widehat{r}(X''(\omega))$, as defined in (9). This leads to the desired results. □

As an illustration, if the p-value process is for the form $p_h(X) = f_h(X_h)$ for some family $\{f_h(\cdot)\}_h$ of measurable functions, $(p_h(X'))_h$ and $(p_h(X''))_h$ are modifications of each other as soon as X' is a modification of X''. As a consequence, Lemma A.4 applies for Examples 2.2 and 2.3 of Section 2.4, which shows that the resulting FDRs do not depend of the (measurable) modification chosen.

B Finite approximation of step-up procedures

The result presented in this section is used to prove Theorem 4.1 (see Section 5.3) and to establish measurability issues related to step-up procedures (see Section 5.2). We describe here how to derive the continuous step-up procedure defined in Definition 3.2 from a limit of finite step-up procedures. As usual, to lighten notation $R, p_h, L_\Delta(r), \widehat{r}$ denote the random quantities $R(X), p_h(X), L_\Delta(X, r), \widehat{r}(X)$. The following result holds:

Lemma B.1. *Consider the step-up procedure $R = L_\Delta(\widehat{r})$ on $\mathcal{H}$ using Λ and with $\widehat{r}$ defined in Definition 3.2. Then there exists a sequence of finitely supported measures Λ_n on $\mathcal{H}$ such that, denoting*

$$\widehat{r}_{n,k} = \max\{r \geq 0 \,:\, \Lambda_n(L_\Delta(r))) \geq r - k^{-1}\},$$

we have

$$\widehat{r} = \lim_{k\to\infty} \left(\limsup_{n\to\infty} \widehat{r}_{n,k} \right) = \lim_{k\to\infty} \left(\liminf_{n\to\infty} \widehat{r}_{n,k} \right) \text{ almost surely.}$$

Proof. We start with the following observation. Consider (Λ_n) some sequence of measures on $\mathcal{H}$ such that $\Lambda_n(\mathcal{H}) \equiv M$. For a fixed realization $x \in X(\Omega)$ of X, we consider $f : r \in \mathbb{R}^+ \mapsto \Lambda(L_\Delta(x, r))$ and $f_{\Lambda_n} : r \in \mathbb{R}^+ \mapsto \Lambda_n(L_\Delta(x, r)))$. Clearly, f and f_{Λ_n} are nondecreasing right-continuous functions. Using Lemma D.3, we conclude that the desired result holds provided that, for P-almost all $x \in \mathcal{X}$, f_{Λ_n} converges uniformly to f over $[0, M + 1]$. It remains thus to prove that there exists a sequence of finitely supported measures Λ_n on $\mathcal{H}$ such that for P-almost all $x \in \mathcal{X}$,

$$\limsup_{n\to\infty} \left\{ \sup_{r\in[0,M+1]} |\Lambda_n(L_\Delta(x, r)) - \Lambda(L_\Delta(x, r))| \right\} = 0\,. \tag{29}$$

Denote $\mathcal{Y}$ the product space $\mathcal{H}^{\mathbb{N}}$, endowed with the product sigma-algebra. For $y := (h_i)_{i\geq 1} \in \mathcal{Y}$ some sequence of hypotheses, denote $\Lambda_n^{[y]} = Mn^{-1} \sum_{i=1}^n \delta_{h_i}$ the suitably scaled uniform atomic measure on $(h_1, \ldots, h_n)$.

Consider now $Y := (H_i)_{i\geq 1} \in \mathcal{Y}$ an i.i.d. sequence of hypotheses drawn independently of X according to the probability distribution Λ/M on $\mathcal{H}$. Observe that for any fixed $x \in X(\Omega)$, $L_\Delta(x,r) = \{h \in \mathcal{H} : p_h(x) \leq \alpha\pi(h)\beta(r)\} = \{h \in \mathcal{H} : q(h,x) \leq \alpha\beta(r)\}$, where

$$q(h,x) := \begin{cases} p_h(x)/\pi(h) & \text{if } \pi(h) > 0; \\ 0 & \text{if } \pi(h) = 0 \text{ and } p_h(x) = 0; \\ \alpha\beta(M+1)+1 & \text{if } \pi(h) = 0 \text{ and } p_h(x) > 0\,. \end{cases}$$

Thus, applying the Glivenko-Cantelli theorem to the i.i.d. variables $(q(H_i,x))_i$, we deduce that for any $x \in \mathcal{X}(\Omega)$, $\zeta(x,y) = \limsup_{n\to\infty} \sup_{r\in[0,M+1]} \left|\Lambda_n^{[y]}(L_\Delta(x,r)) - \Lambda(L_\Delta(x,r))\right| = 0$ for P_Y-almost all realizations y of Y. Observe furthermore that for any fixed r, the function

$$(\omega,y) \in \Omega \times \mathcal{H}^{\mathbb{N}} \mapsto \Lambda_n^{[y]}(L_\Delta(X(\omega),r)) = Mn^{-1}\sum_{i=1}^{n} \mathbf{1}\{p_{h_i}(X(\omega)) \leq \alpha\pi(h_i)\beta(r)\}$$

is a (jointly) measurable function of (ω,y) by assumption **(A2)**. The inside supremum in (29) can be restricted to rational numbers since the functions involved are right-continuous. Therefore, $(\omega,y) \mapsto \zeta(X(\omega),y)$ is a jointly measurable function in its variables. By Fubini's theorem, this implies that $\mathbb{E}_{X,Y}[\zeta(X,Y)] = 0$; and thus also, for P_Y-almost all $y \in \mathcal{Y}$, $\zeta(x,y) = 0$ for P-almost all $x \in \mathcal{X}$. Since an event of probability 1 is non-empty, there exists a fixed $y \in \mathcal{Y}$ such that $\zeta(x,y) = 0$ for P-almost all $x \in \mathcal{X}$, which gives rise to a sequence of finitely supported measures Λ_n satisfying (29). □

C PRDS statements

Lemma C.1. *The p-value process $\mathbf{p}(X) = \{p_t(X), t \in I\}$ defined by (4) is finite dimensional weak PRDS (on any subset).*

Proof. Let us consider a finite subset $(t_j)_{0\leq j\leq N-1}$ of I and D a non-decreasing measurable subset of $[0,1]^N$. Let us prove that the function $u \mapsto \mathbb{P}[\mathbf{p}(X) \in D \mid p_{t_0}(X) \leq u]$ is non-decreasing on $\{u \in [0,1] : \mathbb{P}(p_{t_0}(X) \leq u) > 0\}$. If $F(t_0) \in \{0,1\}$, the result is trivial. We thus assume that $F(t_0) \in (0,1)$, so that $\mathcal{U}_{t_0} = \{G_{t_0}(k), k = m, m-1, \ldots, 0\}$ contains only increasing points of $(0,1]$. Without loss of generality, we only have to prove the non-decreasing property for $u \in \mathcal{U}_{t_0}$. Since G_{t_0} is decreasing from $\{0,\ldots,m\}$ to $\mathcal{U}_{t_0}$, we have $p_{t_0}(X) \leq G_{t_0}(k) \iff m\mathbb{F}_m(X,t_0) \geq k \iff X_{(k)} \leq t_0$ (letting $X_{(0)} = -\infty$). We thus have to prove that for any k, $1 \leq k \leq m$,

$$\mathbb{P}\left[(X_{(1)},\ldots,X_{(m)}) \in D' \mid X_{(k-1)} \leq t_0\right] \geq \mathbb{P}\left[(X_{(1)},\ldots,X_{(m)}) \in D' \mid X_{(k)} \leq t_0\right], \tag{30}$$

where $D' = \{x \in \mathbb{R}^m : (p_{t_j}(x))_{0\leq j\leq N-1} \in D\}$ is a nondecreasing subset of $\mathbb{R}^m$ (because $\mathbf{p}$ is coordinate wise nondecreasing, i.e., $x \leq x' \Rightarrow \forall t, p_t(x) \leq p_t(x')$). Using that the family of order statistics $\{X_{(i)}\}_i$ has positive regression dependency (see Lemma D.1), we derive that the function $f(a,b) = \mathbb{E}\left[(X_{(1)},\ldots,X_{(m)}) \in D' \mid X_{(k-1)} = a, X_{(k)} = b\right]$ is nondecreasing in a and b. Therefore, denoting $\gamma = \mathbb{P}\left[X_{(k)} \leq t_0 \mid X_{(k-1)} \leq t_0\right]$, we get

$$\begin{aligned}\mathbb{P}\left[(X_{(1)},\ldots,X_{(m)}) \in D' \mid X_{(k-1)} \leq t_0\right] &= \gamma\mathbb{E}\left[f(X_{(k-1)},X_{(k)}) \mid X_{(k-1)} \leq t_0, X_{(k)} \leq t_0\right] \\ &\quad + (1-\gamma)\mathbb{E}\left[f(X_{(k-1)},X_{(k)}) \mid X_{(k-1)} \leq t_0 < X_{(k)}\right] \\ &\geq \mathbb{E}\left[f(X_{(k-1)},X_{(k)}) \mid X_{(k-1)} \leq t_0, X_{(k)} \leq t_0\right],\end{aligned}$$

which provides (30) and concludes the proof. □

Lemma C.2. *The p-value process* $\mathbf{p}(X) = \{p_t(X), t \in [0,1]\}$ *defined by* (5) *is finite dimensional strong PRDS (on any subset).*

Proof. Let $M_t = N_{(t+\eta)\wedge 1} - N_{(t-\eta)\vee 0}$ for any $t \in [0,1]$. Fix $(t_j)_{0\le j\le q-1} \in [0,1]^q$ and assume $t_0 \in [\eta, 1-\eta]$ (the other case can be proved similarly). Take a nondecreasing measurable set $D \subset [0,1]^q$ and consider the set $D' = \{(M_{t_j})_{0\le j\le q-1} \in \mathbb{N}^q : (G_{t_j}(M_{t_j}))_{0\le j\le q-1} \in D\}$, which is nonincreasing on $\mathbb{N}^q$ and measurable. We thus aim to prove that for any $n \ge 0$,

$$\mathbb{P}\left[(M_{t_j})_{0\le j\le q-1} \in D' \mid M_{t_0} = n+1\right] \le \mathbb{P}\left[(M_{t_j})_{0\le j\le q-1} \in D' \mid M_{t_0} = n\right]. \tag{31}$$

Denote by $X_1 < \ldots < X_{k_X}$, $Y_1 < \ldots < Y_{k_Y}$ and $Z_1 < \ldots < Z_{k_Z}$ the jump times of the process $(N_t)_{t\in[0,1]}$ within the (disjoint) subsets $[0, t_0-\eta)$, $[t_0-\eta, t_0+\eta]$ and $(t_0+\eta, 1]$, respectively. Remark that $k_Y = M_{t_0}$ with our notation. Since $(N_t)_{t\in[0,1]}$ is a Poisson process, the family $\{(X_i, 1\le i\le k_X, k_X), (Y_i, 1\le i\le k_Y, k_Y), (Z_i, 1\le i\le k_Z, k_Z)\}$, contains mutually independent elements. Furthermore, the distribution of $(Y_1,\ldots,Y_{k_Y})$ conditionally on $k_Y = n$ is equal to the distribution of the order statistics of a sample $(Y'_1,\ldots,Y'_n)$ of i.i.d. random variables with common density $t \mapsto \lambda(t)/\int_{[t_0-\eta,t_0+\eta]}\lambda(s)ds$ on $[t_0-\eta, t_0+\eta]$ (w.r.t. the Lebesgue measure). Next, denoting $I_t = [(t-\eta)\vee 0, (t+\eta)\wedge 1]$, for any $t \in [0,1]$, we can write:

$$\begin{aligned}
&\mathbb{P}\left[(M_{t_j})_{0\le j\le q-1} \in D' \mid M_{t_0} = n+1\right]\\
&= \mathbb{P}\left[\left(\sum_{i=1}^{k_X}\mathbf{1}\{X_i \in I_{t_j}\} + \sum_{i=1}^{n+1}\mathbf{1}\{Y'_i \in I_{t_j}\} + \sum_{i=1}^{k_Z}\mathbf{1}\{Z_i \in I_{t_j}\}\right)_{0\le j\le q-1} \in D'\right]\\
&= \mathbb{P}\left[\left(\sum_{i=1}^{k_X}\mathbf{1}\{X_i \in I_{t_j}\} + \sum_{i=1}^{n}\mathbf{1}\{Y'_i \in I_{t_j}\} + \sum_{i=1}^{k_Z}\mathbf{1}\{Z_i \in I_{t_j}\}\right)_{0\le j\le q-1} \in D' - (\mathbf{1}\{Y'_{n+1} \in I_{t_j}\})_j\right]\\
&\le \mathbb{P}\left[\left(\sum_{i=1}^{k_X}\mathbf{1}\{X_i \in I_{t_j}\} + \sum_{i=1}^{n}\mathbf{1}\{Y'_i \in I_{t_j}\} + \sum_{i=1}^{k_Z}\mathbf{1}\{Z_i \in I_{t_j}\}\right)_{0\le j\le q-1} \in D'\right]\\
&= \mathbb{P}\left[(M_{t_j})_{0\le j\le q-1} \in D' \mid M_{t_0} = n\right],
\end{aligned}$$

the inequality coming from $D' - (\mathbf{1}\{Y'_{n+1} \in I_{t_j}\})_j \subset D'$, because D' is nonincreasing. This proves (31) and concludes the proof. □

D Technical lemmas

Lemma D.1. *Let* $X_1,\ldots,X_m$ *be a sequence of i.i.d. real random variables of common continuous c.d.f.* F*. Then, the family of order statistics* $\{X_{(i)}\}_i$ *has positive regression dependency, that is, for any non-decreasing measurable set* $D \subset \mathbb{R}^m$*, for any* $\{i_1,\ldots,i_j\} \subset \{1,\ldots,m\}$*,*

$$\mathbb{P}\left[(X_{(1)},\ldots,X_{(m)}) \in D \mid X_{(i_1)} = x_1,\ldots,X_{(i_j)} = x_j\right]$$

is non-decreasing in $(x_1,\ldots,x_j)$*.*

Proof. From Proposition 3.2 of Hu et al. (2006) (for instance), it is sufficient to prove that the family is multivariate total positive of order 2 (MTP2), that is, for every $x, y \in \mathbb{R}^m$,

$$g(x)g(y) \le g(x\vee y)g(x\wedge y),$$

where g is the density of $\{X_{(i)}\}_i$ with respect to the m-dimensional Lebesgue measure of $\mathbb{R}^m$, and where the minimum and the maximum are evaluated coordinate-wise. We merely check this condition: denoting $\mathcal{E} = \{z \in \mathbb{R}^m : z_1 < z_2 < \ldots < z_m\}$, and $f = F'$,

$$\begin{aligned} g(x_1, \ldots, x_n) g(y_1, \ldots, y_n) &= (m!)^2 \prod_{i=1}^{m} \big(f(x_i) f(y_i)\big) \mathbf{1}\{x \in \mathcal{E}\} \mathbf{1}\{y \in \mathcal{E}\} \\ &\le (m!)^2 \prod_{i=1}^{m} \big(f(x_i \vee y_i) f(x_i \wedge y_i)\big) \mathbf{1}\{x \vee y \in \mathcal{E}\} \mathbf{1}\{x \wedge y \in \mathcal{E}\} \\ &= g(x \vee y) g(x \wedge y). \end{aligned}$$

□

The next lemmas are elementary.

Lemma D.2. *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a bounded, nondecreasing and right-continuous function and let $\rho := \max\{r \ge 0 : f(r) \ge r\}$. For any $\varepsilon > 0$, the quantities $\rho, \rho_\varepsilon := \max\{r \ge 0 : f(r) \ge r - \varepsilon\}$ and $\rho'_\varepsilon := \sup\{r \in \mathbb{Q}^+ : f(r) \ge r - \varepsilon\}$ are well-defined and we have*

$$\rho = \inf_{\varepsilon > 0} \rho_\varepsilon = \inf_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \rho'_\varepsilon.$$

Proof. Note that the sets entering in the definition of $\rho, \rho_\varepsilon, \rho'_\varepsilon$ contain 0 and are upper bounded by assumption on f. Therefore ρ'_ε is well-defined. First defining ρ, ρ_ε as respective suprema, we have $f(\rho) \ge \rho$ and $f(\rho_\varepsilon) \ge \rho_\varepsilon - \varepsilon$ because f is nondecreasing, so that these suprema are maxima. Also note that $\rho_\varepsilon \ge \rho'_\varepsilon$ and that these functions are nondecreasing in ε. We first prove $\rho \le \inf_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \rho'_\varepsilon$: fixing $\varepsilon > 0$, since f is right-continuous at ρ, there is a $\delta > 0$, such that $f(\rho + \delta) \ge f(\rho) - \varepsilon/2$. Moreover, we can suppose that $\delta < \varepsilon/2$ and that $\rho + \delta \in \mathbb{Q}$ (because $\mathbb{Q}$ is dense in $\mathbb{R}$). Therefore,

$$f(\rho + \delta) \ge f(\rho) - \varepsilon/2 \ge \rho - \varepsilon/2 \ge (\rho + \delta) - \varepsilon,$$

so that $\rho + \delta \le \rho'_\varepsilon$, by definition of ρ'_ε, and because $\rho + \delta \in \mathbb{Q}$. This proves

$$\rho \le \inf_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \rho'_\varepsilon \le \inf_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \rho_\varepsilon = \inf_{\varepsilon > 0} \rho_\varepsilon.$$

To conclude the proof, it is enough now to show that $\rho \ge \inf_{\varepsilon > 0} \rho_\epsilon$. For this, observe that $\varepsilon \mapsto \rho_\varepsilon$ and $\varepsilon \mapsto f(\rho_\varepsilon)$ are nondecreasing, so that their limits exist in 0. By letting ε converge to 0 in expression $\rho_\varepsilon - \varepsilon \le f(\rho_\varepsilon)$, we get

$$\rho_{0^+} := \lim_{\varepsilon \to 0} \rho_\varepsilon = \lim_{\varepsilon \to 0} \{\rho_\varepsilon - \varepsilon\} \le \lim_{\varepsilon \to 0} f(\rho_\varepsilon) = f(\rho_{0^+}),$$

the last equality coming because f is right-continuous By definition of ρ we deduce $\rho \ge \rho_{0^+} = \inf_{\varepsilon > 0} \rho_\epsilon$. □

Lemma D.3. *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a nondecreasing right-continuous function, with $f \le c$, and let $\rho = \max\{r \ge 0 : f(r) \ge r\} = \max\{r \in [0, c] : f(r) \ge r\}$. Suppose that there exists $f_n : \mathbb{R}^+ \to \mathbb{R}^+$ a sequence of nondecreasing right-continuous functions, with $f_n \le c$, which converges uniformly to f on $[0, c+1]$. By letting for any $\varepsilon > 0$ ($\varepsilon < 1$), $\rho_{n,\varepsilon} = \max\{r \ge 0 : f_n(r) \ge r - \varepsilon\} = \max\{r \in [0, c+1] : f_n(r) \ge r - \varepsilon\}$, $\rho^+_\varepsilon = \limsup_n \rho_{n,\varepsilon}$ and $\rho^-_\varepsilon = \liminf_n \rho_{n,\varepsilon}$, we have*

$$\rho = \lim_{\varepsilon \to 0} \rho^+_\varepsilon = \lim_{\varepsilon \to 0} \rho^-_\varepsilon.$$

Proof. Fix $\varepsilon > 0$ and let us first prove $\rho \leq \rho_\varepsilon^-$. Let $\eta_n = \sup_{r\in[0,c+1]} |f_n(r) - f(r)|$, so that $\eta_n \to 0$. Next, for n large enough, we have $\varepsilon > \eta_n$, and thus

$$f_n(\rho) \geq f(\rho) - \eta_n \geq \rho - \varepsilon,$$

so that by definition of $\rho_{n,\varepsilon}$ we get $\rho \leq \rho_{n,\varepsilon}$. Hence $\rho \leq \rho_\varepsilon^-$, and then $\rho \leq \liminf_\varepsilon \rho_\varepsilon^-$.

Conversely, let us now prove $\rho \geq \limsup_\varepsilon \rho_\varepsilon^+$, which will conclude the proof. For any n and ε, we have $f(\rho_{n,\varepsilon}) \geq f_n(\rho_{n,\varepsilon}) - \eta_n \geq \rho_{n,\varepsilon} - \eta_n - \varepsilon$. By taking in the latter expression the supremum limit in n and then in ε, we derive

$$\limsup_{\varepsilon\to 0} \rho_\varepsilon^+ \leq \limsup_{\varepsilon\to 0} \limsup_{n\to\infty} f(\rho_{n,\varepsilon}) \leq \limsup_{\varepsilon\to 0} f(\rho_\varepsilon^+) \leq f(\limsup_{\varepsilon\to 0} \rho_\varepsilon^+),$$

where we used in the two last inequalities that f is nondecreasing and right-continuous Finally, by definition of ρ, we get $\rho \geq \limsup_\varepsilon \rho_\varepsilon^+$. □

Acknowledgements

This work was supported in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778. The third 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). The three authors where supported by the French ministry of foreign and european affairs (EGIDE - PROCOPE project number 21887 NJ).

References

Baraud, Y., Huet, S., and Laurent, B. (2005). Testing convex hypotheses on the mean of a Gaussian vector. Application to testing qualitative hypotheses on a regression function. *Ann. Statist.*, 33(1):214–257.

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. and Hochberg, Y. (1997). Multiple hypotheses testing with weights. *Scand. J. Statist.*, 24(3):407–418.

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.

Billingsley, P. (1999). *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition. A Wiley-Interscience Publication.

Blanchard, G., Delattre, S., and Roquain, E. (2011). Supplement to: "testing over a continuum of null hypotheses".

Blanchard, G. and Roquain, E. (2008). Two simple sufficient conditions for FDR control. *Electron. J. Stat.*, 2:963–992.

Blanchard, G. and Roquain, E. (2009). Adaptive false discovery rate control under independence and dependence. *J. Mach. Learn. Res.*, 10:2837–2871.

Butucea, C. and Tribouley, K. (2006). Nonparametric homogeneity tests. *J. Statist. Plann. Inference*, 136(3):597–639.

Delattre, S. and Roquain, E. (2011). On the false discovery proportion convergence under gaussian equi-correlation. *Statistics & Probability Letters*, 81(1):111–115.

Dellacherie, C. and Meyer, P.-A. (1975). *Probabilités et potentiel.* Hermann, Paris. Chapitres I à IV, Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372.

Durot, C. and Rozenholc, Y. (2006). An adaptive test for zero mean. *Math. Methods Statist.*, 15(1):26–60.

Finner, H., Dickhaus, R., and Roters, M. (2009). On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Statist.*, 37(2):596–618.

Finner, H., Dickhaus, T., and Roters, M. (2007). Dependency and false discovery rate: asymptotics. *Ann. Statist.*, 35(4):1432–1455.

Giné, E. and Nickl, R. (2010). Confidence bands in density estimation. *Ann. Statist.*, 38(2):1122–1170.

Horowitz, J. L. and Spokoiny, V. G. (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica*, 69(3):599–631.

Hu, T., Chen, J., and Xie, C. (2006). Regression dependence in latent variable models. *Probab. Engrg. Inform. Sci.*, 20(2):363–379.

Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, 37:1137–1153.

Low, M. G. (1997). On nonparametric confidence intervals. *Ann. Statist.*, 25(6):2547–2554.

Nickl, R. and Hoffmann, M. (2011). On adaptive inference and confidence bands. *Ann. Statist.*, 27.

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.

Picard, D. and Tribouley, K. (2000). Adaptive confidence interval for pointwise curve estimation. *Ann. Statist.*, 28(1):298–335.

Revuz, D. and Yor, M. (1991). *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin.

Robin, S. (2002). A compound Poisson model for word occurrences in DNA sequences. *J. Roy. Statist. Soc. Ser. C*, 51(4):437–451.

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., Gavrilov, Y., and Adler, R. J. (2010). Peak detection as multiple testing. Biostatistics working paper series., Harvard University. available at http://www.bepress.com/harvardbiostat/paper120.

Seeger, P. (1968). A note on a method for the analysis of significances en masse. *Technometrics*, 10(3):586–593.

Supplement to: "Testing over a continuum of null hypotheses"

Gilles Blanchard Sylvain Delattre Étienne Roquain

Abstract

This report is a supplement for the paper Blanchard et al. (2011) (denoted [BDR] below). This additional study introduces the so-called general PRDS condition, which is a stronger assumption (in general) than the finite dimensional PRDS condition considered in [BDR]. This condition is useful to prove FDR control for procedures which are not necessarily of the step-up type. We also study some conditions under which the finite dimensional PRDS condition is sufficient to ensure the general PRDS condition.

In this supplement we use the setting and the notation defined in Sections 2 and 3 of [BDR].

1 General PRDS condition

We investigate an extension of the PRDS condition to the continuous setting, called "general" PRDS condition. Its definition is different from the one of the "finite-dimensional" PRDS condition introduced in [BDR]. Although it raises some delicate measurability issues, its form is totally analogous to the finite hypothesis case and hence it seems very natural. Let us denote by $[0,1]^{\mathcal{H}}$ the set of measurable functions from $\mathcal{H}$ to $[0,1]$, which is identified with a subset of $[0,1]^{\mathcal{H}}$ in the sequel.

We also extend the definition of a nondecreasing subset as follows: a subset D is said nondecreasing in $[0,1]^{\mathcal{H}}$ if $D \subset [0,1]^{\mathcal{H}}$ and for all $\mathbf{z}, \mathbf{z}' \in \mathcal{L}^0(\mathcal{H}, [0,1])$, if $\forall h \in \mathcal{H}$, $z_h \leq z'_h$, we have $\mathbf{z} \in D$ implies $\mathbf{z}' \in D$.

Definition 1.1. (General PRDS condition) Let $X : \Omega \to \mathcal{X}$ be a random variable with distribution P and $\mathbf{p} : \mathcal{X} \to [0,1]^{\mathcal{H}}$ a p-value functional. For $\mathcal{H}'$ a subset of $\mathcal{H}$, the p-value process $\mathbf{p}(X) = (p_h(X))_{h \in \mathcal{H}}$ is said to be *general weak PRDS on* $\mathcal{H}'$, if for any $h \in \mathcal{H}'$, for any nondecreasing set D in $[0,1]^{\mathcal{H}}$ such that the preimage $\mathbf{p}^{-1}(D)$ is a measurable set of $\mathcal{X}$, the function $u \in [0,1] \mapsto \mathbb{P}(\mathbf{p}(X) \in D \,|\, p_h(X) \leq u)$ is nondecreasing on $\{u \in [0,1] : \mathbb{P}(p_h(X) \leq u) > 0\}$; it is said *general strong PRDS* if the function $u \mapsto \mathbb{P}(\mathbf{p}(X) \in D \,|\, p_h(X) = u)$ is nondecreasing.

It is important to underline that for the definition of the general strong PRDS condition, we do not require the existence of a regular conditional probability distribution $\mathcal{D}(\mathbf{p}(X) \,|\, p_h(X) = u)$, which would demand additional assumptions on the underlying space. Rather, the conditional probability is to be interpreted as the simple conditional expectation $\mathbb{E}\left[\mathbf{1}\{\mathbf{p}(X) \in D\} \,|\, p_h(X) = u\right]$, which for any fixed measurable D is always well-defined as a $p_h(X)$-measurable random variable. The general strong PRDS condition is thus the requirement that there exists a nondecreasing function $f : [0,1] \to [0,1]$ such that $\mathbb{P}$-almost surely, $f(p_h(X)) = \mathbb{E}\left[\mathbf{1}\{\mathbf{p}(X) \in D\} \,|\, p_h(X)\right]$. For the weak general PRDS condition, we are only interested in values u such that $\mathbb{P}(p_h(X) \leq u) > 0$, so that the conditional probability on this range can be defined via a quotient of probabilities, and the above discussion is not needed.

Clearly, the general PRDS condition implies the finite dimensional PRDS one. Conversely, finite dimensional PRDS implies general PRDS for observation spaces satisfying some properties, see Section 3.

The weak general PRDS condition is easy to check for instance in the c.d.f. testing example, as well as the Poisson intensity example. Additionally, we can check that the strong PRDS condition holds for testing the mean of a continuous Gaussian process with positive covariance function; this is a consequence of Proposition 3.1 and Lemma 3.6 in [BDR]. This can also be proved by reproducing the argument of Benjamini and Yekutieli (2001) (for this, we use that the σ-field on the Wiener space is generated by cylinders).

2 FDR control for a general PRDS p-value process

In the following result, we establish FDR control for a class of multiple testing procedures which are more general than step-up procedures considered in [BDR]. For this, the following stronger conditions are assumed:

- general PRDS instead of the finite dimensional PRDS;
- measurability condition **(A2')** instead of **(A2)**.

Theorem 2.1. *Assume that the hypothesis space $\mathcal{H}$ satisfies* (**A1**) *in [BDR] and is endowed with the finite measure Λ. Let $\mathbf{p}(X) = (p_h(X))_{h\in\mathcal{H}}$ be a p-value process satisfying the conditions* **(A2')** *and* **(A3)** *in [BDR]. Suppose that the p-value process $\mathbf{p} = (p_h(X))_{h\in\mathcal{H}}$ is general weak PRDS on $\mathcal{H}_0$, and consider a multiple testing procedure R which can be written under the form $\forall x \in X(\Omega)$, $R(x) = \widetilde{R}(\mathbf{p}(x))$ where $\widetilde{R} : \mathcal{L}^0(\mathcal{H}, [0, 1]) \to \mathfrak{H}$ is a function (without any measurability requirement) such that $\Lambda(\widetilde{R}(\mathbf{p}))$ is nonincreasing in each p-value, that is, for any $\mathbf{z}, \mathbf{z}' \in \mathcal{L}^0(\mathcal{H}, [0, 1])$, if $\forall h \in \mathcal{H}$, $z_h \leq z'_h$, we have $\Lambda(\widetilde{R}(\mathbf{z})) \geq \Lambda(\widetilde{R}(\mathbf{z}'))$. Assume moreover that R satisfies $\mathbf{SC}(\alpha, \pi, \beta)$ in [BDR] with $\alpha \in (0, 1)$, $\beta(x) = x$ and π a probability density function on $\mathcal{H}$ with respect to Λ. Then for any $P \in \mathcal{P}$, we have the inequality*

$$FDR(R, P) \leq \alpha\Pi(\mathcal{H}_0(P)) \quad (\leq \alpha), \tag{1}$$

where $\Pi(\mathcal{H}_0(P)) := \int_{h\in\mathcal{H}_0(P)} \pi(h)d\Lambda(h)$.

Proof. The proof of Theorem 2.1is simpler than the one of Theorem 4.1 in [BDR], because the general PRDS assumption allows us to skip the construction of a finite dimensional approximation of the non-decreasing event "$\Lambda(R) < t$", as we did in the proof of Proposition 5.2 in [BDR] for step-up procedures. Namely, from the methodology advocated in Section 5 of [BDR], Theorem 2.1 is proved as soon as we show that for any $h \in \mathcal{H}_0$, the couple of variables $(U, V) = (p_h, \Lambda(R(X)))$ satisfies (22) in [BDR]. For this, we consider the set $D = \left\{ \mathbf{z} \in \mathcal{L}^0(\mathcal{H}, [0, 1]) \mid \Lambda(\widetilde{R}(\mathbf{z})) < r \right\}$ which is nondecreasing from the nonincreasing property of $\mathbf{z} \mapsto \Lambda(\widetilde{R}(\mathbf{z}))$. Then, we may check that the preimage of D through $\mathbf{p}$ is $\mathbf{p}^{-1}(D) = \{x \in \mathcal{X} : \Lambda(R(x)) < r\}$, which is a measurable set of $\mathcal{X}$, by the measurability assumptions **(A2')** on the p-value process and Fubini's theorem. From the general weak PRDS property, the function $u \mapsto \mathbb{P}(\mathbf{p} \in D \mid p_h \leq u)$ is nondecreasing, which shows (22) in [BDR] and completes the proof of Theorem 2.1.

□

Of course, we may easily check that a step-up procedure with $\beta(x) = x$ satisfies the assumptions of Theorem 2.1; it satisfies ($\mathbf{SC}(\alpha, \pi, \beta)$) with $\beta(x) = x$ from Section 5.2 in [BDR]. Moreover, it is of the form $R(x) = \widetilde{R}(\mathbf{p}(x))$, by letting for any $\mathbf{z} \in \mathcal{L}^0(\mathcal{H}, [0,1])$, $\widetilde{R}(\mathbf{z}) = \{h \in \mathcal{H} : z_h \leq \alpha\pi(h)\widetilde{r}(\mathbf{z})\}$ where $\widetilde{r}(\mathbf{z}) := \max\{r \geq 0 : \Lambda(\{h \in \mathcal{H} : z_h \leq \alpha\pi(h)r\}) \geq r\}$. We easily check that $\widetilde{R}$ is well defined (remember that no measurability property are required for $\widetilde{R}$, except that $\widetilde{R}$ is valued in $\mathfrak{H}$).

Since the step-up procedures maximize the rejection volume within the self-consistent procedures and thus are more powerful for the same level of FDR control, we may legitimately ask whether considering any other self-consistent procedure is relevant and therefore, whether the general PRDS condition is useful (at least through Theorem 2.1). What we want to emphasize here is that in some "constrained" cases, the procedure of interest may not be of the step-up form while it stays self-consistent. For instance, considering some discrete $\mathfrak{D}$ subset of $\mathbb{R}^+$ containing 0, we can consider $\widetilde{R}^{\mathfrak{D}}(\mathbf{z}) = \{h \in \mathcal{H} : z_h \leq \alpha\pi(h)\widetilde{r}^{\mathfrak{D}}(\mathbf{z})\}$ where $\widetilde{r}^{\mathfrak{D}}(\mathbf{z}) := \max\{r \in \mathfrak{D} : \Lambda(\{h \in \mathcal{H} : z_h \leq \alpha\pi(h)r\}) \geq r\}$. This can be useful in practice if no explicit expression stands for the step-up procedure. In that situation, Theorem 2.1 can be applied to control the FDR in replacement to Theorem 4.1 of [BDR], by proving the general PRDS condition (note that the control is on the *continuous* FDR even if this procedure is of a discrete nature).

3 Case where finite dimensional PRDS implies general PRDS

Of course, the general PRDS condition implies the finite dimensional PRDS one. Whether finite dimensional PRDS implies general PRDS for arbitrary spaces stays an open problem. As a matter of fact, even if the σ-field on $\mathcal{X}$ is the product σ-field, while any element of the product σ-field can easily be approached by cylinders, we were not able to state that a *non-decreasing* element of the product σ-field can be approached by *non-decreasing* cylinders.

Nevertheless, under some additional topological assumptions, that are for instance satisfied for continuous p-value processes, the latter reasoning works out rigorously and we may prove that the finite dimensional PRDS condition is equivalent to the general PRDS condition. For this, we assume that $\mathcal{X}$ is a complete separable metric space, endowed with a metric d and the corresponding Borel σ-field. The index set $\mathcal{H}$ is assumed metric and endowed with the corresponding Borel σ-field. As usual, we suppose that the p-value functional $(p_h(x))$ is measurable, that is, satisfies (**A2**) in [BDR].

Proposition 3.1. *Assume that the p-value functional $\mathbf{p} = \{p_h, h \in \mathcal{H}\}$ is separable, i.e. that there exists a countable, dense subset T of $\mathcal{H}$ such that for every $y \in \mathcal{X}$, the following property holds:*

$$\forall h \in \mathcal{H}, \quad \exists (h_1, h_2 ...), \quad h_n \in T, \quad s.t. \;\; h_n \to h, \quad p_{h_n}(y) \to p_h(y). \tag{2}$$

Assume the two following topological properties: for all $h \in \mathcal{H}$ the coordinate projections $x \in \mathcal{X} \mapsto p_h(x)$ are continuous functions and

$$\begin{aligned} &\textit{for any } B(y,\varepsilon) = \{x \in \mathcal{X} : d(y,x) < \varepsilon\}, \\ &\textit{the set } \{z \in B(y,\varepsilon) : \forall h \in \mathcal{H}, p_h(z) \geq p_h(y)\} \textit{ is of non-empty interior.} \end{aligned} \tag{3}$$

Then, for a given subset $\mathcal{H}' \subset \mathcal{H}$, the p-value process $\mathbf{p}(X) = \{p_h(X), h \in \mathcal{H}\}$ is finite dimensional weak (resp. strong) PRDS on $\mathcal{H}'$ if and only if it is general weak (resp. strong) PRDS on $\mathcal{H}'$.

As illustration, the assumption of the above proposition are satisfied for any continuous p-value process on $\mathcal{H} = [0,1]^d$: in that case, we may directly take $\mathcal{X}$ as the set of continuous functions from $[0,1]^d$ to $[0,1]$, endowed with $d(z,w) = ||z-w||_\infty$ and $p_h(x) = x_h$ as the identity function. It is easy to check that the process $\mathbf{p} = \{p_h, h \in \mathcal{H}\}$ is separable, because $[0,1]^d$ is separable and because the p-value process is continuous. Finally, condition (3) holds; for any open L^∞-ball $B(y,\varepsilon)$, we can consider $w = y + \varepsilon/2$ (defined as $\forall h$, $w_h = y_h + \varepsilon/2$), so that any $z \in B(w, \varepsilon/2)$ satisfies $z > w - \varepsilon/2 = y$ and $d(y,z) \leq d(y,w) + d(w,z) < \varepsilon$, hence $B(w,\varepsilon/2) \subset \{z \in B(y,\varepsilon)\, :\, z > y\}$.

For a càdlàg p-value process, the above result can not be applied because the Skorohod topology does not satisfy the required topological assumptions. From an intuitive point of view, the time rescaling of the Skorohod distance is not compatible with the coordinate-wise property of non-decreasing sets.

We now prove Proposition 3.1.

Proof. Let us prove the result for the strong PRDS property (the weak PRDS case is similar). Assume that the finite dimensional PRDS property is valid. Let $h \in \mathcal{H}'$, let $0 \leq u \leq u' \leq 1$, and prove that for any nondecreasing set $\overline{D} \subset [0,1]^{\mathcal{H}}$ such that $D = \mathbf{p}^{-1}(\overline{D})$ is measurable, we have

$$\mu_u(D) \leq \mu_{u'}(D), \tag{4}$$

where we have denoted $\mu_v(D) := \mathbb{P}(X \in D \,|\, p_h(X) = v)$ for $v \in \{u, u'\}$.

First, we prove that it is sufficient to show (4) only for closed sets of the form $D = \mathbf{p}^{-1}(\overline{D})$ with $\overline{D}$ nondecreasing. For this, consider any *measurable* set of the form $D = \mathbf{p}^{-1}(\overline{D})$ with $\overline{D}$ nondecreasing. Since the space $\mathcal{X}$ is assumed to be a complete separable Borel space, μ_u and $\mu_{u'}$, defined at the first place as conditional expectations, can be also defined as conditional probabilities on $\mathcal{X}$. Therefore, the regularity property of probability measures implies that, for $v \in \{u, u'\}$, for a fixed $\varepsilon > 0$, there exists a compact set $K_v \subset D = \mathbf{p}^{-1}(\overline{D})$ such that

$$\mu_v(D) - \mu_v(K_v) < \varepsilon.$$

Letting $\overline{K}_v = \mathbf{p}(K_v) \subset \overline{D}$ and $\overline{K}'_v = \cup_{x \in K_v}\{q \in [0,1]^{\mathcal{H}}\, :\, q \geq \mathbf{p}(x)\} = \cup_{p \in \overline{K}_v}\{q \in [0,1]^{\mathcal{H}}\, :\, q \geq p\}$, since $\overline{D}$ is nondecreasing, we get $\overline{K}_v \subset \overline{K}'_v \subset \overline{D}$ and thus letting $K'_v = \mathbf{p}^{-1}(\overline{K}'_v) \subset D$, we get $K_v \subset \mathbf{p}^{-1}(\mathbf{p}(K_v)) \subset K'_v$ and thus $\mu_v(D) - \mu_v(K'_v) < \varepsilon$. Since K_v is compact, we easily check that K'_v is closed (for a sequence $y^n \in K'_v$ converging to $y \in \mathcal{X}$, there exists a sequence $x^n \in K_v$ with $p_h(x^n) \leq p_h(y^n)$ for any $h \in \mathcal{H}$. Up to consider a subsequence, we get by continuity of the coordinate projection on h that there exists $x \in K_v$ with $p_h(x) \leq p_h(y)$ for any $h \in \mathcal{H}$). Next, we consider $F = K'_u \cup K'_{u'} \subset D$ which is a closet set satisfying $\mu_v(D) - \mu_v(F) < \varepsilon$ for $v \in \{u, u'\}$ and such that $F = \mathbf{p}^{-1}(\overline{K}'_u \cup \overline{K}'_{u'})$. Note that $\overline{K}'_u \cup \overline{K}'_{u'}$ is nondecreasing. As a consequence, if (4) holds for F, we have

$$\mu_u(D) - \varepsilon \leq \mu_u(F) \leq \mu_{u'}(F) \leq \mu_{u'}(D),$$

for any $\varepsilon > 0$, which implies (4) for any *measurable* set of the form $D = \mathbf{p}^{-1}(\overline{D})$ with $\overline{D}$ nondecreasing.

Second, we consider a closed set $F = \mathbf{p}^{-1}(\overline{F})$ with $\overline{F}$ nondecreasing, and we denote $O = F^c$ and $\overline{O} = (\overline{F})^c$ (such that $O = \mathbf{p}^{-1}(\overline{O})$). From the first point above, the proof of the proposition will be finished if we show that $\mu_u(F) \leq \mu_{u'}(F)$, or equivalently

$$\mu_u(O) \geq \mu_{u'}(O). \tag{5}$$

Note that the set $\overline{O}$ is nonincreasing, that is, for all x, y with $x \leq y$, $y \in \overline{O}$ implies $x \in \overline{O}$. Let us prove now that

$$O = \bigcup_{z \in O_0} \{y \in \mathcal{X} \,:\, \mathbf{p}(y) \leq \mathbf{p}(z)\}, \tag{6}$$

where $O_0 \subset O$ is any countable and dense set in O. If y is such that there exists $z \in O_0$ with $\mathbf{p}(y) \leq \mathbf{p}(z)$, we have $y \in O$ since $\overline{O}$ is nonincreasing. Conversely, take $y \in O$; since O is an open set, there exists $\varepsilon > 0$ such that the ball $B(y, \varepsilon)$ is included in O. Applying (3) and since O_0 is dense in O, the set $\{z \in B(y, \varepsilon) \,:\, \mathbf{p}(y) \leq \mathbf{p}(z)\}$ contains at least one element of O_0. This implies that y is included in the right-hand of relation (6). Finally, rewritting O_0 as $\{z^k\}_{k\geq 1}$, expression (6) is equivalent to

$$O = \bigcup_{K\geq 1} \bigcup_{k \leq K} \{y \in \mathcal{X} \,:\, \mathbf{p}(y) \leq \mathbf{p}(z^k)\}. \tag{7}$$

Next, because the process is separable, we can fix a countable set T such that (2) holds for any $y \in \mathcal{X}$. Letting $T = \{h_n\}_{n\geq 1}$ and $A_{n,k} = \{y \in \mathcal{X} \,:\, \forall i = 1, ..., n,\ p_{h_i}(y) \leq p_{h_i}(z^k)\}$, we get for any fixed $z^k \in O_0$,

$$\bigcup_{k \leq K} \{y \in \mathcal{X} \,:\, \mathbf{p}(y) \leq \mathbf{p}(z^k)\} = \bigcup_{k \leq K} \bigcap_{n \geq 1} A_{n,k} = \bigcap_{n\geq 1} \bigcup_{k \leq K} A_{n,k}. \tag{8}$$

The last equality comes from the following argument: fix y such that $\forall n \geq 1$, $\exists k_n \leq K$ with $\forall i = 1, ..., n$, $p_{h_i}(y) \leq p_{h_i}(z^{k_n})$. Then, since the sequence $\{k_n\}_n$ takes values in the finite set $\{1, ..., K\}$, there exists a subsequence $\{\sigma_n\}$ and a $k \geq 1$ such that $k_{\sigma_n} = k$ for large n. Therefore, for large n we get $\forall i = 1, ..., \sigma_n$, $p_{h_i}(y) \leq p_{h_i}(z^k)$. Hence, for all $n \geq 1$ we have $\forall i = 1, ..., n$, $p_{h_i}(y) \leq p_{h_i}(z^k)$, which gives (8).

Then for any $n \geq 1$ and $K \geq 1$, since the event $X \in \bigcup_{k\leq K} A_{n,k}$ only involves the finite subset of p-values $\{p_{h_i}(X), i \leq n\}$ and equals $\{p_{h_i}(X), i \leq n\} \in \bigcup_{k\leq K} \{q \in [0,1]^{\{h_i, i\leq n\}} \,:\, \forall i = 1, ..., n,\ q_{h_i} \leq p_{h_i}(z^k)\}$, which is a nonincreasing set of $[0,1]^{\{h_i, i\leq n\}}$, we get from the finite dimensional PRDS property that $\mathbb{P}(X \in \bigcup_{k\leq K} A_{n,k} \,|\, p_h = u) \geq \mathbb{P}(X \in \bigcup_{k\leq K} A_{n,k} \,|\, p_h = u')$. Since $\bigcup_{k\leq K} A_{n,k} \subset \bigcup_{k\leq K} A_{n-1,k}$, by letting $n \to \infty$, we thus have for any $K \geq 1$,

$$\mathbb{P}(X \in \bigcap_{n\geq 1} \bigcup_{k\leq K} A_{n,k} \,|\, p_h = u) \geq \mathbb{P}(X \in \bigcap_{n\geq 1} \bigcup_{k \leq K} A_{n,k} \,|\, p_h = u').$$

Hence, using (8), we get for any $K \geq 1$,

$$\mathbb{P}(X \in \bigcup_{k\leq K} \{y \in \mathcal{X} \,:\, \mathbf{p}(y) \leq \mathbf{p}(z^k)\} \,|\, p_h = u) \geq \mathbb{P}(X \in \bigcup_{k\leq K} \{y \in \mathcal{X} \,:\, \mathbf{p}(y) \leq \mathbf{p}(z^k)\} \,|\, p_h = u'),$$

which finally implies (5), by letting this time $K \to \infty$ and from (7). □

References

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., Delattre, S., and Roquain, E. (2011). Testing over a continuum of null hypotheses.