Bernoulli **22**(1), 2016, 242–274
DOI: 10.3150/14-BEJ657

Non-asymptotic detection of two-component mixtures with unknown means

BÉATRICE LAURENT*, CLÉMENT MARTEAU** and
CATHY MAUGIS-RABUSSEAU†

Institut de Mathématiques de Toulouse, INSA de Toulouse, Université de Toulouse INSA de Toulouse, 135 avenue de Rangueil, 31077 Toulouse Cedex 4, France. E-mail: *beatrice.laurent@insa-toulouse.fr; **clement.marteau@insa-toulouse.fr; †cathy.maugis@insa-toulouse.fr

This work is concerned with the detection of a mixture distribution from a $\mathbb{R}$-valued sample. Given a sample $X_1, \dots, X_n$ and an even density ϕ, our aim is to detect whether the sample distribution is $\phi(\cdot - \mu)$ for some unknown mean μ, or is defined as a two-component mixture based on translations of ϕ. We propose a procedure which is based on several spacings of the order statistics, which provides a level-α test for all n. Our test is therefore a multiple testing procedure and we prove from a theoretical and practical point of view that it automatically adapts to the proportion of the mixture and to the difference of the means of the two components of the mixture under the alternative. From a theoretical point of view, we prove the optimality of the power of our procedure in various situations. A simulation study shows the good performances of our test compared with several classical procedures.

Keywords: Higher Criticism; mixtures; non-asymptotic testing procedure; order statistics; separation rates

1. Introduction

In this paper, the detection problem of a mixture distribution from a $\mathbb{R}$-valued sample is considered. Let $(X_1, \dots, X_n)$ be i.i.d. random variables from an unknown distribution F. All along the paper, F is assumed to admit a density f w.r.t. the Lebesgue measure on $\mathbb{R}$. The sample is said to be distributed from a mixture when f belongs to the set

$$\mathcal{F}_1 = \{x \in \mathbb{R} \mapsto (1-\varepsilon)\phi(x-\mu_1) + \varepsilon\phi(x-\mu_2); \varepsilon \in]0,1[, (\mu_1, \mu_2) \in \mathbb{R}^2, \mu_1 < \mu_2\}, \quad (1.1)$$

where $\phi(\cdot)$ denotes a density. In this paper, $\phi(\cdot)$ is assumed to be an even known density, and when Gaussian mixtures are considered, $\phi(\cdot) = \phi_G(\cdot)$ with

$$\phi_G(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \qquad \forall x \in \mathbb{R}.$$

For a complete introduction about mixtures, we refer to [18]. The two-component mixtures are often encountered in practice, for instance, in biology and health science. They allow to model situations where a population can be discriminated into two different groups. The first subpopulation is then assumed to be distributed following the density $\phi(\cdot - \mu_1)$ while the second one follows the density $\phi(\cdot - \mu_2)$. The probability that an observation X_i arises from the first (resp. the second) subpopulation is then modeled by $1-\varepsilon$ (resp. ε).

1350-7265

This model has been intensively studied and many paths have been explored in order to provide a satisfying inference. In particular, the detection problem has attracted a lot of attention in the last two decades. The main goal is not to provide the best estimation of the parameters of interest $(\varepsilon, \mu_1, \mu_2)$ but rather to decide whether the incoming observations are following a mixture distribution or not. In other words, one wants to detect if the sample of interest comes from a homogeneous or heterogeneous population. Let $\mathcal{F}_0$ be the density set defined as

$$\mathcal{F}_0 = \{x \in \mathbb{R} \mapsto \phi(x - \mu); \mu \in \mathbb{R}\}. \tag{1.2}$$

Formally, one wants to test

$$\text{“}f \in \mathcal{F}_0\text{”} \quad \text{against} \quad \text{“}f \in \mathcal{F}_1\text{”}. \tag{1.3}$$

In various testing problems involving finite mixtures, the properties of the likelihood ratio test have been widely investigated. We can mention for instance [2,10,11,14] among others. In all these papers, the main challenge is to determine the asymptotic behaviour of the likelihood ratio under the alternative hypothesis in order to investigate the power of the related test. Alternative methods have also been considered: modified likelihood ratio test [8], estimation of the L^2 distance between the densities associated to the null and the alternative hypotheses [7], EM approach [9] or tests based on the empirical characteristic function [17].

The main challenge related to the problem (1.3) is to find (optimal) conditions on $(\varepsilon, \mu_1, \mu_2)$ for which a prescribed second kind error can be achieved. The first study in this way is due to Ingster [15], in the particular case where the mean μ under the null hypothesis is known, the term μ_1 in the alternative is equal to μ, and $\phi(\cdot)$ corresponds to a Gaussian density. Similar results have also been obtained in [12]. In this last paper, the so-called Higher Criticism has been investigated. This algorithm is very powerful in the sense that it is easy to implement, and provides similar power than the usual likelihood ratio test. The asymptotic detection regions have been carefully investigated in two different asymptotic regimes:

- the *sparse regime* where $\varepsilon \underset{n\to+\infty}{\sim} n^{-\delta}$ and $\mu_2 - \mu_1 \underset{n\to+\infty}{\sim} \sqrt{2r\log(n)}$ with $\frac{1}{2} < \delta < 1$ and $0 < r < 1$. In this case, it is proved that the two hypotheses can be asymptotically separated if
$$\begin{cases} r > \delta - \frac{1}{2} & \text{when } \frac{1}{2} < \delta \leq \frac{3}{4}, \\ r > (1 - \sqrt{1-\delta})^2 & \text{when } \frac{3}{4} < \delta < 1; \end{cases}$$
- the *dense regime* where $\varepsilon \underset{n\to+\infty}{\sim} n^{-\delta}$ and $\mu_2 - \mu_1 \underset{n\to+\infty}{\sim} n^{-r}$ with $0 < \delta \leq \frac{1}{2}$ and $0 < r < \frac{1}{2}$. In this framework, the separation is asymptotically possible if $r < \frac{1}{2} - \delta$.

In the equations above, the notation $a_n \underset{n\to+\infty}{\sim} b_n$ means that $\lim_{n\to+\infty} a_n/b_n = 1$. We refer for more details to [15] and [12]. Jager and Wellner [16] proposed a family of tests based on the Renyi divergences which generalizes the procedure based on the Higher Criticism. We also mention that generalizations of this procedure to heteroscedastic mixtures have been proposed by Cai *et al.* in [4] while the problems of estimation and construction of confidence sets in sparse mixture models are considered in [5]. Addario-Berry *et al.* [1] determine non-asymptotic separation rates of testing for the contamination of a standard Gaussian vector in $\mathbb{R}^n$ by non-zero mean components when the alternatives have particular combinatorial and geometric structures. More recently, Cai and Wu [6] consider the detection of sparse mixtures in the situation where the density of the observations under the null hypothesis is fixed, but not necessarily Gaussian.

In this paper, we consider a testing problem where the null hypothesis does not correspond to a fixed density but rather to the set of densities $\mathcal{F}_0$ defined by (1.2) which corresponds to a translation model. Thus the mean parameter μ under the null hypothesis is not assumed to be known. The considered alternative $\mathcal{F}_1$ corresponds to the set of densities that are mixtures of two densities of $\mathcal{F}_0$. Our aim is to decide whether the density f of the observations belongs to $\mathcal{F}_0$ or $\mathcal{F}_1$. To this end, we introduce a new testing procedure based on the order statistics. Contrary to the Higher Criticism algorithm [12], the main advantage of this procedure is that the mean μ under H_0 is not fixed. Since one can find densities in $\mathcal{F}_1$ that are arbitrary close to $\mathcal{F}_0$, it is impossible to build a level-α test that achieves a prescribed power on the whole set $\mathcal{F}_1$. Hence, we introduce subsets of $\mathcal{F}_1$ over which our level-α test has a power greater than $1-\beta$. The construction of such subsets more or less amounts to find conditions on $(\varepsilon, \mu_1, \mu_2)$ which ensure that both hypotheses H_0 and H_1 are separable. To this end, we consider as in [12] and [4] two different regimes: the *dense* case where $|\mu_2 - \mu_1|$ is assumed to be bounded and $\varepsilon \geq C/\sqrt{n}$ for all $n \in \mathbb{N}^*$ and for some positive constant C, and the *sparse* regime where ε is allowed to be much smaller than $1/\sqrt{n}$.

The paper is organized as follows. In Section 2, a testing procedure based on the order statistics is introduced. The Section 3 is dedicated to the *dense* regime: we provide non-asymptotic lower and upper bounds for our testing problem in the Gaussian case. Then, we investigate the *sparse* regime in Section 4 for both Gaussian and Laplace distributions. Some numerical simulations, providing a comparison with existing procedures are displayed in Section 5. Proofs are gathered in Section 6 and technical lemmas in the Appendix.

2. The testing procedure

2.1. A test based on the order statistics

Recall that given an i.i.d. sample $X_1, \dots, X_n$ having a common density f w.r.t. the Lebesgue measure on $\mathbb{R}$, our aim is to consider the testing problem $H_0 : f \in \mathcal{F}_0$ against $H_1 : f \in \mathcal{F}_1$, namely to decide whether f corresponds to a given even density function ϕ (up to a translation) or is defined as a two-components mixture of translations of ϕ.

In this context, one of the most popular testing procedures is the Higher Criticism introduced in [12], whose asymptotic behaviour has been widely investigated (see also references above). Nevertheless, there exists up to our knowledge no description of the non-asymptotic performances of this algorithm. Moreover, this procedure heavily depends on the knowledge of the mean under H_0. In this paper, we work in a slightly different framework in the sense that a translation model under H_0 is considered.

In this section, a new testing procedure based on spacings of the order statistics is proposed. The order statistics are denoted by $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$. The main underlying idea is that the spacing of these order statistics are free with respect to the mean under H_0: for some $k < l \in \{1, \dots, n\}$, the mean value affects the spatial position of a given $X_{(k)}$, but not $X_{(l)} - X_{(k)}$.

Moreover, the distribution of the variables $X_{(l)} - X_{(k)}$ is known under H_0 and has a different behavior under H_1, provided k and l are well-chosen.

Let $\alpha \in]0, 1[$ be a fixed level, $\mathbb{P}_f$ the distribution of $X_1, \dots, X_n$ having common density f, and $\mathbb{E}_f$ the corresponding expectation. In the following, a level-α test function T_α denotes a measurable function of $(X_1, \dots, X_n)$ with values in $\{0, 1\}$, such that the null hypothesis is rejected if $T_\alpha = 1$ and $\sup_{f \in \mathcal{F}_0} \mathbb{P}_f(T_\alpha = 1) \leq \alpha$. Assume that $n \geq 2$ and consider the subset $\mathcal{K}_n$ of $\{1, 2, \dots, n/2\}$ defined by

$$\mathcal{K}_n = \left\{2^j, 0 \leq j \leq \left[\log_2(n/2)\right]\right\}.$$

Our test statistics is defined as

$$\Psi_\alpha := \sup_{k \in \mathcal{K}_n} \{\mathbb{1}_{X_{(n-k+1)} - X_{(k)} > q_{\alpha_n,k}}\}, \tag{2.1}$$

where, for all $u \in]0, 1[$, $q_{u,k}$ is the $(1-u)$-quantile of $X_{(n-k+1)} - X_{(k)}$ under the null hypothesis and

$$\alpha_n = \sup\left\{u \in]0, 1[, \mathbb{P}_{H_0}(\exists k \in \mathcal{K}_n, X_{(n-k+1)} - X_{(k)} > q_{u,k}) \leq \alpha\right\}.$$

Note that, by construction, $\alpha_n \leq \alpha$. Since the distribution of $X_{(n-k+1)} - X_{(k)}$ under the null hypothesis is independent of the mean value μ of the X_i's, $q_{\alpha_n,k}$ and α_n can be approximated (via Monte-Carlo simulations for instance) under the assumption that the X_i's have common density ϕ. Below (see in particular Section 6.1), we also provide explicit upper bounds for the quantiles, which can be used instead of the true $q_{\alpha,k}$ if necessary.

2.2. First and second kind errors

By definition, the test statistics Ψ_α introduced in (2.1) is exactly of level α, namely

$$\mathbb{P}_{H_0}(\Psi_\alpha = 1) = \mathbb{P}_{H_0}(\exists k \in \mathcal{K}_n, X_{(n-k+1)} - X_{(k)} > q_{\alpha_n,k}) \leq \alpha,$$

thanks to the definition of α_n. We point out that $\alpha_n \geq \alpha/|\mathcal{K}_n|$, where $|\mathcal{K}_n|$ denotes the cardinality of $\mathcal{K}_n$. Indeed,

$$\begin{aligned}\mathbb{P}_{H_0}(\exists k \in \mathcal{K}_n, X_{(n-k+1)} - X_{(k)} > q_{\alpha/|\mathcal{K}_n|,k}) &\leq \sum_{k \in \mathcal{K}_n} \mathbb{P}_{H_0}(X_{(n-k+1)} - X_{(k)} > q_{\alpha/|\mathcal{K}_n|,k}) \\ &\leq \sum_{k \in \mathcal{K}_n} \frac{\alpha}{|\mathcal{K}_n|} \leq \alpha.\end{aligned}$$

In practice, the choice of α_n, instead of the so-called Bonferroni correction $\alpha/|\mathcal{K}_n|$, allows a numerical improvement of the performances of Ψ_α. We refer to [13] for an extended discussion on this subject.

Now, we turn our attention to the control of the second kind error. We emphasize that the test Ψ_α is a multiple testing procedure: we combine $|\mathcal{K}_n|$ different tests, which correspond to different spacing for the order statistics. We can remark that, for any $f \in \mathcal{F}_1$

$$\begin{aligned}\mathbb{P}_f(\Psi_\alpha = 0) &= \mathbb{P}_f\Big(\sup_{k\in\mathcal{K}_n}\{\mathbb{1}_{X_{(n-k+1)}-X_{(k)}>q_{\alpha_n,k}}\} = 0\Big)\\ &= \mathbb{P}_f\Big(\bigcap_{k\in\mathcal{K}_n}\{\mathbb{1}_{X_{(n-k+1)}-X_{(k)}>q_{\alpha_n,k}} = 0\}\Big)\\ &\le \inf_{k\in\mathcal{K}_n}\mathbb{P}_f(\mathbb{1}_{X_{(n-k+1)}-X_{(k)}>q_{\alpha_n,k}} = 0).\end{aligned}$$

Hence, the second kind error of Ψ_α is close to the smallest one in the collection $\mathcal{K}_n$. In some sense, the "optimal" choice of $k \in \mathcal{K}_n$ is data-driven. The only price to pay for adaptation relies in the "level" α_n, which is smaller than α.

From now on, our aim is to evaluate precisely the power of the test for different kinds of alternatives: dense mixtures (Section 3) or sparse mixtures (Section 4). A general non-asymptotic result is provided in Section 6.1.

3. Dense mixtures

In this section, we assume that the difference between the means μ_1 and μ_2 of the two components of the mixture is bounded. We will see that the settings of interest correspond to the case where $\varepsilon \ge C/\sqrt{n}$ for some constant $C > 0$. In the literature, this regime is called the dense case.

We consider the set of alternatives

$$\mathcal{F}_1[M] = \big\{f(\cdot) = (1-\varepsilon)\phi(\cdot-\mu_1) + \varepsilon\phi(\cdot-\mu_2), \varepsilon \in]0,1[, 0 < \mu_2 - \mu_1 \le M\big\}$$

with $M > 0$. When the density of the standard normal distribution is considered ($\phi = \phi_G$), this set is denoted $\mathcal{F}_{1,G}[M]$.

The aim of this section is to provide explicit conditions on the triplet $(\varepsilon, \mu_1, \mu_2)$ that guarantee a prescribed power for a test of mixture detection, provided that $f \in \mathcal{F}_1[M]$. More precisely, we measure the distance to the null hypothesis by the quantity $d(\varepsilon, \mu_1, \mu_2) = \varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2$ and we assume that $d(\varepsilon, \mu_1, \mu_2) \ge \rho$ for some $\rho > 0$. The question can be therefore formulated as follows: what is the minimal value of ρ to be able to detect the mixture? Under this condition, is the test proposed in Section 2 powerful? We address these two questions for Gaussian mixture models. We also provide a simple test based on the estimation of the variance which is powerful (not only for Gaussian mixtures) in the framework considered in this section.

3.1. Lower bound for the detection of a Gaussian mixture model

In this section, we consider the same definition of non-asymptotic lower bounds for hypotheses testing problems than the ones introduced in [3] for signal detection in a Gaussian regression model or a Gaussian sequence model. Let us recall these definitions. Given $\beta \in]0,1[$, the class of alternatives $\mathcal{F}_1[M]$, and a level-α test T_α with values in $\{0,1\}$ (rejecting H_0 when $T_\alpha = 1$), we define the uniform separation rate $\rho(T_\alpha, \mathcal{F}_1[M], \beta)$ of T_α over the class $\mathcal{F}_1[M]$ as the smallest positive number ρ such that the test has a second kind error at most equal to β for all alternatives f in $\mathcal{F}_1[M]$ such that $d(\varepsilon, \mu_1, \mu_2) = \varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2 \geq \rho$. More precisely,

$$\rho\big(T_\alpha, \mathcal{F}_1[M], \beta\big) = \inf\Big\{\rho > 0, \sup_{f \in \mathcal{F}_1[M], d(\varepsilon,\mu_1,\mu_2)\geq\rho} \mathbb{P}_f(T_\alpha = 0) \leq \beta\Big\}. \tag{3.1}$$

Then, we introduce the (α, β)-minimax separation rate over $\mathcal{F}_1[M]$ defined as

$$\underline{\rho}\big(\mathcal{F}_1[M], \alpha, \beta\big) = \inf_{T_\alpha} \rho\big(T_\alpha, \mathcal{F}_1[M], \beta\big), \tag{3.2}$$

where the infimum is taken over all level-α tests T_α.

We provide in the next theorem a non-asymptotic lower bound for $\underline{\rho}(\mathcal{F}_1[M], \alpha, \beta)$ in the case where ϕ corresponds to the standard Gaussian density.

Theorem 3.1. *Let $\alpha \in]0,1[$ and $\beta \in]0, 1-\alpha[$. Let*

$$\rho^\star = \frac{1}{C(M)}\left(\sqrt{\frac{-2\log[c(\alpha,\beta)]}{n}}\sqrt{1 + \frac{\log[c(\alpha,\beta)]}{2n}}\right),$$

with $c(\alpha, \beta) = 1 - \frac{(1-\alpha-\beta)^2}{2}$ and $C(M) = \sqrt{\frac{1}{2} + \frac{2M^2}{3}\mathrm{e}^{M^2/4}}$. Then for all $\rho < \rho^\star$,

$$\inf_{T_\alpha} \sup_{f \in \mathcal{F}_{1,G}[M], d(\varepsilon,\mu_1,\mu_2)\geq\rho} \mathbb{P}_f(T_\alpha = 0) > \beta,$$

where the infimum is taken over all level-α test T_α. This implies that

$$\underline{\rho}\big(\mathcal{F}_{1,G}[M], \alpha, \beta\big) \geq \rho^\star.$$

Theorem 3.1 implies that whatever the level-α test T_α, if $\rho < \rho^\star$, there exists a density $f \in \mathcal{F}_{1,G}[M]$ for which $\mathbb{P}_f(T_\alpha = 0) > \beta$. In particular, testing is not possible if $\mu_2 - \mu_1$ is too small with respect to $\varepsilon(1-\varepsilon)$. We will show in Section 3.3 that this condition on $(\varepsilon, \mu_1, \mu_2)$ is optimal (up to constant).

3.2. Upper bound for the testing procedure Ψ_α in the Gaussian case

The goal of this section is to give explicit conditions on $(\varepsilon, \mu_1, \mu_2)$ that ensure a prescribed power for the test Ψ_α defined in (2.1), when ϕ is the standard Gaussian density.

Theorem 3.2. *Let $X_1, \ldots, X_n$ be i.i.d. real random variables with common density f. Let $\alpha \in]0,1[$ and consider the level-α test Ψ_α defined by (2.1). Let $\beta \in]0, 1-\alpha[$ and $M > 0$. Assume that n fulfills $n \geq 3$ and $8.25 \times \log(4\log_2(n/2)/\alpha)/n \leq \int_M^\infty \phi_G(x)\,\mathrm{d}x$.*

Then, there exists a positive constant $C(\alpha, \beta, M)$ depending only on α, β and M, such that if

$$\rho \geq C(\alpha, \beta, M)\sqrt{\frac{\log\log(n)}{n}}, \tag{3.3}$$

then,

$$\sup_{f\in\mathcal{F}_{1,G}[M],d(\varepsilon,\mu_1,\mu_2)\geq\rho}\mathbb{P}_f(\Psi_\alpha=0)\leq\beta.$$

Comments. The technical condition on n to get the result of Theorem 3.2 is satisfied for $n\geq 107$ when $M=1/10$ and $\alpha=0.05$.

Note that the value of ρ proposed in (3.3) differs from the lower bound $\rho^\star$ by a term of order $\sqrt{\log\log n}$. This log log term is due to the multiple (adaptive) testing procedure: the optimal value for $k\in\mathcal{K}_n$ in the test Ψ_α is chosen from the data. Hence, this $\sqrt{\log\log(n)}$ term corresponds to the price to pay in such a setting. This kind of logarithmic loss is quite classical in test theory: see for instance [19] or [13] in slightly different settings.

Instead of considering the test statistics Ψ_α defined by (2.1), we could introduce the statistics

$$\mathbb{1}_{X_{(n-k^*+1)}-X_{(k^*)}>q_{\alpha,k^*}},$$

where k^* has to be suitably chosen and depends on M. By this way, we would avoid the logarithmic loss in the minimax separation rate over the set $\mathcal{F}_{1,G}[M]$ and obtain a rate that coincides (up to constants) with the lower bound given in Theorem 3.1 (see the proof of Theorem 3.2). In practice, using the test statistics Ψ_α is more satisfactory since it does not depend on M.

3.3. A testing procedure based on the variance

In this section, we do not assume that the X_i's are Gaussian random variables. We are interested in a simple test based on the variance of the X_i's. We will prove that this test allows us to achieve the lower bound obtained in Theorem 3.1.

Remark that under H_0, $\mathrm{Var}(X_i)=\sigma^2$, where $\sigma^2=\int_{\mathbb{R}}x^2\phi(x)\,\mathrm{d}x$, while under H_1, $\mathrm{Var}(X_i)=\sigma^2+\varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2$. Hence, we consider the test ψ_α defined by

$$\psi_\alpha=\mathbb{1}_{\{S_n^2>v_{\alpha,n}\}},\qquad\text{where } S_n^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X}_n)^2,\tag{3.4}$$

and $v_{\alpha,n}$ denotes the $(1-\alpha)$-quantile of the variable S_n^2 under H_0. Then the following proposition holds.

Proposition 3.1. *Let $\alpha\in]0,1[$ and $\beta\in]0,1-\alpha[$. Assume that the density function ϕ has a finite fourth moment: $\int_{\mathbb{R}}x^4\phi(x)\,\mathrm{d}x\leq B$. There exists a positive constant $C(\alpha,\beta,M,B)$ depending on (α,β,M,B) such that if*

$$\rho\geq C(\alpha,\beta,M,B)/\sqrt{n},\tag{3.5}$$

then

$$\sup_{f\in\mathcal{F}_1[M],d(\varepsilon,\mu_1,\mu_2)\geq\rho}\mathbb{P}_f(\psi_\alpha=0)\leq\beta.$$

In the Gaussian case, $\int_{\mathbb{R}} x^4 \phi_G(x)\,\mathrm{d}x = 3$. Hence, Proposition 3.1 assesses the optimality of the lower bound given in Theorem 3.1. Note that the value of ρ proposed in (3.5) differs from $\rho^\star$ by constant. Finding the optimal constant for our testing problem is a very difficult question that is out of the scope of this paper. For interested reader, we mention the work of [15] in a slightly different (asymptotic) setting.

The result given in Proposition 3.1 seems more efficient than the one stated in Theorem 3.2 since the condition to control by β the second kind error is $\varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2 > C/\sqrt{n}$ instead of $C\sqrt{\log\log(n)}/\sqrt{n}$. Nevertheless, the test based on the variance would fail in the asymptotic sparse regime (see Sections 4 and 4.3 for more details). This is not satisfactory from a practical point of view since our aim is to provide a testing procedure which adapts to all possible situations.

3.4. An asymptotic study

The results stated in Theorems 3.1 and 3.2 are non-asymptotic. In this section, we will adopt an asymptotic point of view for our testing problem in the Gaussian setting. As in [12], we will work with the following parametrization

$$\varepsilon \underset{n\to+\infty}{\sim} n^{-\delta} \quad \text{and} \quad \mu_2-\mu_1 \underset{n\to+\infty}{\sim} n^{-r} \qquad \text{with } 0<\delta\le \tfrac{1}{2} \text{ and } 0<r<\tfrac{1}{2}. \tag{3.6}$$

Corollary 3.1. *The detection boundary in the dense regime* (3.6) *is* $r^*(\delta)=\frac{1}{4}-\frac{\delta}{2}$: *the detection is possible when* $r<r^*(\delta)=\frac{1}{4}-\frac{\delta}{2}$ *and impossible if* $r>r^*(\delta)$.

In particular, setting $f(\cdot)=(1-\varepsilon)\phi_G(\cdot-\mu_1)+\varepsilon\phi_G(\cdot-\mu_2)$, *we have, for n large enough,*

$$\mathbb{P}_f(\Psi_\alpha=0)\le\beta \quad \textit{and} \quad \mathbb{P}_f(\psi_\alpha=0)\le\beta,$$

provided $r<r^*(\delta)$, *where the tests* Ψ_α *and* ψ_α *are respectively, defined in* (2.1) *and* (3.4)

The proof of Corollary 3.1 is omitted since it can be obviously deduced from Theorems 3.1 and 3.2. These results are therefore different from the one obtained in a dense regime in a contamination framework where one wants to test $H_0: f=\phi_G(\cdot)$ against $H_1: f\in\{(1-\varepsilon)\phi_G(\cdot)+\varepsilon\phi_G(\cdot-\mu);\, \varepsilon\in]0,1[,\, \mu\in\mathbb{R}\}$. In this case, as mentioned in introduction, the detection is possible in the dense regime for $r<\frac{1}{2}-\delta$ (see [12,15]). This difference is due to the fact that the mean under H_0 is unknown, which makes the testing problem harder.

4. Sparse mixtures

In the previous part, we have considered the case where the term $\mu_2-\mu_1$ is bounded under the alternative hypothesis. In this section, we will consider the situation where this quantity is allowed to tend to infinity as n increases. It appears that in such a framework, the most interesting cases correspond to the situation where $\varepsilon \ll \frac{1}{\sqrt{n}}$ as $n\to+\infty$. In the literature, this regime is called the sparse case.

This setting has been considered for several different kinds of distributions. In particular, optimal separation conditions on the behavior of $\mu_2 - \mu_1$ as $n \to +\infty$ have been displayed in various situations. In the following, we prove that our testing procedure provides a satisfying behavior in this sparse setting: in particular, we prove that it reaches the optimal separation conditions established in [12] in both the Gaussian and the Laplace cases.

4.1. The Gaussian case

Let $\mathcal{F}_0$ and $\mathcal{F}_1$ be the sets defined by (1.2) and (1.1) respectively. Given an i.i.d. sample $X_1, \ldots, X_n$ having a common density f, we test in this part

$$\text{“}f \in \mathcal{F}_0\text{”} \quad \text{against} \quad \text{“}f \in \mathcal{F}_1\text{”},$$

in the particular case where $\phi(\cdot) = \phi_G(\cdot)$, the standard Gaussian density. In this setting, the so-called *sparse* regime introduced in [12] is characterized by

$$\varepsilon \underset{n\to+\infty}{\sim} n^{-\delta} \quad \text{and} \quad \mu_2 - \mu_1 \underset{n\to+\infty}{\sim} \sqrt{2r\log(n)} \qquad \text{with } \tfrac{1}{2} < \delta < 1 \text{ and } 0 < r < 1. \tag{4.1}$$

Below, we analyze the performances of our testing procedure (2.1) in this *sparse* regime. The corresponding proof is provided in Section 6.6.

Theorem 4.1. *Let $X_1, \ldots, X_n$ be i.i.d. real random variables with common density f. Let $\alpha \in]0,1[$ and consider the level-α test Ψ_α defined by* (2.1). *We consider the case where $\phi = \phi_G$.*

We assume that the behavior of $(\varepsilon, \mu_1, \mu_2)$ is governed by (4.1) *and that $r > r^*(\delta)$ with*

$$r^*(\delta) = \begin{cases} \delta - \frac{1}{2} & \text{if } \frac{1}{2} < \delta < \frac{3}{4}, \\ (1 - \sqrt{1-\delta})^2 & \text{if } \frac{3}{4} \le \delta < 1. \end{cases}$$

Then, setting $f(\cdot) = (1-\varepsilon)\phi_G(\cdot - \mu_1) + \varepsilon\phi_G(\cdot - \mu_2)$, we have, for n large enough,

$$\mathbb{P}_f(\Psi_\alpha = 0) \le \beta.$$

In the sparse regime, we exactly recover the separation boundaries that are already known in the case where the null hypothesis is reduced to a standard normal density, and the alternative is the mixture $(1-\varepsilon)\phi_G(\cdot) + \varepsilon\phi_G(\cdot - \mu)$. Hence, the fact that the mean under H_0 is unknown does not affect the difficulty of the related testing problem in this specific framework.

This proves the optimality of our procedure in the sparse regime. Indeed, the lower bounds established by [4,15] in the case where the null hypothesis is reduced to the standard Gaussian density also provide lower bounds for our testing problem. This comes from the fact that

- a level-α test for our testing problem is also a level-α test for testing the null hypothesis “$f = \phi_G$”,
- the case where the null hypothesis is reduced to the centered Gaussian density is included in our setting.

4.2. The Laplace case

In this section, we address the testing problem (1.3) in the particular case where ϕ corresponds to the Laplace density, namely $\phi = \phi_L$ where

$$\phi_L(x) = \tfrac{1}{2}\mathrm{e}^{-|x|}, \qquad \forall x \in \mathbb{R}.$$

In other words, given a sample $X_1, \dots, X_n$, our aim is to test whether the underlying density is $\phi_L(\cdot - \mu)$ for some unknown parameter μ or $(1-\varepsilon)\phi_L(\cdot - \mu_1) + \varepsilon\phi_L(\cdot - \mu_2)$ in the particular case where $\varepsilon = \mathrm{o}(1/\sqrt{n})$ as $n \to +\infty$.

In this context, [12] have proved that the cases of interest in the *sparse* regime correspond to the following parametrization

$$\varepsilon \underset{n\to+\infty}{\sim} n^{-\delta} \quad \text{and} \quad \mu_2 - \mu_1 \underset{n\to+\infty}{\sim} r\log(n) \qquad \text{with } \tfrac{1}{2} < \delta < 1 \text{ and } 0 < r < 1. \tag{4.2}$$

The performances of our testing procedure (2.1) are described in the following theorem, whose proof is given in Section 6.7.

Theorem 4.2. *Let $X_1, \dots, X_n$ be i.i.d. real random variables with common density f. Let $\alpha \in]0,1[$ and consider the level-α test Ψ_α defined by (2.1). We consider the case where $\phi = \phi_L$.*

We assume that the behavior of $(\varepsilon, \mu_1, \mu_2)$ is governed by (4.2) and that $r > r^*(\delta)$ with

$$r^*(\delta) = 2\delta - 1.$$

Then, setting $f(\cdot) = (1-\varepsilon)\phi_L(\cdot - \mu_1) + \varepsilon\phi_L(\cdot - \mu_2)$, we have, for n large enough,

$$\mathbb{P}_f(\Psi_\alpha = 0) \le \beta.$$

Remark that the detection boundary $r^*(\delta)$ is the same that have been exhibited by [12]. Once again, these lower bounds remain valid since:

- a level-α test for our testing problem is also a level-α test for testing the null hypothesis "$f = \phi_L$",
- the case where the null hypothesis is reduced to the centered Laplace density is included in our setting.

4.3. The variance test for sparse mixtures: A heuristic discussion

We point out that the testing procedure introduced in Section 3.3 will not be convenient in this asymptotic sparse setting. Indeed, we can remark that

$$\mathrm{Var}_\phi(X_i) = \int_{\mathbb{R}} x^2 \phi(x)\,\mathrm{d}x,$$

while, for any $f = (1-\varepsilon)\phi(\cdot - \mu_1) + \varepsilon\phi(\cdot - \mu_2)$

$$\mathrm{Var}_f(X_i) = \int_{\mathbb{R}} x^2\phi(x)\,\mathrm{d}x + \varepsilon(1-\varepsilon)(\mu_1 - \mu_2)^2.$$

For both Gaussian and Laplace mixtures, in the respective asymptotic schemes (4.1) and (4.2), we get that

$$\mathrm{Var}_f(X_i) - \mathrm{Var}_\phi(X_i) = \varepsilon(1-\varepsilon)(\mu_1 - \mu_2)^2 \ll \frac{1}{\sqrt{n}}, \qquad \text{as } n \to +\infty.$$

Since the variance is estimated at a parametric "rate" $1/\sqrt{n}$, the test ψ_α introduced in (3.4) will fail in this setting: it will not be able to separate H_0 from H_1 with an appropriate power.

5. Simulation study

In this section, we provide some numerical experiments in order to enhance the performances of our testing procedure Ψ_α. Comparisons with the Higher Criticism and the Kolmogorov–Smirnov test are provided. Since these both procedures are not designed for the considered framework (translated model with unknown mean), straightforward modifications are proposed. We have also included in these numerical experiments the test based on the variance defined in Section 3.3.

5.1. Contamination of ϕ_G

In this section, we deal with the framework considered in [12]: the mean under H_0 is assumed to be known (equal to 0) and equal to μ_1. More formally, given $(X_1, \dots, X_n)$, i.i.d. random variables with an unknown density function f, our aim is to test

$$H_0 : f(\cdot) = \phi_G(\cdot) \text{ against } H_1 : f \in \bigl\{x \mapsto (1-\varepsilon)\phi_G(x) + \varepsilon\phi_G(x-\mu);\, \mu \in \mathbb{R}, \varepsilon \in]0,1[\bigr\}. \quad (5.1)$$

In this case, our testing procedure Ψ_α described in (2.1) can be easily adapted as follows:

$$\tilde{\Psi}_\alpha = \sup_{k \in \mathcal{K}_n} \{\mathbb{1}_{X_{(n-k+1)} > q_{\alpha_n,k}}\},$$

where $q_{\alpha,k}$ is the $(1-\alpha)$-quantile of $X_{(n-k+1)}$ under the null hypothesis, $\mathcal{K}_n = \{2^j; 0 \le j \le [\log_2(n/2)]\}$ and

$$\alpha_n = \sup\bigl\{u \in]0,1[, \mathbb{P}_{H_0}(\exists k \in \mathcal{K}_n, X_{(n-k+1)} > q_{u,k}) \le \alpha\bigr\}.$$

For the sake of brevity, we do not exhibit a theoretical study of the performances of this procedure for the testing problem (5.1). Indeed, the methodology is rather close to the one proposed in this paper, up to some technical modifications. It is possible to see that this procedure achieves the optimal asymptotic separation set in both the *dense* and *sparse* regimes, as described in [12].

The power of our testing procedure is compared with the one of

- Kolmogorov–Smirnov test: The level-α test function is $\psi_{\mathrm{KS},\alpha} = \mathbb{1}_{T_{\mathrm{KS}} > q_{\mathrm{KS},\alpha}}$ where

$$T_{\mathrm{KS}} = \sup_{x \in \mathbb{R}} \sqrt{n} \left| F_n(x) - \Phi_G(x) \right|$$

  with the empirical distribution function $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$, Φ_G the cumulative distribution function of the standard Gaussian variable, and $q_{\mathrm{KS},\alpha}$ is the $(1-\alpha)$ quantile of T_{KS} under H_0.
- Higher Criticism [12]: Let $p_i = \mathbb{P}(Z > X_i)$ where $Z \sim \mathcal{N}(0,1)$ for all $i \in \{1, \dots, n\}$ and $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(n)}$. This test is based on

$$\mathrm{HC} = \max_{1 \leq i \leq n} \frac{\sqrt{n}(i/n - p_{(i)})}{\sqrt{p_{(i)}(1 - p_{(i)})}}.$$

  The level-α test function is $\psi_{\mathrm{HC},\alpha} = \mathbb{1}_{\mathrm{HC} > q_{\mathrm{HC},\alpha}}$ where $q_{\mathrm{HC},\alpha}$ is the $(1-\alpha)$ quantile of HC under H_0.
- The test based on the variance (see Section 3.3).

In order to study the power of these testing procedures, a Monte-Carlo procedure is considered with $N = 100\,000$ samples of size $n = 100$ from a mixture distribution $(1-\varepsilon)\phi_G(\cdot) + \varepsilon\phi_G(\cdot - \mu)$ with $\varepsilon \in \{0.05, 0.15, 0.25, 0.35, 0.45\}$ and $\mu \in [0, 10]$. The power functions of these testing procedures in the different scenarios are reported in Figure 1.

It appears that our procedure performs as well as the Higher Criticism when ε is small w.r.t. the size of the sample, while the Kolmogorov–Smirnov test possesses a bad behavior. Such a setting is close to the *sparse* regime. Nevertheless, the performances of the Higher Criticism deteriorates as ε increases while the power of our test $\tilde{\Psi}_\alpha$ remains stable. In this setting, the test based on the variance does not perform very well. The main reason is that, in this case, the mean under H_0 is known. Hence, a test based on the empirical mean of the observations would be more appropriate.

5.2. Gaussian mixtures with unknown means

In this section, we deal with our testing problem. A simulation study is proposed in order to investigate the power of our testing procedure Ψ_α described by (2.1). Our testing procedure is compared with the following adaptations of Kolmogorov–Smirnov test and Higher Criticism:

- Kolmogorov–Smirnov test: The level-α test function is $\hat{\psi}_{\mathrm{KS},\alpha} = \mathbb{1}_{\hat{T}_{\mathrm{KS}} > \hat{q}_{\mathrm{KS},\alpha}}$ where

$$\hat{T}_{\mathrm{KS}} = \sup_{x \in \mathbb{R}} \sqrt{n} \left| F_n(x) - \Phi_G(x - \bar{X}) \right|$$

  with the empirical mean $\bar{X}$, the empirical distribution function $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$, and $\hat{q}_{\mathrm{KS},\alpha}$ is the $(1-\alpha)$ quantile of $\hat{T}_{\mathrm{KS}}$ under H_0.

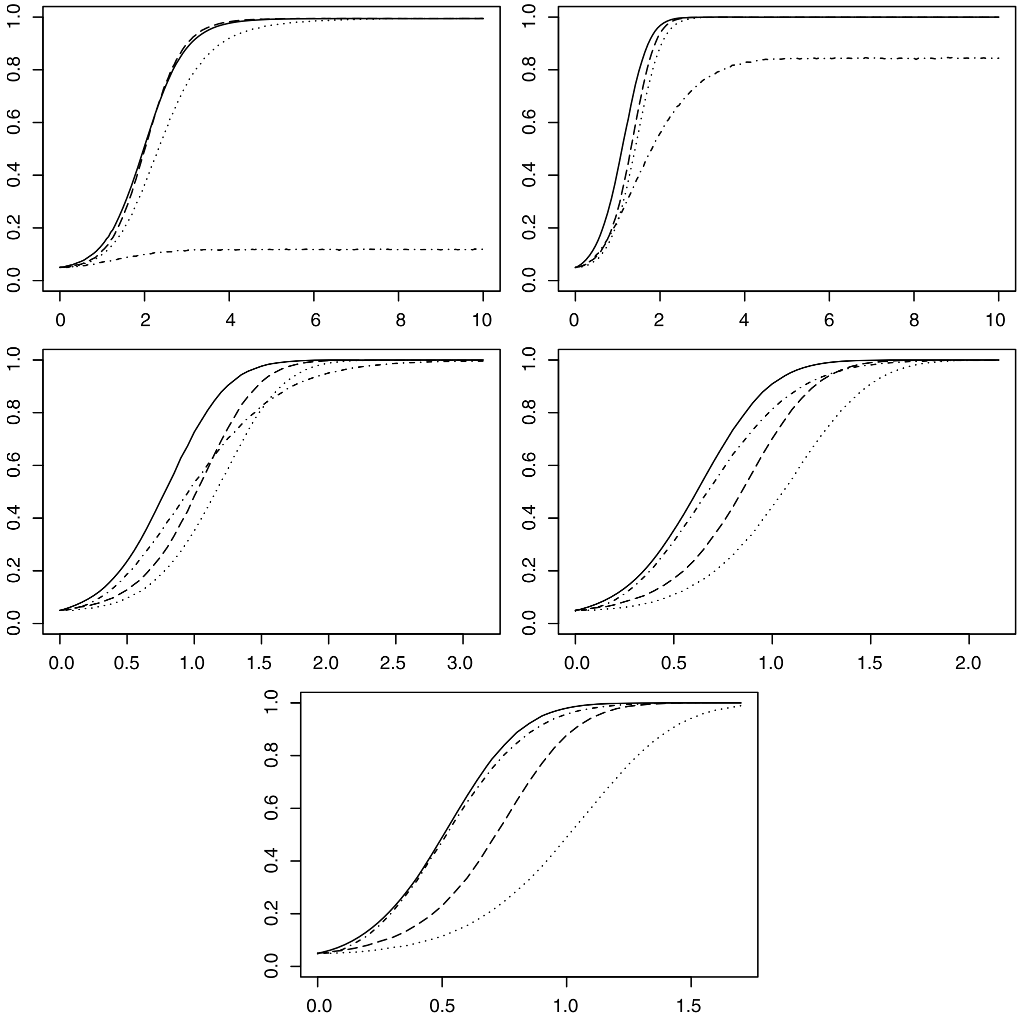

Figure 1. Power function of the three considered testing procedures (continuous line for our test $\tilde{\Psi}_\alpha$, dashed line for Higher Criticism, dashed/dotted line for the Kolmogorov–Smirnov test and dotted line for the test based on the variance) according to μ, for $\varepsilon = 0.05$ (top left), 0.15 (top right), 0.25 (middle left), 0.35 (middle right) and 0.45 (bottom) in a contamination framework.

- Higher Criticism [12]: Let $\hat{p}_i = \mathbb{P}(Z - \bar{X} > X_i)$ where $Z \sim \mathcal{N}(0, 1)$ for all $i \in \{1, \dots, n\}$ and $\hat{p}_{(1)} \le \hat{p}_{(2)} \le \cdots \le \hat{p}_{(n)}$. This test is based on

$$\hat{\mathrm{HC}} = \max_{1 \le i \le n} \frac{\sqrt{n}(i/n - \hat{p}_{(i)})}{\sqrt{\hat{p}_{(i)}(1 - \hat{p}_{(i)})}}.$$

The level-α test function is $\hat{\psi}_{\mathrm{HC},\alpha} = \mathbb{1}_{\hat{\mathrm{HC}} > \hat{q}_{\mathrm{HC},\alpha}}$ where $\hat{q}_{\mathrm{HC},\alpha}$ is the $(1-\alpha)$ quantile of $\hat{\mathrm{HC}}$ under H_0.

- The test based on the variance (see Section 3.3).

In order to study the power of these testing procedures, a Monte-Carlo procedure is considered with $N = 100\,000$ samples of size $n = 100$ from a mixture distribution $(1-\varepsilon)\phi_G(\cdot) + \varepsilon\phi_G(\cdot - \mu_2)$ with $\varepsilon \in \{0.05, 0.15, 0.25, 0.35, 0.45\}$. We deal with $\mu_1 = \mu = 0$ and $\mu_2 \in [0, 10]$. The power functions of these testing procedures in the different scenarios are reported in Figure 2.

Once again, our testing procedure appears to be competitive w.r.t. the existing procedures, and even offers better performances in some particular cases. As in the previous experiment, the behavior of the Higher Criticism deteriorates w.r.t. our procedure as ε increases, namely when we leave the *sparse* regime to the *dense* one. In this setting, the test based on the variance is quite competitive.

Remark that the considered setting is not asymptotic at all since the sample size is 100. As explained in Section 4.3, one can expect that the performances of the test based on the variance will deteriorate in a sparse asymptotic regime. In order to illustrate this discussion, we have compared the test based on the variance and our procedure in a very sparse context where $n = 1000$ and $\varepsilon = 0.001$. The corresponding values of the power are displayed in Table 1.

5.3. Laplace mixtures with unknown means

Since our test Ψ_α is adapted for an even density function ϕ, a Laplace distribution is here considered: $\phi_L(x) = \frac{1}{2}\exp(-|x|)$. As in Section 5.2, the power of Ψ_α is compared with the one of Kolmogorov–Smirnov test and Higher Criticism. Note that these two last tests are adapted as in Section 5.2 but where Φ and Z are now associated to the Laplace distribution. The variance-based test introduced in Section 3.3 is also included in these simulations.

A Monte-Carlo procedure is proposed with $N = 100\,000$ samples of size $n = 100$ from a mixture distribution $(1-\varepsilon)\phi(\cdot) + \varepsilon\phi(\cdot - \mu_2)$ with $\varepsilon \in \{0.05, 0.15, 0.25, 0.35, 0.45\}$ and $\mu \in [0, 10]$. The power functions of these testing procedures in the different scenarios are reported in Figure 3.

Apart in the case where $\varepsilon = 0.05$, our test outperforms Higher Criticism, Kolmogorov–Smirnov and variance-based tests in all other conditions. As previously, the power of Higher Criticism is deteriorated as ε increases.

6. Proofs

6.1. A preliminary result

In this section, we provide a general result that emphasizes the non-asymptotic performances of our testing procedure.

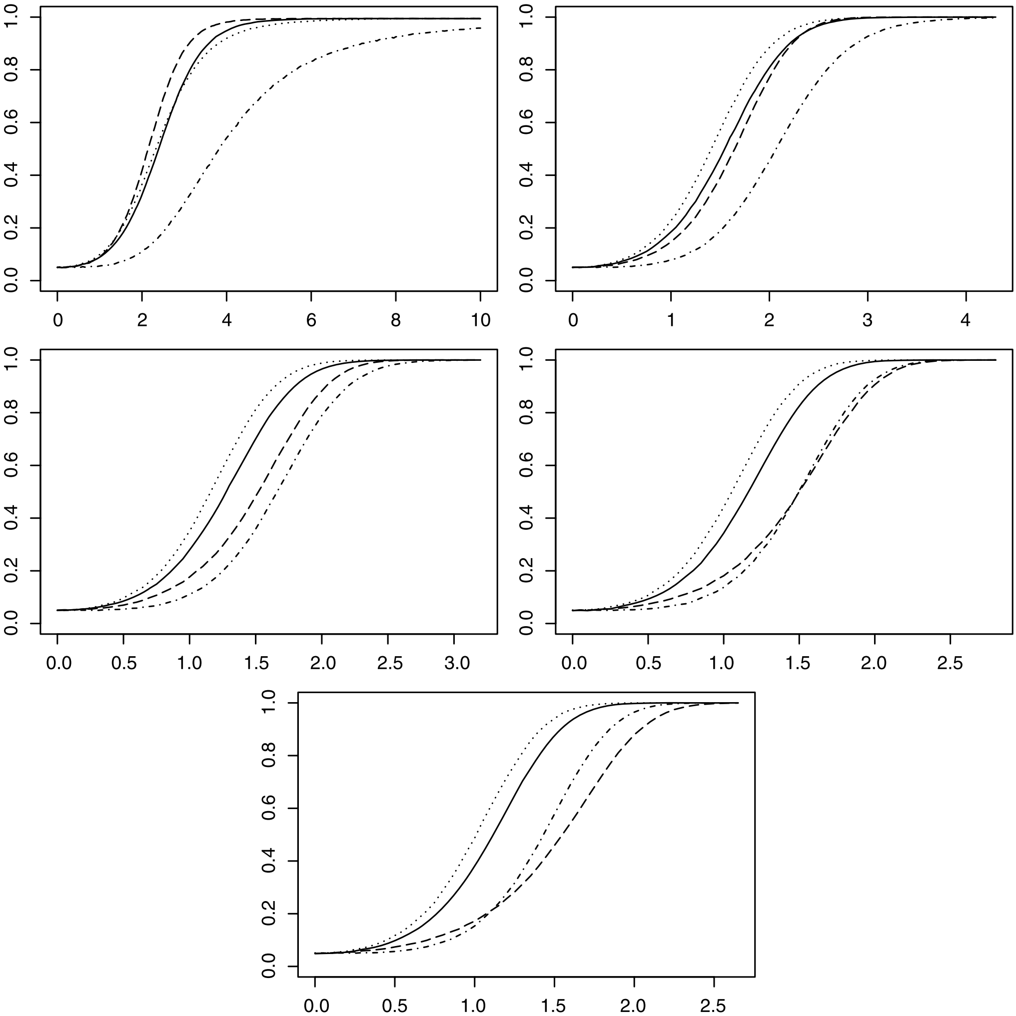

Figure 2. Power function of the three considered testing procedures (continuous line for our test Ψ_α, dashed line for Higher Criticism, dashed/dotted line for the Kolmogorov–Smirnov test and dotted line for the test based on the variance) according to μ_2, for $\varepsilon = 0.05$ (top left), 0.15 (top right), 0.25 (middle left), 0.35 (middle right) and 0.45 (bottom) in the Gaussian mixture framework.

Let $\bar{\Phi}(x) = 1 - \Phi(x)$, where Φ is the cumulative distribution function associated to the density function ϕ. For all $\alpha \in]0, 1[$ and $k \in \{1, 2, \dots, n/2\}$, let $t_{\alpha,k}$ be a positive real number defined by

$$\bar{\Phi}\left(\frac{t_{\alpha,k}}{2}\right) = \frac{k}{n}\left[1 - \sqrt{\frac{2\log(4/\alpha)}{k}}\right] \tag{6.1}$$

Table 1. Comparison of the power of the variance based test (VB) and our procedure (LMM) for $\varepsilon = 0.001$ and $n = 1000$

μ_2	2	4	6	8
LMM	0.0642	0.3006	0.6131	0.6513
VB	0.0596	0.1147	0.2445	0.405

if $k > 2\log(\frac{4}{\alpha})$, and $t_{\alpha,k} = +\infty$ otherwise. For all $\alpha \in]0,1[$, $\rho > 0$, and $k \in \{1, 2, \dots, n/2\}$, we consider the subset $\bar{\mathcal{S}}(\alpha, \rho, k)$ of $\mathbb{R}^3$ defined by:

$$\bar{\mathcal{S}}(\alpha, \rho, k) = \left\{ \begin{array}{l} (\varepsilon, \mu_1, \mu_2) \in]0,1[\times \mathbb{R}^2, \mu_2 > \mu_1; \exists c \in \mathbb{R} \text{ such that:} \\ (1-\varepsilon)\bar{\Phi}\big(t_{\alpha,k} - c + \varepsilon(\mu_2 - \mu_1)\big) + \varepsilon\bar{\Phi}\big(t_{\alpha,k} - c - (1-\varepsilon)(\mu_2 - \mu_1)\big) > \rho \\ (1-\varepsilon)\bar{\Phi}\big(c - \varepsilon(\mu_2 - \mu_1)\big) + \varepsilon\bar{\Phi}\big(c + (1-\varepsilon)(\mu_2 - \mu_1)\big) > \rho \end{array} \right\}. \tag{6.2}$$

When $t_{\alpha,k} = +\infty$, we use the convention $\bar{\mathcal{S}}(\alpha, \rho, k) = \varnothing$ for all $\rho > 0$.

The following proposition highlights the non-asymptotic performances of the test Ψ_α.

Theorem 6.1. *Let $\alpha \in]0,1[$ and $\beta \in]0, 1-\alpha[$. Consider the test Ψ_α described in (2.1). We assume that $n \geq 8\log(4/\alpha_n)$. Consider the alternative sets*

$$\bar{\mathcal{F}}_1[n, \alpha, \beta] = \left\{ f(\cdot) = (1-\varepsilon)\phi(\cdot - \mu_1) + \varepsilon\phi(\cdot - \mu_2); (\varepsilon, \mu_1, \mu_2) \in \bigcup_{k \in \mathcal{K}_n} \bar{\mathcal{S}}\big(\alpha_n, \rho(k,n), k\big) \right\}$$

where, for all $k \in \mathcal{K}_n$, $\bar{\mathcal{S}}(\alpha_n, \rho(k,n), k)$ is defined by (6.2) with

$$\rho(k,n) = \frac{k}{n} + \frac{1 + \sqrt{1 + 2k\beta}}{n\beta}.$$

Then Ψ_α is a level-α test and

$$\sup_{f \in \bar{\mathcal{F}}_1[n,\alpha,\beta]} \mathbb{P}_f(\Psi_\alpha = 0) \leq \beta.$$

In this theorem, we have defined a set $\bar{\mathcal{F}}_1[n, \alpha, \beta]$ over which the level-α test statistics Ψ_α has a power greater than $1-\beta$. This result holds for all n, provided that $n \geq 8\log(4/\alpha_n)$, it is non-asymptotic. The definition of the set $\bar{\mathcal{S}}(\alpha, \rho, k)$ is quite rough. Nevertheless, it will allow us to describe several situations for which the power of our testing procedure will be assessed, in both asymptotic and non-asymptotic cases.

The condition $n \geq 8\log(4/\alpha_n)$ ensures that there exists $k \in \mathcal{K}_n$ such that $k > 2\log(4/\alpha_n)$. Since $\alpha_n \geq \alpha/|\mathcal{K}_n|$, and $|\mathcal{K}_n| \leq \log_2(n/2)$, this condition is satisfied if $n \geq 8\log(4\log_2(n/2)/\alpha)$. For $\alpha = 0.05$, this condition holds at least for $n \geq 49$.

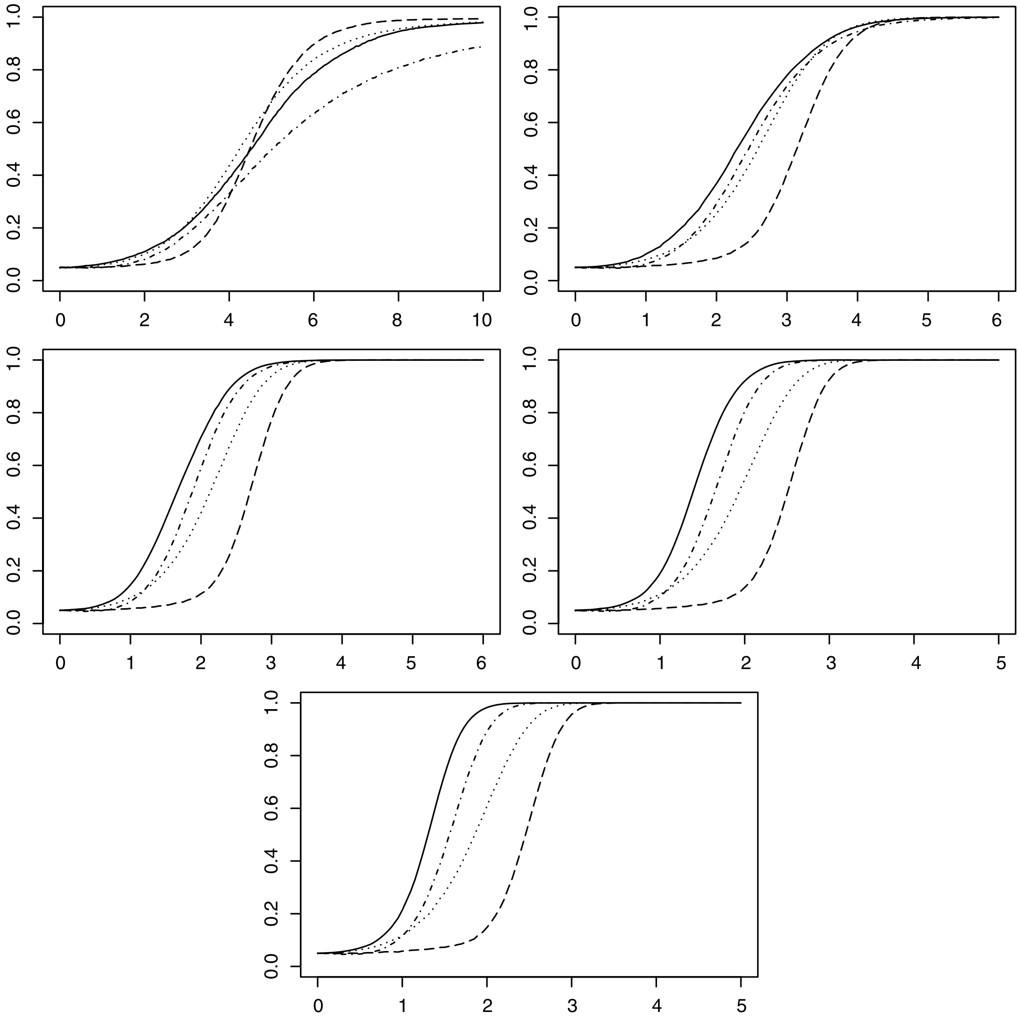

Figure 3. Power function of the three considered testing procedures (continuous line for our test Ψ_α, dashed line for Higher Criticism, dashed/dotted line for the Kolmogorov–Smirnov test and dotted line for the test based on the variance) according to μ_2, for $\varepsilon = 0.05$ (top left), 0.15 (top right), 0.25 (middle left), 0.35 (middle right) and 0.45 (bottom) in the Laplace mixture framework.

6.2. Proof of Theorem 6.1

Following the definition of α_n, Ψ_α is ensured to be a level-α test. In order to control the second kind error of the test Ψ_α, we first give an upper bound for $q_{\alpha_n,k}$. Under the null hypothesis, there exists $\mu \in \mathbb{R}$ such that $f(\cdot) = \phi(\cdot - \mu)$. Thus $X_{(n-k+1)} - X_{(k)}$ is distributed as $Y_{(n-k+1)} - Y_{(k)}$ where $(Y_1, \ldots, Y_n)$ is a n sample from the density $\phi(\cdot)$. Hence, if we find $c_{\alpha_n,k}$ such that

$\mathbb{P}(Y_{(n-k+1)} - Y_{(k)} > c_{\alpha_n,k}) \le \alpha_n$ then $q_{\alpha_n,k} \le c_{\alpha_n,k}$. For all $d \in \mathbb{R}$,

$$\mathbb{P}(Y_{(n-k+1)} - Y_{(k)} > c_{\alpha_n,k}) \le \mathbb{P}(Y_{(n-k+1)} > c_{\alpha_n,k} + d) + \mathbb{P}(Y_{(k)} \le d).$$

According to Lemma A.1, if d fulfills $\Phi(d) \le \frac{k}{n}[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}]$ then $\mathbb{P}(Y_{(k)} \le d) \le \frac{\alpha_n}{2}$. Moreover, by the same lemma, if $c_{\alpha_n,k}$ is chosen such that $\bar{\Phi}(c_{\alpha_n,k} + d) \le \frac{k}{n}[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}]$ then $\mathbb{P}(Y_{(n-k+1)} \ge c_{\alpha_n,k} + d) \le \frac{\alpha_n}{2}$. Choosing d and $c_{\alpha_n,k}$ exactly such that

$$\Phi(d) = \bar{\Phi}(c_{\alpha_n,k} + d) = \frac{k}{n}\left[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right]$$

and since $\phi(\cdot)$ is an even continuous function, we obtain that $d = -\frac{c_{\alpha_n,k}}{2}$. Finally, choosing $c_{\alpha_n,k} = t_{\alpha_n,k}$ where $\bar{\Phi}(\frac{t_{\alpha_n,k}}{2}) = \frac{k}{n}[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}]$, $\mathbb{P}_{H_0}(X_{(n-k+1)} - X_{(k)} > t_{\alpha_n,k}) \le \alpha_n$ and thus $q_{\alpha_n,k} \le t_{\alpha_n,k}$.

Considering $f \in \bar{\mathcal{F}}_1[n, \alpha, \beta]$, we want to control the second kind error of the test:

$$\begin{aligned}\mathbb{P}_f(\Psi_\alpha = 0) &= \mathbb{P}_f(\forall k \in \mathcal{K}_n, X_{(n-k+1)} - X_{(k)} \le q_{\alpha_n,k}) \\ &\le \inf_{k\in\mathcal{K}_n} \mathbb{P}_f(X_{(n-k+1)} - X_{(k)} \le q_{\alpha_n,k}).\end{aligned} \tag{6.3}$$

Since $f \in \bar{\mathcal{F}}_1[n, \alpha, \beta]$, there exist $\varepsilon \in]0, 1[$ and $(\mu_1, \mu_2) \in \mathbb{R}^2$, $\mu_1 < \mu_2$ such that

$$\forall x \in \mathbb{R}, \qquad f(x) = (1-\varepsilon)\phi(x - \mu_1) + \varepsilon\phi(x - \mu_2)$$

and for some $k \in \mathcal{K}_n$, there exists a real c such that $(\varepsilon, \mu_1, \mu_2)$ fulfills the two following conditions:

$$(1-\varepsilon)\bar{\Phi}\big(t_{\alpha_n,k} - c + \varepsilon(\mu_2 - \mu_1)\big) + \varepsilon\bar{\Phi}\big(t_{\alpha_n,k} - c - (1-\varepsilon)(\mu_2 - \mu_1)\big) > \rho(k, n), \tag{6.4}$$

$$(1-\varepsilon)\bar{\Phi}\big(c - \varepsilon(\mu_2 - \mu_1)\big) + \varepsilon\bar{\Phi}\big(c + (1-\varepsilon)(\mu_2 - \mu_1)\big) > \rho(k, n), \tag{6.5}$$

with $\rho(k, n) = \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta}$. Using (6.3) and the fact that $q_{\alpha_n,k} \le t_{\alpha_n,k}$,

$$\begin{aligned}\mathbb{P}_f(X_{(n-k+1)} - X_{(k)} \le q_{\alpha_n,k}) &\le \mathbb{P}_f(X_{(n-k+1)} - X_{(k)} \le t_{\alpha_n,k}) \\ &\le \mathbb{P}_f\big(X_{(n-k+1)} \le t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\big) \\ &\quad + \mathbb{P}_f\big(X_{(k)} > \mathbb{E}_f[X_1] - c\big).\end{aligned} \tag{6.6}$$

For the first term in the right-hand side of (6.6),

$$\begin{aligned}\mathbb{P}_f\big(X_{(n-k+1)} \le t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\big) &\le \mathbb{P}_f\left(\sum_{i=1}^n \mathbb{1}_{\{X_i \le t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\}} > n - k\right) \\ &\le \mathbb{P}_f\left(\sum_{i=1}^n \{\mathbb{1}_{\{X_i \le t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\}} - q_1\} > n(1 - q_1) - k\right)\end{aligned}$$

with

$$\begin{aligned} q_1 &= \mathbb{P}_f\big(X_1 \le t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\big) \\ &= (1-\varepsilon)\Phi\big(t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c - \mu_1\big) + \varepsilon\Phi\big(t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c - \mu_2\big) \\ &= (1-\varepsilon)\Phi\big(t_{\alpha_n,k} - c + \varepsilon(\mu_2-\mu_1)\big) + \varepsilon\Phi\big(t_{\alpha_n,k} - c - (1-\varepsilon)(\mu_2-\mu_1)\big) \end{aligned}$$

since $\mathbb{E}_f[X_1] = (1-\varepsilon)\mu_1 + \varepsilon\mu_2$. Condition (6.4) gives that $n(1-q_1) - k > 0$ and using Markov's inequality,

$$\mathbb{P}_f\big(X_{(n-k+1)} < t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\big) \le \frac{nq_1(1-q_1)}{[n(1-q_1)-k]^2} \le \frac{n(1-q_1)}{[n(1-q_1)-k]^2}.$$

Note that the inequality $\frac{nx}{(nx-k)^2} \le \frac{\beta}{2}$ is fulfilled if and only if $x \notin [\frac{k}{n} + \frac{1}{\beta n} \pm \frac{\sqrt{1+2k\beta}}{\beta n}]$. Then, since condition (6.4) ensures us that $1-q_1 \notin [\frac{k}{n} + \frac{1}{n\beta} \pm \frac{\sqrt{1+2k\beta}}{n\beta}]$,

$$\mathbb{P}_f\big(X_{(n-k+1)} < t_{\alpha_n,k} + \mathbb{E}_f[X_1] - c\big) \le \frac{\beta}{2}.$$

For the second term in the right-hand side of (6.6),

$$\mathbb{P}_f\big(X_{(k)} > \mathbb{E}_f[X_1] - c\big) \le \mathbb{P}_f\left(\sum_{i=1}^n \{\mathbb{1}_{\{X_i > \mathbb{E}_f[X_1]-c\}} - q_2\} > n(1-q_2) - k\right)$$

with

$$\begin{aligned} q_2 &= \mathbb{P}_f\big(X_1 > \mathbb{E}_f[X_1] - c\big) \\ &= (1-\varepsilon)\bar{\Phi}\big(\mathbb{E}_f[X_1] - c - \mu_1\big) + \varepsilon\bar{\Phi}\big(\mathbb{E}_f[X_1] - c - \mu_2\big) \\ &= (1-\varepsilon)\bar{\Phi}\big(-c + \varepsilon(\mu_2-\mu_1)\big) + \varepsilon\bar{\Phi}\big(-c - (1-\varepsilon)(\mu_2-\mu_1)\big) \\ &= (1-\varepsilon)\Phi\big(c - \varepsilon(\mu_2-\mu_1)\big) + \varepsilon\Phi\big(c + (1-\varepsilon)(\mu_2-\mu_1)\big). \end{aligned}$$

Condition (6.5) gives that $n(1-q_2) - k > 0$ and using Markov's inequality,

$$\mathbb{P}_f\big(X_{(k)} > \mathbb{E}_f[X_1] - c\big) \le \frac{nq_2(1-q_2)}{[n(1-q_2)-k]^2} \le \frac{n(1-q_2)}{[n(1-q_2)-k]^2}.$$

According to condition (6.5), $1-q_2 \notin [\frac{k}{n} + \frac{1}{n\beta} \pm \frac{\sqrt{1+2k\beta}}{n\beta}]$, thus

$$\mathbb{P}_f\big(X_{(k)} > \mathbb{E}_f[X_1] - c\big) \le \frac{\beta}{2}.$$

Finally, $\mathbb{P}_f(\Psi_\alpha = 0) \le \beta$.

6.3. Proof of Theorem 3.1

We define

$$\mathcal{F}_{1,G}[\rho, M] = \{f \in \mathcal{F}_{1,G}[M], \varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2 \geq \rho\}.$$

Let T_α be a level-α test. For all $f \in \mathcal{F}_{1,G}[\rho, M]$,

$$\begin{aligned}\mathbb{P}_f(T_\alpha = 0) &= \mathbb{P}_{\phi_G}(T_\alpha = 0) + \mathbb{P}_f(T_\alpha = 0) - \mathbb{P}_{\phi_G}(T_\alpha = 0)\\ &\geq 1-\alpha - [\mathbb{P}_{\phi_G}(T_\alpha = 0) - \mathbb{P}_f(T_\alpha = 0)].\end{aligned}$$

Thus for a density $\tilde{f} \in \mathcal{F}_{1,G}[\rho, M]$ which has to be specified after,

$$\begin{aligned}\sup_{f\in\mathcal{F}_{1,G}[\rho,M]} \mathbb{P}_f(T_\alpha = 0) &\geq 1-\alpha - [\mathbb{P}_{\phi_G}(T_\alpha = 0) - \mathbb{P}_{\tilde{f}}(T_\alpha = 0)]\\ &\geq 1-\alpha - \|\mathbb{P}_{\phi_G} - \mathbb{P}_{\tilde{f}}\|_{\mathrm{TV}},\end{aligned}$$

where $\|P-Q\|_{\mathrm{TV}}$ denotes the total variation distance between two probability distributions P and Q. Since $\|\mathbb{P}_{\phi_G} - \mathbb{P}_{\tilde{f}}\|_{\mathrm{TV}} \leq \sqrt{2[1-A(\phi_G, \tilde{f})^n]}$ where $A(\phi_G, \tilde{f}) = \int_{\mathbb{R}} \sqrt{\phi_G(x)\tilde{f}(x)}\,\mathrm{d}x$ is the Hellinger affinity between the two density functions ϕ_G and $\tilde{f}$,

$$\beta(\mathcal{F}_{1,G}[\rho, M]) := \inf_{T_\alpha} \sup_{f\in\mathcal{F}_{1,G}[\rho,M]} \mathbb{P}_f(T_\alpha = 0) \geq 1-\alpha - \sqrt{2[1-A(\phi_G, \tilde{f})^n]}.$$

If we specify a density $\tilde{f} \in \mathcal{F}_{1,G}[\rho, M]$ such that $A(\phi_G, \tilde{f}) \geq c(\alpha, \beta)^{1/n}$ then $\beta(\mathcal{F}_{1,G}[\rho, M]) \geq 1-\alpha-(1-\alpha-\beta) = \beta$. Moreover, since

$$A(\phi_G, \tilde{f}) \geq 1 - \frac{1}{2}\mathbb{E}_\phi\left[\left(\frac{\tilde{f}(X) - \phi_G(X)}{\phi_G(X)}\right)^2\right],$$

$A(\phi_G, \tilde{f}) \geq c(\alpha, \beta)^{1/n}$ is obtained if $\mathbb{E}_{\phi_G}[(\frac{\tilde{f}(X)-\phi_G(X)}{\phi_G(X)})^2] \leq 2[1-c(\alpha,\beta)^{1/n}]$.

In the sequel, we consider the density $\tilde{f} = (1-\varepsilon)\phi(\cdot-\mu_1) + \varepsilon\phi(\cdot-\mu_2)$, with

$$(1-\varepsilon)\mu_1 = -\varepsilon\mu_2, \tag{6.7}$$

$$\max(\mu_1^2, \mu_2^2, |\mu_1\mu_2|) \leq \nu^2 = \frac{M^2}{4}, \tag{6.8}$$

$$\varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2 = \rho. \tag{6.9}$$

In particular, $\tilde{f} \in \mathcal{F}_{1,G}[\rho, M]$ since $(\mu_2-\mu_1)^2 \leq M^2$.

For this choice,

$$\begin{aligned}
&\mathbb{E}_{\phi_G}\left[\left(\frac{\tilde{f}(X)-\phi_G(X)}{\phi_G(X)}\right)^2\right] \\
&\quad = \int_{\mathbb{R}} \frac{[\tilde{f}(x)-\phi_G(x)]^2}{\phi_G(x)}\,\mathrm{d}x \\
&\quad = \int_{\mathbb{R}} \frac{\{(1-\varepsilon)[\phi_G(x-\mu_1)-\phi_G(x)]+\varepsilon[\phi_G(x-\mu_2)-\phi_G(x)]\}^2}{\phi_G(x)}\,\mathrm{d}x \\
&\quad = (1-\varepsilon)^2\left[\int_{\mathbb{R}} \frac{\phi_G(x-\mu_1)^2}{\phi_G(x)}\,\mathrm{d}x - 1\right] + \varepsilon^2\left[\int_{\mathbb{R}} \frac{\phi_G(x-\mu_2)^2}{\phi_G(x)}\,\mathrm{d}x - 1\right] \\
&\qquad + 2\varepsilon(1-\varepsilon)\left[\int_{\mathbb{R}} \frac{\phi_G(x-\mu_1)\phi_G(x-\mu_2)}{\phi_G(x)}\,\mathrm{d}x - 1\right].
\end{aligned}$$

We have $\int_{\mathbb{R}} \frac{\phi_G(x-\mu_1)\phi_G(x-\mu_2)}{\phi_G(x)}\,\mathrm{d}x = \exp(\mu_1\mu_2)$, for all $\mu_1, \mu_2 \in \mathbb{R}$, hence

$$\mathbb{E}_{\phi_G}\left[\left(\frac{\tilde{f}(X)-\phi_G(X)}{\phi_G(X)}\right)^2\right] = (1-\varepsilon)^2\left[\mathrm{e}^{\mu_1^2}-1\right] + \varepsilon^2\left[\mathrm{e}^{\mu_2^2}-1\right] + 2\varepsilon(1-\varepsilon)\left[\mathrm{e}^{\mu_1\mu_2}-1\right].$$

Next, using that $|\mathrm{e}^u - 1 - u - \frac{1}{2}u^2| \le \frac{\mathrm{e}^{U^2}}{3!}|u|^3$ for all $|u| < U$ with condition (6.8),

$$\begin{aligned}
\mathbb{E}_{\phi_G}\left[\left(\frac{\tilde{f}(X)-\phi_G(X)}{\phi_G(X)}\right)^2\right] &\le (1-\varepsilon)^2\left[\mu_1^2 + \frac{1}{2}\mu_1^4 + \frac{\mathrm{e}^{\nu^2}}{3!}\mu_1^6\right] \\
&\quad + \varepsilon^2\left[\mu_2^2 + \frac{1}{2}\mu_2^4 + \frac{\mathrm{e}^{\nu^2}}{3!}\mu_2^6\right] \\
&\quad + 2\varepsilon(1-\varepsilon)\left[\mu_1\mu_2 + \frac{1}{2}\mu_1^2\mu_2^2 + \frac{\mathrm{e}^{\nu^2}}{3!}|\mu_1\mu_2|^3\right] \\
&\le \left[(1-\varepsilon)\mu_1 + \varepsilon\mu_2\right]^2 + \frac{1}{2}\left[(1-\varepsilon)\mu_1^2 + \varepsilon\mu_2^2\right]^2 \\
&\quad + \frac{\mathrm{e}^{\nu^2}}{3!}\left[(1-\varepsilon)|\mu_1|^3 + \varepsilon|\mu_2|^3\right]^2.
\end{aligned}$$

The parameters of $\tilde{f}$ are constrained such that $(1-\varepsilon)\mu_1 + \varepsilon\mu_2 = 0$ thus

$$\begin{aligned}
&\mathbb{E}_{\phi_G}\left[\left(\frac{\tilde{f}(X)-\phi_G(X)}{\phi_G(X)}\right)^2\right] \\
&\quad \le \frac{1}{2}\left[(1-\varepsilon)\varepsilon(\mu_2-\mu_1)^2\right]^2 + \frac{\mathrm{e}^{\nu^2}}{3!}\left\{(1-\varepsilon)\varepsilon|\mu_2-\mu_1|^3\left[\varepsilon^2 + (1-\varepsilon)^2\right]\right\}^2
\end{aligned}$$

$$\leq (1-\varepsilon)^2\varepsilon^2(\mu_2-\mu_1)^4\left[\frac{1}{2}+\frac{\mathrm{e}^{\nu^2}}{3!}(\mu_2-\mu_1)^2\big[\varepsilon^2+(1-\varepsilon)^2\big]^2\right]$$
$$\leq C^2(M)\big[(1-\varepsilon)\varepsilon(\mu_2-\mu_1)^2\big]^2 = C^2(M)\rho^2$$

with $C^2(M)=\frac{1}{2}+\frac{2}{3}M^2\mathrm{e}^{M^2/4}$. Moreover, if $u<0$, $1-\mathrm{e}^u \geq -u-\frac{1}{2}u^2$ thus $1-c(\alpha,\beta)^{1/n} \geq -\frac{1}{n}\log c(\alpha,\beta)-\frac{1}{2}(\frac{\log c(\alpha,\beta)}{n})^2$. Then, the condition

$$\rho=(1-\varepsilon)\varepsilon(\mu_2-\mu_1)^2 < \frac{1}{C(M)}\sqrt{-\frac{2}{n}\log c(\alpha,\beta)-\left(\frac{\log c(\alpha,\beta)}{n}\right)^2} := \rho^\star$$

implies that $\beta(\mathcal{F}_{1,G}[\rho,M]) > \beta$.

6.4. Proof of Theorem 3.2

Let $f(\cdot)=(1-\varepsilon)\phi_G(\cdot-\mu_1)+\varepsilon\phi_G(\cdot-\mu_2)\in\mathcal{F}_{1,G}[\rho,M]$ where ρ satisfies (3.3). We will prove that $f\in\bar{\mathcal{F}}_1[n,\alpha,\beta]$ and the result will be a consequence of Theorem 6.1. In the following, we consider $k\in\mathcal{K}_n$ such that

$$\frac{0.99}{2}\bar{\Phi}_G(M)\leq\frac{k}{n}\leq 0.99\bar{\Phi}_G(M).$$

Note that this is possible since, under the assumptions of Theorem 3.2, $0.99\bar{\Phi}_G(M)n\geq 1$. Note that $|\mathcal{K}_n|\leq\log_2(n/2)$, hence $\alpha_n\geq\alpha/|\mathcal{K}_n|\geq\alpha/\log_2(n/2)$. We will show that $(\varepsilon,\mu_1,\mu_2)\in\bar{\mathcal{S}}(\alpha_n,\rho(k,n),k)$: Considering $c=t_{\alpha_n,k}/2$ and denoting $\tau=\mu_2-\mu_1$, we want to prove that

$$(1-\varepsilon)\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}+\varepsilon\tau\right)+\varepsilon\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}-(1-\varepsilon)\tau\right)>\rho(k,n), \tag{6.10}$$
$$(1-\varepsilon)\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}-\varepsilon\tau\right)+\varepsilon\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}+(1-\varepsilon)\tau\right)>\rho(k,n) \tag{6.11}$$

hold, with $\rho(k,n)=\frac{k}{n}+\frac{1}{n\beta}+\frac{\sqrt{1+2k\beta}}{n\beta}$.

We use a Taylor expansion at the order 2, the terms of order 1 vanish and this leads to:

$$(1-\varepsilon)\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}+\varepsilon\tau\right)+\varepsilon\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}-(1-\varepsilon)\tau\right)$$
$$=\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right)+\frac{1}{2}(1-\varepsilon)\varepsilon\tau^2\big[\varepsilon\big(-\phi'_G(a)\big)+(1-\varepsilon)\big(-\phi'_G(b)\big)\big],$$

where a (resp. b) belongs to the interval $]\frac{t_{\alpha_n,k}}{2},\frac{t_{\alpha_n,k}}{2}+\varepsilon\tau[$ (resp. $]\frac{t_{\alpha_n,k}}{2}-(1-\varepsilon)\tau,\frac{t_{\alpha_n,k}}{2}[$).

We recall that $\bar{\Phi}_G(\frac{t_{\alpha_n,k}}{2}) = \frac{k}{n}[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}]$. Hence, in order to prove that (6.10) holds, we just have to show that

$$(1-\varepsilon)\varepsilon\tau^2\big\{\varepsilon\big[-\phi_G'(a)\big] + (1-\varepsilon)\big[-\phi_G'(b)\big]\big\} \geq \frac{2}{n\beta} + \frac{\sqrt{k}}{n}\sqrt{2\log(4/\alpha_n)}. \tag{6.12}$$

Next, we want to prove that $[\frac{t_{\alpha_n,k}}{2} - (1-\varepsilon)\tau, \frac{t_{\alpha_n,k}}{2} + \varepsilon\tau]$ remains included in a fixed interval $[c_1(M), c_2(M)]$ with $c_1(M) > 0$.

On one hand, we have

$$\frac{t_{\alpha_n,k}}{2} \geq \bar{\Phi}_G^{-1}\left(\frac{k}{n}\right) \geq \bar{\Phi}_G^{-1}\big(0.99\bar{\Phi}_G(M)\big)$$

and

$$\frac{t_{\alpha_n,k}}{2} - M \geq \bar{\Phi}_G^{-1}\big(0.99\bar{\Phi}_G(M)\big) - M := c_1(M) > 0.$$

Moreover,

$$\begin{aligned}\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right) &\geq \frac{0.99}{2}\bar{\Phi}_G(M) - \sqrt{\frac{2\log(4/\alpha_n)}{\sqrt{n}}}\sqrt{0.99\bar{\Phi}_G(M)} \\ &\geq \frac{\bar{\Phi}_G(M)}{200}\end{aligned}$$

since $(8.25)\log(4\log_2(n/2)/\alpha)/n \leq \bar{\Phi}_G(M)$. This implies that

$$\frac{t_{\alpha_n,k}}{2} + \tau \leq \bar{\Phi}_G^{-1}\left(\frac{\bar{\Phi}_G(M)}{200}\right) + M := c_2(M).$$

Finally, the function $-\phi_G'$ is bounded from below on this interval by some positive constant $C(M) = \min_{x\in[c_1(M),c_2(M)]}(-\phi_G'(x))$. This implies that (6.12) is satisfied if $\varepsilon(1-\varepsilon)\tau^2 \geq C(\alpha,\beta,M)\sqrt{\log\log(n)}/\sqrt{n}$ for some suitable constant $C(\alpha,\beta,M)$. This concludes the proof of (6.10). The proof of (6.11) follows the same arguments.

Remark. If we choose $k^* \in \mathcal{K}_n$ such that

$$\frac{0.99}{2}\bar{\Phi}_G(M) \leq \frac{k^*}{n} \leq 0.99\bar{\Phi}_G(M)$$

and consider the test statistics

$$\mathbb{1}_{X_{(n-k^*+1)} - X_{(k^*)} > q_{\alpha,k^*}}$$

then it is easy to prove that (6.10) and (6.11) are satisfied for $k = k^*$ if $\varepsilon(1-\varepsilon)\tau^2 \geq C'(\alpha,\beta,M)/\sqrt{n}$ for some suitable constant $C'(\alpha,\beta,M)$ since in this case α_n is replaced by α and we do no more have the logarithmic loss in the rate of convergence.

6.5. Proof of Proposition 3.1

Following the definition of the threshold $v_{\alpha,n}$, it is easy to see that ψ_α defined in (3.4) is a level-α test. Now, our aim is to upper bound the term

$$\mathbb{P}_f(\psi_\alpha = 0) = \mathbb{P}_f\big(S_n^2 \leq v_{\alpha,n}\big)$$

when $f \in \mathcal{F}_1[\rho, M]$ where, as previously,

$$\mathcal{F}_1[\rho, M] = \big\{f \in \mathcal{F}_1[M], \varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2 \geq \rho\big\}.$$

In a first time, a control of $v_{\alpha,n}$ is required. If a real number $c_{\alpha,n}$ is determined such that $\mathbb{P}_{H_0}(S_n^2 > c_{\alpha,n}) \leq \alpha$, then $v_{\alpha,n} \leq c_{\alpha,n}$. According to [20], page 200, if $Y_1, \dots, Y_n$ are i.i.d. random variables such that $\mathbb{E}[(Y_1 - \mathbb{E}[Y_1])^4] < +\infty$, then

$$\mathrm{Var}\left(\frac{1}{n-1}\sum_{i=1}^n (Y_i - \bar{Y}_n)^2\right) \leq \frac{1}{n}\left\{\mathbb{E}\big[\big(Y_1 - \mathbb{E}[Y_1]\big)^4\big] - \frac{n-3}{n-1}\mathrm{Var}(Y_1)^2\right\}. \tag{6.13}$$

Hence, since $\mathbb{E}_\phi[X_1^4] < B$ and $\mathbb{E}_\phi[S_n^2] = \sigma^2$,

$$\mathbb{P}_{H_0}\big(S_n^2 > c_{\alpha,n}\big) = \mathbb{P}_{H_0}\big(S_n^2 - \sigma^2 > c_{\alpha,n} - \sigma^2\big) \leq \frac{\mathrm{Var}_\phi(S_n^2)}{(c_{\alpha,n} - \sigma^2)^2} \leq \frac{B}{n(c_{\alpha,n} - \sigma^2)^2}.$$

In particular $\mathbb{P}_{H_0}(S_n^2 > c_{\alpha,n}) \leq \alpha$ with $c_{\alpha,n} = \sigma^2 + \sqrt{\frac{B}{n\alpha}}$, and thus

$$v_{\alpha,n} \leq \sigma^2 + \sqrt{\frac{B}{n\alpha}}.$$

Note that $\mathbb{E}_f[S_n^2] = \mathrm{Var}_f(X_1) = \sigma^2 + \varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2$. Hence, for all $f \in \mathcal{F}_1[\rho, M]$,

$$\begin{aligned}
\mathbb{P}_f(\psi_\alpha = 0) &\leq \mathbb{P}_f\left(S_n^2 \leq \sigma^2 + \sqrt{\frac{B}{n\alpha}}\right) \\
&= \mathbb{P}_f\left(S_n^2 - \mathbb{E}_f\big[S_n^2\big] \leq \sigma^2 + \sqrt{\frac{B}{n\alpha}} - \mathbb{E}_f\big[S_n^2\big]\right) \\
&\leq \mathbb{P}_f\left(\big|S_n^2 - \mathbb{E}_f\big[S_n^2\big]\big| \geq \varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2 - \sqrt{\frac{B}{n\alpha}}\right) \\
&\leq \frac{\mathrm{Var}_f(S_n^2)}{[\varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2 - \sqrt{B/(n\alpha)}]^2}
\end{aligned}$$

if $\varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2 > \sqrt{\frac{B}{n\alpha}}$. Using equation (6.13), we get

$$\mathbb{P}_f(\psi_\alpha = 0) \leq \frac{\mathbb{E}_f[(X_1 - \mathbb{E}_f[X_1])^4]}{n[\varepsilon(1-\varepsilon)(\mu_2 - \mu_1)^2 - \sqrt{B/(n\alpha)}]^2}.$$

In order to conclude, just remark that

$$\begin{aligned}
\mathbb{E}_f\big[\big(X_1-\mathbb{E}[X_1]\big)^4\big] &= (1-\varepsilon)\int_{\mathbb{R}}\big[x-(1-\varepsilon)\mu_1-\varepsilon\mu_2\big]^4\phi(x-\mu_1)\,\mathrm{d}x \\
&\quad+\varepsilon\int_{\mathbb{R}}\big[x-(1-\varepsilon)\mu_1-\varepsilon\mu_2\big]^4\phi(x-\mu_2)\,\mathrm{d}x \\
&= (1-\varepsilon)\int_{\mathbb{R}}\big[y-\varepsilon(\mu_2-\mu_1)\big]^4\phi(y)\,\mathrm{d}y \\
&\quad+\varepsilon\int_{\mathbb{R}}\big[y+(1-\varepsilon)(\mu_2-\mu_1)\big]^4\phi(y)\,\mathrm{d}y \\
&= \mathbb{E}_\phi\big[Z^4\big]+6\varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2\mathbb{E}_\phi\big[Z^2\big] \\
&\quad+\big[\varepsilon(1-\varepsilon)^4+\varepsilon^4(1-\varepsilon)\big](\mu_2-\mu_1)^4 \\
&\le B+\frac{6}{4}\sqrt{B}M^2+M^4\le\big(M^2+\sqrt{B}\big)^2.
\end{aligned}$$

Thus

$$\mathbb{P}_f(\psi_\alpha=0)\le\frac{(M^2+\sqrt{B})^2}{n[\varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2-\sqrt{B/(n\alpha)}]^2}\le\beta$$

as soon as

$$\varepsilon(1-\varepsilon)(\mu_2-\mu_1)^2\ge\frac{C(\alpha,\beta,M,B)}{\sqrt{n}},$$

for some positive constant $C(\alpha,\beta,M,B)$. This concludes the proof of Proposition 3.1.

6.6. Proof of Theorem 4.1

We will prove that, under the assumptions of Theorem 4.1, $f\in\bar{\mathcal{F}}_1[n,\alpha,\beta]$ and the result will be a consequence of Theorem 6.1. We recall that $|\mathcal{K}_n|\le\log_2(n)$, hence $\alpha\ge\alpha_n\ge\alpha/|\mathcal{K}_n|\ge\alpha/\log_2(n)$. We set $\tau=\mu_2-\mu_1$ and we have to prove that there exists $k\in\mathcal{K}_n$ and $c\in\mathbb{R}$ such that

$$(1-\varepsilon)\bar{\Phi}_G(t_{\alpha_n,k}-c+\varepsilon\tau)+\varepsilon\bar{\Phi}_G\big(t_{\alpha_n,k}-c-(1-\varepsilon)\tau\big)>\rho(k,n), \tag{6.14}$$

$$(1-\varepsilon)\bar{\Phi}_G(c-\varepsilon\tau)+\varepsilon\bar{\Phi}_G\big(c+(1-\varepsilon)\tau\big)>\rho(k,n), \tag{6.15}$$

with $\rho(k,n)=\frac{k}{n}+\frac{1}{n\beta}+\frac{\sqrt{1+2k\beta}}{n\beta}$. Note that $\rho(k,n)\le\frac{k}{n}+C_\beta\frac{\sqrt{k}}{n}$ with $C_\beta=\frac{2}{\beta}+\sqrt{\frac{2}{\beta}}$. We recall that $t_{\alpha_n,k}$ is defined by

$$\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right)=\frac{k}{n}\left[1-\sqrt{\frac{2\log(4/\alpha_n)}{k}}\right].$$

In the following, we set $C_{\alpha_n} = \sqrt{2\log(4/\alpha_n)}$. Since $\alpha_n \geq \alpha/\log_2(n)$, note that $0 < C_{\alpha_n} \leq C(\alpha)\sqrt{\log\log(n)}$ for some constant $C(\alpha)$ depending only on α. We choose $k \in \mathcal{K}_n$ such that

$$\lim_{n\to+\infty} \frac{k}{\log(n)\log\log(n)} = +\infty \quad \text{and} \quad \lim_{n\to+\infty} \frac{n}{k} = +\infty \tag{6.16}$$

and we define

$$c = \frac{t_{\alpha_n,k}}{2} - \sqrt{\frac{2}{k}} C_{\alpha_n}. \tag{6.17}$$

For the sake of simplicity, we omit the dependency with respect to n in the notation of k and c. Let us first show that (6.15) holds for n large enough. First, note that

$$(1-\varepsilon)\bar{\Phi}_G(c-\varepsilon\tau) + \varepsilon\bar{\Phi}_G\big(c+(1-\varepsilon)\tau\big) > (1-\varepsilon)\bar{\Phi}_G(c).$$

With the assumptions on k, we have that $c > 0$ for n large enough since $t_{\alpha_n,k} \to +\infty$ and $C_{\alpha_n}/\sqrt{k} \to 0$ as $n \to +\infty$. Hence

$$\bar{\Phi}_G(c) \geq \bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right) + \sqrt{\frac{2}{k}} C_{\alpha_n} \phi_G\left(\frac{t_{\alpha_n,k}}{2}\right).$$

Moreover, for all $u > 0$,

$$\bar{\Phi}_G(u) \leq \frac{1}{2}\exp\big(-u^2/2\big) = \sqrt{\frac{\pi}{2}}\phi_G(u),$$

hence

$$\phi_G\left(\frac{t_{\alpha_n,k}}{2}\right) \geq \sqrt{\frac{2}{\pi}}\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right).$$

This leads to

$$(1-\varepsilon)\bar{\Phi}_G(c) > (1-\varepsilon)\left(1 + \frac{2C_{\alpha_n}}{\sqrt{\pi k}}\right)\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right).$$

After some obvious computations, condition (6.15) is satisfied as soon as

$$(1-\varepsilon)C_{\alpha_n}\left(\frac{2}{\sqrt{\pi}} - 1\right)\frac{\sqrt{k}}{n} > \varepsilon\frac{k}{n} + C_\beta\frac{\sqrt{k}}{n} + \frac{2C_{\alpha_n}^2}{\sqrt{\pi}n}.$$

Since $\varepsilon < 1/\sqrt{n}$ and $k \leq n$, we have $\varepsilon k < \sqrt{k}$. We recall that $C_{\alpha_n} \to +\infty$ as $n \to +\infty$ and with the assumptions on k, we have that $\sqrt{k}/C_{\alpha_n} \to +\infty$ as $n \to +\infty$, and the above inequality holds for n large enough.

It remains to prove that (6.14) is satisfied with the conditions on k imposed by (6.16) and the value of c defined by (6.17). Let Δ satisfy $0 < r < \Delta \leq 1$, we choose $k \in \mathcal{K}_n$ satisfying (6.16) and such that $n^{1-\Delta} \leq k \leq 2n^{1-\Delta}\log^2(n)$. Note that such values of k exist for n large enough. It follows from Lemma A.2 that $t_{\alpha_n,k}/2 \leq \sqrt{2\Delta\log(n)}$. First,

$$\begin{aligned}
\bar{\Phi}_G(t_{\alpha_n,k} - c + \varepsilon\tau) &= \bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2} + \sqrt{\frac{2}{k}}C_{\alpha_n} + \varepsilon\tau\right) \\
&\geq \bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right) - \left(\sqrt{\frac{2}{k}}C_{\alpha_n} + \varepsilon\tau\right)\phi_G\left(\frac{t_{\alpha_n,k}}{2}\right) \\
&\geq \frac{k}{n}\left[1 - \frac{C_{\alpha_n}}{\sqrt{k}}\right] - \left(\sqrt{\frac{2}{k}}C_{\alpha_n} + \varepsilon\tau\right)\phi_G\left(\frac{t_{\alpha_n,k}}{2}\right).
\end{aligned}$$

We have to give an upper bound for $\phi_G(\frac{t_{\alpha_n,k}}{2})$. We use the inequality

$$\forall u > 0, \qquad \bar{\Phi}_G(u) \geq \left(\frac{1}{u} - \frac{1}{u^3}\right)\phi_G(u),$$

this leads to

$$\forall u > 0, \qquad \phi_G(u) \leq \frac{u^3}{u^2 - 1}\bar{\Phi}_G(u) \leq u^3\bar{\Phi}_G(u),$$

provided that $u^2 - 1 \geq 1$. This is the case, for n large enough for $u = t_{\alpha_n,k}/2$, hence we have

$$\begin{aligned}
\phi_G\left(\frac{t_{\alpha_n,k}}{2}\right) &\leq \left[\frac{t_{\alpha_n,k}}{2}\right]^3 \bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right) \\
&\leq \left[\sqrt{2\Delta\log(n)}\right]^3 \frac{k}{n} \\
&\leq 4\sqrt{2}\left[\log(n)\right]^{7/2} n^{-\Delta}.
\end{aligned}$$

Finally, we obtain that

$$\bar{\Phi}_G(t_{\alpha_n,k} - c + \varepsilon\tau) \geq \frac{k}{n} - C_{\alpha_n}\frac{\sqrt{k}}{n} - \left(\frac{\sqrt{2}C_{\alpha_n}}{\sqrt{k}} + \varepsilon\tau\right)4\sqrt{2}\left[\log(n)\right]^{7/2} n^{-\Delta}.$$

Second, we want to lower bound $\bar{\Phi}_G(t_{\alpha_n,k} - c - (1-\varepsilon)\tau)$. We have that

$$\begin{aligned}
\bar{\Phi}_G\left(t_{\alpha_n,k} - c - (1-\varepsilon)\tau\right) &= \bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2} + \sqrt{\frac{2}{k}}C_{\alpha_n} - (1-\varepsilon)\tau\right) \\
&\geq \bar{\Phi}_G\left(\sqrt{2\Delta\log(n)} - \tau + \sqrt{\frac{2}{k}}C_{\alpha_n} + \varepsilon\tau\right) \\
&\geq \bar{\Phi}_G\left(\sqrt{2\Delta\log(n)} - \sqrt{2r\log(n)}\right) \\
&\quad - \left(\varepsilon\tau + \frac{\sqrt{2}C_{\alpha_n}}{\sqrt{k}}\right)\phi_G\left(\sqrt{2\Delta\log(n)} - \sqrt{2r\log(n)}\right)
\end{aligned}$$

since $\tau = \sqrt{2r\log(n)}$. Moreover, since $\phi_G(\sqrt{2\Delta\log(n)} - \sqrt{2r\log(n)}) = (\sqrt{2\pi})^{-1}n^{-(\sqrt{\Delta}-\sqrt{r})^2}$, and using again the inequality $\bar{\Phi}_G(u) \geq (\frac{1}{u} - \frac{1}{u^3})\phi_G(u)$ which holds for all $u > 0$, we obtain that

$$\bar{\Phi}_G\big(t_{\alpha_n,k} - c - (1-\varepsilon)\tau\big) \geq Cn^{-(\sqrt{\Delta}-\sqrt{r})^2}\left(\frac{1}{\sqrt{\log(n)}} - \varepsilon\tau - \frac{\sqrt{2}C_{\alpha_n}}{\sqrt{k}}\right),$$

for some positive constant C depending on Δ and r. Condition (6.14) is thus fulfilled if

$$C\varepsilon n^{-(\sqrt{\Delta}-\sqrt{r})^2}\left(\frac{1}{\sqrt{\log(n)}} - \varepsilon\tau - \sqrt{\frac{2}{k}}C_{\alpha_n}\right)$$
$$> \varepsilon\frac{k}{n} + (C_{\alpha_n} + C_\beta)\frac{\sqrt{k}}{n} + \left(\sqrt{\frac{2}{k}}C_{\alpha_n} + \varepsilon\tau\right)4\sqrt{2}\big[\log(n)\big]^{7/2}n^{-\Delta}.$$

By (6.16), $C_{\alpha_n}/\sqrt{k} = o(1/\sqrt{\log(n)})$, and the left-hand side of this inequality is equivalent as $n \to +\infty$ to $C\varepsilon n^{-(\sqrt{\Delta}-\sqrt{r})^2}/\sqrt{\log(n)}$ and the right-hand side is equivalent as $n \to +\infty$ to $8C_{\alpha_n}(\log(n))^{7/2}n^{-\Delta}/\sqrt{k}$. Hence, the condition (6.14) will be satisfied asymptotically if for some $\Delta \in]0,1]$,

$$\delta + (\sqrt{\Delta} - \sqrt{r})^2 < \frac{1+\Delta}{2}.$$

- If $\frac{1}{2} < \delta \leq \frac{3}{4}$ and $0 < r \leq \frac{1}{4}$, we set $\Delta = 4r$ and the above condition becomes $r > \delta - \frac{1}{2}$.
- If $\frac{1}{2} < \delta \leq \frac{3}{4}$ and $r > \frac{1}{4}$, the above condition is satisfied with $\Delta = 1$ and no additional condition is required.
- If $\delta > \frac{3}{4}$, we set $\Delta = 1$ and the above condition becomes $r > (1 - \sqrt{1-\delta})^2$.

This concludes the proof of Theorem 4.1.

6.7. Proof of Theorem 4.2

We first provide an upper bound for the quantile $q_{\alpha_n,k}$ for all $k \in \{1,\dots,n/2\}$. We have seen in the proof of Theorem 6.1 that

$$q_{\alpha_n,k} \leq t_{\alpha_n,k},$$

where

$$\bar{\Phi}_L\left(\frac{t_{\alpha_n,k}}{2}\right) = \frac{k}{n}\left(1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right). \tag{6.18}$$

This leads to

$$\frac{1}{2}e^{-t_{\alpha_n,k}/2} = \frac{k}{n}\left(1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right).$$

Hence,

$$\frac{t_{\alpha_n,k}}{2} = \log\left(\frac{n}{k}\right) - \log\left(1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right) - \log(2). \tag{6.19}$$

Then, applying Theorem 6.1 with $c = t_{\alpha_n,k}/2$, we get that if, for some $k \in \mathcal{K}_n$,

$$\begin{aligned}(1-\varepsilon)\bar{\Phi}_L\left(\frac{t_{\alpha_n,k}}{2} + \varepsilon(\mu_2-\mu_1)\right) + \varepsilon\bar{\Phi}_L\left(\frac{t_{\alpha_n,k}}{2} - (1-\varepsilon)(\mu_2-\mu_1)\right) \\ > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta}\end{aligned} \tag{6.20}$$

and

$$\begin{aligned}(1-\varepsilon)\bar{\Phi}_L\left(\frac{t_{\alpha_n,k}}{2} - \varepsilon(\mu_2-\mu_1)\right) + \varepsilon\bar{\Phi}_L\left(\frac{t_{\alpha_n,k}}{2} + (1-\varepsilon)(\mu_2-\mu_1)\right) \\ > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta},\end{aligned}$$

then our test is powerful. For the sake of convenience, we will concentrate our attention to the first inequality, the control of the second one following essentially the same lines.

From now on, we will only deal with possible values of k satisfying

$$\frac{t_{\alpha_n,k}}{2} > \mu_2 - \mu_1. \tag{6.21}$$

Using the properties of the Laplace distribution and the equation (6.21), the condition (6.20) becomes

$$\begin{aligned}&(1-\varepsilon)\times\frac{1}{2}e^{-(t_{\alpha_n,k}/2)-\varepsilon(\mu_2-\mu_1)} + \varepsilon\times\frac{1}{2}e^{-(t_{\alpha_n,k}/2)+(1-\varepsilon)(\mu_2-\mu_1)} > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta} \\
\Leftrightarrow\quad &\varepsilon\times\frac{1}{2}e^{-t_{\alpha_n,k}/2+(1-\varepsilon)(\mu_2-\mu_1)} > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta} \\
&\qquad - (1-\varepsilon)\times\frac{1}{2}e^{-(t_{\alpha_n,k}/2)-\varepsilon(\mu_2-\mu_1)} \\
\Leftrightarrow\quad &\varepsilon\times\frac{1}{2}e^{-(t_{\alpha_n,k}/2)+(1-\varepsilon)(\mu_2-\mu_1)} > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta} \\
&\qquad - (1-\varepsilon)\phi_L\left(-\frac{t_{\alpha_n,k}}{2}\right)\times e^{-\varepsilon(\mu_2-\mu_1)}.\end{aligned}$$

Since $\phi_L(x) = \bar{\Phi}_L(x)$ for all $x \geq 0$ and thanks to (6.18), we get that

$$(1-\varepsilon) \times \frac{1}{2}\mathrm{e}^{-(t_{\alpha_n,k}/2)-\varepsilon(\mu_2-\mu_1)} + \varepsilon \times \frac{1}{2}\mathrm{e}^{-(t_{\alpha_n,k}/2)+(1-\varepsilon)(\mu_2-\mu_1)} > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta}$$

$$\Leftrightarrow \quad \varepsilon \times \frac{1}{2}\mathrm{e}^{-(t_{\alpha_n,k}/2)+(1-\varepsilon)(\mu_2-\mu_1)} > \frac{k}{n} + \frac{1+\sqrt{1+2k\beta}}{n\beta} - (1-\varepsilon)\frac{k}{n}\left[1-\sqrt{\frac{2\log(4/\alpha_n)}{k}}\right] \times \bigl(1-\varepsilon(\mu_2-\mu_1)+V_n\bigr),$$

where $V_n \leq C\varepsilon^2(\mu_2-\mu_1)^2$ for some $C > 0$. As in the proof of Theorem 6.1, we will deal with values of k having the parametrization $k/n = n^{-\Delta}$ for some $\Delta \in]0,1[$. In particular,

$$\sqrt{k} = n^{(1-\Delta)/2} \quad \text{and} \quad \frac{\sqrt{k}}{n} = n^{-(1+\Delta)/2}.$$

A short investigation of the asymptotics of the term in the right-hand side of the previous inequality indicates that the dominating term is of order $\sqrt{k}/n$. Indeed, thanks to the parametrization of k, ε an $\mu_2 - \mu_1$, we get that

$$\varepsilon\frac{k}{n}(\mu_2-\mu_1) = \mathrm{o}\left(\frac{\sqrt{k}}{n}\right) \quad \text{and} \quad \frac{1}{n} = \mathrm{o}\left(\frac{\sqrt{k}}{n}\right) \quad \text{as } n \to +\infty.$$

Hence, in order to guarantee that our test is powerful, we have to ensure that

$$\begin{aligned} &\varepsilon \times \frac{1}{2}\mathrm{e}^{-t_{\alpha_n,k}/2+(1-\varepsilon)(\mu_2-\mu_1)} > C(\alpha,\beta)\frac{\sqrt{k}}{n} \\ \Leftrightarrow \quad &\varepsilon \times \frac{1}{2}\mathrm{e}^{-t_{\alpha_n,k}/2+(\mu_2-\mu_1)}\bigl(1-\mathrm{o}(1)\bigr) > C(\alpha,\beta)\frac{\sqrt{k}}{n}, \end{aligned} \tag{6.22}$$

for some positive constant $C(\alpha,\beta)$, as $n \to +\infty$. Thanks to (6.19), the inequality (6.22) becomes

$$\begin{aligned} \frac{1}{n^\delta} \times \frac{1}{n^\Delta} \times n^r > n^{-(1+\Delta)/2} \quad &\Leftrightarrow \quad \delta + \Delta - r < \frac{1+\Delta}{2} \\ &\Leftrightarrow \quad r > \delta + \frac{\Delta}{2} - \frac{1}{2}. \end{aligned}$$

In practice, the smallest possible parameter Δ will provide the less restrictive separation condition. In the same time, we have to ensure that the condition (6.21) is satisfied. It follows from (6.19) that $t_{\alpha_n,k}/2 \sim \Delta\log(n)$ as $n \to \infty$, and (6.21) holds for n large enough as soon as $\Delta > r$. Hence, choosing $\Delta = r + r_0$ for some positive r_0, we can remark that

$$r > \delta + \frac{\Delta}{2} - \frac{1}{2} \quad \Leftrightarrow \quad r > 2(\delta - 1/2) + r_0,$$

which is satisfied as soon as

$$r > 2(\delta - 1/2),$$

provided r_0 is small enough. This concludes the proof.

Appendix: Lemmas for the upper-bound

Lemma A.1. *Let $Y_1, \dots, Y_n$ be n random variables with a cumulative distribution function F and the order statistics are denoted $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$. Let $\alpha \in]0,1[$ and let $k \in \{1, \dots, n\}$ such that $k > 2\log(\frac{2}{\alpha})$. Let c and d be two real numbers such that*

$$F(d) \vee \big(1 - F(c)\big) \le \frac{k}{n}\left[1 - \sqrt{\frac{2\log(2/\alpha)}{k}}\right]. \tag{A.1}$$

Then $\mathbb{P}(Y_{(n-k+1)} \ge c) \le \alpha$ and $\mathbb{P}(Y_{(k)} \le d) \le \alpha$.

Proof.

$$\begin{aligned}\mathbb{P}(Y_{(n-k+1)} \ge c) &= \mathbb{P}\left(\sum_{i=1}^{n} \mathbb{1}_{\{Y_i \ge c\}} \ge k\right)\\ &= \mathbb{P}\left(\sum_{i=1}^{n} \big\{\mathbb{1}_{\{Y_i \ge c\}} - \big[1 - F(c)\big]\big\} \ge k - n\big[1 - F(c)\big]\right).\end{aligned}$$

According to condition (A.1),

$$k - n\big[1 - F(c)\big] \ge k\sqrt{\frac{2\log(2/\alpha)}{k}} > 0.$$

Using a Bernstein's inequality, we get

$$\mathbb{P}(Y_{(n-k+1)} \ge c) \le 2\exp\left[-\frac{1}{2}\frac{(k - n[1 - F(c)])^2}{v + (1/3)(k - n[1 - F(c)])}\right]$$

with $v = \sum_{i=1}^{n} \mathbb{E}[(\mathbb{1}_{\{Y_i \ge c\}} - [1 - F(c)])^2] = \sum_{i=1}^{n} \operatorname{Var}(\mathbb{1}_{Y_i \ge c}) = nF(c)[1 - F(c)] \le n[1 - F(c)]$. Thus, $3v + k - n[1 - F(c)] \le 2n[1 - F(c)] + k \le 3k - 2k\sqrt{\frac{2\log(2/\alpha)}{k}} \le 3k$. This implies that

$$\mathbb{P}(Y_{(n-k+1)} \ge c) \le 2\exp\left[-\frac{3}{2}\frac{(k - n[1 - F(c)])^2}{3k}\right] \le 2\exp\left[-\log\left(\frac{2}{\alpha}\right)\right] = \alpha.$$

In the same way,

$$\begin{aligned}\mathbb{P}(Y_{(k)} \le d) &= \mathbb{P}\left(\sum_{i=1}^{n} \mathbb{1}_{\{Y_i \ge d\}} \le n - k\right)\\ &= \mathbb{P}\left(\sum_{i=1}^{n} \big\{\mathbb{1}_{\{Y_i \ge d\}} - \big[1 - F(d)\big]\big\} \le nF(d) - k\right).\end{aligned}$$

Since $nF(d) - k < 0$ according to condition (A.1), a Bernstein's inequality implies that

$$\mathbb{P}(Y_{(k)} \le d) \le \mathbb{P}\left(\left|\sum_{i=1}^{n}\left\{\mathbb{1}_{\{Y_i \ge d\}} - \left[1 - F(d)\right]\right\}\right| \ge k - nF(d)\right)$$
$$\le 2\exp\left[-\frac{1}{2}\frac{[nF(d)-k]^2}{v + (1/3)[k - nF(d)]}\right]$$

with $v = \sum_{i=1}^{n} \mathbb{E}[(\mathbb{1}_{\{Y_i \ge d\}} - [1 - F(d)])^2] = \sum_{i=1}^{n} \mathrm{Var}(Y_i \ge d) = nF(d)[1 - F(d)] \le nF(d)$. Thus, $3v + k - nF(d) \le 2nF(d) + k \le 3k - 2k\sqrt{\frac{2\log(2/\alpha)}{k}} \le 3k$. This implies that

$$\mathbb{P}(Y_{(k)} \le d) \le 2\exp\left[-\frac{3}{2}\frac{[nF(d)-k]^2}{3k}\right] \le 2\exp\left[-\log\left(\frac{2}{\alpha}\right)\right] = \alpha.$$

□

Lemma A.2. *If* $k \ge 8\log(4/\alpha_n)$ *and* $\frac{k}{n} \ge n^{-\Delta}$ *with* $\Delta \in]0, 1[$, *then*

$$t_{\alpha_n,k} \le 2\sqrt{2\Delta\log(n)}.$$

Proof.

$$\bar{\Phi}_G\left(\frac{t_{\alpha_n,k}}{2}\right) = \frac{k}{n}\left[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right]$$
$$\le \frac{1}{2}\exp\left[-\frac{1}{2}\left(\frac{t_{\alpha_n,k}}{2}\right)^2\right],$$

thus

$$\exp\left[\frac{1}{2}\left(\frac{t_{\alpha_n,k}}{2}\right)^2\right] \le \frac{1}{2}\left[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right]^{-1} n^{\Delta}.$$

If $k \ge 8\log(4/\alpha_n)$, then

$$2\left[1 - \sqrt{\frac{2\log(4/\alpha_n)}{k}}\right] \ge 1$$

which leads to $t_{\alpha_n,k} \le 2\sqrt{2\Delta\log(n)}$. □

Acknowledgements

The authors would like to thank the associate editor and the two referees for their constructive remarks that have helped to improve the paper. They also acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant MixStatSeq (ANR-13-JS01-0001-01).

References

[1] Addario-Berry, L., Broutin, N., Devroye, L. and Lugosi, G. (2010). On combinatorial testing problems. *Ann. Statist.* **38** 3063–3092. MR2722464
[2] Azaïs, J.-M., Gassiat, É. and Mercadier, C. (2009). The likelihood ratio test for general mixture models with or without structural parameter. *ESAIM Probab. Stat.* **13** 301–327. MR2528086
[3] Baraud, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** 577–606. MR1935648
[4] Cai, T.T., Jeng, X.J. and Jin, J. (2011). Optimal detection of heterogeneous and heteroscedastic mixtures. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** 629–662. MR2867452
[5] Cai, T.T., Jin, J. and Low, M.G. (2007). Estimation and confidence sets for sparse normal mixtures. *Ann. Statist.* **35** 2421–2449. MR2382653
[6] Cai, T.T. and Wu, Y. (2014). Optimal detection of sparse mixtures against a given null distribution. *IEEE Trans. Inform. Theory* **60** 2217–2232. MR3181520
[7] Charnigo, R. and Sun, J. (2004). Testing homogeneity in a mixture distribution via the L^2 distance between competing models. *J. Amer. Statist. Assoc.* **99** 488–498. MR2062834
[8] Chen, H., Chen, J. and Kalbfleisch, J.D. (2001). A modified likelihood ratio test for homogeneity in finite mixture models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63** 19–29. MR1811988
[9] Chen, J. and Li, P. (2009). Hypothesis test for normal mixture models: The EM approach. *Ann. Statist.* **37** 2523–2542. MR2543701
[10] Chernoff, H. and Lander, E. (1995). Asymptotic distribution of the likelihood ratio test that a mixture of two binomials is a single binomial. *J. Statist. Plann. Inference* **43** 19–40. MR1314126
[11] Dacunha-Castelle, D. and Gassiat, E. (1999). Testing the order of a model using locally conic parametrization: Population mixtures and stationary ARMA processes. *Ann. Statist.* **27** 1178–1209. MR1740115
[12] Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* **32** 962–994. MR2065195
[13] Fromont, M. and Laurent, B. (2006). Adaptive goodness-of-fit tests in a density model. *Ann. Statist.* **34** 680–720. MR2281881
[14] Garel, B. (2007). Recent asymptotic results in testing for mixtures. *Comput. Statist. Data Anal.* **51** 5295–5304. MR2370872
[15] Ingster, Y. (1999). Minimax detection of a signal for l^n-balls. *Math. Methods Statist.* **7** 401–428.
[16] Jager, L. and Wellner, J.A. (2007). Goodness-of-fit tests via phi-divergences. *Ann. Statist.* **35** 2018–2053. MR2363962
[17] Klar, B. and Meintanis, S.G. (2005). Tests for normal mixtures based on the empirical characteristic function. *Comput. Statist. Data Anal.* **49** 227–242. MR2129175
[18] McLachlan, G. and Peel, D. (2000). *Finite Mixture Models. Wiley Series in Probability and Statistics: Applied Probability and Statistics*. New York: Wiley. MR1789474
[19] Spokoiny, V.G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist.* **24** 2477–2498. MR1425962
[20] Wilks, S.S. (1962). *Mathematical Statistics. A Wiley Publication in Mathematical Statistics*. New York: Wiley. MR0144404

Received April 2013 and revised February 2014