Adaptive test on components of densities mixture

F. Autin[1] and C. Pouet[2]

[1]Aix-Marseille Université, Laboratoire d'Analyse, Topologie, Probabilités,
39 rue F. Joliot-Curie, 13453 Marseille cedex 13, France.
E-mail: autin@cmi.univ-mrs.fr

[2]Ecole Centrale Marseille, Laboratoire d'Analyse, Topologie, Probabilités,
38 rue F. Joliot-Curie, 13451 Marseille cedex 20, France.
E-mail: cpouet@centrale-marseille.fr

Summary: We are interested in testing whether two independent samples of n independent random variables are based on the same mixing-components or not. We provide a test procedure to detect if at least two mixing-components are distinct when their difference is a smooth function. Our test procedure is proved to be optimal according to the minimax adaptive setting that differs from the minimax setting by the fact that the smoothness is not known. Moreover, we show that the adaptive minimax rate suffers from a loss of order $(\sqrt{\ln(\ln(n))})^{-1}$ compared to the minimax rate.

Key words: Adaptation, Besov spaces, hypothesis testing, minimax theory, mixture models, nonparametric tests, wavelet decomposition.

2000 *Mathematics Subject Classification:* 62G05, 62G10, 62G20.

1 Introduction

Mixture distributions have been successful in modeling heterogeneity of data in many different domains such as medicine, physics, chemistry, social sciences, marketing and texture modeling (see McLachlan and Peel [10] for instance).

A mixture-model involves two sets of parameters: the *mixing-weights* and the *mixing-components*. Definitions of these parameters are recalled in Definition 2.1. In our study we assume that the mixing-weights are entirely known and we are interested in the mixing-components. Our goal is to provide an homogeneity test procedure for the two-samples set-up, i.e. a test procedure designed to detect whether two independent samples of n independent random variables are based on different mixtures of M components or not.
For the case $M = 1$, the classical answer to this testing problem is the Kolmogorov-Smirnov test procedure and can be found in standard textbooks. Butucea and Tribouley [3] have studied this problem from the minimax point of view (see Ingster and Suslina [9] or Spokoiny [13] for a complete review of the minimax approach in the testing problem). In particular, by considering the $\mathbb{L}_2$-loss function, they prove that the minimax rate of testing is in order of $n^{-\frac{2s}{4s+1}}$ when dealing with densities that belong to the Besov space $\mathcal{B}^s_{2,\infty}$ where $s > \frac{1}{4}$ (see Härdle et al. [8] for the precise definition of Besov spaces and their well-known properties of embeddings). Moreover, Butucea and Tribouley [3] also provided a test procedure that automatically adapts to the smoothness s of the underlying densities and that attains the rate $(n(\ln(\ln(n)))^{-\frac{1}{2}})^{-\frac{2s}{4s+1}}$, which also corresponds to the minimax rate up to a loss of order $l_n := (\ln(\ln(n)))^{-\frac{1}{2}}$. These authors believe that this loss is the least possible but let the readers wonder about it.

Here we go further. First we extend the study of Butucea and Tribouley [3] to a general mixture-model ($M > 1$) with varying mixing-weights (see Definition 2.2). Our homogeneity test procedure on the mixing-components of the two samples is shown to be adaptive, contrary to the one provided by Autin and Pouet [1]. We refer to Section 2 and to Autin and Pouet [1] and [2] for the interest to consider testing problem for such kind of models. Second, we prove that our test procedure is optimal in the minimax sense since it attains the adaptive minimax rate that is proven to be $(nl_n)^{-\frac{2s}{4s+1}}$. Hence, we prove that the loss of order l_n is unavoidable for adaptation.

The paper is organized as follows: the set-up of the study and the description of the testing problem are presented in Section 2. Section 3 deals with the description of our test procedure. In Section 4 we prove its adaptive minimax optimality and we provide the adaptive minimax rate (see Theorem 1 and Theorem 2). Technical lemmas and their proofs are postponed in Section 5.

2 Set-up and testing problem under interest

Let us now describe the set-up of our study and then the testing problem we are interested in.

The mixture-model under consideration takes into account two samples of size n ($n > 2$): $Y = (Y_1, \ldots, Y_n)$ is a sample of independent random variables with unknown marginal densities

$$f_i := \sum_{l=1}^{M} \omega_l(i) p_l, \; 1 \leq i \leq n$$

and $Z = (Z_1, \ldots, Z_n)$ is another sample of independent random variables for which the unknown marginal densities are

$$g_i := \sum_{l=1}^{M} \omega_l(i) q_l, \; 1 \leq i \leq n.$$

In addition to this, we assume that the two samples are independent from each other and that:

$$\omega_l(i) \geq 0 \quad \forall (l,i) \in \{1, \ldots, M\} \times \{1, \ldots, n\},$$
$$\sum_{l=1}^{M} \omega_l(i) = 1, \quad \forall i \in \{1, \ldots, n\}.$$

Definition 2.1. *Dealing with the mixture-model,*

1. *the vectors* ω_l $(1 \leq l \leq M)$ *are called the mixing-weights,*
2. *the densities* p_l $(1 \leq l \leq M)$ *are called the mixing-components of* Y*,*
3. *the densities* q_l $(1 \leq l \leq M)$ *are called the mixing-components of* Z*.*

Before going further, we give an example taken from medicine where such kind of mixture-models with $M = 2$ and same varying mixing-weights between the two samples is useful.

Motivated by the article of Catalona et al [4], suppose that we are interested in a screening test for prostate cancer called Prostate Specific Antigen (PSA) and imagine that we want to evaluate the effect of a new treatment on prostate cancer. Each patient of the survey is diagnosed by PSA. As this screening test is not accurate, we do not rely on the issue of the test itself but on the positive predictive value equal to 0.4 and the negative predicted value equal to 0.89 (see Catalona et al [4], Table 4). The sample is divided into a study-group containing patients who take the treatment and a comparison-group containing patients who do not take the treatment (those patients take a placebo or the usual treatment instead). These two groups are chosen such that

- the number of positive PSA screening test in the study-group is the same as the one in the comparison-group,
- the number of negative PSA screening test in the study-group is the same as the one in the comparison-group.

The distribution of the biological feature under study can be modeled as a mixture distribution. In each group, the mixing-weights are varying and they are the same between the two groups. Indeed for the study-group the distribution is

- $f_i = 0.4\ p_1 + 0.6\ p_2$ if the screening test for patient i is true,
- $f_i = 0.11\ p_1 + 0.89\ p_2$ if the screening test for patient i is false.

Similarly for the comparison group the distribution under interest is

- $g_j = 0.4\ q_1 + 0.6\ q_2$ if the screening test for patient j is true,
- $g_j = 0.11\ q_1 + 0.89\ q_2$ if the screening test for patient j is false.

Such a mixture-model can be interesting to study a possible difference between the distributions of the biological feature in the two subpopulations of patients.

Now let us go back to our theoretical study. In the sequel,

- $\overrightarrow{p} := (p_1, \ldots, p_M)$ and $\overrightarrow{q} := (q_1, \ldots, q_M)$ will characterize the mixing-components of the two samples Y and Z,
- $\mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}$ will denote the distribution of (Y, Z),

- $\mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}(\cdot)$ and $\mathbb{V}ar_{\overrightarrow{p},\overrightarrow{q}}(\cdot)$ will respectively denote the expected value and the variance under this distribution.

Definition 2.2. *Let $\Gamma^{(n)}$ be the $(M \times M)$-matrix with general term*

$$\Gamma^{(n)}_{jj'} = \frac{1}{n}\sum_{i=1}^{n} \omega_j(i)\omega_{j'}(i).$$

We say that the mixture-model is associated with varying mixing-weights as soon as the rank of $\Gamma^{(n)}$ is invertible.

For any $R > 0$, let $\mathcal{D}(R)$ be the set of all probability densities such that both their $\mathbb{L}_2$-norm and their $\mathbb{L}_\infty$-norm are bounded by R. Let $\mathcal{B}^s_{2,\infty}(R)$ be the ball of the Besov space $\mathcal{B}^s_{2,\infty}$ as defined in (3.1). We consider two subspaces containing vectors of $(\mathcal{D}(R))^{2M}$, namely $\Theta_0\,(R)$ and $\Theta_1\,(R, C, v_n, s)\,$, that are respectively defined by:

$$\Theta_0\,(R) := \{(\overrightarrow{p}, \overrightarrow{q})\ : \forall l \in \{1, \ldots, M\}, \quad p_l = q_l\}\,,$$

$$\Theta_1\,(R, C, v_n, s) := \Big\{(\,\overrightarrow{p}, \overrightarrow{q}): \ \forall l \in \{1, \ldots, M\}, \ \ \nabla_l := p_l - q_l \in \mathcal{B}^s_{2,\infty}(R), \\ \exists u \in \{1, \ldots, M\}, (p_u, q_u) \in \Lambda(C, v_n, s)\Big\},$$

where $\Lambda(C, v_n, s) := \left\{(p, q), \|p - q\|_2 \geq C v_n^{-\frac{2s}{4s+1}}\right\}$, C is a positive constant and v_n is a sequence of positive numbers tending to infinity when n goes to infinity.

For two given real numbers $s_\star$ and $s^\star$, such that $\frac{1}{4} < s_\star < s^\star$, we consider the testing problem $\mathcal{P}_n(\mathcal{H}_0, \mathcal{H}_1)$ defined just below.

Definition 2.3. *Let $\mathcal{P}_n(\mathcal{H}_0, \mathcal{H}_1)$ be the testing problem such that*

- *the null hypothesis is*

$$\mathcal{H}_0 :\ (\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0\,(R)\,,$$

- *the alternative hypothesis is*

$$\mathcal{H}_1 :\ (\overrightarrow{p}, \overrightarrow{q}) \in \Theta_1^\star\,(R, C, nl_n) := \bigcup_{s \in [s_\star,\ s^\star]} \Theta_1\,(R, C, nl_n, s)$$

  where the sequence l_n (resp. nl_n) goes to zero (resp. infinity) when n goes to infinity.

More details on the values of the sequence l_n will be given in Section 3.

Autin and Pouet [1] have considered the same testing problem but in the particular case where $s_\star = s^\star$ and $v_n = n$ instead of $v_n = nl_n$. This case is called the *non adaptive case* and involves a priori knowledge of the regularity on the functions ∇_l $(1 \leq l \leq M)$. These authors provided an optimal test procedure and proved that the minimax rate of testing was of order $n^{-\frac{2s_\star}{4s_\star+1}}$, when the smallest eigenvalue of $\Gamma^{(n)}$ is greater than a constant $K \in (0,1)$ that does not depend on n.
In this previous work, a strong link between the mixing-weights of the model and the performances of the optimal test procedure has been underlined and extensively discussed. In particular, it has been shown that the smaller K, the worse the performance of the test procedure.

In this study we consider that $s_\star < s^\star$. This case is called the *adaptive case* and does not involve a priori knowledge of the regularity on functions ∇_l $(1 \leq l \leq M)$. We recall that our goal is twofold: we aim at providing a test procedure which is optimal in the adaptive minimax sense and at exhibiting the *optimal loss* of rate of testing in comparison to the minimax setting.

3 Adaptive test procedure $\Delta_t^\star$

From now on, we consider the wavelet setting and thus an orthonormal wavelet basis of the functional space $\mathbb{L}_2$, $\{\phi_{jk}, \psi_{j'k};\ j' \geq j, k \in \mathbb{Z}\}$. We assume that the supports of the scaling function ϕ and the mother wavelet ψ are both included in $[-L, L)$ for some $L > 0$. We refer the reader to Daubechies [5] for some examples of such bases.
To derive optimal results, we suppose that any function $p_u - q_u$ $(1 \leq u \leq M)$ is regular enough that is to say the energy of its wavelet coefficients with large level is decreasing fast enough. More precisely, any function $p_u - q_u$ is supposed to belong to the Besov ball $\mathcal{B}^s_{2,\infty}(R)$ defined as follows

$$\mathcal{B}^s_{2,\infty}(R) := \left\{ h \in \mathbb{L}_2;\ \sup_{j \in \mathbb{N}} 2^{2js} \sum_{j' \geq j} \sum_{k \in \mathbb{Z}} \beta^2_{j'k} \leq R^2 \right\}. \tag{3.1}$$

Definition 3.1. *[Inverse mixing-weights]. Let us consider the mixture-model presented in Section 2. The vectors* $a_l := (a_l(1), \ldots, a_l(n))$ $(1 \leq l \leq M)$ *are called the inverse mixing-weights of the mixture-model and are such that*

$$a_l(i) := \frac{1}{det(\Gamma^{(n)})} \sum_{u=1}^{M} (-1)^{l+u} \gamma_{l_u} \omega_u(i) \quad (1 \leq i \leq n), \tag{3.2}$$

where γ_{l_u} *is the minor* (l, u) *of the matrix* $\Gamma^{(n)}$.

Remark 3.1. *Following Maiboroda [11] and Pokhyl'ko [12] and by denoting* $\delta_{..}$ *the Kronecker delta, the inverse mixing-weights are such that, for any* $(k, l) \in \{1, \ldots, M\}^2$, $\frac{1}{n} \sum_{i=1}^{n} \omega_k(i) a_l(i) = \delta_{kl}$.

As discussed in Section 3.1 of Autin and Pouet [1] for the case $s_\star = s^\star$, a natural way to test whether the two independent samples come from the same mixing-components or not is to consider the test statistic

$$T_j := \frac{2}{n^2} \sum_{i_1=2}^{n} \sum_{i_2=1}^{i_1-1} \sum_{l=1}^{M} \sum_{k \in \mathbb{Z}} a_l(i_1) a_l(i_2) (\phi_{jk}(Y_{i_1}) - \phi_{jk}(Z_{i_1}))(\phi_{jk}(Y_{i_2}) - \phi_{jk}(Z_{i_2}))$$

for a judicious choice of level j and to use the test $\Delta_j(t_n) := \mathbf{1}_{\{T_j > t_n\}}$, where $\mathbf{1}_{\{.\}}$ denotes the indicator function and t_n is a threshold value that is of order $n^{-\frac{4s_\star}{4s_\star+1}}$.
As proved by these authors, the level j that must be considered - and also the test $\Delta_j(t_n)$ - depends on the regularity $s_\star = s^\star$ that contributes in the definition of the alternative hypothesis (see Theorem 3.6 in Autin and Pouet [1]). We stress that the context is different here: the alternative hypothesis $\mathcal{H}_1$ we consider is weaker since we suppose that $s_\star < s^\star$.

For any integer $n > 2$, let us introduce the quantity $l_n := \left(\sqrt{\ln(\ln(n))}\right)^{-1}$ and the integers $j_\star$ and $j^\star$ such that

$$2^{-j_\star} \leq (nl_n)^{-\frac{2}{4s_\star+1}} < 2^{1-j_\star} \quad \text{and} \quad 2^{-j^\star} \leq (nl_n)^{-\frac{2}{4s^\star+1}} < 2^{1-j^\star}. \tag{3.3}$$

We define the set $\mathcal{J}_n$ of all the levels with value in $[j^\star, j_\star]$, i.e.

$$\mathcal{J}_n = \{j \in \mathbb{N} : j^\star \leq j \leq j_\star\}. \tag{3.4}$$

Remark 3.2. *According to Lemma 1, we notice that the cardinality of* $\mathcal{J}_n$ *does not exceed* $\ln(nl_n)$ *up to a constant that only depends on* $s_\star$ *and* $s^\star$.

Analogously to Pokhyl'ko [12] and Autin and Pouet [1], the following assumption is made on the mixture-model:

Assumption 1. *[Behavior of the smallest eigenvalue] For any* $n > 2$, *the smallest eigenvalue of the* $(M \times M)$*-matrix* $\Gamma^{(n)}$ *is greater than or equal to* K, *with* $0 < K < 1$.

Note that in that case, according to (3.2),

$$\sup_{1 \le i \le n} |a_l(i)| \le \frac{(M-1)!}{K^M}. \tag{3.5}$$

We now present our test procedure.

Definition 3.2. *[Adaptive test* $\Delta_t^\star$*]. Let* $n > 2$. *For any* $j \in \mathcal{J}_n$, *we denote by* s_j *the real number such that* $2^{-j} = (nl_n)^{-\frac{2}{4s_j+1}}$ *and for* $t > 0$, *we define a threshold* $t_n(j)$ *as follows*

$$t_n(j) := t\,(nl_n)^{-\frac{4s_j}{4s_j+1}}.$$

To solve the testing problem $\mathcal{P}_n(\mathcal{H}_0, \mathcal{H}_1)$ we consider the following test

$$\Delta_t^\star := \max\left\{\Delta_j(t_n(j)),\ j \in \mathcal{J}_n\right\}, \tag{3.6}$$

where, for any $j \in \mathcal{J}_n$, $\Delta_j(t_n(j)) := \mathbf{1}_{\{T_j > t_n(j)\}}$.

We note two main differences when the test $\Delta_t^\star$ is compared to the test proposed in the non adaptive case in paragraph 3.1 of Autin and Pouet [1]. Indeed,

1. the decision rule depends on the results of many test statistics T_j,
2. the threshold values $t_n(j)$ depend on j and take into account l_n in their definition.

4 Adaptive minimax optimality of $\Delta_t^\star$

We describe the behavior of the test procedure $\Delta_t^\star$. According to the minimax setting, it is consistent for a loss of order l_n as we shall prove it thanks to Theorems 1 and 2.

Theorem 1 (Upper bound)**.** *Let $R > 0$. There exists a non negative constant $C_{\scriptscriptstyle T} = C_{\scriptscriptstyle T}(L, R, \|\phi\|_\infty)$ such that for any $C > \left(2M\left(R^2 + \frac{\sqrt{2C_{\scriptscriptstyle T}}}{K}\right)\right)^{\frac{1}{2}}$ and any $MK^{-1}(2C_{\scriptscriptstyle T})^{\frac{1}{2}} < t < \left(\frac{C^2}{2} - MR^2\right)$, the test $\Delta_t^\star$ defined in (3.6) satisfies*

$$\lim_{n\to+\infty}\left(\sup_{(\overrightarrow{p},\overrightarrow{q})\in\Theta_0(R)} \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_t^\star = 1) \ + \sup_{(\overrightarrow{p},\overrightarrow{q})\in\Theta_1^\star(R,C,nl_n)} \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_t^\star = 0)\right) = 0.$$

Proof. The proof of Theorem 1 is a direct consequence of two propositions: Proposition 4.1 deals with the control of the first-type error and Proposition 4.2 deals with the control of the second-type error.

Proposition 4.1. *Let $R > 0$. There exists a non negative constant $C_{\scriptscriptstyle T} = C_{\scriptscriptstyle T}(L, R, \|\phi\|_\infty)$ such that, for any $t > MK^{-1}(2C_{\scriptscriptstyle T})^{\frac{1}{2}}$, the test $\Delta_t^\star$ satisfies*

$$\lim_{n\to+\infty}\sup_{(\overrightarrow{p},\overrightarrow{q})\in\Theta_0(R)} \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_t^\star = 1) = 0.$$

Proof. From the definition of $\Delta_t^\star$ one gets for any $(\overrightarrow{p},\overrightarrow{q}) \in \Theta_0(R)$,

$$\begin{aligned}\mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_t^\star = 1) &\leq \sum_{j\in\mathcal{J}_n} \mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}(\Delta_j(t_n(j)) = 1)\\ &= \sum_{j\in\mathcal{J}_n} \mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}(T_j > t_n(j)).\end{aligned}$$

According to Remark 3.2, if, for any $j \in \mathcal{J}_n$, $\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}(T_j > t_n(j))$ converges to zero faster than $(\ln(nl_n))^{-1}$ as n tends to infinity, then the proof of Proposition 4.1 will be ended. Let us now prove that result of convergence.

Let $j \in \mathcal{J}_n$. Under the null hypothesis $\mathcal{H}_0$, the test statistic T_j is clearly centered. Let $\sigma_n^2(j)$ denote the variance of T_j when $(\overrightarrow{p},\overrightarrow{q}) \in \Theta_0(R)$ and let $\tilde{T}_j = \sigma_n^{-1}(j)T_j$ be the normalized test statistic which can be rewritten as follows

$$\tilde{T}_j = \sum_{i_1=2}^{n} \sum_{i_2=1}^{i_1-1} h_j(i_1, i_2),$$

where, for any (i_1, i_2),

$$h_j(i_1, i_2) := \frac{2}{n^2\sigma_n(j)} \sum_{l=1}^{M} \sum_{k\in\mathbb{Z}} a_l(i_1)a_l(i_2) \left(\phi_{jk}(Y_{i_1}) - \phi_{jk}(Z_{i_1})\right)\left(\phi_{jk}(Y_{i_2}) - \phi_{jk}(Z_{i_2})\right). \quad (4.1)$$

Notice that the normalized test statistic $\tilde{T}_j$ is a zero-mean martingale under the null hypothesis $\mathcal{H}_0$ and that for any chosen $q > 0$, $\mathcal{J}_n$ can be split into two subsets:

$$\mathcal{J}_{0,n} = \left\{ j \in \mathcal{J}_n : \sigma_n^2(j) \leq t_n^2(j)(\ln(n))^{-(1+q)} \right\},$$
$$\mathcal{J}_{1,n} = \left\{ j \in \mathcal{J}_n : \sigma_n^2(j) > t_n^2(j)(\ln(n))^{-(1+q)} \right\}.$$

For the case of small variance ($j \in \mathcal{J}_{0,n}$), we use Bienayme-Chebyshev's inequality (see Devroye and Lugosi [6] - chapter 2) to get the following bounds

$$\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}\left(T_j > t_n(j)\right) \leq \frac{\sigma_n^2(j)}{t_n^2(j)} \leq \frac{1}{(\ln(n))^{1+q}}.$$

According to the right-hand side of the last inequality, we deduce that $\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}\left(T_j > t_n(j)\right)$ converges to zero faster than $(\ln(nl_n))^{-1}$ as n tends to infinity. Therefore it remains to handle the case of large variance ($j \in \mathcal{J}_{1,n}$). Here we apply Theorem 3.9 from Hall and Heyde [7] and the concentration properties of the Gaussian distribution to get the following bound

$$\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}\left(T_j > t_n(j)\right) \leq M_{1,n,j} + C_1 \left(M_{2,n,j} + M_{3,n,j}\right)^{\frac{1}{5}} \quad (4.2)$$

with

$$M_{1,n,j} \quad := \quad \exp\left(-\frac{t_n^2(j)}{2\sigma_n^2(j)}\right),$$

$$M_{2,n,j} := \mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(\sum_{i_1=2}^{n}\left(\sum_{i_2=1}^{i_1-1} h_j(i_1,i_2)\right)^4\right),$$

$$M_{3,n,j} := \mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(\left(\sum_{i_1=2}^{n}\left(\sum_{i_2=1}^{i_1-1} h_j(i_1,i_2)\right)^2 - 1\right)^2\right),$$

and where C_1 is an absolute constant.

Following Proposition 3.4 from Autin and Pouet [1], there exists a non negative constant C_T depending on L, R and $\|\phi\|_\infty$ such that, for any $(\overrightarrow{p},\overrightarrow{q}) \in (\mathcal{D}(R))^{2M}$ and any level j,

$$\mathbb{V}ar_{\overrightarrow{p},\overrightarrow{q}}(T_j) \leq \frac{C_T M^2}{K^2}\left(\frac{2^j}{n^2}+\frac{1}{n}\sum_{l=1}^{M}\|p_l-q_l\|_2^2+\sqrt{\frac{2^j}{n^3}}\sum_{l=1}^{M}\|p_l-q_l\|_2\right). \quad (4.3)$$

Therefore, under $\mathcal{H}_0$, the variance of the test statistics T_j, namely $\sigma_n^2(j)$, is smaller than or equal to $C_T M^2 K^{-2} n^{-2} 2^j$.

So,

$$\begin{aligned} M_{1,n,j} &= \exp\left(-\frac{t_n^2(j)}{2\sigma_n^2(j)}\right) \\ &\leq \exp\left(-\frac{t^2K^2n^2\,(nl_n)^{-\frac{8s_j}{4s_j+1}}}{C_T\ 2^{j+1}M^2}\right) \\ &= \exp\left(-\frac{t^2K^2}{2C_T M^2 l_n^2}\right) \\ &= \exp\left(-\frac{t^2K^2}{2C_T M^2}\ln\ln(n)\right) \\ &= \big(\ln(n)\big)^{-\frac{t^2K^2}{2C_T M^2}}. \end{aligned}$$

Hence $M_{1,n,j}$ goes to zero faster than $(\ln(nl_n))^{-1}$ when n tends to infinity, provided $t^2 > 2C_T M^2 K^{-2}$.

From Lemma 4 and the definition of the constant F within, the following bounds can be obtained for any $j \in \mathcal{J}_{1,n}$:

$$\begin{aligned} M_{2,n,j} &\leq F\ \frac{2^{2j}(1+6n)}{n^4\sigma_n^2(j)} \\ &\leq \frac{F(nl_n)^{\frac{4}{4s_j+1}}(1+6n)(\ln(n))^{1+q}}{n^4 t_n^2(j)} \\ &= \frac{F(nl_n)^{\frac{8s_j+4}{4s_j+1}}(1+6n)(\ln(n))^{1+q}}{n^4 t^2} \\ &\leq \frac{7F}{t^2}\frac{(\ln(n))^{1+q}}{n^{\frac{4s_\star-1}{4s_\star+1}}}. \end{aligned}$$

Since $s_\star > \frac{1}{4}$, $M_{2,n,j}$ goes to zero faster than any power of $(\ln(nl_n))^{-1}$ as n tends to infinity.

From Lemma 5 and the definition of the constant G within, the following bounds can be obtained for any $j \in \mathcal{J}_{1,n}$:

$$\begin{aligned} M_{3,n,j} &\leq M_{2,n,j} + G\ \frac{2^j}{n^3\sigma_n^2(j)}\left(2^j + \frac{1}{n\sigma_n^2(j)}\right) \\ &\leq \frac{7F}{t^2}\frac{(\ln(n))^{1+q}}{n^{\frac{4s_\star-1}{4s_\star+1}}} + \frac{G(nl_n)^{\frac{2}{4s_j+1}}(\ln(n))^{1+q}}{n^3 t_n^2(j)}\left((nl_n)^{\frac{2}{4s_j+1}} + \frac{(\ln(n))^{1+q}}{nt_n^2(j)}\right) \\ &= \frac{7F}{t^2}\frac{(\ln(n))^{1+q}}{n^{\frac{4s_\star-1}{4s_\star+1}}} + \frac{G(\ln(n))^{1+q}}{n^3 t^2}\left((nl_n)^{\frac{8s_j+4}{4s_j+1}} + \frac{(\ln(n))^{1+q}(nl_n)^{\frac{16s_j+2}{4s_j+1}}}{nt^2}\right) \\ &\leq \frac{7F}{t^2}\frac{(\ln(n))^{1+q}}{n^{\frac{4s_\star-1}{4s_\star+1}}} + \frac{G(\ln(n))^{1+q}}{t^2}\left(\frac{1}{n^{\frac{4s_\star-1}{4s_\star+1}}} + \frac{(\ln(n))^{1+q}}{n^{\frac{2}{4s^\star+1}}\ t^2}\right). \end{aligned}$$

Since $s_\star > \frac{1}{4}$, $M_{3,n,j}$ goes to zero faster than any power of $(\ln(nl_n))^{-1}$ as n tends to infinity.

Finally, looking at (4.2) and the bounds of $M_{1,n,j}$, $M_{2,n,j}$ and $M_{3,n,j}$, we conclude that, even for $j \in \mathcal{J}_{1,n}$, $\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}(T_j \geq t_n(j))$ converges to zero faster than $(\ln(nl_n))^{-1}$ as n tends to infinity. Proposition 4.1 is also proved. $\square$

Proposition 4.2. *Let $R > 0$ and C_T be the constant appearing in Proposition 4.1. For any $C > \left(2M\left(R^2 + \frac{\sqrt{2C_T}}{K}\right)\right)^{\frac{1}{2}}$ and any $t < \left(\frac{c^2}{2} - MR^2\right)$,* *the test $\Delta_t^\star$ satisfies*

$$\lim_{n\to+\infty} \sup_{(\overrightarrow{p},\overrightarrow{q})\in\Theta_1^\star(R,C,nl_n)} \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_t^\star = 0) = 0.$$

Proof. Let $(\overrightarrow{p},\overrightarrow{q}) \in \Theta_1^\star(R,C,nl_n)$, let us bound the second-type error $\mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_t^\star = 0)$. From the definition of $\Delta_t^\star$, we remark that it suffices to prove that $\lim_{n\to+\infty} \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_j(t_n(j)) = 0) = 0$ for one $j \in \mathcal{J}_n$.
Let $s \in [s_\star, s^\star]$ be such that $(\overrightarrow{p},\overrightarrow{q}) \in \Theta_1(R,C,nl_n,s)$ and j_s be the integer such that $2^{-j_s} \leq (nl_n)^{\frac{-2}{4s+1}} < 2^{1-j_s}$. We use the expectation of the test statistic with some approximation argument.

$$\begin{aligned}\mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}(\Delta_{j_s}(t_n(j_s)) = 0) &= \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}\left(T_{j_s} \leq t_n(j_s)\right) \\ &= \mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}\left(\mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}(T_{j_s}) - T_{j_s} \geq \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}(T_{j_s}) - t_n(j_s)\right).\end{aligned}$$

Following Corollary 3.3 of Autin and Pouet [1], the wavelet expansion in the Besov ball $\mathcal{B}^s_{2,\infty}(R)$ leads, for any n large enough, to

$$\begin{aligned}\mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}(T_{j_s}) - t_n(j_s) = &\sum_{l=1}^M \|p_l - q_l\|_2^2 - \sum_{l=1}^M \sum_{j=j_s}^\infty \sum_{k\in\mathbb{Z}} \left(\int_{\mathbb{R}} (p_l - q_l)\psi_{jk}\right)^2 \\ &- \frac{1}{n^2}\sum_{l=1}^M \sum_{k\in\mathbb{Z}} \sum_{i=1}^n \left(\int_{\mathbb{R}} \left(a_l(i)f_i - a_l(i)g_i\right)\phi_{j_s k}\right)^2 - t_n(j_s) \\ \geq &\sum_{l=1}^M \|p_l - q_l\|_2^2 - M\ R^2\ 2^{-2sj_s} - \frac{8LMR^2}{Kn} - t_n(j_s) \\ \geq &\frac{1}{2}\sum_{l=1}^M \|p_l - q_l\|_2^2 - M\ R^2\ 2^{-2sj_s} - t_n(j_s).\end{aligned}$$

Since $t < \left(\frac{C^2}{2} - MR^2\right)$ the right-hand side of the inequality is necessarily positive. According to (4.3) and applying Bienayme-Chebychev's inequality we get:

$$\begin{aligned}&\mathbb{P}_{\overrightarrow{p},\overrightarrow{q}}\left(\mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}(T_{j_s}) - T_{j_s} \geq \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}(T_{j_s}) - t_n(j_s)\right) \\ \leq\ &\frac{C_{\scriptscriptstyle T} M^2\left(2^{j_s} + n\sum_{l=1}^M \|p_l - q_l\|_2^2 + \sqrt{2^{j_s}n}\sum_{l=1}^M \|p_l - q_l\|_2\right)}{n^2\ K^2\left(\frac{1}{2}\sum_{l=1}^M \|p_l - q_l\|_2^2 - M\ R^2\ 2^{-2sj_s} - t_n(j_s)\right)^2}.\end{aligned}$$

Still according to the definition of j_s one gets for any n large enough:

$$\begin{aligned}
&\mathbb{P}_{\vec{p},\vec{q}}\left(\Delta_{j_s}(t_n(j_s))=0\right)\\
\leq\ & \frac{C_T M^2\left(2^{j_s}+n\sum_{l=1}^{M}\|p_l-q_l\|_2^2+\sqrt{2^{j_s}n}\sum_{l=1}^{M}\|p_l-q_l\|_2\right)}{n^2\ K^2\ \left(\frac{1}{2}-\frac{MR^2}{C^2}-\frac{t}{C^2}\right)^2\left(\sum_{l=1}^{M}\|p_l-q_l\|_2^2\right)^2}\\
\leq\ & \frac{C_T}{\left(\frac{C^2}{2M}-R^2-\frac{t}{M}\right)^2K^2}\left(2l_n^2+\frac{C^2\,(nl_n)^{\frac{4s}{4s+1}}}{n}+C\sqrt{\frac{2M}{n^3}}\,(nl_n)^{\frac{6s+1}{4s+1}}\right).
\end{aligned}$$

The right-hand side of the last inequality does not depend on $\vec{p}$ and $\vec{q}$ and goes to zero when n tends to infinity. Proposition 4.2 is also proved. □

□

The adaptive minimax optimality of $\Delta_t^\star$ is given by the statement of a lower bound. The next theorem proves that the price to pay for adaptation is the one expected and conjectured by Butucea and Tribouley [3], that is to say $l_n = \left(\sqrt{\ln(\ln(n))}\right)^{-1}$.

Theorem 2 (Lower bound)**.** *There exists $c_0 := c_0(L, M, R, s^\star)$ such that for any $C < c_0$,*

$$\lim_{n\to+\infty}\inf_{\Delta}\left(\sup_{(\vec{p},\vec{q})\in\Theta_0(R)}\mathbb{P}_{\vec{p},\vec{q}}(\Delta=1)\ +\sup_{(\vec{p},\vec{q})\in\Theta_1^\star(R,C,nl_n)}\mathbb{P}_{\vec{p},\vec{q}}(\Delta=0)\right)=1.$$

Remark 4.1. *Although we are not interested in the exact separation constant, a careful study of the proof of Theorem 2 shows that the expression of the constant c_0 can be improved with an extra assumption on the behavior of the smallest eigenvalues of matrices $\Gamma^{(n)}$, $n > 2$. Indeed an upper bound involving the constant K that appears in Assumption 1 leads to an expression of c_0 depending also on K. In this case the smallest K, the largest the constant c_0.*

Proof. Here we decide to consider the Haar wavelet basis in order to avoid technicalities. For this wavelet basis the constant c_0 will be such that

$$c_0 := \frac{2^{-s^\star}}{\sqrt{M}}\min\left(R, 2^{-\frac{1}{4}}\right).$$

As we are not interested in the exact separation constant, this result can easily be extended to the case of compactly supported wavelets. A parameter L connected to the supports of the scaling function and the mother wavelet must be introduced in this case.

Let $\mathcal{S}_n$ denote a net on the smoothness space with cardinality N of order $\ln(n)(\ln(\ln(n)))^{-1}$ and such that

$$\forall s,t \in \mathcal{S}_n : \ \ s_\star \leq s,t \leq s^\star, \ \ |s-t| \geq \frac{\ln(\ln(n))}{\ln(n)}.$$

Let $\mathcal{H}_n$ denote the associated net on the level space and, for any $s \in \mathcal{S}_n$, let j_s be the level of $\mathcal{H}_n$ such that $2^{-j_s} \leq (nl_n)^{-\frac{2}{4s+1}} < 2^{1-j_s}$.

We consider $\overrightarrow{p} \in \mathcal{D}(R)$ such that

$$\inf_{1 \leq l \leq M} \inf_{z \in [0,1[} p_l(z) > \frac{1}{2}. \tag{4.4}$$

For any given $s \in \mathcal{S}_n$ and any $l \in \{1,\dots,M\}$, we introduce $\overrightarrow{q} \in \mathcal{D}(R)$ as follows:

$$q_l = p_l + \theta_l r_{n,j_s} \sum_{k \in \mathcal{T}_{j_s}} \zeta_{j_s k} \psi_{j_s k},$$

where $\mathcal{T}_{j_s} = \left\{0,\dots,2^{j_s}-1\right\}$, $\theta = (\theta_1,\dots,\theta_M)$ is an eigenvector associated to the smallest eigenvalue λ_n $(0 < \lambda_n < 1)$ of the matrix $\Gamma^{(n)}$ with l_2-norm equal to 1 and

$$r_{n,j_s} \quad = \quad C\sqrt{M}\ 2^s\ 2^{-j_s s - \frac{j_s}{2}}. \tag{4.5}$$

The prior probability is defined as follows: the random variables $\zeta_{j_s k}$ $(k \in \mathcal{T}_{j_s})$ are independent Rademacher random variables. Let π_{j_s} denote the prior probability concentrated on level j_s. The level is chosen uniformly on $\mathcal{H}_n$.

Therefore, for any $1 \leq i \leq n$, the densities of the random variable Y_i and Z_i are respectively

$$f_i := \sum_{l=1}^{M} \omega_l(i) p_l \quad \text{ and } \quad g_i := f_i + \sum_{l=1}^{M} \omega_l\,(i)\,(q_l - p_l)\,.$$

Direct calculations left to the reader entail that $(\overrightarrow{p}, \overrightarrow{q})$ belongs to the set $\Theta_1(R, C, nl_n, s)$ provided $C \leq 2^{-s} R M^{-\frac{1}{2}}$.

Unless explicitly specified, the expectation $\mathbb{E}(\cdot)$ is taken regarding the random variables $Y_1, \ldots, Y_n, Z_1, \ldots, Z_n$ and $\zeta_{j_s k}, k \in \mathcal{T}_{j_s}, j_s \in \mathcal{H}_n$.

The general calculation follows the usual path (see Ingster and Suslina [9]). The sum of the errors is bounded below as follows:

$$\inf_{\Delta} \left(\mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}}(\Delta = 1) + \mathbb{P}_{\overrightarrow{p}, \overrightarrow{q}}(\Delta = 0) \right) \geq 1 - \frac{1}{2} \left\| \frac{1}{N} \sum_{s \in \mathcal{S}_n} \mathbb{P}_{\pi_{j_s}} - \mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}} \right\|_1 .$$

In order to show indistinguishability, we prove that the $\mathbb{L}_1$-distance distance goes to zero. An upper bound of the $\mathbb{L}_1$-distance is given by

$$\left\| \frac{1}{N} \sum_{s \in \mathcal{S}_n} \mathbb{P}_{\pi_{j_s}} - \mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}} \right\|_1 \leq \sqrt{\mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left(\left(\frac{1}{N} \sum_{s \in \mathcal{S}_n} \frac{d\mathbb{P}_{\pi_{j_s}}}{d\mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}}} - 1 \right)^2 \right)}$$

$$\leq \sqrt{\frac{1}{N^2} \sum_{s \in \mathcal{S}_n} \mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left(\left(\frac{d\mathbb{P}_{\pi_{j_s}}}{d\mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}}} - 1 \right)^2 \right) + \frac{1}{N^2} \sum_{s \neq s' \in \mathcal{S}_n} \left[\mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left(\frac{d\mathbb{P}_{\pi_{j_s}}}{d\mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}}} \frac{d\mathbb{P}_{\pi_{j_{s'}}}}{d\mathbb{P}_{\overrightarrow{p}, \overrightarrow{p}}} \right) - 1 \right]}.$$

Remark 4.2. *Within the square root, the left term (we call it the square term) is handled as in the non-adaptive case while the right term (we call it the cross term) only appears in the adaptive case and leads to technical calculations.*

Lemma 6 shows that there exists an universal upper bound for each summand in the cross term and it goes to zero when n tends to infinity.

Let us now pay attention to the square term. It is handled as in the non-adaptive case. Analogously to Autin and Pouet [1], because of (4.4) we get the following inequalities:

$$\begin{aligned} \mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left(\left(\frac{dP_{\pi_{j_s}}}{dP_{\overrightarrow{p}, \overrightarrow{p}}} - 1 \right)^2 \right) &\leq \exp\left(2^{4s+1} C^4 M^2 \left| \mathcal{T}_{j_s} \right| (n\lambda_n)^2 \, 2^{-4j_s s - 2j_s} \right) - 1 \\ &\leq \exp\left(2^{4s^\star + 1} C^4 M^2 l_n^{-2} \right) - 1. \end{aligned}$$

Because of $N = \mathcal{O}\left(\frac{\ln(n)}{\ln(\ln(n))}\right)$, the following limit holds

$$\lim_{n\to+\infty} N^{-1} \ (\ln(n))^{2^{4s^\star+1}M^2C^4} = 0,$$

provided C is such that $C < 2^{-s^\star-\frac{1}{4}}M^{-\frac{1}{2}}$. This result entails that the square term goes to zero when n tends to infinity.

Gathering the results for the square and the cross terms, we conclude that, if $C < c_0 = 2^{-s^\star}M^{-\frac{1}{2}}\min\left(R, 2^{-\frac{1}{4}}\right)$ the lower bound goes to one as n tends to infinity and that the loss l_n is unavoidable. □

5 Appendix

In this last section, we present the technical lemmas used to prove Theorem 1 and Theorem 2.

Lemma 1. *Let $n > 2$. Consider the set $\mathcal{J}_n$ defined in (3.4). Then its cardinality can be bounded as follows:*

$$Card(\mathcal{J}_n) \le \left(\frac{2}{4s_\star+1} - \frac{2}{4s^\star+1} + 1\right)\frac{\ln(nl_n)}{\ln(2)}.$$

Proof. From (3.3) and according to the definition of $\mathcal{J}_n$ we deduce that for any $j \in \mathcal{J}_n$,

$$j \ge \frac{2\ln(nl_n)}{(4s^\star+1)\ln(2)} \text{ and } j < \frac{2\ln(nl_n)}{(4s_\star+1)\ln(2)} + 1.$$

Hence,

$$\begin{aligned} Card(\mathcal{J}_n) &\le \left(\frac{2\ln(nl_n)}{(4s_\star+1)\ln(2)} - \frac{2\ln(nl_n)}{(4s^\star+1)\ln(2)} + 1\right) \\ &\le \left(\frac{2}{4s_\star+1} - \frac{2}{4s^\star+1} + 1\right)\frac{\ln(nl_n)}{\ln(2)}. \end{aligned}$$

□

Lemma 2. *Let $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$. For any level j, any $i \in \{1,\dots,n\}$ and any $(k_1,k_2) \in \mathbb{Z}^2$,*

$$\mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}\left|\left(\phi_{jk_1}(Y_i) - \phi_{jk_1}(Z_i)\right)\left(\phi_{jk_2}(Y_i) - \phi_{jk_2}(Z_i)\right)\right| \le 2R.$$

Proof. Let us consider $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and a level j. For any $i \in \{1, \dots, n\}$ and any $(k_1, k_2) \in \mathbb{Z}^2$,

$$\begin{aligned}
& \mathbb{E}_{\overrightarrow{p}, \overrightarrow{q}} \left| (\phi_{jk_1}(Y_i) - \phi_{jk_1}(Z_i)) (\phi_{jk_2}(Y_i) - \phi_{jk_2}(Z_i)) \right| \\
\leq \; & \frac{1}{2} \left[\mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left((\phi_{jk_1}(Y_i) - \phi_{jk_1}(Z_i))^2 \right) + \mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left((\phi_{jk_2}(Y_i) - \phi_{jk_2}(Z_i))^2 \right) \right] \\
\leq \; & \frac{1}{2} \left(\|f_i\|_\infty \left(\int_{\mathbb{R}} \phi_{jk_1}^2 + \int_{\mathbb{R}} \phi_{jk_2}^2 \right) + \|g_i\|_\infty \left(\int_{\mathbb{R}} \phi_{jk_1}^2 + \int_{\mathbb{R}} \phi_{jk_2}^2 \right) \right) \\
= \; & \|f_i\|_\infty + \|g_i\|_\infty \\
\leq \; & 2R.
\end{aligned}$$

□

Lemma 3. *Let $(i_1, i_2, i_3, i_4) \in \{1, \dots, n\}^4$ such that $i_k \neq i_l$ for any $k \neq l$. There exists $D = D(K, L, M, R, \|\phi\|_\infty)$ such that, for any $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and any level j,*

$$\mathbb{E}_{\overrightarrow{p}, \overrightarrow{q}} \left(h_j(i_1, i_2) h_j(i_1, i_4) h_j(i_3, i_2) h_j(i_3, i_4) \right) \leq D \; \frac{2^j}{n^8 \sigma_n^4(j)},$$

with $h_j(.)$ as defined in (4.1) and $\sigma_n^2(j)$ as defined in Section 4.

Proof. Consider $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and a level j. From Lemma 2 and (3.5)

$$\begin{aligned}
& \mathbb{E}_{\overrightarrow{p}, \overrightarrow{q}} \left(h_j(i_1, i_2) h_j(i_1, i_4) h_j(i_3, i_2) h_j(i_3, i_4) \right) \\
\leq \; & \frac{16}{n^8 \; \sigma_n^4(j)} (2R)^3 M^4 (K^{-M} (M-1)!)^8 (4L)^2 \\
& \times \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} \mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left| (\phi_{jk_1}(Y_{i_1}) - \phi_{jk_1}(Z_{i_1})) (\phi_{jk_2}(Y_{i_1}) - \phi_{jk_2}(Z_{i_1})) \right| \\
\leq \; & \frac{16}{n^8 \; \sigma_n^4(j)} (2R)^3 M^4 (K^{-M} (M-1)!)^8 (4L)^3 \sum_{k \in \mathbb{Z}} \mathbb{E}_{\overrightarrow{p}, \overrightarrow{p}} \left(\phi_{jk}(Y_{i_1}) - \phi_{jk}(Z_{i_1}) \right)^2 \\
\leq \; & \frac{32}{n^8 \; \sigma_n^4(j)} (2R)^3 M^4 (K^{-M} (M-1)!)^8 L (4L)^3 (2^{\frac{j}{2}+1} \|\phi\|_\infty)^2 \\
= \; & 2^{16} \; K^{-8M} L^4 M^4 ((M-1)!)^8 R^3 \|\phi\|_\infty^2 \; 2^j \; \frac{1}{n^8 \; \sigma_n^4(j)} \\
:= \; & D \; \frac{2^j}{n^8 \; \sigma_n^4(j)}.
\end{aligned}$$

□

Lemma 4. *There exists a positive constant $F = F(K, L, M, \|\phi\|_\infty)$ such that, for any $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and any level j:*

$$M_{2,n,j} \leq F \; \frac{2^{2j}(1+6n)}{n^4 \sigma_n^2(j)}.$$

Proof. Let us consider $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and a level j. The quantity $M_{2,n,j}$ can be decomposed as follows:

$$\begin{aligned}
M_{2,n,j} = &\quad \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}} \left(\sum_{i_1=2}^{n} \left(\sum_{i_2=1}^{i_1-1} h_j(i_1, i_2) \right)^4 \right) \\
= &\quad \sum_{i_1=2}^{n} \sum_{i_2=1}^{i_1-1} \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}} \left(h_j^4(i_1, i_2) \right) + 6 \sum_{i_1=3}^{n} \sum_{i_2=2}^{i_1-1} \sum_{i_3=1}^{i_2-1} \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}} \left(h_j^2(i_1, i_2) h_j^2(i_1, i_3) \right) \\
:= &\quad A_{2,n,j} + 6B_{2,n,j}.
\end{aligned}$$

Clearly, using (3.5) we have:

$$\begin{aligned}
A_{2,n,j} &= \sum_{i_1=2}^{n} \sum_{i_2=1}^{i_1-1} \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}} \left(h_j^4(i_1, i_2) \right) \\
&\leq \frac{4M^2(K^{-M}(M-1)!)^4}{n^4 \sigma_n^2(j)} (4L)^2 \left(2^{\frac{j}{2}+1} \|\phi\|_\infty \right)^4 \sum_{i_1=2}^{n} \sum_{i_2=1}^{i_1-1} \mathbb{E}_{\overrightarrow{p},\overrightarrow{p}} \left(h_j^2(i_1, i_2) \right) \\
&= 2^{10} M^2 L^2 (K^{-M}(M-1)!)^4 \|\phi\|_\infty^4 \; \frac{2^{2j}}{n^4 \; \sigma_n^2(j)} \; \mathbb{V}ar_{\overrightarrow{p},\overrightarrow{p}}(\tilde{T}_j) \\
&= 2^{10} K^{-4M} L^2 M^2 ((M-1)!)^4 \|\phi\|_\infty^4 \; \frac{2^{2j}}{n^4 \; \sigma_n^2(j)} \\
&:= F \; \frac{2^{2j}}{n^4 \; \sigma_n^2(j)}.
\end{aligned}$$

In the same way,

$$
\begin{aligned}
B_{2,n,j} &= \sum_{i_1=3}^{n}\sum_{i_2=2}^{i_1-1}\sum_{i_3=1}^{i_2-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}\left(h_j^2(i_1,i_2)h_j^2(i_1,i_3)\right)\\
&\leq \frac{4M^2(K^{-M}(M-1)!)^4}{n^3\sigma_n^2(j)}(4L)^2\left(2^{\frac{j}{2}+1}\|\phi\|_\infty\right)^4\sum_{i_1=2}^{n}\sum_{i_2=1}^{i_1-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(h_j^2(i_1,i_2)\right)\\
&= 2^{10}M^2L^2(K^{-M}(M-1)!)^4\|\phi\|_\infty^4\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}\ \mathbb{V}ar_{\overrightarrow{p},\overrightarrow{p}}(\tilde{T}_j)\\
&= 2^{10}L^2M^2(K^{-M}(M-1)!)^4\|\phi\|_\infty^4\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}\\
&= F\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}.
\end{aligned}
$$

Looking at the bounds of $A_{2,n,j}$ and $B_{2,n,j}$, we conclude that

$$
M_{2,n,j} \leq F\ \frac{2^{2j}(1+6n)}{n^4\sigma_n^2(j)}.
$$

□

Lemma 5. *There exists a constant $G = G(K, M, R, L, \|\phi\|_\infty)$ such that, for any $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and any level j:*

$$
M_{3,n,j} \leq M_{2,n,j} + G\ \frac{2^j}{n^3\sigma_n^2(j)}\left(2^j + \frac{1}{n\sigma_n^2(j)}\right).
$$

Proof. Let us consider $(\overrightarrow{p}, \overrightarrow{q}) \in \Theta_0(R)$ and a level j.

$$
\begin{aligned}
M_{3,n,j} &= \mathbb{E}_{\overrightarrow{p},\overrightarrow{q}}\left[\left(\sum_{i_1=2}^{n}\left(\sum_{i_2=1}^{i_1-1}h_j(i_1,i_2)\right)^2-1\right)^2\right]\\
&= \mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left[\left(\sum_{i_1=2}^{n}\left(\sum_{i_2=1}^{i_1-1}h_j(i_1,i_2)\right)^2\right)^2\right] - 2\mathbb{V}ar_{\overrightarrow{p},\overrightarrow{p}}(\tilde{T}_j)+1\\
&= \mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left[\left(\sum_{i_1=2}^{n}\left(\sum_{i_2=1}^{i_1-1}h_j(i_1,i_2)\right)^2\right)^2\right] - 1\\
&:= M_{2,n,j} + 2A_{3,n,j} - 1,
\end{aligned}
$$

with $A_{3,n,j} = \sum_{i_1=3}^{n}\sum_{i_3=2}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(\left(\sum_{i_2=1}^{i_1-1} h_j(i_1,i_2)\right)^2 \left(\sum_{i_2=1}^{i_3-1} h_j(i_2,i_3)\right)^2 \right)$.

We focus on the term $A_{3,n,j}$.

$$
\begin{aligned}
A_{3,n,j} = & \sum_{i_1=3}^{n}\sum_{i_3=2}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(\left(\sum_{i_2=1}^{i_1-1} h_j(i_1,i_2)\right)^2 \left(\sum_{i_2=1}^{i_3-1} h_j(i_2,i_3)\right)^2 \right) \\
= & \sum_{i_1=3}^{n}\sum_{i_3=2}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(\left(\sum_{i_2=1}^{i_1-1} h_j^2(i_1,i_2)\right) \left(\sum_{i_2=1}^{i_3-1} h_j^2(i_2,i_3)\right) \right) \\
& + 2\sum_{i_1=3}^{n}\sum_{i_2=2}^{i_1-1}\sum_{i_3=1}^{i_2-1} \mathbb{E}_{\vec{p},\vec{p}} \left(h_j(i_1,i_2)h_j(i_1,i_3)h_j^2(i_2,i_3) \right) \\
& + 4\sum_{i_1=4}^{n}\sum_{i_3=3}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(\left(\sum_{i_2=2}^{i_1-1}\sum_{i_4=1}^{i_2-1} h_j(i_1,i_2)h_j(i_1,i_4)\right) \left(\sum_{i_2=2}^{i_3-1}\sum_{i_4=1}^{i_2-1} h_j(i_3,i_2)h_j(i_3,i_4)\right) \right).
\end{aligned}
$$

The first term of the right-hand side can be bounded as follows:

$$
\begin{aligned}
& \sum_{i_1=3}^{n}\sum_{i_3=2}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(\left(\sum_{i_2=1}^{i_1-1} h_j^2(i_1,i_2)\right) \left(\sum_{i_2=1}^{i_3-1} h_j^2(i_2,i_3)\right) \right) \\
\leq & \frac{1}{2} \left(\mathbb{E}_{\vec{p},\vec{p}} \left(\sum_{i_1=2}^{n}\sum_{i_2=1}^{i_1-1} h_j^2(i_1,i_2) \right) \right)^2 + \sum_{i_1=3}^{n}\sum_{i_2=1}^{i_1-1}\sum_{i_3=1}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(h_j^2(i_1,i_2)h_j^2(i_2,i_3) \right) \\
= & \frac{1}{2}\, \mathbb{V}ar_{\vec{p},\vec{p}}(\tilde{T}_j) + \sum_{i_1=3}^{n}\sum_{i_2=1}^{i_1-1}\sum_{i_3=1}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(h_j^2(i_1,i_2)h_j^2(i_2,i_3) \right) \\
= & \frac{1}{2} + \sum_{i_1=3}^{n}\sum_{i_2=1}^{i_1-1}\sum_{i_3=1}^{i_1-1} \mathbb{E}_{\vec{p},\vec{p}} \left(h_j^2(i_1,i_2)h_j^2(i_2,i_3) \right).
\end{aligned}
$$

Since

$$\begin{aligned}
&\sum_{i_1=3}^{n}\sum_{i_2=1}^{i_1-1}\sum_{i_3=1}^{i_1-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(h_j^2(i_1,i_2)h_j^2(i_2,i_3)\right)\\
\leq\ & \frac{2^{10}M^2(K^{-M}(M-1)!)^4}{n^3\sigma_n^2(j)}L^2 2^{2j}\|\phi\|_\infty^4\sum_{i_1=3}^{n}\sum_{i_2=1}^{i_1-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(h_j^2(i_1,i_2)\right)\\
\leq\ & 2^{10}K^{-4M}L^2M^2((M-1)!)^4\|\phi\|_\infty^4\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}\ \ \mathbb{V}ar_{\overrightarrow{p},\overrightarrow{p}}(\tilde{T}_j)\\
=\ & F\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)},
\end{aligned}$$

we deduce that a bound for the first term is

$$\sum_{i_1=3}^{n}\sum_{i_3=2}^{i_1-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(\left(\sum_{i_2=1}^{i_1-1}h_j^2(i_1,i_2)\right)\left(\sum_{i_2=1}^{i_3-1}h_j^2(i_2,i_3)\right)\right)\leq\frac{1}{2}+F\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}.$$

Let us now focus on the second term. From similar calculus as before,

$$\begin{aligned}
&2\sum_{i_1=3}^{n}\sum_{i_2=2}^{i_1-1}\sum_{i_3=1}^{i_2-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(h_j(i_1,i_2)h_j(i_1,i_3)h_j^2(i_2,i_3)\right)\\
\leq\ & \frac{8M^2(K^{-M}(M-1)!)^4}{n^3\sigma_n^2(j)}(4L)^2\left(2^{\frac{j}{2}+1}\|\phi\|_\infty\right)^4\sum_{i_1=2}^{n}\sum_{i_2=1}^{i_1-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(h_j^2(i_1,i_2)\right)\\
\leq\ & 2^{11}K^{-4M}M^2L^2((M-1)!)^4\|\phi\|_\infty^4\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}\ \ \mathbb{V}ar_{\overrightarrow{p},\overrightarrow{p}}(\tilde{T}_j)\\
=\ & 2F\ \frac{2^{2j}}{n^3\ \sigma_n^2(j)}.
\end{aligned}$$

We deal with the third term and we use Lemma 3. We have:

$$\begin{aligned}
&4\sum_{i_1=4}^{n}\sum_{i_3=3}^{i_1-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(\left(\sum_{i_2=2}^{i_1-1}\sum_{i_4=1}^{i_2-1}h_j(i_1,i_2)h_j(i_1,i_4)\right)\left(\sum_{i_2=2}^{i_3-1}\sum_{i_4=1}^{i_2-1}h_j(i_3,i_2)h_j(i_3,i_4)\right)\right)\\
=\ & 4\sum_{i_1=4}^{n}\sum_{i_3=3}^{i_1-1}\sum_{i_2=2}^{i_3-1}\sum_{i_4=1}^{i_2-1}\mathbb{E}_{\overrightarrow{p},\overrightarrow{p}}\left(h_j(i_1,i_2)h_j(i_1,i_4)h_j(i_3,i_2)h_j(i_3,i_4)\right)\\
\leq\ & 4D\ \frac{2^j}{n^4\sigma_n^4(j)}.
\end{aligned}$$

Finally, gathering all the bounds obtained, we conclude that

$$M_{3,n,j} \leq M_{2,n,j} + G\ \frac{2^j}{n^3\sigma_n^2(j)}\left(2^j + \frac{1}{nv_n^2(j)}\right),$$

with $G := \max(6F, 8D)$. □

Lemma 6. *For $s' < s$ in $\mathcal{S}_n$, i.e. $j_{s'} > j_s$, the following bound holds uniformly*

$$\mathbb{E}\left(\frac{d\mathbb{P}_{\pi_{j_s}}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}}\frac{d\mathbb{P}_{\pi_{j_{s'}}}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}}\right) - 1 \leq 2\left(C^2 2^{1+3s^\star} M\right)^2 l_n^{-2}\exp\left(-\frac{4\ln(nl_n)\ln(\ln(n))}{(4s^\star+1)^2\ln(n)}\right).$$

Before proving this lemma, let us remark that this upper bound goes to zero when n goes to infinity.

Proof. For convenience, for $j \in \{j_s, j_{s'}\}$ we consider $r_{n,j}$ as defined in (4.5) and we denote

$$\tilde{r}_{n,j}(u) = r_{n,j} f_u(Z_u)^{-1}, \quad \theta(\omega; u) = \sum_{l=1}^{M} \omega_l(u)\,\theta_l,$$

$$\frac{d\mathbb{P}_{\pi_j}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}} = \prod_{u=1}^{n}\left[1 + \theta(\omega;u)\tilde{r}_{n,j}(u)\sum_{k\in\mathcal{T}_j}\zeta_{jk}\psi_{jk}(Z_u)\right].$$

To avoid cumbersome notations, we denote $j := j_s$ and $j' := j_{s'}$ in the sequel. We evaluate the quantity

$$\mathbb{E}\left(\frac{d\mathbb{P}_{\pi_j}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}}\frac{d\mathbb{P}_{\pi_{j'}}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}} - 1\right).$$

The products can be expanded and the compactness of Haar wavelets leads to

$$\begin{aligned}\mathbb{E}\left(\frac{d\mathbb{P}_{\pi_j}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}}\frac{d\mathbb{P}_{\pi_{j'}}}{d\mathbb{P}_{\overrightarrow{p},\overrightarrow{p}}} - 1\right) \quad &= \quad \sum_{a=1}^{n}\sum_{1\leq u_1<\ldots<u_a\leq n}\theta(\omega;u_1)^2\ldots\theta(\omega;u_a)^2 \\ &\times\sum_{k_1,\ldots,k_a\in\mathcal{T}_{j'}}\mathbb{E}\left(\prod_{b=1}^{a}\tilde{r}_{n,j'}(u_b)\zeta_{j'k_b}\psi_{j'k_b}(Z_{u_b})\,\tilde{r}_{n,j}(u_b)\zeta_{j\lfloor 2^{j-j'}k_b\rfloor}\psi_{j\lfloor 2^{j-j'}k_b\rfloor}(Z_{u_b})\right).\end{aligned}$$

A careful study of all the summands shows that many of them are equal to zero.

First, the number a has to be even. If it is odd, the expectation term with respect to the random variables $\zeta_{jk}, \zeta_{j'k}$ will be zero. Moreover in order to get a non-zero summand we need a condition on $k_1, \dots, k_a$. They have to be paired. Here we mean that the number of k_b being equal is also an even number. The expectation term with respect to the random variables $\zeta_{jk}, \zeta_{j'k'}$ $(k \in \mathcal{T}_j, k' \in \mathcal{T}_{j'})$ is zero if this condition is not fulfilled, otherwise the product $\prod_{b=1}^{a} \zeta_{j\lfloor 2^{j-j'}k_b \rfloor}\zeta_{j'k_b}$ is equal to one.

These two remarks imply that the sum over index a starts at $a = 2$ and the number of products $\tilde{r}_{n,j'}(\cdot)\tilde{r}_{n,j}(\cdot)$ is at least equal to two. We have

$$\begin{aligned} \left| \mathbb{E}\left(\frac{\psi_{j\lfloor 2^{j-j'}k \rfloor}(Z_u)\,\psi_{j'k}(Z_u)}{f_u^2(Z_u)} \right) \right| &= \left| \int_{\mathbb{R}} \frac{\psi_{j\lfloor 2^{j-j'}k \rfloor}\ \psi_{j'k}}{f_u} \right| \\ &\leq 2^{\frac{j-j'}{2}+1}. \end{aligned}$$

We go back to the sum over index a. First note that the number of terms in the sum over $u_1, \dots, u_a$ is at most $n^a/a!$ and the one involved in the sum over $k_1, \dots, k_a$ is at most $|\mathcal{T}_{j'}|^{\frac{a}{2}}\, a!/\left((a/2)!\ 2^{\frac{a}{2}}\right)$. Thus, gathering all the results, we obtain for any n large enough:

$$\begin{aligned} &\Big| \sum_{a=1}^{n} \sum_{1\leq u_1<u_2<\ldots<u_a\leq n} \big[\theta(\omega;u_1)^2 \ldots \theta(\omega;u_a)^2 \\ &\times \sum_{k_1,\dots,k_a\in\mathcal{T}_{j'}} \mathbb{E}\left(\prod_{b=1}^{a} \tilde{r}_{n,j'}(u_b)\zeta_{j'k_b}\psi_{j'k_b}(Z_{u_b})\,\tilde{r}_{n,j}(u_b)\zeta_{j(2^{i-j}k_b)}\psi_{j\lfloor 2^{j-j'}k_b\rfloor}(Z_{u_b}) \right) \big] \Big| \\ &\leq \sum_{a=2}^{n} \left(C^2 2^s 2^{s'} M \right)^a \frac{n^a\, |\mathcal{T}_{j'}|^{\frac{a}{2}}}{\left(\frac{a}{2}\right)! 2^{\frac{a}{2}}} \left(2^{-js-\frac{j}{2}} \right)^a \left(2^{-j's'-\frac{j'}{2}} \right)^a \left(2^{1+\frac{j-j'}{2}} \right)^a \\ &\leq \sum_{a=2}^{n} \left(C^2 2^{1+2s^\star} M \right)^a n^a \left(2^{j's'-js} 2^{-j'(2s'+\frac{1}{2})} \right)^a \\ &\leq \sum_{a=2}^{n} \left(C^2 2^{1+3s^\star} M \right)^a l_n^{-a} (nl_n)^{a\left(\frac{2s'}{4s'+1}-\frac{2s}{4s+1}\right)} \\ &\leq 2 \left(C^2 2^{1+3s^\star} M \right)^2 l_n^{-2} (nl_n)^{2\left(\frac{2s'}{4s'+1}-\frac{2s}{4s+1}\right)} \\ &\leq 2 \left(C^2 2^{1+3s^\star} M \right)^2 l_n^{-2} \exp\left(-\frac{4\ln(nl_n)}{(4s^\star+1)^2 \ln(n)}\ \ln(\ln(n)) \right). \end{aligned}$$

The last inequality comes from the fact that the minimal distance between s' and s in $\mathcal{S}_n$ is such that: $s' - s < -\ln(\ln(n))/\ln(n)$. Since the right-hand side of the last inequality goes to zero when n tends to infinity, the lemma is now proved. □

References

[1] Autin, F., and Pouet, C. (2011). Test on components of densities mixture. *Stat. & Risk Modeling*, **28**, 389-410.

[2] Autin, F., and Pouet, C. (2012). Testing means from sampling populations with undefined labels. Available on arXiv.org. Article id.:1205.2679.

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

[4] Catalona, W.J., Smith, D.S., Ratliff, T.L., Dodds, K.M., Coplen, D.E., Yuan, J.J., Petros, J.A., and Andriole, G.L. (1991). Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N. Engl. J. Med.,* **324**, 1156-1161.

[5] Daubechies, I. (1992). *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

[6] Devroye, L., and Lugosi, G. (2000). *Combinatorial Methods in Density Estimation*, Springer, New York.

[7] Hall, P., and Heyde, C.C. (1980). *Martingale Limit Theory and its Application*, Probability and Mathematical Statistics. Academic Press, New York.

[8] Härdle, W., Kerkyacharian, G., Picard, D., and Tsybakov, A.B. (1998). *Wavelets, Approximation and Statistical Applications*, Springer Verlag, New York.

[9] Ingster, Y., and Suslina, I. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*, Springer-Verlag, New York.

[10] McLachlan, G.J., and Peel, P. (2000). *Finite Mixture Models*, Wiley, New York.

[11] Maiboroda, R.E. (1996). Estimates for distributions of components of mixtures with varying concentrations. *Ukr. Math. Journ.*, **48**, 618-622.

[12] Pokhyl'ko, D. (2005). Wavelet estimators of a density constructed from observations of a mixture. *Theor. Prob. and Math. Statist.*, **70**, 135-145.

[13] Spokoiny, V.G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist.*, **24**, 2477-2498.