



ELSEVIER

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Minimax rates over Besov spaces in ill-conditioned mixture-models with varying mixing-weights

Florent Autin^{a,*}, Christophe Pouet^{b,**}

^a Aix-Marseille Université, 39, rue F. Joliot Curie, 13453 Marseille Cedex 13, France

^b Ecole Centrale de Marseille, 38, rue F. Joliot Curie, 13013 Marseille, France

ARTICLE INFO

Article history:

Received 27 June 2013

Received in revised form

17 August 2013

Accepted 13 September 2013

Keywords:

Besov spaces

Minimax theory

Mixture-models

Nonparametric tests

Wavelet decomposition

ABSTRACT

We consider ill-conditioned mixture-models with varying mixing-weights. We study the classical homogeneity testing problem in the minimax setup and try to push the model to its limits, that is to say to let the mixture model to be really ill-conditioned. We highlight the strong connection between the mixing-weights and the expected rate of testing. This link is characterized by the behavior of the smallest eigenvalue of a particular matrix computed from the varying mixing-weights. We provide optimal testing procedures and we exhibit a wide range of rates that are the minimax and minimax adaptive rates for Besov balls.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

For many years, the mixture-models have gained a lot of interests in statistics, particularly because of their large fields of application as in finance, economy, biology, astronomy among many others. Most of the theoretical results found in the literature deal with the problem of estimation of the mixing-weights as in the work of Hall (1981), Titterton (1983), Hall and Titterton (1984) and Qin (1999), or of the mixing-components as in Maiboroda (1996, 2000a), Pokhyl'ko (2005) and Lodatko and Maiboroda (2007). Facing with the problem of homogeneity tests for mixture-models, theoretical results are fewer and more recent. For the case of fixed mixing-weights we refer the reader to the papers of Garel (2001, 2005) and Delmas (2003).

Here, we are interested in studying mixture-models with varying mixing-weights. Such models are popular nowadays especially since the works published by Maiboroda (1996, 2000a) and Pokhyl'ko (2005) that deal with the estimation of the mixing-components, and since the work of Maiboroda (2000b) that deals with testing problems on the mixing-components.

Many applications of the varying mixing-weights model have been proposed. The first example is taken from Lodatko and Maiboroda (2007). It deals with the difference between atypical pneumonia patients and healthy patients. Social science data were also considered. Autin and Pouet (2011) modeled the annual working time for technical executive staff and commercial and administrative executive staff in two French regions. The difference in travel time between the State of New York and the State of California for several means of transportation was investigated in Autin and Pouet (2013). The varying mixing-weights model was also successfully applied to medical data by Maiboroda and Sugakova (2012). They

* Corresponding author. Tel.: +33 4 13 55 10 15; fax: +33 4 13 55 14 02.

** Corresponding author. Tel.: +33 4 91 05 45 41; fax: +33 4 13 55 14 02.

E-mail addresses: florent.autin@univ-amu.fr (F. Autin), cpouet@centrale-marseille.fr (C. Pouet).

studied DNA microarray data collected on tumors. At last, successful applications to simulated and real data have also been done. Parametric hypotheses testing was conducted by [Autin and Pouet \(2013\)](#) for simulated data.

As in [Autin and Pouet \(2011, 2012\)](#), we propose here to study a testing problem for mixture-models with varying mixing-weights through the minimax adaptive approach as introduced by [Spokoiny \(1996\)](#). From two independent samples of n independent random variables ($n > 2$), we are interested in testing whether these two samples can be considered as samples with random variables built from the same mixture of M ($M \in \mathbb{N}^*$) components or not. In particular, we aim at providing the minimax rates of testing over the Besov spaces.

In [Autin and Pouet \(2011, 2012\)](#), the smallest eigenvalues of the Gram matrix of the mixing-weights vectors were fixed and larger than or equal to $k \in (0, 1)$; it was proved that

- the minimax rate of testing over the balls of the Besov space $\mathcal{B}_{2,\infty}^s$ (as defined later in (3)) was of order $n^{-2s/(1+4s)}$,
- the minimax adaptive rate of testing over such spaces suffers from an unavoidable loss of order $(\ln \ln n)^{-1/2}$ compared to the minimax one.

Hence, the minimax rate of testing in this case is of the same order as the one in the case $M = 1$ proved by [Butucea and Tribouley \(2006\)](#). Nevertheless, [Autin and Pouet \(2011\)](#) noted that the smaller the parameter k the worse the expected exact separation constant. Although this result was a surprise, it is logical as the smallest eigenvalue k is directly connected with the matrix of the mixing-weights, that is to say with the design of the statistical experiment. We also recall that if the mixing-weights are the same for all the observations, we are in the usual mixture-model (see [McLachlan and Peel, 2000](#)) and the homogeneity test with two samples is meaningless. Therefore here we address the natural question that arises from these previous works: what are the minimax and minimax adaptive rates when the smallest eigenvalue of the Gram matrix of the mixing-weights vectors, namely k_n , is decreasing to zero as n is increasing? More than that, we exhibit for the first time a wide range of new rates that correspond to minimax and minimax adaptive rates over the Besov spaces in statistical models lying on indirect observations. Although there are some similarities between our underlying models and the well-known ill-posed inverse problems, here the deterioration of speed due to the ill-conditioning aspect is not located in the power anymore.

After introducing the setting in [Section 2](#), we prove in [Section 3](#) that the faster the decrease, the worse the minimax rate of testing over the Besov spaces. More precisely we prove that

- the minimax rate of testing over the balls of the Besov space $\mathcal{B}_{2,\infty}^s$ is of order $(nk_n)^{-2s/(1+4s)}$,
- the minimax adaptive rate of testing over such spaces suffers from an unavoidable loss of order $(\ln \ln n)^{-1/2}$ compared to the minimax one, independently of the sequence $(k_n)_{n \in \mathbb{N}}$ that tends to zero not too fast.

In addition to the link between the mixing-weights and the expected rates, we provide both a minimax (see [Section 3.1](#)) and a minimax adaptive testing procedure (see [Section 3.2](#)) that are based on the comparison between a test statistic inspired from the ones used in [Autin and Pouet \(2011, 2012\)](#) and a well-chosen threshold value. The proofs of the results are given in [Appendix A](#).

2. Setting of the study

2.1. Mixture-model with varying mixing-weights

For any integer $n > 2$, let us consider two independent samples $Y = (Y_1, \dots, Y_n)$ and $Z = (Z_1, \dots, Z_n)$ of independent random variables with unknown marginal densities. For each sample, assume that the marginal densities are mixtures of $M \in \mathbb{N}^*$ common densities: the mixing-components. We denote respectively by

$$f_i(\cdot) = \sum_{l=1}^M \omega_l(i) p_l(\cdot) \quad \text{and} \quad g_i(\cdot) = \sum_{l=1}^M \omega_l(i) q_l(\cdot)$$

the marginal densities of Y_i and Z_i ($i \in \{1, \dots, n\}$) and by Γ_n the Gram matrix of the mixing-weights which is the symmetric matrix of order M with $(1/n) \sum_{i=1}^n \omega_l(i) \omega_u(i)$ as element at line l and column u .

Here, the mixing-components p_u and q_u ($1 \leq u \leq M$) are unknown but the mixing-weights $(\omega_l(i), 1 \leq l \leq M, 1 \leq i \leq n)$ are supposed to be known and satisfy

$$\begin{aligned} \forall (l, i) \in \{1, \dots, M\} \times \{1, \dots, n\}, \quad \omega_l(i) &\geq 0, \\ \forall i \in \{1, \dots, n\}, \quad \sum_{l=1}^M \omega_l(i) &= 1. \end{aligned}$$

We consider that the mixture-model is the one with varying mixing-weights corresponding to the case of an invertible matrix Γ_n . We denote by k_n the smallest eigenvalue of the matrix Γ_n . In addition, we consider [Assumption 2.1](#) on our mixture-model.

Assumption 2.1. The sequence of positive real numbers $(k_n)_{n > 2}$ is such that

1. $\lim_{n \rightarrow +\infty} k_n = 0$,
2. $\lim_{n \rightarrow +\infty} nk_n = +\infty$.

2.2. Hypotheses testing problem

We aim at studying whether the two samples of random variables come from the same mixing-components or not. In the sequel,

- $\vec{p} := (p_1, \dots, p_M)$ and $\vec{q} := (q_1, \dots, q_M)$ will characterize the mixing-components of the two samples Y and Z ,
- $\mathbb{P}_{\vec{p}, \vec{q}}$ will denote the distribution of (Y, Z) ,
- $\mathbb{E}_{\vec{p}, \vec{q}}(\cdot)$ will denote the expected value and $\text{Var}_{\vec{p}, \vec{q}}(\cdot)$ the variance under this distribution.

For any $R > 0$, let $\mathcal{D}(R)$ be the set of all probability densities such that their \mathbb{L}_2 -norm and their \mathbb{L}_∞ -norm are bounded by R . For $s > 0$, let $\mathcal{B}_{2,\infty}^s(R)$ be the ball of the Besov space $\mathcal{B}_{2,\infty}^s$ as defined in (3). We consider two subspaces containing vectors of $(\mathcal{D}(R))^{2M}$, namely $\Theta_0(R)$ and $\Theta_1(R, C, r_n, s)$, which are respectively defined by

$$\begin{aligned} \Theta_0(R) &= \{(\vec{p}, \vec{q}) : \forall l \in \{1, \dots, M\}, p_l = q_l\}, \\ \Theta_1(R, C, r_n, s) &= \{(\vec{p}, \vec{q}) : \forall l \in \{1, \dots, M\}, \forall i := p_l - q_l \in \mathcal{B}_{2,\infty}^s(R), \\ &\exists u \in \{1, \dots, M\}, (p_u, q_u) \in \Lambda(C, r_n, s)\}, \end{aligned}$$

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

We are interested in the two kinds of hypotheses testing problems presented below.

Non-adaptive case (s is known and $s > \frac{1}{4}$)

$$\begin{aligned} \mathcal{H}_0 &: (\vec{p}, \vec{q}) \in \Theta_0(R), \\ \mathcal{H}_1 &: (\vec{p}, \vec{q}) \in \Theta_1(R, C, r_n, s). \end{aligned} \tag{1}$$

Adaptive case (s belongs to $[s_*, s^*]$ with $s_* > \frac{1}{4}$ but s is unknown)

$$\begin{aligned} \mathcal{H}_0 &: (\vec{p}, \vec{q}) \in \Theta_0(R), \\ \mathcal{H}_1 &: (\vec{p}, \vec{q}) \in \Theta_1^*(R, C, r_n l_n) = \bigcup_{s \in [s_*, s^*]} \Theta_1(R, C, r_n l_n, s), \end{aligned} \tag{2}$$

where l_n (resp. $r_n l_n$) is a sequence of positive numbers that goes to zero (resp. infinity) when n tends to infinity.

In our study we shall consider the minimax approach and the minimax adaptive approach which are often used to evaluate the performances of testing procedures.

We recall that the minimax (resp. minimax adaptive) approach aims at providing testing procedures that achieve the *minimax rate* $r_n^{-2s/(1+4s)}$ (resp. *minimax adaptive rate* $(r_n l_n)^{-2s/(4s+1)}$) that, in context (1) (resp. in context (2)), corresponds to the best possible rate separating at least one of the M pairs of mixing-components p_u and q_u . We refer the interested reader to [Spokoiny \(1996\)](#) for a precise presentation of the minimax and minimax adaptive approaches in testing problems. Nevertheless we recall the two steps necessary to prove that $r_n^{-2s/(1+4s)}$ is the minimax rate (resp. $r_n l_n^{-2s/(1+4s)}$ is the minimax adaptive rate) for the testing problem (1) (resp. for the testing problem (2)).

First step: [Upper bound]. For any $\gamma \in]0, 1[$ there exist $C_\gamma > 0$ and a testing procedure Δ_s built from the samples such that, for any $C \geq C_\gamma$,

$$\lim_{n \rightarrow \infty} \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_s = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1(R, C, r_n, s)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_s = 0) \right) \leq \gamma$$

(resp. there exist $C_* > 0$ and a testing procedure Δ_* built from the samples such that, for any $C > C_*$,

$$\lim_{n \rightarrow \infty} \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_* = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1^*(R, C, r_n l_n)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_* = 0) \right) = 0.)$$

Second step: [Lower bound]. There exists $c_\gamma > 0$ such that, for any $C < c_\gamma$,

$$\liminf_{n \rightarrow \infty} \inf_{\Delta} \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1(R, C, r_n, s)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta = 0) \right) > \gamma$$

(resp. there exists $c_* > 0$ such that, for any $C < c_*$,

$$\liminf_{n \rightarrow \infty} \inf_{\Delta} \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1^*(R, C, r_n l_n)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta = 0) \right) = 1).$$

where the infimum is taken over all testing procedures.

In that context, we shall say that $r_n^{-2s/(1+4s)}$ (resp. $(r_n l_n)^{-2s/(1+4s)}$) is the minimax rate of testing (resp. the minimax adaptive rate of testing) over the Besov ball $\mathcal{B}_{2,\infty}^s(R)$ (as defined in (3)) of the testing problem (1) (resp. (2)) and that Δ_s (resp. Δ_*) is minimax (resp. minimax adaptive).

2.3. Wavelet setting

Wavelets bases offer the advantage of being localized both in the frequency and time domains. Therefore they are often used in many mathematical fields such as approximation theory, signal analysis and statistics.

Let us recall that such bases are built from functions ϕ_{jk} and $\psi_{jk}(j \in \mathbb{N}, k \in \mathbb{Z})$ that are dilatations and translations of a chosen scaling function ϕ and a chosen wavelet function ψ . For any $j \in \mathbb{N}$, any h in \mathbb{L}_2 can be decomposed in the wavelet basis built from (ϕ, ψ) as follows:

$$h(\cdot) = \sum_{k \in \mathbb{Z}} \alpha_{jk} \phi_{jk}(\cdot) + \sum_{j' \geq j} \sum_{k \in \mathbb{Z}} \beta_{j'k} \psi_{j'k}(\cdot),$$

where $\alpha_{jk} = \int h(t) \phi_{jk}(t) dt$ and $\beta_{j'k} = \int h(t) \psi_{j'k}(t) dt$.

For the sake of simplicity, we shall consider the Haar wavelet basis in the sequel that is built from the following scaling and wavelet functions:

$$\phi(\cdot) = \mathbf{1}_{[0,1]}(\cdot), \quad \psi(\cdot) = \mathbf{1}_{[0,\frac{1}{2}]}(\cdot) - \mathbf{1}_{[\frac{1}{2},1]}(\cdot).$$

Nevertheless our results could be easily generalized for any choice of compactly supported wavelet basis.

To derive optimal results, we have considered that any function $p_u - q_u (1 \leq u \leq M)$ is regular enough, it means that the energy of their wavelet coefficients is decreasing fast enough. More precisely, any function $p_u - q_u$ is supposed to belong to the ball of a Besov space $\mathcal{B}_{2,\infty}^s(R)$ defined as follows:

$$\mathcal{B}_{2,\infty}^s(R) = \left\{ h \in \mathbb{L}_2; \sup_{j \in \mathbb{N}} 2^{2js} \sum_{j' \geq j} \sum_{k \in \mathbb{Z}} \beta_{j'k}^2 \leq R^2 \right\}. \tag{3}$$

3. Optimal testing procedures

In this section we provide procedures that are based on the test statistics T_j defined for any $j \in \mathbb{N}$ by

$$T_j = \frac{1}{n^2} \sum_{i_1 \neq i_2}^n \sum_{k \in \mathbb{Z}} \sum_{l=1}^M a_l(i_1) a_l(i_2) [\phi_{jk}(Y_{i_1}) - \phi_{jk}(Z_{i_1})][\phi_{jk}(Y_{i_2}) - \phi_{jk}(Z_{i_2})], \tag{4}$$

where, for all $(l, u) \in \{1, \dots, M\}^2$, $(1/n) \sum_{i=1}^n a_l(i) \omega_u(i) = \delta_{lu}$ and $\langle a_l, a_l \rangle_2$ is minimal.

We recall the explicit formulas of $a_l(i)$ given in Maiboroda (2000a):

$$a_l(i) = \frac{1}{\det(\Gamma_n)} \sum_{u=1}^M (-1)^{l+u} \gamma_{lu} \omega_u(i), \tag{5}$$

where γ_{lu} is the minor (l, u) of the matrix Γ_n . Both motivations for the choice of these test statistics and similarities between varying mixing-weights models and inverse problems are provided in Autin and Pouet (2011, 2012). In particular in those previous works, the test statistics T_j is shown to be a natural estimator of $\sum_{l=1}^M \|p_l - q_l\|_2^2$, provided that j is large enough. Following Autin and Pouet (2012) we recall that according to (5),

$$\sup_{1 \leq i \leq n} |a_l(i)| \leq \frac{(M-1)!}{k_n^M}. \tag{6}$$

3.1. Non-adaptive case

We assume that the smoothness parameter s for the differences $p_u - q_u (1 \leq u \leq M)$ is known. Hence we are interested in the hypotheses testing problem (1).

3.1.1. Testing procedure $\Delta_{s,t}$

For any $s > \frac{1}{4}$ let $j_{n,s}$ be the smallest integer such that $2^{-j_{n,s}} \leq (nk_n)^{-2/(1+4s)}$. For any positive real number t , we denote by $\Delta_{s,t}$ the testing procedure defined as follows:

$$\Delta_{s,t} = \begin{cases} 1 & \text{if } (nk_n)^{4s/(1+4s)} T_{j_{n,s}} > t \\ 0 & \text{otherwise} \end{cases}$$

3.1.2. Optimality of $\Delta_{s,t}$ and minimax rate

Theorem 3.1 (Upper bound). Let $\gamma \in]0, 1[$. There exist $C_\gamma = C_\gamma(M, R, \gamma)$ and $t = t(M, R, \gamma)$ such that, for any $C \geq C_\gamma$,

$$\lim_{n \rightarrow \infty} \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta_{s,t} = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1(R, C, nk_n, s)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta_{s,t} = 0) \right) \leq \gamma.$$

Theorem 3.2 (Lower bound). Let $\gamma \in]0, 1[$. There exists $c_\gamma = c_\gamma(M, R, \gamma, s)$ such that, for any $C < c_\gamma$,

$$\liminf_{n \rightarrow \infty} \Delta \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1(R, C, nk_n, s)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta = 0) \right) > \gamma.$$

From Theorems 3.1 and 3.2, we deduce the following corollary:

Corollary 3.1. For any $s > \frac{1}{4}$, the testing procedure $\Delta_{s,t}$ is minimax. Moreover, the minimax rate of testing over the Besov ball $\mathcal{B}_{2,\infty}^s(R)$ is $(nk_n)^{-(2s/(1+4s))}$.

We note that the smaller the k_n , the worse the minimax rate of testing. Surprisingly the minimax rate does not only depend on the size of the sample and the regularity of the underlying functions but also on the sequence $(k_n)_{n > 2}$.

3.2. Adaptive case

Here, we assume that the parameter s of regularity is not known anymore. Nevertheless, we suppose that it belongs to the interval $[s_*, s^*]$ with $s_* > \frac{1}{4}$. Hence we are now interested in the hypotheses testing problem (2).

Assumption 3.1. The sequence of positive real numbers $(k_n)_{n > 2}$ is such that

$$\ln\left(\frac{1}{k_n}\right) = o((\ln n)^\delta) \quad \text{where } \delta \in]0, 1[.$$

Remark 3.1. This assumption will be sufficient to derive the minimax adaptive results. It deals with mixture-models for which the rate of decrease of the sequence of parameters $(k_n)_n$ is not polynomially fast but faster than any logarithmic rate.

3.2.1. Testing procedure $\Delta_{*,t}$

For $s \in [s_*, s^*]$, let $j_{n,s}^*$ be the smallest integer such that $2^{-j_{n,s}^*} \leq (nk_n(\sqrt{\ln \ln n})^{-1})^{-2/(1+4s)}$. For any positive real number t , we denote by $\Delta_{*,t}$ the testing procedure defined as follows:

$$\Delta_{*,t} = \begin{cases} 1 & \text{if } \max_{s \in [s_*, s^*]} (nk_n l_n)^{4s/(1+4s)} T_{j_{n,s}^*}^* > t \\ 0 & \text{otherwise} \end{cases}$$

where $(l_n)_{n > 2}$ is the sequence of positive real numbers defined, for any $n > 2$, by $l_n = (\sqrt{\ln \ln n})^{-1}$.

3.2.2. Optimality of $\Delta_{*,t}$ and minimax adaptive rate

Theorem 3.3 (Upper bound). There exist $C_* = C_*(M, R)$ and $t = t(M, R)$ such that, for any $C > C_*$,

$$\lim_{n \rightarrow \infty} \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta_{*,t} = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1^*(R, C, nk_n l_n)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta_{*,t} = 0) \right) = 0.$$

Theorem 3.4 (Lower bound). There exists $c_* = c_*(M, R, \delta, s^*)$ such that for any $C < c_*$,

$$\liminf_{n \rightarrow \infty} \Delta \left(\sup_{(\vec{p}, \vec{q}) \in \Theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta = 1) + \sup_{(\vec{p}, \vec{q}) \in \Theta_1^*(R, C, nk_n l_n)} \mathbb{P}_{\vec{p}, \vec{q}} \rightarrow \rightarrow (\Delta = 0) \right) = 1.$$

From Theorems 3.3 and 3.4, we deduce the following corollary:

Corollary 3.2. The testing procedure $\Delta_{*,t}$ is minimax adaptive. Moreover, for any $s > \frac{1}{4}$, the minimax adaptive rate of testing over the Besov ball $\mathcal{B}_{2,\infty}^s(R)$ is $(nk_n l_n)^{-2s/(1+4s)}$.

When we compare the adaptive case to the non-adaptive one, we deduce that there is a price to pay for adaptation and it is of order l_n . We also note that this cost does not depend on k_n .

4. Conclusion

This work is a contribution to nonparametric hypotheses testing problems from indirect observations through the varying mixing-weights. Optimal testing procedures have been proposed and new rates have been exhibited as the minimax and minimax adaptive rates for Besov spaces in the underlying testing models. We proved that the minimax and minimax adaptive rates strongly depend on the smallest eigenvalue k_n of the Gram matrix of the mixing-weights : the bigger the k_n the better the minimax and minimax adaptive rates.

At first glance, it could be surprising that the minimax and minimax adaptive rates differ from the ones in ill-posed inverse problems. We claim that the reason for this difference is the following: in the context of mixture-models the dimension of the matrix Γ_n involved in the determination of the rates is finite, whatever n , contrary to ill-posed inverse problem for which the matrix involved in the determination of the rates has infinite dimension.

Here the threshold values of our testing methods are calibrated to entail the theoretical result, that is to say the asymptotic distinguishability of the two hypotheses. In practical cases, the threshold values can be computed by Monte-Carlo methods. Other popular approaches can also be applied and are related to the computation of the p -value such as bootstrap tests or permutation tests (see Efron and Tibshirani, 1993).

Appendix A. Proofs of theorems

In this section, we prove the main theorems. The proofs follow some results previously given in Autin and Pouet (2011, 2012). In Proposition A.1 we recall the properties of the test statistic T_j given in Corollary 3.3 and Proposition 3.4 of Autin and Pouet (2011).

Proposition A.1. *Let $j \in \mathbb{N}$ and T_j be the test statistic defined in (4). There exists a positive constant C_R that depends on R such that, for any $n > 2$ and any $(\vec{p}, \vec{q}) \in \mathcal{D}(R)^{2M}$:*

$$\left| \mathbb{E}_{\vec{p}, \vec{q}} (T_j) - \frac{1}{n^2} \sum_{l=1}^M \sum_{k \in \mathbb{Z}} \sum_{i=1}^n \left(a_l(i) \int_{\mathbb{R}} (f_i - g_i) \phi_{jk} \right) \right| \leq \frac{8LMR^2}{nk_n}, \tag{7}$$

$$\mathbb{V}ar_{\vec{p}, \vec{q}} (T_j) \leq \frac{C_R M^2}{k_n^2} \left(\frac{2^j}{n^2} + \frac{1}{n} \sum_{l=1}^M \|p_l - q_l\|_2^2 + \sqrt{\frac{2^j}{n^3}} \sum_{l=1}^M \|p_l - q_l\|_2 \right). \tag{8}$$

Remark A.1. For any $j \in \mathbb{N}$ and any $(\vec{p}, \vec{q}) \in \Theta_0(R)$, the test statistic is centered and its variance is such that $\mathbb{V}ar_{\vec{p}, \vec{q}} (T_j) \leq C_R M^2 (nk_n)^{-2} 2^j$.

A.1. Proof of Theorem 3.1

Let us fix $0 < \gamma < 1$ and $s > 1/4$. By the Bienayme–Chebyshev inequality and following (8), for any $(\vec{p}, \vec{q}) \in \Theta_0(R)$:

$$\begin{aligned} \mathbb{P}_{\vec{p}, \vec{q}} (\Delta_{s,t} = 1) &= \mathbb{P}_{\vec{p}, \vec{q}} (T_{j_{n,s}} > t(nk_n)^{-4s/(1+4s)}) \\ &\leq t^{-2} (nk_n)^{8s/(1+4s)} \mathbb{V}ar_{\vec{p}, \vec{q}} (T_{j_{n,s}}) \\ &\leq C_R M^2 t^{-2} 2^{j_{n,s}} (nk_n)^{-2/(1+4s)}. \end{aligned}$$

According to the definition of the level $j_{n,s}$ and for the choice $t = 2M\sqrt{C_R/\gamma}$, it entails that

$$\mathbb{P}_{\vec{p}, \vec{q}} (\Delta_{s,t} = 1) \leq \frac{\gamma}{2}.$$

We now focus on the alternative and we put $t_{n,s} = t(nk_n)^{-4s/(1+4s)}$. The second type error can be rewritten as follows:

$$\mathbb{P}_{\vec{p}, \vec{q}} (\Delta_{s,t} = 0) = \mathbb{P}_{\vec{p}, \vec{q}} (-T_{j_{n,s}} + \mathbb{E}_{\vec{p}, \vec{q}} (T_{j_{n,s}}) \geq -t_{n,s} + \mathbb{E}_{\vec{p}, \vec{q}} (T_{j_{n,s}})).$$

According to (7), the wavelet expansion in the Besov ball $\mathcal{B}_{2,\infty}^s(R)$ leads to

$$\begin{aligned} \mathbb{E}_{\vec{p}, \vec{q}} (T_{j_{n,s}}) - t_{n,s} &= \sum_{l=1}^M \|p_l - q_l\|_2^2 - \sum_{l=1}^M \sum_{j \geq j_{n,s}} \sum_{k \in \mathbb{Z}} \left(\int_{\mathbb{R}} (p_l - q_l) \psi_{jk} \right)^2 \\ &\quad - \frac{1}{n^2} \sum_{l=1}^M \sum_{k \in \mathbb{Z}} \sum_{i=1}^n \left(a_l(i) \int_{\mathbb{R}} (f_i - g_i) \phi_{j_{n,s}k} \right)^2 - t_{n,s} \end{aligned}$$

$$\begin{aligned} &\geq \sum_{l=1}^M \|p_l - q_l\|_2^2 - MR^2 2^{-2s_{j_{n,s}}} - \frac{8LMR^2}{nk_n} - t_{n,s} \\ &\geq \frac{1}{2} \sum_{l=1}^M \|p_l - q_l\|_2^2 - MR^2 2^{-2s_{j_{n,s}}} - t_{n,s}, \end{aligned}$$

for any n large enough. As a consequence, applying the Bienayme–Chebyshev inequality and (8), one gets

$$\begin{aligned} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{s,t} = 0) &= \mathbb{P}_{\vec{p}, \vec{q}}\left(-T_{j_{n,s}} + \mathbb{E}_{\vec{p}, \vec{q}}(T_{j_{n,s}}) \geq -t_{n,s} + \mathbb{E}_{\vec{p}, \vec{q}}(T_{j_{n,s}})\right) \\ &\leq \frac{C_R M^2 (2^{j_{n,s}} + n \sum_{l=1}^M \|p_l - q_l\|_2^2 + \sqrt{2^{j_{n,s}} n} \sum_{l=1}^M \|p_l - q_l\|_2)}{n^2 k_n^2 \left(\frac{1}{2} \sum_{l=1}^M \|p_l - q_l\|_2^2 - MR^2 2^{-2s_{j_{n,s}}} - t_{n,s}\right)^2}. \end{aligned}$$

The choice of $j_{n,s}$ and the fact that the functions are in the alternative lead to the following upper bounds for any n large enough:

$$\begin{aligned} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{s,t} = 0) &\leq \frac{C_R M^2 (2^{j_{n,s}} + n \sum_{l=1}^M \|p_l - q_l\|_2^2 + \sqrt{2^{j_{n,s}} n} \sum_{l=1}^M \|p_l - q_l\|_2)}{n^2 k_n^2 (\frac{1}{2} \sum_{l=1}^M \|p_l - q_l\|_2^2 - MR^2 2^{-2s_{j_{n,s}}} - t_{n,s})} \\ &\leq \frac{C_R M^2 (2^{j_{n,s}} + n \sum_{l=1}^M \|p_l - q_l\|_2^2 + \sqrt{2^{j_{n,s}} n} \sum_{l=1}^M \|p_l - q_l\|_2)}{n^2 k_n^2 (2^{-1} - MR^2 C^{-2} - t C^{-2})^2 (\sum_{l=1}^M \|p_l - q_l\|_2^2)^2} \\ &\leq 3C_R \left(\frac{C^2}{2M} - R^2 - \frac{t}{M}\right)^{-2}. \end{aligned}$$

For all $C \geq C_{\gamma,t} = \sqrt{2M}(\sqrt{6\gamma^{-1}}C_R + R^2 + tM^{-1})^{1/2}$, we finally obtain

$$\mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{s,t} = 0) \leq \frac{\gamma}{2}.$$

The results on the first-type and second-type errors prove that if $C \geq C_{\gamma,t}$ the sum of the errors associated to the testing procedure $\Delta_{j_{n,s},t}$ is asymptotically less than γ . Therefore Theorem 3.1 is proved.

A.2. Proof of Theorem 3.2

For any $\vec{p} \in \mathcal{D}(R)^M$,

$$\inf_{\Delta} \left(\sup_{(\vec{p}', \vec{q}') \in \Theta_0(R)} \mathbb{P}_{\vec{p}', \vec{q}'}(\Delta = 1) + \sup_{(\vec{p}', \vec{q}') \in \Theta_1(R, C, nk_n, s)} \mathbb{P}_{\vec{p}', \vec{q}'}(\Delta = 0) \right) \geq 1 - \frac{1}{2} \|\mathbb{P}_{\vec{p}, \vec{p}} - \mathbb{P}_{\pi}\| \tag{9}$$

where $\|\cdot\|$ is the \mathbb{L}_1 -distance and π is an a priori probability measure concentrated on the set $\Lambda(C, nk_n, s)$.

Therefore, it suffices to prove that for judicious choices of \vec{p} and π we get

$$\|\mathbb{P}_{\vec{p}, \vec{p}} - \mathbb{P}_{\pi}\| < 2(1 - \gamma).$$

Let \vec{p} be such that for any $l \in \{1, \dots, M\}$,

$$\left\{x; p_l(x) > \frac{1}{2}\right\} \supset [0, 1].$$

Let $\theta = (\theta_1, \dots, \theta_M)$ denote an eigenvector associated with the smallest eigenvalue k_n of Γ_n such that $\|\theta\|_2 = 1$.

Let $\mathcal{T}_s = \{0, 1, \dots, 2^{j_{n,s}} - 1\}$ be the subset of \mathbb{Z} of cardinality $T = 2^{j_{n,s}}$. The following parametric family of functions is considered:

$$q_{l,\zeta}(\cdot) = p_l(\cdot) + 2^s C \sqrt{M} \theta_l \sum_{k \in \mathcal{T}_s} \zeta_k 2^{-s_{j_{n,s}} - j_{n,s}/2} \psi_{j_{n,s},k}(\cdot),$$

where, for any $k \in \mathcal{T}_s$, $\zeta_k \in \{-1, 1\}$.

Assume that the marginal densities of Y_i and $Z_i (1 \leq i \leq n)$ are respectively

$$f_i(\cdot) = \sum_{l=1}^M \omega_l(i) p_l(\cdot) \quad \text{and} \quad g_{i,\zeta}(\cdot) = \sum_{l=1}^M \omega_l(i) q_{l,\zeta}(\cdot).$$

The functions $q_{l,\zeta}$ are clearly densities that satisfy for large n :

$$\left\{x; q_{l,\zeta}(x) > \frac{1}{2}\right\} \supset [0, 1].$$

Moreover, if $C < RM^{-1/2} 2^{-s}$, then $q_{l,\zeta} - p_l$ belongs to the Besov ball $\mathcal{B}_{2,\infty}^s(\mathbb{R})$. We notice that for at least one $l \in \{1, \dots, M\}$, $M\theta_l^2 \geq 1$. Hence

$$\|q_{l,\zeta} - p_l\|_2^2 \geq C^2 (nk_n)^{-4s/(1+4s)}.$$

We consider the probability measure π such that the ζ_k 's are independent Rademacher random variables with parameter 1/2. Clearly,

$$\left\| \mathbb{P}_{\vec{p}, \vec{p}} - \mathbb{P}_\pi \right\| \leq \sqrt{\mathbb{E}_{\vec{p}, \vec{p}} \left[\left(\mathbb{E}_\pi \left(\prod_{i=1}^n \frac{g_{i,\zeta}(Z_i)}{f_i(Z_i)} \right) \right)^2 \right]} - 1. \tag{10}$$

We introduce the following random variables:

$$\tilde{Z}_{ik} = 2^s C \sqrt{M} 2^{-s j_{n,s} - j_{n,s}} / 2^{j_{n,s} k} \frac{g_{i,\zeta}(Z_i)}{f_i(Z_i)} \sum_{l=1}^M \theta_l \omega_l(i).$$

We only need to evaluate the second-order moment of the likelihood ratio:

$$\mathbb{E}_{\vec{p}, \vec{p}} \left[\left(\mathbb{E}_\pi \left(\prod_{i=1}^n \frac{g_{i,\zeta}(Z_i)}{f_i(Z_i)} \right) \right)^2 \right] = \mathbb{E}_{\vec{p}, \vec{p}} \left[\left(\prod_{k \in \mathcal{T}_s} \int \prod_{i=1}^n (1 + \zeta_k \tilde{Z}_{ik}) d\pi(\zeta_1, \dots, \zeta_T) \right)^2 \right].$$

We have

$$\begin{aligned} & \mathbb{E}_{\vec{p}, \vec{p}} \left[\left(\prod_{k \in \mathcal{T}_s} \int \prod_{i=1}^n (1 + \zeta_k \tilde{Z}_{ik}) d\pi(\zeta_1, \dots, \zeta_T) \right)^2 \right] \\ &= \mathbb{E}_{\vec{p}, \vec{p}} \left[\prod_{k \in \mathcal{T}_s} \frac{1}{4} \left[\prod_{i=1}^n (1 + \tilde{Z}_{ik}) + \prod_{i=1}^n (1 - \tilde{Z}_{ik}) \right]^2 \right] \\ &= \mathbb{E}_{\vec{p}, \vec{p}} \left[\prod_{k \in \mathcal{T}_s} \frac{1}{2} \left(\prod_{i=1}^n (1 + \tilde{Z}_{ik}^2) + \prod_{i=1}^n (1 - \tilde{Z}_{ik}^2) \right) + \sum_{k \in \mathcal{T}_s} \sum_{i=1}^n \tilde{Z}_{ik} \tilde{h}_i(k) \right], \end{aligned}$$

where functions $\tilde{h}_i(k)$ are sums of products of random variables $\tilde{Z}_{j\kappa}$ where the pairs (j, κ) are in the set $\{1, \dots, n\} \times \mathcal{T}_s \setminus \{(i, k)\}$.

Since $\mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}) = 0$ and $\tilde{Z}_{ik}\tilde{Z}_{ik'} = 0$ for $k \neq k'$, the last term vanishes. Thus only the first term remains. As $\tilde{Z}_{ik}\tilde{Z}_{ik'} = 0$ for $k \neq k'$ and the random variables \tilde{Z}_{ik} and $\tilde{Z}_{i'k}$ for $i \neq i'$ are independent, we have

$$\begin{aligned} & \mathbb{E}_{\vec{p}, \vec{p}} \left[\prod_{k \in \mathcal{T}_s} \frac{1}{2} \left(\prod_{i=1}^n (1 + \tilde{Z}_{ik}^2) + \prod_{i=1}^n (1 - \tilde{Z}_{ik}^2) \right) \right] \\ & \leq \prod_{k \in \mathcal{T}_s} \left[\frac{1}{2} \left(\prod_{i=1}^n (1 + \mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2)) + \prod_{i=1}^n (1 - \mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2)) \right) \right] \\ & \leq \prod_{k \in \mathcal{T}_s} \cosh \left(\sum_{i=1}^n \mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2) \right) \\ & \leq \exp \left(\frac{1}{2} \sum_{k \in \mathcal{T}_s} \left(\sum_{i=1}^n \mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2) \right) \right). \end{aligned}$$

Each expectation $\mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2)$ is bounded as follows:

$$\mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2) \leq C^2 M 2^{2s+1-2s j_{n,s} - j_{n,s}} \left(\sum_{l=1}^M \theta_l \omega_l(i) \right)^2.$$

Therefore this bound entails

$$\begin{aligned} & \mathbb{E}_{\vec{p}, \vec{p}} \left[\left(\mathbb{E}_\pi \left(\prod_{i=1}^n \frac{g_{i,\zeta}(Z_i)}{f_i(Z_i)} \right) \right)^2 \right] \\ & \leq \exp \left(\frac{1}{2} \sum_{k \in \mathcal{T}_s} \left(\sum_{i=1}^n \mathbb{E}_{\vec{p}, \vec{p}}(\tilde{Z}_{ik}^2) \right) \right) \\ & \leq \exp \left(\frac{1}{2} \sum_{k \in \mathcal{T}_s} 2^{4s+2-4s j_{n,s} - 2j_{n,s}} C^4 M^2 \left(\sum_{i=1}^n \sum_{l,m=1}^M \theta_l \theta_m \omega_l(i) \omega_m(i) \right) \right) \end{aligned}$$

$$\begin{aligned}
 &= \exp\left(\sum_{k \in \mathcal{I}_s} 2^{4s+1} 2^{-4sj_{ns} - 2j_{ns}} C^4 M^2 (t \theta n \Gamma_n \theta)^2\right) \\
 &= \exp\left(\sum_{k \in \mathcal{I}_s} 2^{4s+1} 2^{-4sj_{ns} - 2j_{ns}} C^4 M^2 (nk_n)^2\right) \\
 &\leq \exp(2^{4s+1} C^4 M^2).
 \end{aligned} \tag{11}$$

Inequalities (10) and (11) lead to

$$\left\| \mathbb{P}_{\vec{p}, \vec{q}} - \mathbb{P}_\pi \right\| \leq \sqrt{\exp(2^{4s+1} C^4 M^2) - 1}. \tag{12}$$

Hence, according to inequalities (9) and (12), the choice of any constant $C < c_\gamma = M^{-1/2} (2R^4 \wedge \ln(1 + 4(1 - \gamma)^2))^{1/4} 2^{-(s+(1/4))}$ entails that

$$\inf_{\Delta} \left(\sup_{(\vec{p}, \vec{q}) \in \theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta = 1) + \sup_{(\vec{p}, \vec{q}) \in \theta_1(R, C, nk_n, s)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta = 0) \right) > \gamma.$$

Therefore Theorem 3.2 is proved.

A.3. Proof of Theorem 3.3

The proof of Theorem 3.3 is a direct consequence of two propositions taking $C_\star = \sqrt{2M(R^2 + \sqrt{2C_R})}$, where C_R is the constant introduced in Proposition A.1, and choosing t such that $M\sqrt{2C_R} < t < C^2/2 - MR^2$. Proposition A.2 deals with the control of the first-type error and Proposition A.3 deals with the control of the second-type error.

Proposition A.2. Let $R > 0$, C_R as in Proposition A.1 and $t > M\sqrt{2C_R}$. Then,

$$\lim_{n \rightarrow +\infty} \sup_{(\vec{p}, \vec{q}) \in \theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{\star, t} = 1) = 0.$$

Proof. We denote by \mathcal{J}_n and \mathcal{S}_n the following sets:

$$\begin{aligned}
 \mathcal{J}_n &= \{j \in \mathbb{N}^\star; j = j_{n,s}^\star, s \in [s_\star, s_\star^\star]\}, \\
 \mathcal{S}_n &= \{s_{n,j} \in [s_\star, s_\star^\star]; 2^{-j} = (nk_n l_n)^{-2/(1+4s_{n,j})}, j \in \mathcal{J}_n\}.
 \end{aligned}$$

According to the definition of $\Delta_{\star, t}$ and Bonferroni inequality, one gets for any $(\vec{p}, \vec{q}) \in \theta_0(R)$:

$$\mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{\star, t} = 1) \leq \sum_{j \in \mathcal{J}_n} \mathbb{P}_{\vec{p}, \vec{q}}(T_j > t(nk_n l_n)^{-4s_{n,j}/(1+4s_{n,j})}).$$

Analogous to Lemma 1 in Autin and Pouet (2012) we easily check that, for any n large enough,

$$\text{card } \mathcal{J}_n \leq \left(\frac{2}{1+4s_\star} - \frac{2}{1+4s_\star^\star} + 1 \right) \frac{\ln(nk_n l_n)}{\ln 2}$$

and similar to the proof of Proposition 4.1 given in Autin and Pouet (2012) we can prove that under (3.1), for any $j \in \mathcal{J}_n$, $\mathbb{P}_{\vec{p}, \vec{q}}(T_j > t(nk_n l_n)^{-4s_{n,j}/(1+4s_{n,j})})$ converges to zero faster than $(\ln(nk_n l_n))^{-1}$ as n tends to infinity provided $t > M\sqrt{2C_R}$ and $s_\star > \frac{1}{4}$. Hence,

$$\lim_{n \rightarrow +\infty} \sup_{(\vec{p}, \vec{q}) \in \theta_0(R)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{\star, t} = 1) = 0$$

and the proof of Proposition A.2 is also ended. \square

Proposition A.3. Let $R > 0$ and $C > \sqrt{2M(R^2 + \sqrt{2C_R})}$, where C_R is as in Proposition A.1. Then, for $t < C^2/2 - MR^2$,

$$\lim_{n \rightarrow +\infty} \sup_{(\vec{p}, \vec{q}) \in \theta_1^\star(R, C, nk_n l_n)} \mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{\star, t} = 0) = 0.$$

Proof. Consider $(\vec{p}, \vec{q}) \in \theta_1^\star(R, C, nk_n l_n)$. We aim at getting an upper bound for the second-type error that corresponds to $\mathbb{P}_{\vec{p}, \vec{q}}(\Delta_{\star, t} = 0)$. From the definition of $\Delta_{\star, t}$, we remark that it suffices to prove that, for one $s \in \mathcal{S}_n$,

$$\mathbb{P}_{\vec{p}, \vec{q}}(T_{j_{n,s}}^\star \leq t(nk_n l_n)^{-4s/(1+4s)}) \text{ tends to zero when } n \text{ goes to infinity.}$$

Let $s \in [s_*, s^*]$ be such that $(\vec{p}, \vec{q}) \in \Theta_1(R, C, nk_n l_n, s)$. Analogous to the proof of [Theorem 3.1](#), we get for any n large enough:

$$\begin{aligned} & \mathbb{P}_{\vec{p}, \vec{q}} \left(T_{j_{n,s}^*} \leq t(nk_n l_n)^{-4s/(1+4s)} \right) \\ & \leq \frac{C_R M^2 \left(2^{j_{n,s}^*} + n \sum_{l=1}^M \|p_l - q_l\|_2^2 + \sqrt{2^{j_{n,s}^*}} n \sum_{l=1}^M \|p_l - q_l\|_2 \right)}{n^2 k_n^2 (2^{-1} - MR^2 C^{-2} - tC^{-2})^2 \left(\sum_{l=1}^M \|p_l - q_l\|_2^2 \right)^2} \\ & \leq \frac{C_R}{((2M)^{-1} C^2 - R^2 - tM^{-1})^2} \left(2l_n^2 + \frac{(nk_n l_n)^{4s/(1+4s)}}{nk_n^2} + \frac{1}{k_n^2} \sqrt{\frac{2M}{n^3}} (nk_n l_n)^{(1+6s)/(1+4s)} \right). \end{aligned}$$

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 according to [Assumption 3.1](#). [Proposition A.3](#) is also proved. \square

A.4. Proof of [Theorem 3.4](#)

This proof looks like the proof of [Theorem 3.3](#).

Let \bar{S}_n denote a net on the smoothness space such that

$$\forall s, t \in \bar{S}_n : s_* \leq s, t \leq s^*, \quad |s - t| \geq \frac{1}{N}, \quad N = \text{card } \bar{S}_n = \mathcal{O}((\ln n)^{1-\delta}),$$

where N also satisfies

$$\lim_{n \rightarrow +\infty} \ln \left(\frac{1}{\ln k_n} \right) - \frac{2 \ln(n)}{(4s^* + 1)^2 N} = -\infty.$$

This choice is made possible because of [Assumption 3.1](#).

Let $\bar{\mathcal{J}}_n$ denote the associated net on the level space that is

$$\bar{\mathcal{J}}_n = \{j \in \mathbb{N}^*; j = j_{n,s}^*, s \in \bar{S}_n\}.$$

Let \vec{p} such that for any $l \in \{1, \dots, M\}$,

$$\{x; p_l(x) > \frac{1}{2}\} \supset [0, 1[.$$

For any given $s \in \bar{S}_n$ and any $l \in \{1, \dots, M\}$, we introduce the following densities:

$$q_{l,\zeta,s}(\cdot) = p_l(\cdot) + 2^s C \sqrt{M} \theta_l \sum_{k \in \mathcal{T}_s^*} \zeta_{k,s} 2^{-s j_{n,s}^* - j_{n,s}^*/2} \psi_{j_{n,s}^* k}(\cdot),$$

where $\mathcal{T}_s^* = \{0, \dots, 2^{j_{n,s}^*} - 1\}$ and, for any $k \in \mathcal{T}_s^*$, $\zeta_{k,s} \in \{-1, 1\}$.

Here, $\theta = (\theta_1, \dots, \theta_M)$ is still an eigenvector associated with the smallest eigenvalue k_n of Γ_n such that $\|\theta\|_2 = 1$.

Suppose that the marginal densities of Y_i and Z_i ($1 \leq i \leq n$) are respectively

$$f_i(\cdot) = \sum_{l=1}^M \omega_l(i) p_l(\cdot) \quad \text{and} \quad g_{i,\zeta,s}(\cdot) = \sum_{l=1}^M \omega_l(i) q_{l,\zeta,s}(\cdot).$$

The functions $q_{l,\zeta,s}$ are clearly densities that satisfy for large n :

$$\{x; q_{l,\zeta,s}(x) > \frac{1}{2}\} \supset [0, 1[.$$

Moreover, if $C < RM^{-1/2} 2^{-s}$, then the difference $q_{l,\zeta,s} - p_l$ belongs to the Besov ball $\mathcal{B}_{2,\infty}^s(R)$. Notice that for at least one $l \in \{1, \dots, M\}$, $M\theta_l^2 \geq 1$. Hence

$$\|q_{l,\zeta,s} - p_l\|_2^2 \geq C^2 (nk_n l_n)^{-4s/(1+4s)}.$$

The prior probability π_s on $\Lambda(C, nk_n l_n, s)$ is chosen such that the random variables $\zeta_{k,s} (k \in \mathcal{T}_s^*)$ are independent Rademacher whereas the level $j_{n,s}^*$ is chosen uniformly on $\bar{\mathcal{J}}_n$.

Unless explicitly specified, the expectation $\mathbb{E}(\cdot)$ is taken regarding the random variables $Y_1, \dots, Y_n, Z_1, \dots, Z_n, \zeta_{k,s}, k \in \mathcal{T}_s$ and $j_{n,s}^* \in \bar{\mathcal{J}}_n$.

The general calculation follows the same path as in [Theorem 3.2](#). The sum of the errors is lower-bounded as follows:

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

An upper bound of the \mathbb{L}_1 -distance is given by

$$\left\| \frac{1}{N} \sum_{s \in \bar{S}_n} \mathbb{P}_{\pi_s} - \mathbb{P}_{\vec{p}, \vec{p}} \right\| \leq \sqrt{\mathbb{E} \left(\left(\frac{1}{N} \sum_{s \in \bar{S}_n} \frac{d\mathbb{P}_{\pi_s}}{d\mathbb{P}_{\vec{p}, \vec{p}}} - 1 \right)^2 \right)}$$

$$\leq \sqrt{\frac{1}{N^2} \sum_{s \in \bar{S}_n} \mathbb{E} \left(\left(\frac{d^{\mathbb{P}}_{\pi_s}}{d^{\mathbb{P}}_{\vec{p} \cdot \vec{p}}} - 1 \right)^2 \right) + \frac{1}{N^2} \sum_{s \neq s' \in \bar{S}_n} \left[\mathbb{E} \left(\frac{d^{\mathbb{P}}_{\pi_s}}{d^{\mathbb{P}}_{\vec{p} \cdot \vec{p}}} \frac{d^{\mathbb{P}}_{\pi_{s'}}}{d^{\mathbb{P}}_{\vec{p} \cdot \vec{p}}} \right) - 1 \right]}.$$

The first term is called the *square term* and is handled as in the non-adaptive case. The second term is called the *cross term*. It only appears in the adaptive case and leads to technical calculations.

We state a more general result than Lemma 6 of [Autin and Pouet \(2012\)](#). As it is easily proved by following the proof of Lemma 6 of [Autin and Pouet \(2012\)](#) we decide to omit its proof.

Lemma A.1. For any $s' < s$ in \bar{S}_n , i.e. $j_{s'} < j_s$, the following bound holds:

$$\mathbb{E} \left(\frac{d^{\mathbb{P}}_{\pi_s}}{d^{\mathbb{P}}_{\vec{p} \cdot \vec{p}}} \frac{d^{\mathbb{P}}_{\pi_{s'}}}{d^{\mathbb{P}}_{\vec{p} \cdot \vec{p}}} \right) - 1 \leq 2(C^2 2^{1+s^*} M)^2 (l_n k_n)^{-2} \exp \left(-2 \ln(n l_n k_n) \frac{2(s'-s)}{(4s'+1)(4s+1)} \right).$$

The choice of N and the fact that s' is smaller than s entail that the upper bound of each summand in the cross term is the same and goes to zero when n tends to infinity.

The second part of the proof is the study of the square term. It is handled as in the non-adaptive case but here leads to an upper bound which goes to infinity. This behavior is compensated by the normalizing factor N^{-1} . Similar to [Autin and Pouet \(2012\)](#), we have

$$\mathbb{E} \left(\left(\frac{d^{\mathbb{P}}_{\pi_s}}{d^{\mathbb{P}}_{\vec{p} \cdot \vec{p}}} - 1 \right)^2 \right) \leq \exp \left(2^{4s^*+1} C^4 M^2 l_n^{-2} \right) - 1.$$

If $C^4 < (1-\delta)2^{-4s^*-1}M^{-2}$, then the following limit holds:

$$\lim_{n \rightarrow +\infty} N^{-1} (\ln n)^{2^{4s^*+1} M^2 C^4} = 0.$$

This result entails that the square term goes to zero when n tends to infinity provided $C < c_* = 2^{-s^*} M^{-(1/2)} (R \wedge (2^{-1}(1-\delta))^{1/4})$.

Gathering the results for the square and cross terms, we conclude that the lower bound goes to one as n tends to infinity and that the loss $l_n = (\sqrt{\ln \ln n})^{-1}$ is unavoidable.

References

Autin, F., Pouet, C., 2011. Test on components of mixture densities. *Statistics & Risk Modeling* 28 (4), 389–410. <http://arXiv:0912.0786> (extended version).
 Autin, F., Pouet, C., 2012. Adaptive test on components of mixture densities. *Mathematical Methods of Statistics* 21, 93–108.
 Autin, F., Pouet, C., 2013. Testing means from sampling populations with undefined labels. *Statistical Methods* 10, 151–169.
 Butucea, C., Tribouley, K., 2006. Nonparametric homogeneity tests. *Journal of Statistical Planning and Inference* 136, 597–639.
 Delmas, C., 2003. On likelihood ratio tests in Gaussian mixture models. *Indian Journal of Statistics* 65, 513–531.
 Efron, B., Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.
 Garel, B., 2001. Likelihood ratio test for univariate Gaussian mixture. *Journal of Statistical Planning and Inference* 96, 325–350.
 Garel, B., 2005. Asymptotic theory of the likelihood ratio test for the identification of a mixture. *Journal of Statistical Planning and Inference* 131, 271–296.
 Hall, P., 1981. On the nonparametric estimation of mixture proportions. *Journal of the Royal Statistical Society: Series B* 43, 147–156.
 Hall, P., Titterton, D.M., 1984. Efficient nonparametric estimation of mixture proportions. *Journal of the Royal Statistical Society: Series B* 46, 465–473.
 Lodatko, N., Maiboroda, R.E., 2007. Estimation of the density of a distribution from observations with an admixture. *Theory of Probability and Mathematical Statistics* 73, 99–108.
 Maiboroda, R.E., 1996. Estimates for distribution of components of mixtures with varying concentrations. *Ukrainian Mathematical Journal* 48, 618–622.
 Maiboroda, R.E., 2000a. An asymptotically effective estimate for a distribution from a sample with a varying mixture. *Theory of Probability and Mathematical Statistics* 61, 121–130.
 Maiboroda, R.E., 2000b. A test for the homogeneity of mixtures with varying concentrations. *Ukrainian Mathematical Journal* 52, 1256–1263.
 Maiboroda, R.E., Sugakova, O.V., 2012. Statistics of mixtures with varying concentrations with application to DNA microarray data analysis. *Journal of Nonparametric Statistics* 24 (1), 201–215.
 McLachlan, G.J., Peel, P., 2000. *Finite Mixture Models*. Wiley, New York.
 Pokhyl'ko, D., 2005. Wavelet estimators of a density constructed from observations of a mixture. *Theory of Probability and Mathematical Statistics* 70, 135–145.
 Qin, J., 1999. Empirical likelihood ratio based confidence intervals for mixture proportions. *Annals of Statistics* 27, 1368–1384.
 Spokoiny, V.G., 1996. Adaptive hypothesis testing using wavelets. *Annals of Statistics* 24, 2477–2498.
 Titterton, D.M., 1983. Minimum distance nonparametric estimation of mixture proportions. *Journal of the Royal Statistical Society: Series B* 45, 37–46.