

Minimax goodness-of-fit testing in ill-posed inverse problems with partially unknown operators

Clément Marteau^a and Theofanis Sapatinas^b

^a*Institut Camille Jordan, Université Lyon I – Claude Bernard, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne, France.*

E-mail: marteau@math.univ-lyon1.fr

^b*Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, CY 1678 Nicosia, Cyprus. E-mail: fanis@ucy.ac.cy*

Received 23 November 2015; revised 10 May 2016; accepted 22 May 2016

Abstract. We consider a Gaussian sequence model that contains ill-posed inverse problems as special cases. We assume that the associated operator is partially unknown in the sense that its singular functions are known and the corresponding singular values are unknown but observed with Gaussian noise. For the considered model, we study the minimax goodness-of-fit testing problem. Working with certain ellipsoids in the space of square-summable sequences of real numbers, with a ball of positive radius removed, we obtain lower and upper bounds for the minimax separation radius in the non-asymptotic framework, i.e., for fixed values of the involved noise levels. Examples of mildly and severely ill-posed inverse problems with ellipsoids of ordinary-smooth and super-smooth sequences are examined in detail and minimax rates of goodness-of-fit testing are obtained for illustrative purposes.

Résumé. Nous considérons un modèle séquentiel gaussien incluant des modèles de type « problème inverse » comme cas particulier. Nous supposons que l'opérateur associé est partiellement connu au sens où les fonctions propres associées à la décomposition en valeurs singulières sont connues, mais pas les valeurs propres. Ces dernières sont malgré tout observables en présence d'un bruit gaussien. Dans ce cadre, nous nous concentrons sur l'étude d'un problème de test d'adéquation. En utilisant à la fois une condition d'énergie sur le signal considéré ainsi que des propriétés de régularité, nous établissons des bornes supérieures et inférieures pour la vitesse minimax de séparation dans un cadre asymptotique, i.e. pour des valeurs fixées des niveaux de bruit impliqués dans les observations. Pour finir, des exemples particuliers de problèmes inverses et de régularités sont considérés et les vitesses minimax de séparation correspondantes sont discutées en détail afin d'illustrer les résultats obtenus.

MSC: 62G05; 62K20

Keywords: Ellipsoids; Compact operators; Gaussian sequence model; Gaussian white noise model; Ill-posed inverse problems; Minimax goodness-of-fit testing; Minimax signal detection; Singular value decomposition

1. Introduction

We consider the following Gaussian sequence model (GSM), introduced in [10],

$$\begin{cases} Y_j = b_j \theta_j + \varepsilon \xi_j, & j \in \mathbb{N}, \\ X_j = b_j + \sigma \eta_j, & j \in \mathbb{N}, \end{cases} \quad (1.1)$$

where $\mathbb{N} = \{1, 2, \dots\}$ is the set of natural numbers, $b = \{b_j\}_{j \in \mathbb{N}} > 0$ is an *unknown* sequence, $\theta = \{\theta_j\}_{j \in \mathbb{N}} \in l^2(\mathbb{N})$ is the *unknown* signal of interest, $\xi = \{\xi_j\}_{j \in \mathbb{N}}$ and $\eta = \{\eta_j\}_{j \in \mathbb{N}}$ are sequences of independent standard Gaussian random variables (and independent of each other), and $\varepsilon, \sigma > 0$ are *known* parameters (the noise levels). The observations are given by the sequence $(Y, X) = \{(Y_j, X_j)\}_{j \in \mathbb{N}}$ from the GSM (1.1) and their joint law is denoted by $\mathbb{P}_{\theta, b}$. Here, $l^2(\mathbb{N})$ denotes the space of square-summable sequence of real numbers, i.e., $l^2(\mathbb{N}) = \{\theta \in \mathbb{R}^{\mathbb{N}} : \|\theta\|^2 := \sum_{j \in \mathbb{N}} \theta_j^2 < +\infty\}$.

The GSM (1.1) arises in the case of ill-posed inverse problems with *noisy* operators. Indeed, consider the Gaussian white noise model (GWNM)

$$dY_\varepsilon(t) = Af(t) dt + \varepsilon dW(t), \quad t \in V, \tag{1.2}$$

where A is a linear bounded operator acting on a Hilbert space \mathcal{H}_1 with values on another Hilbert space \mathcal{H}_2 , $f(\cdot) \in \mathcal{H}_1$ is the unknown response function that one wants to detect or estimate, $W(\cdot)$ is a standard Wiener process on $V \subseteq \mathbb{R}$, and $\varepsilon > 0$ is a known parameter (the noise level). For the sake of simplicity, we only consider the case when A is injective (meaning that A has a trivial nullspace) and assume that $V = [0, 1]$, $\mathcal{H}_1 = L^2([0, 1])$, $U \subseteq \mathbb{R}$ and $\mathcal{H}_2 = L^2(U)$. In most cases of interest, A is a compact operator (see, e.g., Chapter 2 of [14]). In particular, it admits a singular value decomposition (SVD) $(b_j, \psi_j, \varphi_j)_{j \in \mathbb{N}}$, in the sense that

$$A\varphi_j = b_j\psi_j, \quad A^*\psi_j = b_j\varphi_j, \quad j \in \mathbb{N}, \tag{1.3}$$

where A^* denotes the adjoint operator of A , and $(b_j^2)_{j \in \mathbb{N}}$ and $(\varphi_j)_{j \in \mathbb{N}}$ are, respectively, the eigenvalues and the eigenfunctions of A^*A . Thus, the (first equation in) GSM (1.1) arises where for all $j \in \mathbb{N}$

$$Y_j = \int_0^1 \psi_j(t) dY_\varepsilon(t), \quad \theta_j = \int_0^1 \varphi_j(t) f(t) dt, \quad \xi_j = \int_0^1 \psi_j(t) dW(t), \quad j \in \mathbb{N},$$

and $b_j^2 > 0$ (since A is injective). In this case, the GWNM (1.2) corresponds to a so-called ill-posed inverse problem since the inversion of A^*A is not bounded. Possible examples of such decompositions arise with, e.g., convolution or Radon-transform operators, see, e.g., [14]. The effect of the ill-posedness of the model is clearly seen in the decay of the singular values b_j as $j \rightarrow +\infty$. As $j \rightarrow +\infty$, $b_j\theta_j$ gets weaker and is then more difficult to perform inference on the sequence $\theta = \{\theta_j\}_{j \in \mathbb{N}}$.

In order to motivate our investigation, we stress that ill-posed inverse problems have been at the core of several contributions over the last two decades. In the early literature, the associated compact operator A (and, hence, its sequence $b = \{b_j\}_{j \in \mathbb{N}}$ of singular values) was supposed to be *fully known*. (Note that, in this case, the second equation in the GSM (1.1) does not appear.) We refer, e.g., to [7–9,11,12] in the minimax estimation context. In the goodness-of-fit paradigm, we mention for instance [18,21] where minimax separation radius/minimax separation rates (signal detection or goodness-of-fit testing) were established, for ill-posed inverse problems with various smoothness conditions on the sequence of interest. Specific situations have been also recently investigated in, e.g., [17] (tomography) or [5] (biophotonic imaging).

The case of an unknown compact operator A that is observed with Gaussian noise has also been recently treated in the estimation literature, especially the situation where A is *partially unknown*, see, e.g., [10,13,15,20]. In these contributions, it is assumed for the corresponding SVD (1.3) that

- the sequence of singular functions $(\psi, \varphi) = (\psi_j, \varphi_j)_{j \in \mathbb{N}}$ is *known*,
- the sequence of singular values $b = \{b_j\}_{j \in \mathbb{N}}$ is *unknown* but observed with some Gaussian noise.

In other words, the following sequence model is considered

$$X_j = b_j + \sigma \eta_j, \quad j \in \mathbb{N},$$

where $\eta = \{\eta_j\}_{j \in \mathbb{N}}$ is a sequence of independent standard Gaussian random variables (and independent of the standard Gaussian sequence $\xi = \{\xi_j\}_{j \in \mathbb{N}}$), and $\sigma > 0$ is a known parameter (the noise level). Therefore, the second equation in the GSM (1.1) is also readily available.

To motivate the GSM (1.1), consider the following deconvolution model (see also [10] for a complete discussion on this subject)

$$dY_\varepsilon(t) = g \star f(t) + \varepsilon dW(t), \quad t \in [0, 1], \tag{1.4}$$

where

$$g \star f(t) = \int_0^1 g(t-x)f(x) dx, \quad t \in [0, 1],$$

is the convolution between $g(\cdot)$ and $f(\cdot)$, $g(\cdot)$ is an unknown 1-periodic (convolution) kernel in $L^2([0, 1])$, $f(\cdot)$ is an unknown 1-periodic signal in $L^2([0, 1])$, $dY_\varepsilon(\cdot)$ is observed, $W(\cdot)$ is a standard Wiener process, and $\varepsilon > 0$ is the noise level. Let $\phi_j(\cdot)$, $j \in \mathbb{N}$, be the usual real trigonometric basis on V . The model (1.4) is equivalent to the (first equation in the) GSM (1.1) by a projection on the trigonometric basis $\phi_j(\cdot)$, $j \in \mathbb{N}$. In the case where the kernel $g(\cdot)$ is unknown (i.e., the sequence $(b_k)_{k \in \mathbb{N}} = ((g, \phi_k))_{k \in \mathbb{N}}$ is unknown), suppose that we can pass the trigonometric basis $\phi_j(\cdot)$, $j \in \mathbb{N}$, through the convolution kernel, i.e., to send *each* $\phi_j(\cdot)$, $j \in \mathbb{N}$, as an input function $f(\cdot)$ and observe the corresponding $dY_{\varepsilon, j}(\cdot)$, $j \in \mathbb{N}$. In other words, we are able to obtain training data for the estimation of the unknown convolution kernel $g(\cdot)$ in this setting. In particular, we obtain exactly the two sequences of observations Y_j and X_j , $j \in \mathbb{N}$, in the GSM (1.1). In this case, the corresponding noise levels coincide, i.e., $\varepsilon = \sigma$.

To the best of our knowledge, there is no research work on minimax goodness-of-fit testing in ill-posed inverse problems with partially unknown operators. Our aim is to fill this gap. In particular, considering the GSM (1.1) and working with certain ellipsoids in the space of squared-summable sequences of real numbers, with a ball of positive radius removed, we obtain lower and upper bounds for the minimax separation radius in the non-asymptotic framework, i.e., for fixed values of ε and σ . Examples of mildly and severely ill-posed inverse problems with ellipsoids of ordinary-smooth and super-smooth sequences are examined in detail and minimax rates of goodness-of-fit testing are obtained for illustrative purposes.

The paper is organized as follows. Section 2 presents the considered statistical setting, a brief overview of the main results. Section 3 is devoted to the construction of the suggested testing procedure. A general upper bound on the maximal second kind error is then displayed and special benchmark examples are presented for illustrative purposes. The corresponding lower bounds are proposed in Section 4. In Section 5 we discuss some open questions and present a short selection of applied work that motivated our investigation. Finally, all proofs and technical arguments are gathered in the [Appendix](#).

Throughout the paper we set the following notations. For all $x, y \in \mathbb{R}$, $\delta_x(y) = 1$ if $x = y$ and $\delta_x(y) = 0$ if $x \neq y$. Also, $x \wedge y := \min\{x, y\}$ and $x \vee y := \max\{x, y\}$. Given two sequences $(c_j)_{j \in \mathbb{N}}$ and $(d_j)_{j \in \mathbb{N}}$ of real numbers, $c_j \sim d_j$ means that there exists $0 < \kappa_0 \leq \kappa_1 < \infty$ such that $\kappa_0 \leq c_j/d_j \leq \kappa_1$ for all $j \in \mathbb{N}$. Let ν be either ε or σ or (ε, σ) , and let \mathcal{V} be either $\mathbb{R}^+ := (0, +\infty)$ or $\mathbb{R}^+ \times \mathbb{R}^+$. Given two collections $(c_\nu)_{\nu \in \mathcal{V}}$ and $(d_\nu)_{\nu \in \mathcal{V}}$ of positive real numbers, $c_\nu \gtrsim d_\nu$ means that there exists $0 < \kappa_0 < +\infty$ such that $c_\nu \geq \kappa_0 d_\nu$ for all $\nu \in \mathcal{V}$. Similarly, $c_\nu \lesssim d_\nu$ means that there exists $0 < \kappa_1 < +\infty$ such that $c_\nu \leq \kappa_1 d_\nu$ for all $\nu \in \mathcal{V}$.

2. Minimax goodness-of-fit testing

2.1. The statistical setting

Given observations $(Y, X) = \{(Y_j, X_j)\}_{j \in \mathbb{N}}$ from the GSM (1.1), the aim is to compare the underlying (unknown) signal $\theta \in l^2(\mathbb{N})$ to a (known) benchmark signal θ_0 , i.e., to test

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta - \theta_0 \in \mathcal{F}, \quad (2.1)$$

for some given θ_0 and a given subspace \mathcal{F} . The statistical setting (2.1) is known as *goodness-of-fit testing* when $\theta_0 \neq 0$ or *signal detection* when $\theta_0 = 0$.

Remark 2.1. Given observations from the GWNM (1.2), the testing problem (2.1) is related to the test

$$H_0 : f = f_0 \quad \text{versus} \quad H_1 : f - f_0 \in \tilde{\mathcal{F}}, \quad (2.2)$$

for a given benchmark function f_0 and a given subspace $\tilde{\mathcal{F}}$. In most cases, $\tilde{\mathcal{F}}$ contains functions $f \in L^2([0, 1])$ that admit a Fourier series expansion with Fourier coefficients θ belonging to \mathcal{F} (see, e.g., [19], Section 3.2). In these cases, the tests (2.1) and (2.2) are equivalent.

The choice of the set \mathcal{F} is important. Indeed, it should be rich enough in order to contain the true θ . At the same time, if it is too rich, it will not be possible to control the performances of a given test due to the complexity of the problem. The common approach for such problems is to impose both a *regularity* condition (which characterizes the smoothness of the underlying signal) and an *energy* condition (which measures the amount of the underlying signal).

Concerning the regularity condition, we will work with certain ellipsoids in $l^2(\mathbb{N})$. In particular, we assume that $\theta \in \mathcal{E}_a(R)$, the set $\mathcal{E}_a(R)$ being defined as

$$\mathcal{E}_a(R) = \left\{ \theta \in l^2(\mathbb{N}), \sum_{j \in \mathbb{N}} a_j^2 \theta_j^2 \leq R \right\}, \quad (2.3)$$

where $a = (a_j)_{j \in \mathbb{N}}$ denotes a non-decreasing sequence of positive real numbers with $a_j \rightarrow +\infty$ as $j \rightarrow +\infty$, and $R > 0$ is a constant. The set $\mathcal{E}_a(R)$ can be seen as a condition on the decay of θ . The cases where a increases very fast correspond to θ with a small amount of non-zero coefficients. In such a case, the corresponding signal can be considered as being ‘smooth.’ Without loss of generality, in what follows, we set $R = 1$, and write \mathcal{E}_a instead of $\mathcal{E}_a(1)$.

Regarding the energy condition, it will be measured in the $l^2(\mathbb{N})$ -norm. In particular, given $r_{\varepsilon, \sigma} > 0$ (called the radius), which is allowed to depend on the noise levels $\varepsilon, \sigma > 0$, we will consider alternatives such that $\|\theta - \theta_0\| > r_{\varepsilon, \sigma}$. Given a smoothness sequence a and a radius $r_{\varepsilon, \sigma} > 0$, the set \mathcal{F} can thus be defined as

$$\mathcal{F} := \Theta_a(r_{\varepsilon, \sigma}) = \{v \in \mathcal{E}_a, \|v\| \geq r_{\varepsilon, \sigma}\}. \quad (2.4)$$

In other words, the set \mathcal{F} is an ellipsoid in $l^2(\mathbb{N})$ with a ball of radius $r_{\varepsilon, \sigma} > 0$ removed. In many cases of interest, the set \mathcal{F} provides constraints on the Fourier coefficients of $f \in L^2([0, 1])$ in the model (1.2) (see, e.g., [19], Section 3.2).

We consider below the hypothesis testing setting (2.1) with $\theta_0 \neq 0$ (i.e., goodness-of-fit testing). More formally, given observations from the GSM (1.1), for any given $\theta_0 \in \mathcal{E}_a \setminus \{0\}$, we will be dealing with the following goodness-of-fit testing problem

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta - \theta_0 \in \Theta_a(r_{\varepsilon, \sigma}), \quad (2.5)$$

where $\Theta_a(r_{\varepsilon, \sigma})$ is defined in (2.4). The sequence a being fixed, the main issue for the problem (2.5) is then to characterize the values of $r_{\varepsilon, \sigma} > 0$ for which both hypotheses H_0 (called the null hypothesis) and H_1 (called the alternative hypothesis) are ‘separable’ (in a sense which will be made precise later on).

Remark 2.2. We would like to stress at this point that in the *standard* GSM (i.e., (1.1) with $\sigma = 0$), signal detection (i.e., $\theta_0 = 0$) and goodness-of-fit testing (i.e., $\theta_0 \neq 0$) problems are equivalent as soon as the involved operator is injective. Indeed, without loss of generality, we can still replace the observed sequence $(Y_j)_{j \in \mathbb{N}}$ by $(\tilde{Y}_j)_{j \in \mathbb{N}} := (Y_j - b_j \theta_{0,j})_{j \in \mathbb{N}}$. This is no more the case in the GSM (1.1) since the sequence $(b_j)_{j \in \mathbb{N}}$ is unknown. Signal detection and goodness-of-fit problems should therefore be treated in a different manner. In this work, we only address the goodness-of-fit testing problem (2.5).

Remark 2.3. In the same spirit, we point out that the benchmark function θ_0 is assumed to belong to $\mathcal{E}_a \setminus \{0\}$ in the testing problem (2.5). Such an assumption is not necessary in the classical setting where $\sigma = 0$. Given observations from the GSM (1.1), the requirement $\theta_0 \in \mathcal{E}_a \setminus \{0\}$ allows a sharp control of the additional uncertainty generated by the noise in the operator. This requirement could be replaced by $\theta_0 \in l^2(\mathbb{N}) \setminus \{0\}$, leading to the minimax separation rates described in [18,21] with $\max(\varepsilon, \sigma)$ instead of ε .

In the following, a (non-randomized) test $\Psi := \Psi(Y, X)$ will be defined as a measurable function of the observation $(Y, X) = (Y_j, X_j)_{j \in \mathbb{N}}$ from GSM (1.1) having values in the set $\{0, 1\}$. By convention, H_0 is rejected if $\Psi = 1$ and H_0 is not rejected if $\Psi = 0$. Then, given a test Ψ , we can investigate

- the first kind (type I) error probability defined as

$$\alpha_{\varepsilon, \sigma}(\Psi) := \mathbb{P}_{\theta_0, b}(\Psi = 1), \quad (2.6)$$

which measures the probability to reject H_0 when H_0 is true (i.e., $\theta = \theta_0$); it is often constrained as being bounded by a prescribed level $\alpha \in]0, 1[$, and

- the maximal second kind (type II) error probability defined as

$$\beta_{\varepsilon, \sigma}(\Theta_a(r_{\varepsilon, \sigma}), \Psi) := \sup_{\theta: \theta - \theta_0 \in \Theta_a(r_{\varepsilon, \sigma})} \mathbb{P}_{\theta, b}(\Psi = 0), \quad (2.7)$$

which measures the worst possible probability not to reject H_0 when H_0 is not true (i.e., when $\theta - \theta_0 \in \Theta_a(r_{\varepsilon,\sigma})$); one would like to ensure that it is bounded by a prescribed level $\beta \in]0, 1[$.

For simplicity in our exposition, we will restrict ourselves to α -level tests, i.e., tests Ψ_α satisfying $\alpha_{\varepsilon,\sigma}(\Psi_\alpha) \leq \alpha$, for any fixed value $\alpha \in]0, 1[$.

Let $\alpha, \beta \in]0, 1[$ be given, and let Ψ_α be an α -level test.

Definition 2.1. The separation radius of the α -level test Ψ_α over the class \mathcal{E}_a is defined as

$$r_{\varepsilon,\sigma}(\mathcal{E}_a, \Psi_\alpha, \beta) := \inf\{r_{\varepsilon,\sigma} > 0 : \beta_{\varepsilon,\sigma}(\Theta_a(r_{\varepsilon,\sigma}), \Psi_\alpha) \leq \beta\},$$

where the maximal second kind error probability $\beta_{\varepsilon,\sigma}(\Theta_a(r_{\varepsilon,\sigma}), \Psi_\alpha)$ is defined in (2.7).

In some sense, the separation radius $r_{\varepsilon,\sigma}(\mathcal{E}_a, \Psi_\alpha, \beta)$ corresponds to the smallest possible value of the available signal $\|\theta - \theta_0\|$ for which H_0 and H_1 can be ‘separated’ by the α -level test Ψ_α with maximal second kind error probability, bounded by a prescribed level $\beta \in]0, 1[$.

Definition 2.2. The minimax separation radius $\tilde{r}_{\varepsilon,\sigma} := \tilde{r}_{\varepsilon,\sigma}(\mathcal{E}_a, \alpha, \beta) > 0$ over the class \mathcal{E}_a is defined as

$$\tilde{r}_{\varepsilon,\sigma} := \inf_{\tilde{\Psi}_\alpha: \alpha_{\varepsilon,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} r_{\varepsilon,\sigma}(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta). \tag{2.8}$$

The minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$ corresponds to the smallest radius $r_{\varepsilon,\sigma} > 0$ such that there exists some α -level test $\tilde{\Psi}_\alpha$ for which the maximal second kind error probability $\beta_{\varepsilon,\sigma}(\Theta_a(r_{\varepsilon,\sigma}), \tilde{\Psi}_\alpha)$ is not greater than β .

2.2. Summary of the results

In this section, we present an overview of the main results obtained in the subsequent sections. The main Theorem 2.1 presented below combines Theorems 3.1 (upper bound) and Theorem 4.1 (lower bound). Moreover, the minimax rates displayed in Tables 1 and 2 correspond, respectively, to the minimax rates given in Theorem 3.2 and Theorem 4.2 for various smoothness and ill-posedness conditions.

We emphasize that the aim of this work is to establish ‘optimal’ separation conditions for the goodness-of-fit testing problem (2.5). This task requires, in particular, precise (non-asymptotic) controls of the first kind error probability $\alpha_{\varepsilon,\sigma}(\Psi_\alpha)$ and the maximal second kind error probability $\beta_{\varepsilon,\sigma}(\Theta_a(r_{\varepsilon,\sigma}), \Psi_\alpha)$ (of a specific test Ψ_α that will be made precise in Section 3) by prescribed levels $\alpha, \beta \in]0, 1[$, respectively. Such controls allow us to derive both upper and lower bounds on the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$, as summarized in the following theorem.

Table 1
Minimax goodness-of-fit testing with *unknown* singular values: *upper bounds* on the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}^2$ for $0 < \varepsilon \leq \varepsilon_0, \varepsilon_0 \in]0, 1[$, and $0 < \sigma \leq \sigma_0, \sigma_0 \in]0, 1[$, for all $t, s > 0$

Goodness-of-fit testing problem	Ordinary-smooth $a_j \sim j^s$	Super-smooth $a_j \sim \exp\{js\}$
Mildly ill-posed $b_j \sim j^{-t}$	$\varepsilon^{4s/(2s+2t+1/2)} \vee [\sigma \ln^{3/4}(1/\sigma)]^{2[(s/t)\wedge 1]}$	$\varepsilon^2 (\ln(1/\varepsilon))^{2t+1/2} \vee \sigma^2 \ln^{3/2}(1/\sigma)$
Severely ill-posed $b_j \sim \exp\{-jt\}$	$(\ln(1/\varepsilon))^{-2s} \vee [\ln(1/\sigma \ln^{-1/2}(1/\sigma))]^{-2s}$	$\varepsilon^{2s/(s+t)} \vee [\sigma \ln^{1/2}(1/\sigma)]^{2[(s/t)\wedge 1]}$

Table 2
 Minimax goodness-of-fit testing with *unknown* singular values: *lower bounds* on the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}^2$ for $0 < \varepsilon \leq \varepsilon_0$, $\varepsilon_0 \in]0, 1[$, and $0 < \sigma \leq \sigma_0$, $\sigma_0 \in]0, 1[$, for all $t, s > 0$

Goodness-of-fit testing problem	Ordinary smooth $a_j \sim j^s$	Super smooth $a_j \sim \exp\{js\}$
Mildly ill-posed $b_j \sim j^{-t}$	$\varepsilon^{4s/(2s+2t+1/2)} \vee \sigma^{2[(s/t)\wedge 1]}$	$\varepsilon^2(\ln \varepsilon^{-1})^{2t+1/2} \vee \sigma^2$
Severely ill-posed $b_j \sim \exp\{-jt\}$	$(\ln(1/\varepsilon))^{-2s} \vee (\ln(1/\sigma))^{-2s}$	$\varepsilon^{2s/(s+t)} \vee \sigma^{2[(s/t)\wedge 1]}$

Theorem 2.1. *Let $\alpha, \beta \in]0, 1[$ be fixed, such that $\alpha \leq \beta$. Consider the goodness-of-fit testing problem (2.5). Then, there exist explicit positive constants¹ $\tilde{C}(\alpha, \beta) > 0$, $C_{\alpha,\beta} > 0$, $c_{\alpha,\beta} > 0$ and $\sigma_0 \in]0, 1[$ such that, for all $0 < \sigma \leq \sigma_0$ and for each $\varepsilon > 0$,*

$$(i) \quad \tilde{r}_{\varepsilon,\sigma}^2 \leq \inf_{D \in \mathbb{N}} \left[\tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{\ln(2/\alpha)}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right],$$

and, for all $\varepsilon, \sigma > 0$,

$$(ii) \quad \tilde{r}_{\varepsilon,\sigma}^2 \geq \left\{ \frac{C_{\alpha,\beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \right\} \vee \left\{ \sup_{D \in \mathbb{N}} \left[c_{\alpha,\beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \right\},$$

where the bandwidths M_0, M_1 and M_2 depend¹ on both $(b_j)_{j \in \mathbb{N}}$ and σ .

Theorem 2.1 provides a precise description on the behavior of the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$ in terms of the sequences $(a_j)_{j \in \mathbb{N}}$ and $(b_j)_{j \in \mathbb{N}}$ and of the noise levels ε and σ . It is worth pointing out that this control is *non-asymptotic*. There is indeed a technical constraint on the value of σ ($0 < \sigma \leq \sigma_0$, $\sigma_0 \in]0, 1[$), but we do not assume its convergence towards 0, i.e., we work with *fixed* values of the noise levels ε and σ . We also note that the restriction $\alpha \leq \beta$ on the levels, is only necessary for the upper bound (item (i)): see Propositions 3.2 and 4.2.

Then, we apply the above result on specific problems. Namely, we consider various behaviors for both sequences $(a_j)_{j \in \mathbb{N}}$ and $(b_j)_{j \in \mathbb{N}}$, and discuss the properties of the associated minimax separation radii $\tilde{r}_{\varepsilon,\sigma}$. Concerning the eigenvalues $(b_j^2)_{j \in \mathbb{N}}$ of the operator A^*A , we will alternatively consider situations where

$$b_j \sim j^{-t} \quad \text{or} \quad b_j \sim \exp(-jt), \quad \forall j \in \mathbb{N}, \text{ for some } t > 0.$$

The first case corresponds to the so-called *mildly ill-posed* problems while the second one corresponds to *severely ill-posed* problems. Concerning the ellipsoids \mathcal{E}_a , i.e., the sequence $(a_j)_{j \in \mathbb{N}}$, two different kinds of smoothness will be investigated, namely,

$$a_j \sim j^s \quad \text{or} \quad a_j \sim \exp(js), \quad \forall j \in \mathbb{N}, \text{ for some } s > 0,$$

the so-called *ordinary-smooth* and *super-smooth* cases, respectively. In the above scenarios, we apply Theorem 2.1 and describe the associated upper and lower bounds on the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$. They are, respectively, displayed in Table 1 and Table 2.

¹For the sake of brevity, these quantities are made precise in the subsequent sections. See, respectively, (3.12), (4.5), (4.11) for the constants $\tilde{C}(\alpha, \beta)$, $C(\alpha, \beta)$, $c(\alpha, \beta)$, and (3.3), (4.4) for the terms M_0, M_1, M_2 .

Table 3
 Minimax goodness-of-fit testing with *known* singular values: the separation rates \tilde{r}_ε^2 for $0 < \varepsilon \leq \varepsilon_0$, $\varepsilon_0 \in]0, 1[$, for all $t, s > 0$

Goodness-of-fit testing problem	Ordinary-smooth $a_j \sim j^s$	Super-smooth $a_j \sim \exp\{js\}$
Mildly ill-posed $b_j \sim j^{-t}$	$\varepsilon^{4s/(2s+2t+1/2)}$	$\varepsilon^2(\ln \varepsilon^{-1})^{2t+1/2}$
Severely ill-posed $b_j \sim \exp\{-jt\}$	$(\ln \varepsilon^{-1})^{-2s}$	$\varepsilon^{2s/(s+t)}$

Looking at these tables, both lower and upper bounds coincide in every considered case, up to a logarithmic term that depends on the noise level σ . Hence, Theorem 2.1 provides a sharp control on the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$ in various settings. The interesting property of such minimax separation radii is that they have the same structure whatever the considered situation: a maximum between two terms depending, respectively, on the noise levels ε and σ . It is also worth pointing out that the first term depending on ε corresponds to the minimax separation radius in the case where the operator is known (i.e., $\sigma = 0$), as displayed in Table 3.

The results displayed in Theorem 2.1 and Tables 1, 2 can also be understood as follows. Two problems are at hand: detection of the underlying signal (with a minimax separation radius that only depends on ε) and detection of the ‘frequencies’ j for which the terms b_j can be replaced by observations X_j without loss of precision (with a minimax separation radius that depends only on σ). The final minimax separation radius is then the maximum of these two terms, i.e., the signal detection hardness is related to the most difficult underlying problem. We stress that such phenomenon has already been discussed in the minimax estimation framework, see, e.g., [13,20].

3. Upper bound on the minimax separation radius

In this section, we first propose an α -level testing procedure. Then, we investigate its maximal second kind error probability and establish a non-asymptotic upper bound on the minimax separation radius (which corresponds to item (i) of Theorem 2.1). Finally, in Section 3.3, we provide a control of the upper bounds for minimax separation radii for the specific cases displayed in Table 1.

3.1. The spectral cut-off test

For a given $\theta_0 \neq 0$, the aim of the goodness-of-fit testing problem (2.5) is to determine whether or not $\theta = \theta_0$. In particular, for any given $j \in \mathbb{N}$, one would like to infer the corresponding value θ_j from the observation $(Y, X) = (Y_j, X_j)_{j \in \mathbb{N}}$ from GSM (1.1). Typically, for any given $j \in \mathbb{N}$, one may use the ‘naive’ estimate $\hat{\theta}_j$ of θ_j , defined by

$$\hat{\theta}_j := \frac{Y_j}{X_j} = \frac{b_j}{X_j} \theta_j + \varepsilon \frac{1}{X_j} \xi_j, \quad j \in \mathbb{N}.$$

In order to ensure a ‘good’ approximation of θ_j by $\hat{\theta}_j$ (in a sense which will be made precise later on), a precise control of the ratio b_j/X_j is required. To this end, we want to avoid coefficients for which $X_j \lesssim \sigma$, namely for which the observation X_j is of the order of the corresponding noise level σ , that does not have ‘discriminatory’ power. Therefore, we will restrict ourselves to coefficients X_j with indices $1 \leq j \leq M$, where the bandwidth M is defined by

$$M := \inf\{j \in \mathbb{N} : |X_j| \leq \sigma h_j\} - 1, \tag{3.1}$$

where, for all $j \in \mathbb{N}$,

$$h_j = 16 \sqrt{\ln\left(\frac{\kappa j^2}{\alpha}\right)} + \sqrt{2 \ln\left(\frac{10}{\alpha}\right)}, \tag{3.2}$$

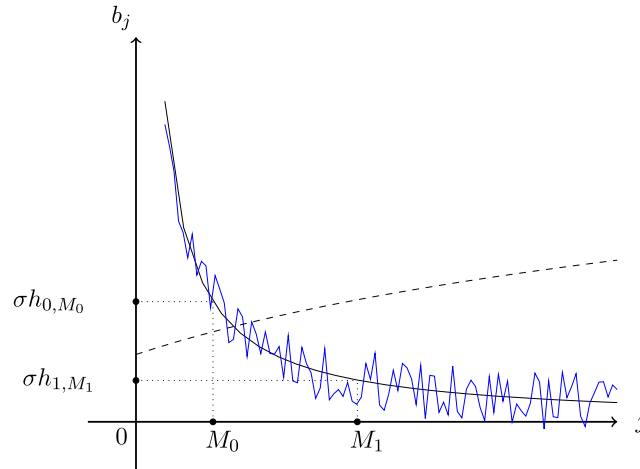


Fig. 1. An illustration of the spatial positions of the bandwidths M_0 and M_1 , defined in (3.3). The decreasing solid curve corresponds to the values of the sequence $b = (b_j)_{j \in \mathbb{N}}$ with respect to the index $j \in \mathbb{N}$, while the oscillating curve demonstrates one realization of the random sequence $X = (X_j)_{j \in \mathbb{N}}$ according to the GSM (1.1). The increasing dashed curve draws the behavior of the sequence σh_j . For the corresponding random ‘bandwidth’ M defined in (3.3), Lemma A.1 shows that $M \in [M_0, M_1[$ with high probability.

for some $\kappa > \exp(1)$. In some sense, the terms σh_j can be hence considered as a threshold for the variables X_j . We stress that unlike classical thresholding procedures, the terms h_j here depend on the level α and allow a control of the first kind error.

Remark 3.1. The value of κ is, in some sense, related to the value of the first kind error probability of the suggested testing procedure. We will see below that the value $\kappa = 5(3\pi^2 + 12)/6$ is convenient to our purpose. We stress that κ is not a regularization parameter: an ‘optimal’ value of κ only allows to get ‘optimal’ constants in the final results but will not change the order of the corresponding minimax separation rates. Finding optimal constants is outside the scope of this work.

The bandwidth M is a random variable but can be controlled in the sense that $M \in [M_0, M_1[$ with high probability (see Lemma A.1 for precise computations and Figure 1 for a graphical illustration), where the bandwidths M_0 and M_1 are defined by

$$\begin{cases} M_0 := \inf\{j \in \mathbb{N} : b_j \leq \sigma h_{0,j}\} - 1, \\ M_1 := \inf\{j \in \mathbb{N} : b_j \leq \sigma h_{1,j}\}, \end{cases} \tag{3.3}$$

and the sequences $h_0 = (h_{0,j})_{j \in \mathbb{N}}$, $h_1 = (h_{1,j})_{j \in \mathbb{N}}$ satisfy

$$h_{0,j} = 18\sqrt{\ln\left(\frac{\kappa j^2}{\alpha}\right)} + \sqrt{2\ln\left(\frac{10}{\alpha}\right)}, \tag{3.4}$$

$$h_{1,j} = 16\sqrt{\ln\left(\frac{\kappa j^2}{\alpha}\right)}, \tag{3.5}$$

for all $j \in \mathbb{N}$. The sequences $h = (h_j)_{j \in \mathbb{N}}$, $h_0 = (h_{0,j})_{j \in \mathbb{N}}$ and $h_1 = (h_{1,j})_{j \in \mathbb{N}}$ in the definition of M_0 , M_1 and M allow a ‘uniform’ control of the standard Gaussian sequence $\eta = (\eta_j)_{j \in \mathbb{N}}$ (associated with $X = (X_j)_{j \in \mathbb{N}}$), for all $1 \leq j \leq M_1$ (see Lemmas A.1, A.2 and A.3 in the Appendix).

We are now in the position to construct a (spectral cut-off) testing procedure. According to the methodology proposed earlier in the literature (see, e.g., [1,19] or [23]), our test will be based on an estimation of $\|\theta - \theta_0\|^2$. For any fixed $D \in \mathbb{N}$, consider the test statistic

$$T_{D,M} := \sum_{j=1}^{D \wedge M} \left(\frac{Y_j}{X_j} - \theta_{j,0} \right)^2. \tag{3.6}$$

Given a prescribed level $\alpha \in]0, 1[$ for the first kind error probability, the associated spectral cut-off test is then defined as

$$\Psi_{D,M} := \mathbf{1}_{\{T_{D,M} > t_{1-\alpha,D}(X)\}}, \tag{3.7}$$

where

$$t_{1-\alpha,D}(X) := \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} + C(\alpha) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4} + (1 + \sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]}, \tag{3.8}$$

and

$$C(\alpha) = 3\sqrt{x_{\alpha/2}} + 2x_{\alpha/2} > 0, \quad x_\gamma := \ln(1/\gamma) \quad \forall \gamma \in]0, 1[. \tag{3.9}$$

In other words, if the ‘estimator’ $T_{D,M}$ of $\|\theta - \theta_0\|^2$ is greater than the random threshold $t_{1-\alpha,D}(X)$, θ and θ_0 are very unlikely to be close to each other, and we will reject H_0 .

Remark 3.2. Under H_0 , $Y_j = b_j \theta_{j,0} + \varepsilon \xi_j$, $j \in \mathbb{N}$, and, hence,

$$T_{D,M} = \sum_{j=1}^{D \wedge M} \left[\left(\frac{b_j}{X_j} - 1 \right) \theta_{j,0} + \varepsilon X_j^{-1} \xi_j \right]^2.$$

Therefore, the law of $T_{D,M}$ is not available and, thus, its corresponding $(1 - \alpha)$ -quantile is not computable in practice, since the sequence $b = (b_j)_{j \in \mathbb{N}}$ is unknown. However, Proposition 3.1 below ensures that the threshold $t_{1-\alpha,D}(X)$ defined in (3.8) provides a computable upper bound on this quantile.

First, we focus on the first kind error probability. The following proposition states that the spectral cut-off test $\Psi_{D,M}$ defined in (3.7)–(3.8), is an α -level test.

Proposition 3.1. *Let $\alpha \in]0, 1[$ be fixed. Consider the goodness-of-fit testing problem (2.5). Then, setting $\kappa = 5(3\pi^2 + 12)/6$, there exists $\sigma_0 \in]0, 1[$ such that, for all $0 < \sigma \leq \sigma_0$ and for each $\varepsilon > 0$, the spectral cut-off test $\Psi_{D,M}$, defined in (3.7)–(3.8), is an α -level test, i.e.,*

$$\alpha_{\varepsilon,\sigma}(\Psi_{D,M}) \leq \alpha. \tag{3.10}$$

The proof is postponed to Section A.2.1.

Remark 3.3. In order to shed light on the term σ_0 , we provide bellow a heuristic argument. Note that, under H_0 , thanks to a (rough) Taylor expansion,

$$T_{D,M} \simeq \sum_{j=1}^{D \wedge M} [\varepsilon b_j^{-1} \xi_j + \sigma b_j^{-1} \theta_{j,0} \eta_j]^2.$$

Compared to the ‘noise-free’ case (i.e., $\sigma = 0$), we have in some sense to deal with the additional term $\sigma b_j^{-1} \theta_{j,0} \eta_j$. Two scenarios are at hand

- If $\sup_j b_j^{-1} a_j^{-1} \leq C_0$, the expected amount of additional signal is

$$\sigma^2 \sum_{j=1}^{D \wedge M} b_j^{-2} \theta_j^2 \leq \sigma^2 C_0 \|\theta\|^2,$$

which is of the order of the classical parametric rate σ^2 . However, since C_0 is unknown, we use a rough standard deviation control on this additional term, which requires a logarithmic term (i.e., $\ln^{3/2}(1/\sigma)$) in the right hand side of (3.8). We stress that this logarithmic term can be removed if the knowledge of C_0 is assumed.

- If $\sup_j b_j^{-1} a_j^{-1} > C_0$, we can prove that $\sigma b_j^{-1} \eta_j$ (see Lemma A.4) is bounded with controlled probability, according to the construction of the bandwidth M given in (3.1). In such case, the additional term can be controlled by the ‘bias’ $a_{D \wedge M}^{-2}$.

Due to the additional logarithmic term mentioned above, the first kind error probability can be controlled as soon as σ is small enough (i.e., $0 < \sigma \leq \sigma_0$ for some $\sigma_0 \in]0, 1[$). Unsurprisingly, it is impossible to retrieve any kind of information on the observations if the noise level σ is too large.

3.2. A non-asymptotic upper bound

We now turn our attention to the maximal second error probability. The following proposition provides, for each noise level $\varepsilon > 0$ and for noise level σ small enough, an upper bound for the separation radius $r_{\varepsilon, \sigma}(\mathcal{E}_a, \Psi_{D, M}, \beta)$ of the spectral cut-off test $\Psi_{D, M}$ defined in (3.6)–(3.8).

Proposition 3.2. *Let $\alpha, \beta \in]0, 1[$ be fixed, such that $\alpha \leq \beta$. Consider the goodness-of-fit testing problem (2.5). Let $\Psi_{D, M}$ be the spectral cut-off test, defined in (3.7)–(3.8). Then, there exists $\sigma_0 \in]0, 1[$ such that, for all $0 < \sigma \leq \sigma_0$ and for each $\varepsilon > 0$,*

$$r_{\varepsilon, \sigma}^2(\mathcal{E}_a, \Psi_{D, M}, \beta) \leq \tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]}, \tag{3.11}$$

where

$$\tilde{C}(\alpha, \beta) = 16(C(\alpha) + 3\sqrt{x_{\beta/2}}). \tag{3.12}$$

The proof is postponed to Section A.2.2.

Note that the upper bound on the separation radius $r_{\varepsilon, \sigma}^2(\mathcal{E}_a, \Psi_{D, M}, \beta)$ given in (3.11) depends on two antagonistic terms, namely, $\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}}$ and $\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}$. Ideally, one would like to make this upper bound as small as possible, i.e., to obtain the weakest possible condition on $\|\theta - \theta_0\|$ such that, for any fixed $\beta \in]0, 1[$, $\beta_{\varepsilon, \sigma}(\Theta_a(r_{\varepsilon, \sigma}), \Psi_{D, M}) \leq \beta$. Therefore, one would like to select $D := D^*$ such that

$$D^* := \arg \min_{D \in \mathbb{N}} \left\{ \tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right\},$$

where $\tilde{C}(\alpha, \beta)$ is defined in (3.12). However, this ‘optimal’ bandwidth D^* is not available in practice since the sequence $b = (b_j)_{j \in \mathbb{N}}$ is not assumed to be known. To this end, we use instead the bandwidth $D := D^\dagger$ defined as

$$D^\dagger := \arg \min_{D \in \mathbb{N}} \left\{ \tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]} \right\}. \tag{3.13}$$

The following theorem illustrates the performances of the corresponding spectral cut-off test $\Psi_{D^\dagger, M}$, defined in (3.7), with $D := D^\dagger$, defined in (3.13).

Theorem 3.1. *Let $\alpha, \beta \in]0, 1[$ be fixed, such that $\alpha \leq \beta$. Consider the goodness-of-fit testing problem (2.5). Let $\Psi_{D^\dagger, M}$ be the spectral cut-off test, defined in (3.7) with $D := D^\dagger$, defined in (3.13). Then, there exists $\sigma_0 \in]0, 1[$ such that, for all $0 < \sigma \leq \sigma_0$ and for each $\varepsilon > 0$,*

$$\alpha_{\varepsilon, \sigma}(\Psi_{D^\dagger, M}) \leq \alpha \tag{3.14}$$

and

$$r_{\varepsilon, \sigma}^2(\mathcal{E}_a, \Psi_{D^\dagger, M}, \beta) \leq \inf_{D \in \mathbb{N}} \left[\tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right], \tag{3.15}$$

where the constant $\tilde{C}(\alpha, \beta)$ has been introduced in (3.12).

The proof of Theorem 3.1 is postponed to Section A.2.3.

Remark 3.4. According to Theorem 3.1, given a radius $r_{\varepsilon, \sigma} > 0$, then

$$\begin{aligned} r_{\varepsilon, \sigma}^2 &\geq \inf_{D \in \mathbb{N}} \left[\tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right] \\ &\Rightarrow \beta_{\varepsilon, \sigma}(\Theta_a(r_{\varepsilon, \sigma}), \Psi_{D^\dagger, M}) \leq \beta, \end{aligned}$$

and, hence,

$$\tilde{r}_{\varepsilon, \sigma}^2 \leq \inf_{D \in \mathbb{N}} \left[\tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right]. \tag{3.16}$$

This upper bound corresponds to item (i) of Theorem 2.1.

Remark 3.5. Although we considered the spectral cut-off test (3.7), we stress that several alternative testing strategies could be investigated (see, e.g., [23]). Since our aim is to propose an accurate description of the minimax separation radius/rates associated to the goodness-of-fit problem (2.1), the proposed testing procedure should be considered as a tool that allows to propose an upper bound. Nevertheless, any testing procedure designed for the goodness-of-fit testing problem (2.5) depends on the sequence $(b_k)_{k \in \mathbb{N}}$. The uncertainty on the operator, however, requires a specific treatment and cannot be managed by an immediate application of classical procedures (see, e.g., [18]).

3.3. Upper bounds: Specific cases

Our aim in this section is to determine an explicit value (in terms of the noise levels ε and σ) for the upper bounds on the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ obtained in Theorem 3.1 above. To this end, we will consider well-known specific cases regarding the behavior of both sequences $(a_j)_{j \in \mathbb{N}}$ and $(b_j)_{j \in \mathbb{N}}$. According to the existing literature, we will essentially deal with mildly and severely ill-posed problems with ellipsoids of ordinary-smooth and super-smooth functions (see also Section 2.2 for formal definitions).

Theorem 3.2. *Consider the goodness-of-fit testing problem (2.5) when observations are given by (1.1), and the signal of interest has smoothness governed by (2.3).*

(i) If $b_j \sim j^{-t}$, $t > 0$, and $a_j \sim j^s$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^{\frac{4s}{2s+2t+1/2}} \vee [\sigma \ln^{3/4}(1/\sigma)]^{2(\frac{s}{t} \wedge 1)}.$$

(ii) If $b_j \sim j^{-t}$, $t > 0$, and $a_j \sim \exp\{js\}$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^2 [\ln(1/\varepsilon)]^{(2t+\frac{1}{2})} \vee \sigma^2 \ln^{\frac{3}{2}}(1/\sigma).$$

(iii) If $b_j \sim \exp\{-jt\}$, $t > 0$, and $a_j \sim j^s$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim [\ln(1/\varepsilon)]^{-2s} \vee \left[\ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \right]^{-2s}.$$

(iv) If $b_j \sim \exp\{-jt\}$, $t > 0$, and $a_j \sim \exp\{js\}$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^{\frac{2s}{s+t}} \vee [\sigma \ln^{1/2}(1/\sigma)]^{2(\frac{s}{t} \wedge 1)}.$$

The minimax rates displayed above are summarized in Table 1. The proof is postponed to Section A.3. The main task is to compute the asymptotic trade-off between both antagonistic terms $\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}}$ and $[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]$ in the upper bounds on the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ displayed in (3.16).

4. Lower bounds on the minimax separation radius

We establish a non-asymptotic lower bound on the minimax separation radius (which corresponds to item (ii) of Theorem 2.1). In order to do this, we consider two special cases of the GSM (1.1), namely the situations where

- (a) $\varepsilon = 0$: the signal is observed without noise but the eigenvalues of the operator at hand are still noisy, and
- (b) $\sigma = 0$: the ‘classical’ model (see, e.g., [18] or [21]) where the eigenvalues of the operator at hand are known.

Both models (a) and (b) correspond to some ‘extreme’ situations but provide, in some sense, a benchmark for the problem at hand. We first establish a lower bound for the case (a) in Section 4.1 and recall the lower bound for the case (b) (that has already been discussed in, e.g., [1,21] or [23]) in Section 4.2. Then, we establish in Section 4.3 that the minimax separation radius associated to goodness-of-fit testing problem (2.5) is always greater than the maximum of the minimax separation radii associated to the cases (a) and (b). Finally, in Section 4.4, we provide a control of the lower bounds for minimax separation radii for the specific cases displayed in Table 2.

4.1. Lower bounds for a GSM with $\varepsilon = 0$

We consider the GSM (1.1) with $b = \bar{b}$ and $\varepsilon = 0$, i.e.,

$$\begin{cases} Y_j = \bar{b}_j \theta_j, & j \in \mathbb{N}, \\ X_j = \bar{b}_j + \sigma \eta_j, & j \in \mathbb{N}. \end{cases} \tag{4.1}$$

For a given sequence $b = (b)_{j \in \mathbb{N}}$, define

$$\mathcal{B}(b) = \{v \in l^2(\mathbb{N}) : C_0 |b_j| \leq |v_j| \leq C_1 |b_j|, j \in \mathbb{N}, 0 < C_0 \leq 1 \leq C_1 < +\infty\}.$$

Given observations from the GSM (4.1), for any given $\theta_0 \in \mathcal{E}_a \setminus \{0\}$ and $\bar{b} \in \mathcal{B}(b)$, we consider the following goodness-of-fit testing problem

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta - \theta_0 \in \Theta_a(r_\sigma), \quad \bar{b} \in \mathcal{B}(b), \tag{4.2}$$

where $\Theta_a(r_\sigma) = \{\mu \in \mathcal{E}_a, \|\mu\| \geq r_\sigma\}$.

Our aim below is to provide a lower bound on the minimax separation radius $\tilde{r}_{0,\sigma}$, defined as

$$\tilde{r}_{0,\sigma} := \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} r_{0,\sigma}(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta),$$

where $r_{0,\sigma}(\mathcal{E}_a, \Psi_\alpha, \beta)$ is the separation radius of any given α -level test Ψ_α , defined as

$$r_{0,\sigma}(\mathcal{E}_a, \Psi_\alpha, \beta) := \inf\{r_\sigma > 0 : \beta_{0,\sigma,b}(\Theta_a(r_\sigma), \mathcal{B}(b), \Psi_\alpha) \leq \beta\},$$

and $\beta_{0,\sigma,b}(\Theta_a(r_\sigma), \mathcal{B}(b), \Psi_\alpha)$ is the associated maximal second kind error probability, defined as

$$\beta_{0,\sigma,b}(\Theta_a(r_\sigma), \mathcal{B}(b), \Psi_\alpha) := \sup_{\substack{\theta: \theta - \theta_0 \in \Theta_a(r_\sigma) \\ \bar{b} \in \mathcal{B}(b)}} \mathbb{P}_{\theta, \bar{b}}(\Psi_\alpha = 0).$$

The following proposition states a lower bound for the minimax separation radius $\tilde{r}_{0,\sigma}$ of the goodness-of-fit testing problem (4.2).

Proposition 4.1. *Assume that $(Y, X) = (Y_j, X_j)_{j \in \mathbb{N}}$ are observations from the GSM (4.1) and consider the goodness-of-fit testing problem (4.2). Let $\alpha \in]0, 1[$ and $\beta \in]0, 1 - \alpha[$ be given. Then, for every $\sigma > 0$, the minimax separation radius $\tilde{r}_{0,\sigma}$ is lower bounded by*

$$\tilde{r}_{0,\sigma} \geq \frac{C_{\alpha,\beta}}{2} \sigma \max_{1 \leq D \leq M_2} [b_D^{-1} a_D^{-1}], \tag{4.3}$$

where

$$M_2 := \sup \left\{ D \in \mathbb{N} : C_{\alpha,\beta} \sigma b_D^{-1} \leq 1 \text{ and } G_D(C_0, C_1) \geq 1 - \frac{1}{2}(1 - \alpha - \beta) \right\} \tag{4.4}$$

with

$$C_{\alpha,\beta} = \ln(1 + (1 - \alpha - \beta)^2) > 0 \quad \text{and} \quad G_D(C_0, C_1) = \frac{1}{\sigma \sqrt{2\pi}} \int_{C_0 b_D}^{C_1 b_D} \exp\left\{-\frac{1}{2\sigma^2}(t - b_D)^2\right\} dt, \tag{4.5}$$

for some constants $0 < C_0 \leq 1 \leq C_1 < +\infty$.

The proof is postponed to Section A.4.1.

Remark 4.1. Note that

$$G_D(C_0, C_1) = \Phi\left((C_1 - 1)\frac{b_D}{\sigma}\right) - \Phi\left((C_0 - 1)\frac{b_D}{\sigma}\right),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard Gaussian distribution. Hence,

$$G_D(C_0, C_1) \geq 1 - \frac{1}{2}(1 - \alpha - \beta) \Leftrightarrow b_D \geq \sigma K,$$

where $K := K(C_0, C_1, \alpha, \beta) > 0$. Then M_2 in (4.4) can be re-expressed as

$$M_2 := \sup\{D \in \mathbb{N} : b_D \geq \sigma[K \vee C_{\alpha,\beta}]\}. \tag{4.6}$$

This expression M_2 in (4.6) can be compared to the respective expressions of M_0 and M_1 defined in (3.3). In particular, we point-out that there is no logarithmic term involved in M_2 .

4.2. Lower bounds for the GSM when $\sigma = 0$

We consider the GSM (1.1) with $\sigma = 0$, i.e.,

$$\begin{cases} Y_j = b_j\theta_j + \varepsilon\xi_j, & j \in \mathbb{N}, \\ X_j = b_j, & j \in \mathbb{N}. \end{cases} \quad (4.7)$$

Note that, in this case, the above model can be re-expressed as

$$Y_j = b_j\theta_j + \varepsilon\xi_j, \quad j \in \mathbb{N}, \quad (4.8)$$

where $b = (b_j)_{j \in \mathbb{N}}$ is a known positive sequence.

The following proposition states a lower bound for the minimax separation radius $\tilde{r}_{\varepsilon,0}$, defined in (2.2) with $\sigma = 0$, of the following goodness-of-fit testing problem

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta - \theta_0 \in \Theta_a(r_{\varepsilon,0}), \quad (4.9)$$

where $\theta_0 \in \mathcal{E}_a \setminus \{0\}$ and $\Theta_a(r_{\varepsilon,0})$ is defined in (2.4) with $\sigma = 0$.

Proposition 4.2. *Assume that $Y = (Y_j)_{j \in \mathbb{N}}$ are observations from the GSM (4.8) and consider the goodness-of-fit testing problem (4.9). Let $\alpha \in]0, 1[$ and $\beta \in]0, 1 - \alpha[$ be given. Then, for every $\varepsilon > 0$, the minimax separation radius $\tilde{r}_{\varepsilon,0}$ is lower bounded by*

$$\tilde{r}_{\varepsilon,0}^2 \geq \sup_{D \in \mathbb{N}} \left[c_{\alpha,\beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right], \quad (4.10)$$

where

$$c_{\alpha,\beta} = (2 \ln(1 + 4(1 - \alpha - \beta)^2))^{1/4} > 0. \quad (4.11)$$

The proof of the Proposition 4.2 with detailed arguments are related discussion can be found in e.g., [1,21] and [23].

4.3. A combined lower bound

The following result provides a lower bound on the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$ for the goodness-of-fit testing problem (2.5). This lower bound corresponds to item (ii) of Theorem 2.1.

Theorem 4.1. *Consider the GSMs (1.1), (4.1) and (4.7). Denote by $\tilde{r}_{\varepsilon,\sigma}$, $\tilde{r}_{0,\sigma}$ and $\tilde{r}_{\varepsilon,0}$ the corresponding minimax separation radii. Then, for every $\varepsilon > 0$ and $\sigma > 0$,*

$$\tilde{r}_{\varepsilon,\sigma} \geq \tilde{r}_{0,\sigma} \vee \tilde{r}_{\varepsilon,0}. \quad (4.12)$$

In particular,

$$\tilde{r}_{\varepsilon,\sigma}^2 \geq \left\{ \frac{C_{\alpha,\beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \right\} \vee \left\{ \sup_{D \in \mathbb{N}} \left[c_{\alpha,\beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \right\}, \quad (4.13)$$

where $C_{\alpha,\beta}$ is given in (4.5), M_2 is given in (4.4) and $c_{\alpha,\beta}$ is given in (4.11).

The proof of Theorem 4.1 is postponed to Section A.4.2.

Remark 4.2. At a first sight, the upper and lower bounds respectively displayed in (i) and (ii) of Theorem 2.1 do not exactly match up. However, a closer look at the involved formulas indicates that both quantities contain terms that have similar behaviors. This is, in some sense, confirmed in Section 4.4 below where specific sequences $(a_j)_{j \in \mathbb{N}}$ and $(b_j)_{j \in \mathbb{N}}$ are treated.

4.4. Lower bounds: Specific cases

Our aim in this section is to determine an explicit value (in terms of the noise levels ε and σ) for the lower bounds on the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ obtained in Theorem 4.1 above for the specific sequences $(a_j)_{j \in \mathbb{N}}$ and $(b_j)_{j \in \mathbb{N}}$ considered in Section 3.3.

Theorem 4.2. Consider the goodness-of-fit testing problem (2.5) when observations are given by (1.1), and the signal of interest has smoothness governed by (2.3).

- (i) If $b_j \sim j^{-t}$, $t > 0$, and $a_j \sim j^s$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim \varepsilon^{\frac{4s}{2s+2t+1/2}} \vee \sigma^{2(\frac{s}{t} \wedge 1)}.$$

- (ii) If $b_j \sim j^{-t}$, $t > 0$, and $a_j \sim \exp\{js\}$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim \varepsilon^2 [\ln(1/\varepsilon)]^{(2t+\frac{1}{2})} \vee \sigma^2.$$

- (iii) If $b_j \sim \exp\{-jt\}$, $t > 0$, and $a_j \sim j^s$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim [\ln(1/\varepsilon)]^{-2s} \vee [\ln(1/\sigma)]^{-2s}.$$

- (iv) If $b_j \sim \exp\{-jt\}$, $t > 0$, and $a_j \sim \exp\{js\}$, $s > 0$, for all $j \in \mathbb{N}$, then, there exist $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim \varepsilon^{\frac{2s}{s+t}} \vee \sigma^{2(\frac{s}{t} \wedge 1)}.$$

The minimax rates displayed above are summarized in Table 2. The proof is postponed to Section A.5. As in the case of the upper bound, the main task is to compute the trade-off between both different antagonistic terms involved in the lower bound on the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ displayed in (4.13).

5. Concluding remarks

5.1. Theoretical outcomes

The main conclusion of this work is that goodness-of-fit testing in an inverse problem setting is ‘feasible,’ even in the specific situation where some uncertainty is observed on the operator at hand in the GSM (1.1). We have established ‘optimal’ separation conditions for the goodness-of-fit testing problem (2.5) via a sharp control of the associated minimax separation radius.

We stress that several outcomes and open questions are still of interest. We can mention, among others:

- *Adaptivity:* As proved in Theorem 3.1, the test $\Psi_{D^\dagger, M}$ introduced in (3.6)–(3.9) with D^\dagger defined in (3.13) is powerful in the sense that its separation radius is equal (up to constant) to the minimax one. However, this test strongly depends on the sequence $a = (a_j)_{j \in \mathbb{N}}$ that characterizes the smoothness of the signal of interest. In practice, this sequence is unknown and *adaptive* procedures are necessary. A classical way to proceed to this adaptation step is

to build an α -level test $\Psi_{\alpha,D}$ for each candidate D in a given set \mathcal{D} . Then, one define the adaptive (Bonferroni type) test Ψ_α^* as

$$\Psi_\alpha^* = \max_{D \in \mathcal{D}} \Psi_{\alpha, \#\mathcal{D}},$$

where $\#\mathcal{D}$ denotes the cardinality of the set \mathcal{D} . It can be proved that, provided the set \mathcal{D} is well-chosen, the test Ψ_α^* is of level α and attains the minimax rate, up to an additional logarithmic term. We refer to, e.g., [19] or [18] for more details on adaptation (in the GSM (1.1) with $\sigma = 0$).

- *Non-Gaussian models:* In the GSM (1.1), the observations are assumed to be Gaussian. However, it may sometimes be of practical (or mathematical) interest to consider alternative distributions. The upper bounds obtained earlier are established using a sharp control on the deviation of the sum of independent squared Gaussian random variables, see Lemma A.4. The main ingredients for the proof of this lemma are refined computations and control of the Laplace transform of such variables (see, e.g., [4,21,22]). We conjecture that such a lemma could be generalized to alternative distributions. On the other hand, the Gaussian structure of the measurements is at the heart of the lower bounds proposed above. Proposing lower bounds with generic conditions on the noise distributions appears to be more challenging.
- *Signal detection:* We have already mentioned in Remark 2.2 that *signal detection* is different from *goodness-of-fit testing* (2.5) when the GSM (1.1) is at hand. In this work, we were concerned with the case where $\theta_0 \neq 0$ (goodness-of-fit testing). However, some attention should also be paid in the future to the case where $\theta_0 = 0$ (signal detection). In particular, testing methodologies and related minimax separation radii are quite different from those presented above.

All these topics require special attention that is beyond the scope of this paper. Nevertheless, they provide an avenue for future research.

5.2. Applied outcomes

As discussed above, minimax testing in an ill-posed inverse problem context has been at the core of several theoretical investigations over the last decade. In this section, we present a short selection of applied examples that involve goodness-of-fit approaches in an inverse problem setting that motivated our investigation.

Blind deconvolution for ophthalmoscopy

In [25], the authors considered images I following the model

$$I = O * PSF + N,$$

where O denotes the object of interest, PSF (not exactly known in practice) the point spread function (which models deformations of the object of interest) and N represents stochastic noise. In practice, the signal-to-noise ratio is not very high: one has to find solution in order to improve the quality of the data. A possible way is to gather several images pointing on the same object and to average them. Nevertheless, one has to be sure that a given image can be added to the existing sample.

Such a situation can be modeled as follows: suppose one has at his/her disposal N images, say I_1, I_2, \dots, I_N , which correspond to observations on the same object f_0 . Writing these observations in the frequency domain

$$I_{jk} = b_k \theta_{k,0} + \varepsilon \xi_{jk}, \quad j \in \{1, 2, \dots, N\}, k \in \mathbb{N}.$$

Summing-up over $j \in \{1, 2, \dots, N\}$, we obtain the average image, say I' , satisfying in the frequency domain

$$I'_k = b_k \theta_{k,0} + \varepsilon \xi'_k \quad \text{with } \xi'_k = \frac{1}{N} \sum_{j=1}^N \xi_{jk}, k \in \mathbb{N}.$$

Then, one might want to compare a new observed image, say I'' , to the average image I' , in order to decide whether or not it can be included in the sample. In particular, the idea is to determine if a non-expected movement has occurred

during the acquisition of the data. In other words, one wants to determine whether or not $\theta = \theta_0$. Provided N is large enough, one can (in an idealized framework) assume that the noise $\varepsilon\xi'_k$ has no real influence on the considered task, and can be thus considered to be equal to 0, in which case we recover the GSM (1.1). A similar discussion is provided in [5] in a biophotonic imaging setting.

The instrumental variable (IV) regression model

This setting deals with a sample $(Y_j, X_j, W_j)_{j=1,2,\dots,n}$ satisfying

$$Y_j = Tg(W_j) + U_j, \quad j \in \{1, 2, \dots, n\}, \tag{5.1}$$

where g is an unknown function of interest, U_j are independent and identically distributed random variables such that $\mathbb{E}[U_j/W_j] = 0$, and T is an unknown operator that can be estimated using the couples (X_j, W_j) .

Econometricians have been interested in testing specific values for the unknown function g . For instance, [16] considered a parametric model. The generalization of our investigations to the this instrumental regression setting may allow to handle a complete non-parametric framework.

ECG data analysis

In [2], shifted observations are at hand. More precisely, one observes N 1-dimensional signals that can be modeled as follows

$$Y_j = f(\cdot - \tau_j) + \varepsilon\xi_j, \quad j \in \{1, 2, \dots, N\},$$

where the τ_j denotes a random sample obtained from a random variable τ having density g . Writing the above observations in the frequency domain and summing-up over $j \in \{1, \dots, N\}$, one gets

$$Y'_k = \theta_k \frac{1}{N} \sum_{j=1}^N e^{-2i\pi k\tau_j} + \frac{\varepsilon}{\sqrt{N}} \zeta_k, \quad k \in \mathbb{N}.$$

The law of large number guarantee that $\frac{1}{N} \sum_{j=1}^N e^{-2i\pi k\tau_j}$ converges to the k th-Fourier coefficient of g . Provided that g is known, we can then recover θ (see, e.g., [3]).

Now, assume that these N observed signals correspond to ‘normal’ ECG measure for a patient in a good health, or in a particular period. Then, imagine that given another hearth cycle measurement, one would like to detect a possible arrhythmic cycle. In this case, we get a new signal $Y' = f(\cdot - \tau') + \varepsilon\xi'$, which in the frequency domain is written as

$$Y'_k = \theta'_k e^{-2i\pi k\tau'} + \varepsilon z'_k, \quad k \in \mathbb{N}.$$

In such case, one might want to test whether or not $\theta = \theta'$. Note, in particular, that the operator at hand (its eigenvalues) is random.

Errors-in-variables model

Density model with measurement errors have been at the core of several statistical studies in the past decades (see, e.g., [24] for an overview). Formally, given a sample of independent and identical distributed random variables $(Y_j)_{j=1,2,\dots,n}$ satisfying

$$Y_j = X_j + \varepsilon_j, \quad j = 1, 2, \dots, n,$$

the aim is to produce some inference on the *unknown* density of the X_i , denoted by f , the ε_i corresponding to some error, with *known* density η . This appears to be an inverse (deconvolution) problem since the Y_i are associated to the convolved density $f * \eta$. In a goodness-of-fit testing task, this model has been discussed in [6] and minimax separation rates (in the asymptotic minimax testing framework) have been established in various settings. In the spirit of our contribution, it could be interesting to propose goodness-of-fit testing methods taking into account some possible uncertainty on the density η at hand.

The above examples are somehow considered in slightly different settings compared to the GSM (1.1) studied in this paper. However, the related questions might provide an avenue for future research. In any case, determining the separation rates in these or similar settings will require at some step a careful investigation of the GSM (1.1).

Appendix: Proofs

A.1. Useful lemmas

The constant $C > 0$ and $0 < \tau < 1$ below will vary from place to place.

The following lemma is inspired by Lemma 6.1 of [10].

Lemma A.1. *Let M, M_0, M_1 be defined as in (3.1)–(3.5) where $\alpha \in]0, 1[$ and $\kappa \geq \exp(1)$ are fixed values. Define the event*

$$\mathcal{M} = \{M_0 \leq M < M_1\}. \quad (\text{A.1})$$

Then, for any $\sigma \in]0, 1[$,

$$\mathbb{P}(\mathcal{M}^c) \leq \frac{\alpha}{10} + \frac{\alpha\pi^2}{6\kappa}. \quad (\text{A.2})$$

Proof. It is easily seen that

$$\begin{aligned} \mathbb{P}(M \geq M_1) &= \mathbb{P}\left(\bigcap_{j=1}^{M_1} \{|X_j| > \sigma h_j\}\right) \leq \mathbb{P}(|X_{M_1}| > \sigma h_{M_1}) \\ &\leq \mathbb{P}(|b_{M_1}| + \sigma |\eta_{M_1}| > \sigma h_{M_1}) \\ &\leq \mathbb{P}(|\eta_{M_1}| > h_{M_1} - h_{1, M_1}) \\ &= \mathbb{P}\left(|\eta_{M_1}| > \sqrt{2 \ln\left(\frac{10}{\alpha}\right)}\right), \end{aligned}$$

where the sequences $(h_j)_{j \in \mathbb{N}}$ and $(h_{1, j})_{j \in \mathbb{N}}$ are defined in (3.2) and (3.5) respectively. Using the bound

$$\frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{s^2}{2}} ds \leq \frac{1}{x} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \quad \forall x > 0, \quad (\text{A.3})$$

we get

$$\mathbb{P}(M \geq M_1) \leq \frac{2}{\sqrt{2\pi}} \frac{\alpha}{10} \frac{1}{\sqrt{2 \ln(10/\alpha)}} \leq \frac{\alpha}{10}, \quad (\text{A.4})$$

since $\sqrt{2 \ln(10/\alpha)} > 1$ for all $\alpha \in]0, 1[$. In the same spirit,

$$\begin{aligned} \mathbb{P}(M < M_0) &= \mathbb{P}\left(\bigcup_{j=1}^{M_0} \{|X_j| \leq \sigma h_j\}\right) \leq \sum_{j=1}^{M_0} \mathbb{P}(|X_j| \leq \sigma h_j) \\ &\leq \sum_{j=1}^{M_0} \mathbb{P}(|b_j| - \sigma |\eta_j| \leq \sigma h_j) \\ &\leq \sum_{j=1}^{M_0} \mathbb{P}(\sigma |\eta_j| \geq |b_j| - \sigma h_j) \\ &\leq \sum_{j=1}^{M_0} \mathbb{P}(|\eta_j| \geq h_{0, j} - h_j). \end{aligned}$$

According to the respective definition of $(h_j)_{j \in \mathbb{N}}$, $(h_{0,j})_{j \in \mathbb{N}}$ (see (3.2) and (3.4)), and using again inequality (A.3), we obtain

$$\begin{aligned} \mathbb{P}(M \geq M_1) &\leq \frac{2}{\sqrt{2\pi}} \sum_{j=1}^{M_0} \frac{1}{h_{0,j} - h_j} \exp\left\{-\frac{1}{2}(h_{0,j} - h_j)^2\right\} \\ &\leq \frac{2}{\sqrt{2\pi}} \sum_{j=1}^{M_0} \frac{1}{2\sqrt{\ln(\frac{\kappa j^2}{\alpha})}} \frac{\alpha}{\kappa j^2} \\ &\leq \frac{\alpha}{\kappa} \sum_{j \in \mathbb{N}} \frac{1}{j^2} = \frac{\alpha\pi^2}{6\kappa}, \end{aligned} \tag{A.5}$$

on noting that $\sum_{j \in \mathbb{N}} \frac{1}{j^2} = \pi^2/6$. Since

$$\mathbb{P}(\mathcal{M}^c) \leq \mathbb{P}(M < M_0) + \mathbb{P}(M \geq M_1),$$

the lemma follows, thanks to (A.4) and (A.5). □

Lemma A.2. *Let M be defined as in (3.1) and (3.4) where $\alpha \in]0, 1[$ and $\kappa \geq \exp(1)$ are fixed values. Define the event*

$$\mathcal{B} = \bigcap_{j=1}^M \left\{ \sigma |\eta_j| \leq \frac{b_j}{2} \right\}. \tag{A.6}$$

Then, for any $\sigma \in]0, 1[$

$$\mathbb{P}(\mathcal{B}^c) \leq \frac{\alpha}{10} + \frac{\alpha\pi^2}{3\kappa}. \tag{A.7}$$

Proof. Using the definitions of \mathcal{M} and M_1 , simple calculations give

$$\begin{aligned} \mathbb{P}(\mathcal{B}^c) &= \mathbb{P}(\mathcal{B}^c \cap \mathcal{M}) + \mathbb{P}(\mathcal{B}^c \cap \mathcal{M}^c) \\ &\leq \mathbb{P}\left(\bigcup_{j=1}^{M_1-1} \left\{ \sigma |\eta_j| > \frac{b_j}{2} \right\}\right) + \mathbb{P}(\mathcal{M}^c) \\ &\leq \sum_{j=1}^{M_1-1} \mathbb{P}\left(|\eta_j| > \frac{1}{2}h_{1,j}\right) + \mathbb{P}(\mathcal{M}^c). \end{aligned}$$

Using (3.5), Lemma A.1 and (A.3), we obtain

$$\mathbb{P}(\mathcal{B}^c) \leq \frac{2}{\sqrt{2\pi}} \sum_{j=1}^{M_1} \frac{1}{\sqrt{8^2 \ln(\frac{\kappa j^2}{\alpha})}} \frac{\alpha}{\kappa j^2} + \frac{\alpha}{10} + \frac{\alpha\pi^2}{6\kappa} \leq \frac{\alpha}{10} + \frac{\alpha\pi^2}{3\kappa}. \tag{A.8}$$

Hence, the lemma holds true. □

Lemma A.3. *Let $\theta \in \mathcal{E}_a$ be given. Let M be defined as in (3.1) and (3.4) where $\alpha \in]0, 1[$ and $\kappa \geq \exp(1)$ are fixed values. Then, for any $\sigma \in]0, 1[$ and for any $D \in \mathbb{N}$,*

$$\mathbb{P}\left(\sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_j^2 \geq \sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}\right) \leq \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2\right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\},$$

for some $C > 0$ and $0 < \tau < 1$.

Proof. Using Lemma A.1, Lemma A.2 and a Taylor expansion as in Lemma 6.6 of [10], we get, for all $j \leq M$,

$$\frac{b_j}{X_j} = \frac{1}{1 + \sigma b_j^{-1} \eta_j} = 1 - \sigma b_j^{-1} \eta_j + \sigma^2 \zeta_j^{-2} \eta_j^2,$$

where $\zeta_j^{-1} \leq 8b_j^{-1}$ on the even \mathcal{B} defined in (A.6). Hence

$$\begin{aligned} & \mathbb{P}\left(\sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_j^2 \geq \sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}\right) \\ &= \mathbb{P}\left(\sum_{j=1}^{D \wedge M} (-\sigma b_j^{-1} \eta_j + \sigma^2 \zeta_j^{-2} \eta_j^2)^2 \theta_j^2 \geq \sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}\right) \\ &\leq \mathbb{P}\left(2\sigma^2 \sum_{j=1}^{D \wedge M} b_j^{-2} \theta_j^2 \eta_j^2 + 2\sigma^4 \sum_{j=1}^{D \wedge M} \zeta_j^{-4} \theta_j^2 \eta_j^4 \geq \sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}\right). \end{aligned}$$

Therefore

$$\begin{aligned} & \mathbb{P}\left(\sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_j^2 \geq \sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}\right) \\ &\leq \mathbb{P}\left(\left\{2\sigma^2 \sum_{j=1}^{D \wedge M} b_j^{-2} \theta_j^2 \eta_j^2 \geq \frac{1}{2}[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\quad + \mathbb{P}\left(\left\{2\sigma^4 \sum_{j=1}^{D \wedge M} \zeta_j^{-4} \theta_j^2 \eta_j^4 \geq \frac{1}{2}[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) + \mathbb{P}((\mathcal{B} \cap \mathcal{M})^c) \\ &:= T_1 + T_2 + \mathbb{P}((\mathcal{B} \cap \mathcal{M})^c). \end{aligned} \tag{A.9}$$

We concentrate below our attention on the term T_1 defined as

$$T_1 := \mathbb{P}\left(\left\{2\sigma^2 \sum_{j=1}^{D \wedge M} b_j^{-2} \theta_j^2 \eta_j^2 \geq \frac{1}{2}[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]\right\} \cap (\mathcal{B} \cap \mathcal{M})\right).$$

We consider the two following possible scenarios: (i) $a_j^{-1} b_j^{-1} \leq C_0$ as $j \rightarrow +\infty$, for some $C_0 > 0$, and (ii) $a_j^{-1} b_j^{-1} \rightarrow +\infty$ as $j \rightarrow +\infty$.

Consider first scenario (i). Then, using again (A.3)

$$\begin{aligned} T_1 &\leq \mathbb{P}\left(\left\{2\sigma^2 \max_{1 \leq j \leq D \wedge M} (b_j^{-2} a_j^{-2} \eta_j^2) \geq \frac{1}{2}[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \mathbb{P}\left(\left\{2C_0^2 \sigma^2 \max_{1 \leq j \leq D \wedge M} (\eta_j^2) \geq \frac{1}{2} \sigma^2 \ln^{3/2}(1/\sigma)\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \sum_{j=1}^{M_1-1} \mathbb{P}\left(|\eta_j| \geq \frac{1}{2C_0} \ln^{3/4}(1/\sigma)\right) \\ &\leq \frac{2M_1}{\sqrt{2\pi}} \frac{2C_0}{\ln^{3/4}(1/\sigma)} \exp\left(-\frac{\ln^{3/2}(1/\sigma)}{8C_0^2}\right) \leq C \exp\{-\ln^{1+\tau}(1/\sigma)\}, \end{aligned} \tag{A.10}$$

for some constants $C, \tau \in \mathbb{R}^+$. A similar bound occurs for the term T_2 for this scenario.

Consider now the second scenario (ii). Then

$$\begin{aligned} T_1 &\leq \mathbb{P}\left(\left\{2b_{D\wedge M}^{-2}a_{D\wedge M}^{-2}\sigma^2 \max_{1\leq j\leq D\wedge M}(\eta_j^2) \geq \frac{1}{2}a_{D\wedge M}^{-2}\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \mathbb{P}\left(\left\{\sigma^2 \max_{1\leq j\leq D\wedge M}(\eta_j^2) \geq \frac{1}{4}b_{D\wedge M}^2\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \sum_{j=1}^{M_1-1} \mathbb{P}\left(\sigma^2 \eta_j^2 \geq \frac{1}{4}b_{M_1-1}^2\right), \end{aligned}$$

since the sequence $(b_j)_{j\in\mathbb{N}}$ is non-increasing. Using (A.3), we get

$$\begin{aligned} T_1 &\leq \sum_{j=1}^{M_1-1} \mathbb{P}\left(|\eta_j| \geq \frac{1}{2}h_{1,M_1-1}\right) \\ &\leq \frac{2M_1}{\sqrt{2\pi}} \frac{2}{h_{1,M_1-1}} \exp\left(-\frac{h_{1,M_1-1}^2}{8}\right) \\ &\leq M_1 \exp\left(-\ln\left(\frac{\kappa M_1^2}{\alpha}\right)\right) \\ &\leq M_1 \times \frac{\alpha}{\kappa M_1^2} \leq \frac{\alpha}{\kappa}. \end{aligned} \tag{A.11}$$

By similar computations, we get

$$\begin{aligned} T_2 &:= \mathbb{P}\left(\left\{2\sigma^4 \sum_{j=1}^{D\wedge M} \zeta_j^{-4}\theta_j^2\eta_j^4 \geq \frac{1}{2}\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D\wedge M}^{-2}\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \mathbb{P}\left(\left\{2 \times 8^4\sigma^4 \sum_{j=1}^{D\wedge M} b_j^{-4}\theta_j^2\eta_j^4 \geq \frac{1}{2}a_{D\wedge M}^{-2}\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \mathbb{P}\left(\left\{2 \times 8^4\sigma^4 \max_{j=1,\dots,D\wedge M} \eta_j^4 \geq \frac{1}{2}b_{D\wedge M}^4\right\} \cap (\mathcal{B} \cap \mathcal{M})\right) \\ &\leq \sum_{j=1}^{M_1-1} \mathbb{P}\left(|\eta_j| \geq \frac{1}{8\sqrt{2}}h_{1,M_1-1}\right) \\ &\leq \frac{2}{\sqrt{2\pi}} \frac{8\sqrt{2}M_1}{h_{1,M_1}} \exp\left(-\frac{1}{4 \times 8^2}4 \times 8^2 \ln\left(\frac{\kappa M_1^2}{\alpha}\right)\right) \leq \frac{\alpha}{\kappa}. \end{aligned} \tag{A.12}$$

Hence, the lemma follows from Lemmas A.1, A.2 and (A.9)–(A.12). □

Lemma A.4. *Let*

$$Z_j = v_j + v_j\omega_j, \quad j \in \mathbb{N},$$

where $\omega = (\omega_j)_{j\in\mathbb{N}}$ is a sequence of independent standard Gaussian random variables, and $(v_j)_{j\in\mathbb{N}}, (v_j)_{j\in\mathbb{N}}$ denote two given real sequences. For all $D \in \mathbb{N}$, define

$$T = \sum_{j=1}^D Z_j^2 \quad \text{and} \quad \Sigma = \sum_{j=1}^D v_j^4 + 2 \sum_{j=1}^D v_j^2 v_j^2.$$

Then, for all $x > 0$,

$$\mathbb{P}\left(T - \mathbb{E}(T) > 2\sqrt{\Sigma x} + 2x \sup_{1 \leq j \leq D} (v_j^2)\right) \leq \exp(-x), \tag{A.13}$$

$$\mathbb{P}(T - \mathbb{E}(T) < -2\sqrt{\Sigma x}) \leq \exp(-x). \tag{A.14}$$

Proof. The proof is given in Lemma 2 of [21]. □

A.2. Non-asymptotic upper bounds

A.2.1. Proof of Proposition 3.1

By definition,

$$\alpha_{\varepsilon, \sigma}(\Psi_{D, M}) := \mathbb{P}_{\theta_0, b}(\Psi_{D, M} = 1) = \mathbb{P}_{\theta_0, b}(T_{D, M} > t_{1-\alpha, D}(X)).$$

Conditionally to the sequence $X = (X_j)_{j \in \mathbb{N}}$, for each $1 \leq j \leq D \wedge M$, the random variable $X_j^{-1}Y_j - \theta_{j,0}$ is Gaussian with mean $v_j = (b_j/X_j - 1)\theta_{j,0}$ and standard deviation $v_j = \varepsilon X_j^{-1}$. In particular, for all $D \in \mathbb{N}$

$$\mathbb{E}_{\theta_0, b}(T_{D, M} | X) := \mathbb{E}_{\theta_0, b}\left[\sum_{j=1}^{D \wedge M} \left(\frac{Y_j}{X_j} - \theta_{j,0}\right)^2 \mid X\right] = \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_{j,0}^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2}. \tag{A.15}$$

For all $D \in \mathbb{N}$, define

$$\Sigma_{D, M} := \varepsilon^4 \sum_{j=1}^{D \wedge M} X_j^{-4} + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_{j,0}^2.$$

Applying Lemma A.4 with $T = T_{D, M}$, $\Sigma = \Sigma_{D, M}$ and $x = x_{\alpha/2} := \ln(2/\alpha)$, we get

$$\mathbb{P}_{\theta_0, b}\left(T_{D, M} - \mathbb{E}_{\theta_0}(T_{D, M} | X) > 2\sqrt{\Sigma_{D, M} x_{\alpha/2}} + 2\varepsilon^2 x_{\alpha/2} \max_{1 \leq j \leq D \wedge M} (X_j^{-2} | X)\right) \leq \frac{\alpha}{2}. \tag{A.16}$$

Using the inequalities $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ and $ab \leq a^2/2 + b^2/2$ for $a, b > 0$, it is easily seen that

$$\begin{aligned} \sqrt{\Sigma_{D, M}} &\leq \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} + \sqrt{\varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_{j,0}^2} \\ &\leq \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} + \sqrt{\varepsilon^2 \max_{1 \leq j \leq D \wedge M} X_j^{-2} \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_{j,0}^2} \\ &\leq \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} + \frac{1}{2}\varepsilon^2 \max_{1 \leq j \leq D \wedge M} X_j^{-2} + \frac{1}{2} \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_{j,0}^2. \end{aligned} \tag{A.17}$$

According to (A.15)–(A.17), we obtain the following bound

$$\mathbb{P}_{\theta_0, b}\left(T_{D, M} > (1 + \sqrt{x_{\alpha/2}}) \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1\right)^2 \theta_{j,0}^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} + C(\alpha)\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4} | X}\right) \leq \frac{\alpha}{2},$$

where the constant $C(\alpha)$ is defined in (3.9). Since $\mathbb{E}[\mathbb{E}(V | W)] = \mathbb{E}(V)$ for any random variables V and W , the previous inequality leads to

$$\mathbb{P}_{\theta_0,b} \left(T_{D,M} > (1 + \sqrt{x_{\alpha/2}}) \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1 \right)^2 \theta_{j,0}^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} + C(\alpha) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} \right) \leq \frac{\alpha}{2}.$$

Then, by defining

$$\mathcal{A} = \left\{ \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1 \right)^2 \theta_{j,0}^2 < \sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2} \right\},$$

and applying Lemma A.3, we immediately get

$$\begin{aligned} \alpha_{\varepsilon,\sigma}(\Psi_{D,M}) &\leq \mathbb{P}_{\theta_0,b}(\{T_{D,M} > t_{1-\alpha,D}(X)\} \cap \mathcal{A}) + \mathbb{P}(\mathcal{A}^c) \\ &\leq \frac{\alpha}{2} + \frac{\alpha}{5} + \frac{\alpha}{6\kappa} (3\pi^2 + 12) + C \exp\{-\ln^{1+\tau}(1/\sigma)\} \\ &= \frac{7\alpha}{10} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\}, \end{aligned}$$

for some $C > 0$ and $0 < \tau < 1$. In particular, setting

$$\kappa = 5 \left(\frac{\pi^2}{2} + 2 \right),$$

there exists $\sigma_0 \in]0, 1[$ such that, for all $\sigma \leq \sigma_0$ and for each $\varepsilon > 0$,

$$\alpha_{\varepsilon,\sigma}(\Psi_{D,M}) \leq \alpha.$$

This concludes the proof of the proposition.

A.2.2. Proof of Proposition 3.2

Let $\theta, \theta_0 \in \mathcal{E}_a$ and $\theta - \theta_0 \in \Theta_a(r_{\varepsilon,\sigma})$. Then

$$\mathbb{P}_{\theta,b}(\Psi_{D,M} = 0) = \mathbb{P}_{\theta,b}(\{\Psi_{D,M} = 0\} \cap (\mathcal{B} \cap \mathcal{M})) + \mathbb{P}_{\theta,b}(\{\Psi_{D,M} = 0\} \cap (\mathcal{B} \cap \mathcal{M})^c) := T_1 + T_2. \tag{A.18}$$

Control of T_2 : Using Lemma A.1, Lemma A.2 and elementary probabilistic arguments, we get

$$\begin{aligned} T_2 &:= \mathbb{P}_{\theta,b}(\{\Psi_{D,M} = 0\} \cap (\mathcal{B} \cap \mathcal{M})^c) \leq \mathbb{P}((\mathcal{B} \cap \mathcal{M})^c) \\ &\leq \mathbb{P}(\mathcal{B}^c) + \mathbb{P}(\mathcal{M}^c) \\ &\leq \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) \leq \frac{\beta}{5} + \frac{\beta}{\kappa} \left(\frac{\pi^2}{2} + 2 \right), \end{aligned} \tag{A.19}$$

since $\beta > \alpha$.

Control of T_1 : Define $t_{\beta/2,D}(\theta, X)$ to be the $\beta/2$ -quantile of $T_{D,M}$, conditionally on X , i.e.,

$$\mathbb{P}_{\theta,b}(T_{D,M} \leq t_{\beta/2,D}(\theta, X) | X) \leq \frac{\beta}{2}.$$

Then, by elementary probabilistic arguments, we get

$$\begin{aligned} T_1 &:= \mathbb{P}_{\theta,b}(\{\Psi_{D,M} = 0\} \cap \{\mathcal{B} \cap \mathcal{M}\}) \\ &= \mathbb{E}[\mathbb{P}_{\theta,b}(\{\Psi_{D,M} = 0\} | X) \mathbf{1}\{\mathcal{B} \cap \mathcal{M}\}] \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E}[\mathbb{P}_{\theta,b}(T_{D,M} \leq t_{1-\alpha,D}(X) \mid X) \mathbf{1}\{\mathcal{B} \cap \mathcal{M}\}] \\
 &\leq \frac{\beta}{2} \mathbb{E}[\mathbf{1}\{t_{1-\alpha,D}(X) \leq t_{\beta/2,D}(\theta, X)\} \mathbf{1}\{\mathcal{B} \cap \mathcal{M}\}] \\
 &\quad + \mathbb{E}[\mathbf{1}\{t_{1-\alpha,D}(X) > t_{\beta/2,D}(\theta, X)\} \mathbf{1}\{\mathcal{B} \cap \mathcal{M}\}] \\
 &\leq \frac{\beta}{2} + \mathbb{E}[\mathbf{1}\{t_{1-\alpha,D}(X) > t_{\beta/2,D}(\theta, X)\} \mathbf{1}\{\mathcal{B} \cap \mathcal{M}\}] \\
 &\leq \frac{\beta}{2} + \mathbb{P}_{\theta,b}(\{t_{1-\alpha,D}(X) > t_{\beta/2,D}(\theta, X)\} \cap \{\mathcal{B} \cap \mathcal{M}\}).
 \end{aligned} \tag{A.20}$$

Our next task is to provide a lower bound for $t_{\beta/2,D}(\theta, X)$. Under H_1 , conditionally to the sequence $X = (X_j)_{j \in \mathbb{N}}$, for each $1 \leq j \leq D \wedge M$, the random variable $X_j^{-1}Y_j - \theta_{j,0}$ is Gaussian with mean v_j and standard deviation v_j defined as

$$v_j = \left(\frac{b_j}{X_j} - 1\right)\theta_j + (\theta_j - \theta_{j,0}) \quad \text{and} \quad v_j = \varepsilon X_j^{-1}.$$

In particular,

$$\begin{aligned}
 \mathbb{E}_{\theta,b}(T_{D,M} \mid X) &= \sum_{j=1}^{D \wedge M} \left[\left(\frac{b_j}{X_j} - 1\right)\theta_j + (\theta_j - \theta_{j,0}) \right]^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} \\
 &= \sum_{j=1}^{D \wedge M} v_j^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2}.
 \end{aligned} \tag{A.21}$$

Let

$$\begin{aligned}
 \tilde{\Sigma}_{D,M} &:= \varepsilon^4 \sum_{j=1}^{D \wedge M} X_j^{-4} + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} \left[\left(\frac{b_j}{X_j} - 1\right)\theta_j + (\theta_j - \theta_{j,0}) \right]^2 \\
 &= \varepsilon^4 \sum_{j=1}^{D \wedge M} X_j^{-4} + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} v_j^2.
 \end{aligned} \tag{A.22}$$

Using Lemma A.4 with $T = T_{D,M}$, $\Sigma = \tilde{\Sigma}_{D,M}$ and $x = x_{\beta/2} := \ln(2/\beta)$, we obtain

$$\begin{aligned}
 &\mathbb{P}_{\theta,b} \left(T_{D,M} < \sum_{j=1}^{D \wedge M} v_j^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} - 2\sqrt{\tilde{\Sigma}_{D,M} x_{\beta/2}} \mid X \right) \leq \frac{\beta}{2} \\
 &\Rightarrow t_{\beta/2,D}(\theta, X) \geq \sum_{j=1}^{D \wedge M} v_j^2 + \varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} - 2\sqrt{\tilde{\Sigma}_{D,M} x_{\beta/2}}.
 \end{aligned} \tag{A.23}$$

Therefore, using (3.8) and (A.23), we get

$$\begin{aligned}
 &\mathbb{P}_{\theta,b}(\{t_{1-\alpha,D}(X) > t_{\beta/2,D}(\theta, X)\} \cap \{\mathcal{B} \cap \mathcal{M}\}) \\
 &\leq \mathbb{P}_{\theta,b} \left(\left\{ \sum_{j=1}^{D \wedge M} v_j^2 < (C(\alpha) + 2\sqrt{x_{\beta/2}}) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} \right. \right.
 \end{aligned}$$

$$\begin{aligned}
 &+ (1 + \sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}] + 2\sqrt{x_{\beta/2}} \sqrt{\varepsilon^2 \sum_{j=1}^{D \wedge M} X_j^{-2} v_j^2} \Big\} \cap \{\mathcal{B} \cap \mathcal{M}\} \\
 &\leq \mathbb{P}_{\theta, b} \left(\left\{ \frac{1}{2} \sum_{j=1}^{D \wedge M} v_j^2 < C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4} + (1 + \sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}]} \right\} \cap \{\mathcal{B} \cap \mathcal{M}\} \right),
 \end{aligned}$$

where

$$C(\alpha, \beta) := C(\alpha) + 3\sqrt{x_{\beta/2}}, \tag{A.24}$$

and $C(\alpha)$ is defined in (3.9). Note that, for any $a, b \in \mathbb{R}$, using the Young inequality $2ab \leq \gamma a^2 + \gamma^{-1} b^2$ for $\gamma = 1/2$ we get $(a + b)^2 \geq a^2/2 - b^2$. Applying the latter inequality with

$$a = \theta_j - \theta_{j,0}, \quad b = \left(\frac{b_j}{X_j} - 1 \right) \theta_j, \quad j = 1, \dots, D \wedge M,$$

and using Lemma A.3, we arrive at

$$\begin{aligned}
 &\mathbb{P}_{\theta, b}(\{t_{1-\alpha, D}(X) > t_{\beta/2, D}(\theta, X)\} \cap \{\mathcal{B} \cap \mathcal{M}\}) \\
 &\leq \mathbb{P}_{\theta, b} \left(\left\{ \sum_{j=1}^{D \wedge M} (\theta_j - \theta_{j,0})^2 < 4C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} \right. \right. \\
 &\quad \left. \left. + 4(1 + \sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}] + 2 \sum_{j=1}^{D \wedge M} \left(\frac{b_j}{X_j} - 1 \right)^2 \theta_j^2 \right\} \cap \{\mathcal{B} \cap \mathcal{M}\} \right) \\
 &\leq \mathbb{P}_{\theta, b} \left(\left\{ \sum_{j=1}^{D \wedge M} (\theta_j - \theta_{j,0})^2 < 4C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} \right. \right. \\
 &\quad \left. \left. + (6 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}] \right\} \cap \{\mathcal{B} \cap \mathcal{M}\} \right) \\
 &\quad + \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\}.
 \end{aligned}$$

Using the fact that $\theta \in \mathcal{E}_a$, we get

$$\begin{aligned}
 &\mathbb{P}_{\theta, b}(\{t_{1-\alpha, D}(X) > t_{\beta/2, D}(\theta, X)\} \cap \{\mathcal{B} \cap \mathcal{M}\}) \\
 &\leq \mathbb{P}_{\theta, b} \left(\left\{ \|\theta - \theta_0\|^2 < 4C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} \right. \right. \\
 &\quad \left. \left. + (6 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}] + \sum_{j > D \wedge M} (\theta_j - \theta_{j,0})^2 \right\} \cap \{\mathcal{B} \cap \mathcal{M}\} \right) \\
 &\quad + \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\}
 \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{P}_{\theta,b} \left(\left\{ \|\theta - \theta_0\|^2 < 4C(\alpha, \beta)\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M} X_j^{-4}} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M}^{-2}] \right\} \cap \{\mathcal{B} \cap \mathcal{M}\} \right) \\ &\quad + \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\}. \end{aligned} \tag{A.25}$$

To conclude the proof, note that on the event $\{\mathcal{B} \cap \mathcal{M}\}$, we have

$$M_0 \leq M < M_1 \quad \text{and} \quad \frac{b_j}{X_j} \in \left[\frac{2}{3}, 2 \right] \quad \forall j = 1, \dots, M. \tag{A.26}$$

Hence, using (A.25) and (A.26)

$$\begin{aligned} &\mathbb{P}_{\theta,b}(\{t_{1-\alpha,D}(X) > t_{\beta/2,D}(\theta, X)\} \cap \{\mathcal{B} \cap \mathcal{M}\}) \\ &\leq \mathbb{P}_{\theta,b} \left(\|\theta - \theta_0\|^2 < 16C(\alpha, \beta)\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}] \right) \\ &\quad + \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\} \\ &= \frac{\alpha}{5} + \frac{\alpha}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\}, \end{aligned}$$

as soon as

$$\|\theta - \theta_0\|^2 \geq \tilde{C}(\alpha, \beta)\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}], \tag{A.27}$$

where $\tilde{C}(\alpha, \beta) = 16C(\alpha, \beta)$ is defined in (3.12). Therefore, for any fixed $\beta \in]\alpha, 1[$, (A.27) implies that, there exists $\sigma_0 \in]0, 1[$ such that, for all $0 < \sigma < \sigma_0$ and for each $\varepsilon > 0$,

$$\mathbb{P}_{\theta,b}(\Psi_{D,M} = 0) \leq \frac{7\beta}{10} + \frac{\beta}{\kappa} \left(\frac{\pi^2}{2} + 2 \right) + C \exp\{-\ln^{1+\tau}(1/\sigma)\} \leq \beta,$$

for some $C > 0$ and $0 < \tau < 1$, which, in turn, implies that (3.11) holds true. The last part of the theorem is a direct consequence of (2.7) and (3.11). This completes the proof of the proposition.

A.2.3. Proof of Theorem 3.1

The validity of (3.14) can be immediately derived from Proposition 3.1 taking into account that Lemma A.3 is still valid with $D := D^\dagger$ (that depends on the sequence $X = (X_j)_{j \in \mathbb{N}}$). For the proof of (3.11), note first that (A.18), (A.19) and (A.20) still holds true with $D := D^\dagger$. In the same spirit, it is easy to see that Lemma A.3 is still valid when the bandwidth D is measurable with respect to the sequence $(X_k)_{k \in \mathbb{N}}$. Hence, the same inequality than (A.25) can be obtained with $D := D^\dagger$, namely

$$\begin{aligned} &\mathbb{P}_{\theta,b}(\{t_{1-\alpha,D^\dagger}(X) > t_{\beta/2,D^\dagger}(\theta, X)\} \cap \{\mathcal{B} \cap \mathcal{M}\}) \\ &\leq \mathbb{P}_{\theta,b} \left(\left\{ \|\theta - \theta_0\|^2 < 4C(\alpha, \beta)\varepsilon^2 \sqrt{\sum_{j=1}^{D^\dagger \wedge M} X_j^{-4}} + (7 + 4\sqrt{x_{\alpha/2}})[\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D^\dagger \wedge M}^{-2}] \right\} \cap \{\mathcal{B} \cap \mathcal{M}\} \right) \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{P}_{\theta,b} \left(\|\theta - \theta_0\|^2 < \inf_{D \in \mathbb{N}} \left[16C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right] \right) \\ &= 0, \end{aligned}$$

as soon as

$$\|\theta - \theta_0\|^2 \geq \inf_{D \in \mathbb{N}} \left[\tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right],$$

where $\tilde{C}(\alpha, \beta)$ is defined in (3.12). Therefore, we immediately get that (3.11) holds true. Finally, the validity of (3.16) follows immediately on noting that

$$\begin{aligned} \tilde{r}_{\varepsilon,\sigma}^2 &:= \inf_{\tilde{\Psi}_\alpha: \alpha_{\varepsilon,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} r_{\varepsilon,\sigma}^2(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta) \\ &\leq r_{\varepsilon,\sigma}^2(\mathcal{E}_a, \Psi_{D^+, M}, \beta) \\ &\leq \inf_{D \in \mathbb{N}} \left[\tilde{C}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + (7 + 4\sqrt{x_{\alpha/2}}) [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right]. \end{aligned}$$

This completes the proof of the theorem.

A.3. Upper bounds: Specific cases

For the sake of convenience, we give the proof of each item (i)–(iv) in Theorem 3.2 in different sections.

A.3.1. Case (i): Mildly ill-posed problems with ordinary smooth functions

Recall that

$$b_j \sim j^{-t}, \quad t > 0, \quad \text{and} \quad a_j \sim j^s, \quad s > 0, \quad j \in \mathbb{N}. \tag{A.28}$$

Proposition A.1. *Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.28). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon,\sigma}$ satisfies*

$$\tilde{r}_{\varepsilon,\sigma}^2 \lesssim \varepsilon^{\frac{4s}{2s+2t+1/2}} \vee [\sigma \ln^{3/4}(1/\sigma)]^{2(\frac{s}{t} \wedge 1)}. \tag{A.29}$$

Proof. In a first time, we determine the order of the bandwidths M_0 and M_1 . Setting

$$\bar{M}_1 := \left(\sigma \sqrt{\frac{1}{t} \ln(1/\sigma)} \right)^{-1/t} \quad \text{and} \quad \tilde{M}_1 := \left(\sigma \sqrt{\frac{1}{2t} \ln(1/\sigma)} \right)^{-1/t}$$

we get

$$\sigma h_{1, \bar{M}_1} \sim \sigma \sqrt{\ln(\bar{M}_1)} = \sigma \sqrt{\frac{1}{t} \ln(1/\sigma) - \frac{1}{2t} \ln\left(\frac{1}{t} \ln(1/\sigma)\right)} \leq \sigma \sqrt{\frac{1}{t} \ln(1/\sigma)} \sim b_{\bar{M}_1},$$

which implies that $M_1 \gtrsim \bar{M}_1$. At the same time

$$\sigma h_{1, \tilde{M}_1} \sim \sigma \sqrt{\ln(\tilde{M}_1)} = \sigma \sqrt{\frac{1}{t} \ln(1/\sigma) - \frac{1}{2t} \ln\left(\frac{1}{2t} \ln(1/\sigma)\right)}$$

$$\begin{aligned}
 &= \sigma \sqrt{\frac{1}{2t} \ln(1/\sigma) + \frac{1}{2t} \ln(1/\sigma) - \frac{1}{2t} \ln\left(\frac{1}{2t} \ln(1/\sigma)\right)} \\
 &\gtrsim \sigma \sqrt{\frac{1}{2t} \ln(1/\sigma)} \sim b_{\tilde{M}_1},
 \end{aligned}$$

which implies that $M_1 \lesssim \tilde{M}_1$. Hence, we can conclude that

$$M_1 \sim (\sigma \sqrt{\ln(1/\sigma)})^{-1/t}.$$

Similarly, we get that

$$M_0 \sim (\sigma \sqrt{\ln(1/\sigma)})^{-1/t}.$$

In order to control the terms involved in the upper bound on the minimax separation radius, we consider the cases $s < t$ and $s \geq t$ separately.

Consider first the case $s < t$. In this case, for all $D \in \mathbb{N}$,

$$a_{D \wedge M_0}^{-2} \gtrsim a_{M_0}^{-2} \sim M_0^{-2s} \sim (\sigma \ln^{1/2}(1/\sigma))^{2s/t} \gtrsim \sigma^2 \ln^{3/2}(1/\sigma).$$

Hence,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}} + [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}] \right] \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}} + a_{D \wedge M_0}^{-2} \right].$$

Define now the value of $J \in \mathbb{N}$ that satisfies the following equation

$$\varepsilon^2 \sqrt{\sum_{j=1}^J b_j^{-4}} \sim a_J^{-2} \Leftrightarrow \varepsilon^2 J^{2t+1/2} \sim J^{2s} \Leftrightarrow J := J^* \sim \varepsilon^{\frac{-2}{2s+2t+1/2}}.$$

We now consider the following situations (see Figure A.1 for a graphical illustration):

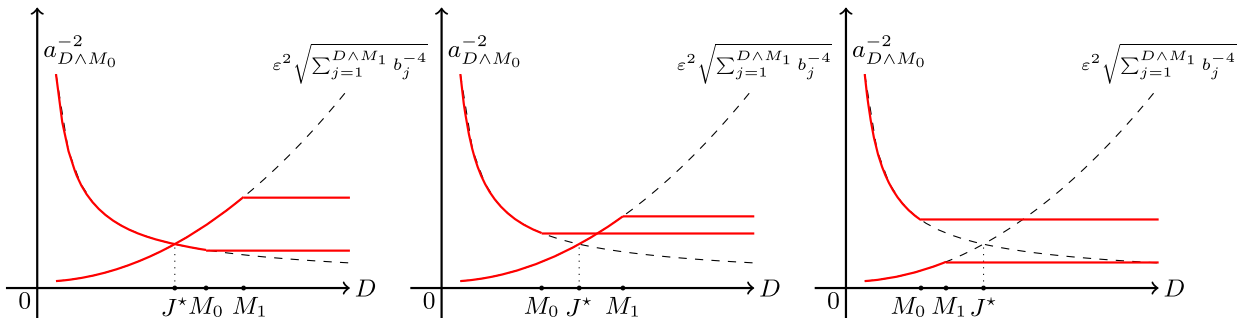


Fig. A.1. [Case I: $a_{D \wedge M_0}^{-2} \gtrsim \sigma^2 \ln^{3/2}(1/\sigma)$]. An illustration of the two resulting two terms (red color), namely $\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}}$ and $a_{D \wedge M_0}^{-2}$, for each $D \in \mathbb{N}$, involved in the (upper bound of the) minimax separation radius $\tilde{r}_{\varepsilon, \sigma}^2$ (see (3.16)), where the bandwidths M_0 and M_1 are defined in (3.3). The bandwidth J^* corresponds to the value $J \in \mathbb{N}$ where the two dashed lines cross, i.e., $J \in \mathbb{N} : \varepsilon^2 \sqrt{\sum_{j=1}^J b_j^{-4}} = a_J^{-2}$. The computation of the separation radius $\tilde{r}_{\varepsilon, \sigma}$, for $0 < \varepsilon \leq \varepsilon_0$, $\varepsilon_0 \in]0, 1[$ and $0 < \sigma \leq \sigma_0$, $\sigma_0 \in]0, 1[$, leads to three different scenarios: $J^* \lesssim M_0 \lesssim M_1$ (left figure), $M_0 \lesssim J^* \lesssim M_1$ (center figure) and $M_0 \lesssim M_1 \lesssim J^*$ (right figure).

- ($J^* \lesssim M_0$) In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim a_{J^*}^{-2} \lesssim \varepsilon^{\frac{4s}{2s+2t+1/2}}.$$

- ($J^* \gtrsim M_1$) In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} [a_{D \wedge M_0}^{-2}] \lesssim a_{M_0}^{-2} \sim [\sigma \ln^{1/2}(1/\sigma)]^{\frac{2s}{t}}.$$

- ($M_0 \lesssim J^* \lesssim M_1$) In this case,

$$\begin{aligned} \tilde{r}_{\varepsilon, \sigma}^2 &\lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \\ &= \left\{ \inf_{D \leq M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \right\} \wedge \left\{ \inf_{D > M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \right\} \\ &\lesssim a_{M_0}^{-2} \wedge \left\{ \inf_{D > M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{M_0}^{-2}} \right] \right\} \\ &\lesssim a_{M_0}^{-2} \wedge \left\{ \varepsilon^2 \sqrt{\sum_{j=1}^{M_0} b_j^{-4} + a_{M_0}^{-2}} \right\} \\ &\lesssim a_{M_0}^{-2} \sim [\sigma \ln^{1/2}(1/\sigma)]^{\frac{2s}{t}}. \end{aligned}$$

Combining the above terms, we immediately get

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^{\frac{4s}{2s+2t+1/2}} \vee [\sigma \ln^{1/2}(1/\sigma)]^{\frac{2s}{t}}. \quad (\text{A.30})$$

Consider now the case $s \geq t$. Define the value of $M \in \mathbb{N}$ that satisfies the following equation

$$a_M^{-2} \sim \sigma^2 \ln^{3/2}(1/\sigma) \quad \Leftrightarrow \quad M := M^* \sim [\sigma \ln^{3/4}(1/\sigma)]^{-\frac{1}{s}}.$$

Hence,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right] \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M^*}^{-2}} \right].$$

Working along the lines of the case $s \leq t$, by replacing M_0 by M^* (see Figure A.2), we get

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^{\frac{4s}{2s+2t+1/2}} \vee [\sigma \ln^{3/4}(1/\sigma)]^2. \quad (\text{A.31})$$

Hence, (A.29) follows thanks to (A.30) and (A.31). This completes the proof of the proposition. \square

A.3.2. Case (ii): Mildly ill-posed problems with super smooth functions

Recall that

$$b_j \sim j^{-t}, \quad t > 0, \quad \text{and} \quad a_j \sim \exp\{js\}, \quad s > 0, \quad j \in \mathbb{N}. \quad (\text{A.32})$$

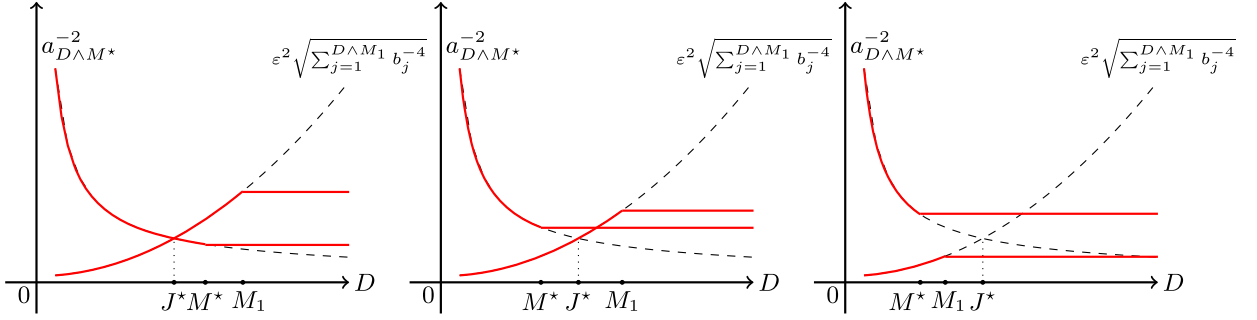


Fig. A.2. [Case II: $a_{D \wedge M_0}^{-2} \sim \sigma^2 \ln^{3/2}(1/\sigma)$]. An illustration of the two resulting two terms (red color), namely $\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}}$ and $a_{D \wedge M^*}^{-2}$, for each $D \in \mathbb{N}$, involved in the (upper bound of the) minimax separation radius $\tilde{r}_{\varepsilon, \sigma}^2$ (see (3.16)), where the bandwidth M_1 is defined in (3.3) and the bandwidth M^* is the value of $M \in \mathbb{N}$ such that the two terms a_M^{-2} and $\sigma^2 \ln^{3/2}(1/\sigma)$ are of the same order, i.e., $M \in \mathbb{N} : a_M^{-2} \sim \sigma^2 \ln^{3/2}(1/\sigma)$. The bandwidth J^* corresponds to the value $J \in \mathbb{N}$ where the two dashed lines cross, i.e., $J \in \mathbb{N} : \varepsilon^2 \sqrt{\sum_{j=1}^J b_j^{-4}} = a_J^{-2}$. The computation of the separation radius $\tilde{r}_{\varepsilon, \sigma}$, for $0 < \varepsilon \leq \varepsilon_0$, $\varepsilon_0 \in]0, 1[$ and $0 < \sigma \leq \sigma_0$, $\sigma_0 \in]0, 1[$, leads to three different scenarios: $J^* \lesssim M^* \lesssim M_1$ (left figure), $M^* \lesssim J^* \lesssim M_1$ (center figure) and $M^* \lesssim M_1 \lesssim J^*$ (right figure).

Proposition A.2. Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.32). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^2 [\ln(1/\varepsilon)]^{(2t+\frac{1}{2})} \vee \sigma^2 \ln^{\frac{3}{2}}(1/\sigma). \quad (\text{A.33})$$

Proof. According to Section A.3.1, we obtain again

$$M_1 \sim (\sigma \sqrt{\ln(1/\sigma)})^{-1/t} \quad \text{and} \quad M_0 \sim (\sigma \sqrt{\ln(1/\sigma)})^{-1/t}.$$

Then, for all $D \in \mathbb{N}$,

$$a_{D \wedge M_0}^{-2} \gtrsim a_{M_0}^{-2} \sim \exp\{-2M_0 s\} \sim \exp\{-2s(\sigma \ln^{1/2}(1/\sigma))^{-1/t}\} \lesssim \sigma^2 \ln^{3/2}(1/\sigma).$$

Define as in the previous case the value $M \in \mathbb{N}$ that satisfies the following equation

$$a_M^{-2} \sim \sigma^2 \ln^{3/2}(1/\sigma) \quad \Leftrightarrow \quad M =: M^* \sim \frac{1}{s} \ln \left[\frac{1}{\sigma \ln^{3/4}(1/\sigma)} \right].$$

Hence,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}} + [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}] \right] \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4}} + a_{D \wedge M^*}^{-2} \right].$$

Define now the value of $J \in \mathbb{N}$ that satisfies the following equation

$$\varepsilon^2 \sqrt{\sum_{j=1}^J b_j^{-4}} \sim a_J^{-2} \quad \Leftrightarrow \quad \varepsilon^2 J^{2t+1/2} \sim \exp\{-2Js\} \quad \Leftrightarrow \quad J := J^* \sim \frac{1}{\sigma} \ln(1/\varepsilon) - \ln \left[\left(\frac{1}{\sigma} \ln(1/\varepsilon) \right)^{2t+\frac{1}{2}} \right].$$

We now consider the following situations:

- ($J^* \lesssim M^*$) In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim a_{J^*}^{-2} \lesssim \varepsilon^2 [\ln(1/\varepsilon)]^{(2t+\frac{1}{2})}.$$

- ($J^* \gtrsim M_1$) In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} [a_{D \wedge M^*}^{-2}] \lesssim a_{M^*}^{-2} \sim \sigma^2 \ln^{\frac{3}{2}}(1/\sigma).$$

- ($M^* \lesssim J^* \lesssim M_1$) In this case,

$$\begin{aligned} \tilde{r}_{\varepsilon, \sigma}^2 &\lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M^*}^{-2}} \right] \\ &= \left\{ \inf_{D \leq M^*} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M^*}^{-2}} \right] \right\} \wedge \left\{ \inf_{D > M^*} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M^*}^{-2}} \right] \right\} \\ &\lesssim a_{M^*}^{-2} \wedge \left\{ \inf_{D > M^*} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{M^*}^{-2}} \right] \right\} \\ &\lesssim a_{M^*}^{-2} \wedge \left\{ \varepsilon^2 \sqrt{\sum_{j=1}^{M^*} b_j^{-4} + a_{M^*}^{-2}} \right\} \\ &\lesssim a_{M^*}^{-2} \sim \sigma^2 \ln^{\frac{3}{2}}(1/\sigma). \end{aligned}$$

Combining the above terms, we immediately get (A.33). This completes the proof of the proposition. \square

A.3.3. Case (iii): Severely ill-posed problems with ordinary smooth functions

Recall that

$$b_j \sim \exp\{-jt\}, \quad t > 0, \quad \text{and} \quad a_j \sim j^s, \quad s > 0, \quad j \in \mathbb{N}. \quad (\text{A.34})$$

Proposition A.3. *Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.34). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies*

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim [\ln(1/\varepsilon)]^{-2s} \vee \left[\ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \right]^{-2s}. \quad (\text{A.35})$$

Proof. In a first time, we determine the order of the bandwidths M_0 and M_1 . Setting

$$\bar{M}_1 := \frac{1}{t} \ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \quad \text{and} \quad \tilde{M}_1 := \frac{1}{t} \ln(1/\sigma),$$

we get

$$\sigma h_{1, \bar{M}_1} \sim \sigma \sqrt{\ln(\bar{M}_1)} = \sigma \sqrt{\ln \left(\frac{1}{t} \ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \right)} \lesssim e^{-\bar{M}_1 t} = \sigma \sqrt{\ln(1/\sigma)} \sim b_{\bar{M}_1},$$

which implies that $M_1 \geq \bar{M}_1$ for σ small enough. At the same time

$$\begin{aligned} \sigma h_{1, \tilde{M}_1} &\sim \sigma \sqrt{\ln(\tilde{M}_1)} = \sigma \sqrt{\ln \left(\frac{1}{t} \ln(1/\sigma) \right)} \\ &\gtrsim b_{\tilde{M}_1} \sim e^{-\tilde{M}_1 t} = \sigma, \end{aligned}$$

which implies that $M_1 \leq \tilde{M}_1$ for σ small enough. Hence, we can conclude that

$$\frac{1}{t} \ln\left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)}\right) \leq M_1 \leq \frac{1}{t} \ln(1/\sigma),$$

for σ small enough. Similarly, we get that

$$\frac{1}{t} \ln\left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)}\right) \leq M_0 \leq \frac{1}{t} \ln(1/\sigma),$$

for σ small enough.

Now, we turn our attention to the proof of (A.35). For all $D \in \mathbb{N}$,

$$a_{D \wedge M_0}^{-2} \gtrsim a_{M_0}^{-2} \gtrsim M_0^{-2s} \sim \left[\ln\left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)}\right) \right]^{-2s} \gtrsim \sigma^2 \ln^{3/2}(1/\sigma).$$

Hence,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right] \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right].$$

Define now the value of $J \in \mathbb{N}$ that satisfies the following equation

$$\varepsilon^2 \sqrt{\sum_{j=1}^J b_j^{-4}} \sim a_J^{-2} \Leftrightarrow \varepsilon^2 \exp\{2tJ\} \sim J^{-2s} \Leftrightarrow J := J^* \sim \frac{1}{t} \ln(1/\varepsilon) - \ln\left[\left(\frac{1}{t} \ln(1/\varepsilon)\right)^{2s}\right].$$

We now consider the following situations:

- $(J^* \lesssim M_0)$ In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim a_{J^*}^{-2} \lesssim [\ln(1/\varepsilon)]^{-2s}.$$

- $(J^* \gtrsim M_1)$ In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} [a_{D \wedge M_0}^{-2}] \lesssim a_{M_0}^{-2} \sim \left[\ln\left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)}\right) \right]^{-2s}.$$

- $(M_0 \lesssim J^* \lesssim M_1)$ In this case,

$$\begin{aligned} \tilde{r}_{\varepsilon, \sigma}^2 &\lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \\ &= \left\{ \inf_{D \leq M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \right\} \wedge \left\{ \inf_{D > M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \right\} \\ &\lesssim a_{M_0}^{-2} \wedge \left\{ \inf_{D > M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{M_0}^{-2}} \right] \right\} \end{aligned}$$

$$\begin{aligned} &\lesssim a_{M_0}^{-2} \wedge \left\{ \varepsilon^2 \sqrt{\sum_{j=1}^{M_0} b_j^{-4} + a_{M_0}^{-2}} \right\} \\ &\lesssim a_{M_0}^{-2} \lesssim \left[\ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \right]^{-2s}. \end{aligned}$$

Combining the above terms, we immediately get (A.35). This completes the proof of the proposition. \square

A.3.4. Case (iv): Severely ill-posed problems with super smooth functions

Recall that

$$b_j \sim \exp\{-jt\}, \quad t > 0, \quad \text{and} \quad a_j \sim \exp\{js\}, \quad s > 0, \quad j \in \mathbb{N}. \quad (\text{A.36})$$

Proposition A.4. *Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.36). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies*

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \varepsilon^{\frac{2s}{s+t}} \vee \left[\sigma \ln^{1/2}(1/\sigma) \right]^{2(\frac{s}{t} \wedge 1)}. \quad (\text{A.37})$$

Proof. According to Section A.3.3, we obtain again that

$$\frac{1}{t} \ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \leq M_1 \leq \frac{1}{t} \ln(1/\sigma)$$

and

$$\frac{1}{t} \ln \left(\frac{1}{\sigma \ln^{1/2}(1/\sigma)} \right) \leq M_0 \leq \frac{1}{t} \ln(1/\sigma),$$

for σ small enough. Now, we consider the cases $s < t$ and $s \geq t$ separately.

Consider first the case $s < t$. In this case, for all $D \in \mathbb{N}$,

$$a_{D \wedge M_0}^{-2} \gtrsim a_{M_0}^{-2} \sim \exp\{-2sM_0\} \gtrsim (\sigma \ln^{1/2}(1/\sigma))^{2s/t} \gtrsim \sigma^2 \ln^{3/2}(1/\sigma).$$

Hence,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right] \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right].$$

Define now the value of $J \in \mathbb{N}$ that satisfies the following equation

$$\varepsilon^2 \sqrt{\sum_{j=1}^J b_j^{-4}} \sim a_J^{-2} \quad \Leftrightarrow \quad \varepsilon^2 \exp\{2tJ\} \sim \exp\{2sJ\} \quad \Leftrightarrow \quad J := J^* \sim \frac{1}{s+t} \ln(1/\varepsilon).$$

We now consider the following situations:

- ($J^* \lesssim M_0$) In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim a_{J^*}^{-2} \lesssim \varepsilon^{\frac{2s}{s+t}}.$$

- ($J^* \gtrsim M_1$) In this case,

$$\tilde{r}_{\varepsilon, \sigma}^2 \lesssim \inf_{D \in \mathbb{N}} [a_{D \wedge M_0}^{-2}] \lesssim a_{M_0}^{-2} \lesssim \left[\sigma \ln^{1/2}(1/\sigma) \right]^{\frac{2s}{t}}.$$

- $(M_0 \lesssim J^* \lesssim M_1)$ In this case,

$$\begin{aligned}
\tilde{r}_{\varepsilon,\sigma}^2 &\lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \\
&= \left\{ \inf_{D \leq M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \right\} \wedge \left\{ \inf_{D > M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{D \wedge M_0}^{-2}} \right] \right\} \\
&\lesssim a_{M_0}^{-2} \wedge \left\{ \inf_{D > M_0} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{M_0}^{-2}} \right] \right\} \\
&\lesssim a_{M_0}^{-2} \wedge \left\{ \varepsilon^2 \sqrt{\sum_{j=1}^{M_0} b_j^{-4} + a_{M_0}^{-2}} \right\} \\
&\lesssim a_{M_0}^{-2} \lesssim [\sigma \ln^{1/2}(1/\sigma)]^{\frac{2s}{t}}.
\end{aligned}$$

Combining the above terms, we immediately get

$$\tilde{r}_{\varepsilon,\sigma}^2 \lesssim \varepsilon^{\frac{2s}{s+t}} \vee [\sigma \ln^{1/2}(1/\sigma)]^{\frac{2s}{t}}. \quad (\text{A.38})$$

Consider now the case $s \geq t$. Define the value $M \in \mathbb{N}$ that satisfies the following equation

$$a_M^{-2} \sim \sigma^2 \ln^{3/2}(1/\sigma) \Leftrightarrow M := M^* \sim \frac{1}{s} \ln \left[\frac{1}{\sigma \ln^{3/4}(1/\sigma)} \right].$$

Hence,

$$\tilde{r}_{\varepsilon,\sigma}^2 \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + [\sigma^2 \ln^{3/2}(1/\sigma) \vee a_{D \wedge M_0}^{-2}]} \right] \lesssim \inf_{D \in \mathbb{N}} \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D \wedge M_1} b_j^{-4} + a_{M^*}^{-2}} \right].$$

Working along the lines of the case $s < t$ by replacing M_0 by M^* , we get

$$\tilde{r}_{\varepsilon,\sigma}^2 \lesssim \varepsilon^{\frac{2s}{s+t}} \vee [\sigma \ln^{1/2}(1/\sigma)]^2. \quad (\text{A.39})$$

Hence, (A.37) follows thanks to (A.38) and (A.39). This completes the proof of the proposition. \square

A.4. Non-asymptotic lower bounds

A.4.1. Proof of Proposition 4.1

Let θ_0, θ be two given sequences (to be made precise below) such that $\theta_0 \in \mathcal{E}_a \setminus \{0\}$ and $\theta - \theta_0 \in \mathcal{E}_a$, and let D be a given bandwidth such that

$$\theta_k = \theta_{0,k} \quad \text{when } k \neq D.$$

For any $t \in \mathbb{R}$, define the sequence $\tilde{b}(t) = (\tilde{b}(t))_{k \in \mathbb{N}}$ as

$$\tilde{b}(t)_k = b_k \quad \forall k \neq D \quad \text{and} \quad \tilde{b}(t)_D = t.$$

Then

$$\begin{aligned} \beta_{0,\sigma,b}(\Theta(r_\sigma), \mathcal{B}(b)) &= \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \sup_{\tilde{\theta}: \tilde{\theta} - \theta_0 \in \Theta_\alpha(r_\sigma)} \mathbb{P}_{\tilde{\theta}, \tilde{b}}(\tilde{\Psi}_\alpha = 0) \\ &\geq \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \sup_{\tilde{b} \in \mathcal{B}(b)} \mathbb{P}_{\theta, \tilde{b}}(\tilde{\Psi}_\alpha = 0) \\ &\geq \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \sup_{t \in [C_0 b_D, C_1 b_D]} \mathbb{P}_{\theta, \tilde{b}(t)}(\tilde{\Psi}_\alpha = 0). \end{aligned}$$

From the last inequality, we get

$$\begin{aligned} \beta_{0,\sigma,b}(\Theta(r_\sigma), \mathcal{B}(b)) &\geq \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \int_{C_0 b_D}^{C_1 b_D} \mathbb{P}_{\theta, \tilde{b}(t)}(\tilde{\Psi}_\alpha = 0) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-b_D)^2}{2\sigma^2}} dt \\ &\geq \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \int_{\mathbb{R}} \mathbb{P}_{\theta, \tilde{b}(t)}(\tilde{\Psi}_\alpha = 0) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-b_D)^2}{2\sigma^2}} dt - \int_{\mathbb{R} \setminus [C_0 b_D; C_1 b_D]} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-b_D)^2}{2\sigma^2}} dt \\ &= \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \int_{\mathbb{R}} \mathbb{P}_{\theta, \tilde{b}(t)}(\tilde{\Psi}_\alpha = 0) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(t-b_D)^2}{2\sigma^2}} dt - (1 - G_D(C_0, C_1)), \end{aligned} \tag{A.40}$$

where

$$G_D(C_0, C_1) := \frac{1}{\sigma \sqrt{2\pi}} \int_{C_0 b_D}^{C_1 b_D} \exp\{- (t - b_D)^2 / (2\sigma^2)\} dt.$$

The inequality (A.40) can be rewritten in a Bayesian form

$$\beta_{0,\sigma,b}(\Theta(r_\sigma), \mathcal{B}(b)) \geq \inf_{\tilde{\Psi}_\alpha: \alpha_{0,\sigma}(\tilde{\Psi}_\alpha) \leq \alpha} \mathbb{P}_\pi(\tilde{\Psi}_\alpha = 0) - (1 - G_D(C_0, C_1)),$$

where π is a (prior) probability measure on both sequences $\tilde{\theta}$ and \tilde{b} defined as

$$\pi = \prod_{j \in \mathbb{N}} \pi_j, \quad \pi_j = \pi_{j,1} \times \pi_{j,2}, \quad j \in \mathbb{N},$$

with

$$\pi_{j,1} = \delta_{\theta_{j,0}} \quad \text{and} \quad \pi_{j,2} = \delta_{b_j}, \quad j \neq D,$$

and

$$\pi_{D,1} = \delta_{\theta_D} \quad \text{and} \quad d\pi_{D,2}(t) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(t - b_D)^2\right\} dt.$$

We stress that using the above product probability measure π , we deal in fact with observations $(Y, X) = (Y_j, X_j)_{j \in \mathbb{N}}$ from the following Bayesian sequence model

$$Y_j = b_j \theta_{j,0}, \quad X_j = b_j + \sigma \eta_j, \quad j \in \mathbb{N} \setminus \{D\},$$

and

$$Y_D = B_D \theta_D, \quad X_D = B_D + \sigma \eta_D, \tag{A.41}$$

where B_D is Gaussian random variable with mean b_D and variance σ^2 , that is independent of the standard Gaussian sequence $\{\eta_j\}_{j \in \mathbb{N}}$.

By standard Bayesian arguments (see, e.g., Section 3.1 of [23]), we arrive at

$$\beta_{0,\sigma,b}(\Theta(r_\sigma), \mathcal{B}(b)) \geq 1 - \alpha - \frac{1}{2}(\mathbb{E}_0[L_\pi^2(Y, X)] - 1)^{1/2} - (1 - G_D(C_0, C_1)), \quad (\text{A.42})$$

In view of the above, it is immediately seen that

$$L_\pi(Y, X) = \prod_{j \in \mathbb{N}} L_{\pi_j}(Y_j, X_j) = L_{\pi_D}(Y_D, X_D).$$

Hence, as before, we arrive at

$$\beta_{0,\sigma,b}(\Theta(r_\sigma), \mathcal{B}(b)) \geq 1 - \alpha - \frac{1}{2}(\mathbb{E}_0[L_{\pi_D}^2(Y_D, X_D)] - 1)^{1/2} - (1 - G_D(C_0, C_1)). \quad (\text{A.43})$$

Our task below is then to provide an upper bound on $\mathbb{E}_0[L_{\pi_D}^2(Y_D, X_D)]$. To this end, it is easily seen from model (A.41) that $Z_D = (X_D, Y_D)$, $D \in \mathbb{N}$, is Gaussian random vector with mean $U_{\theta,D}$ and covariance matrix $\sigma^2 \Sigma_{\theta,D}$, where

$$U_{\theta,D} = \begin{pmatrix} b_D \\ b_D \theta_D \end{pmatrix}, \quad \Sigma_{\theta,D} = \begin{pmatrix} 2 & \theta_D \\ \theta_D & \theta_D^2 \end{pmatrix}.$$

Note that

$$\Sigma_{\theta,D}^{-1} = \frac{1}{\theta_D^2} \begin{pmatrix} \theta_D^2 & -\theta_D \\ -\theta_D & 2 \end{pmatrix},$$

and

$$\begin{aligned} & (Z_D - U_{\theta,D})' \Sigma_{\theta,D}^{-1} (Z_D - U_{\theta,D}) \\ &= \frac{1}{\theta_D^2} (X_D - b_D, Y_D - b_D \theta_D) \begin{pmatrix} \theta_D^2 & -\theta_D \\ -\theta_D & 2 \end{pmatrix} \begin{pmatrix} X_D - b_D \\ Y_D - b_D \theta_D \end{pmatrix} \\ &= \frac{1}{\theta_D^2} (X_D - b_D, Y_D - b_D \theta_D) \begin{pmatrix} \theta_D^2 (X_D - b_D) - \theta_D (Y_D - b_D \theta_D) \\ -\theta_D (X_D - b_D) + 2(Y_D - b_D \theta_D) \end{pmatrix} \\ &= \frac{1}{\theta_D^2} (X_D - b_D, Y_D - b_D \theta_D) \begin{pmatrix} X_D \theta_D^2 - Y_D \theta_D \\ 2Y_D - b_D \theta_D - X_D \theta_D \end{pmatrix} \\ &= \frac{1}{\theta_D^2} ((X_D - b_D)(X_D \theta_D^2 - Y_D \theta_D) + (Y_D - b_D \theta_D)(2Y_D - b_D \theta_D - X_D \theta_D)) \\ &= \frac{1}{\theta_D^2} [(Y_D - X_D \theta_D)^2 + (Y_D - b_D \theta_D)^2]. \end{aligned}$$

Hence,

$$\begin{aligned} L_{\pi_D}(Z_D) &= \exp\left\{ (Z_D - U_{\theta_0,D})' \sigma^{-2} \Sigma_{\theta_0,D}^{-1} (Z_D - U_{\theta_0,D}) - (Z_D - U_{\theta,D})' \sigma^{-2} \Sigma_{\theta,D}^{-1} (Z_D - U_{\theta,D}) \right\} \\ &= \exp\left\{ \frac{1}{\sigma^2} \left[\frac{(Y_D - X_D \theta_{D,0})^2}{\theta_{D,0}^2} + \frac{(Y_D - b_D \theta_{D,0})^2}{\theta_{D,0}^2} - \frac{(Y_D - X_D \theta_D)^2}{\theta_D^2} - \frac{(Y_D - b_D \theta_D)^2}{\theta_D^2} \right] \right\} \\ &= \exp\left\{ \frac{1}{\sigma^2} \left[2Y_D^2 \left(\frac{1}{\theta_{D,0}^2} - \frac{1}{\theta_D^2} \right) - 2X_D Y_D \left(\frac{1}{\theta_{D,0}} - \frac{1}{\theta_D} \right) - 2b_D Y_D \left(\frac{1}{\theta_{D,0}} - \frac{1}{\theta_D} \right) \right] \right\} \\ &= \exp\left\{ \frac{1}{\sigma^2} \left[2Y_D^2 \left(\frac{1}{\theta_{D,0}^2} - \frac{1}{\theta_D^2} \right) - 2Y_D (X_D + b_D) \left(\frac{1}{\theta_{D,0}} - \frac{1}{\theta_D} \right) \right] \right\}. \end{aligned}$$

Under H_0 , $Y_D = B_D \theta_{D,0}$. Therefore, conditionally on B_D ,

$$\begin{aligned} \mathbb{E}_0[L_\pi^2(Z_D)] &= \mathbb{E}\left[\exp\left\{\frac{2}{\sigma^2}\left(2B_D^2\left(1 - \frac{\theta_{D,0}^2}{\theta_D^2}\right) - 2B_D(X_D + b_D)\left(1 - \frac{\theta_{D,0}}{\theta_D}\right)\right)\right\}\right] \\ &= \mathbb{E}\left(\exp\left\{\frac{4B_D^2}{\sigma^2}\left(1 - \frac{\theta_{D,0}^2}{\theta_D^2}\right) - \frac{4B_D b_D}{\sigma^2}\left(1 - \frac{\theta_{D,0}}{\theta_D}\right)\right\}\right. \\ &\quad \left.\times \mathbb{E}\left[\exp\left\{-\frac{4B_D X_D}{\sigma^2}\left(1 - \frac{\theta_{D,0}}{\theta_D}\right)\right\} \mid B_D\right]\right) \\ &:= \mathbb{E}(\mathbb{E}_0[L_\pi^2(Z_D) \mid B_D]). \end{aligned}$$

Using the formula

$$\mathbb{E}[\exp(-(\lambda_1 + \lambda_2 V))] = \exp(-\lambda_1 + \lambda_2^2/2), \quad \lambda_1, \lambda_2 \in \mathbb{R}, \quad (\text{A.44})$$

for any standard Gaussian random variable V , with

$$\lambda_1 = \frac{4B_D^2}{\sigma^2}\left(1 - \frac{\theta_{D,0}}{\theta_D}\right), \quad \lambda_2 = \frac{4B_D}{\sigma}\left(1 - \frac{\theta_{D,0}}{\theta_D}\right),$$

we arrive at

$$\begin{aligned} \mathbb{E}_0[L_\pi^2(Z_D) \mid B_D] &= \exp\left\{\frac{4B_D^2}{\sigma^2}\left(1 - \frac{\theta_{D,0}^2}{\theta_D^2}\right) - \frac{4B_D}{\sigma^2}\left(1 - \frac{\theta_{D,0}}{\theta_D}\right)(B_D + b_D)\right\} \\ &\quad \times \exp\left\{\frac{8B_D^2}{\sigma^2}\left(1 - \frac{\theta_{D,0}}{\theta_D}\right)^2\right\} \\ &= \exp\left\{\sigma^{-2}B_D^2[4(1 - \rho_D^2) - 4(1 - \rho_D) + 8(1 - \rho_D)^2]\right\} \\ &\quad \times \exp\{-4\sigma^{-2}B_D b_D(1 - \rho_D)\}, \end{aligned}$$

where

$$\rho_D = \frac{\theta_{D,0}}{\theta_D}.$$

Using simple algebra, we get

$$\begin{aligned} \mathbb{E}_0[L_\pi^2(Z_D) \mid B_D] &= \exp\left\{\frac{4B_D^2}{\sigma^2}(\rho_D - 1)(\rho_D - 2) - \frac{4B_D b_D}{\sigma^2}(1 - \rho_D)\right\} \\ &= \exp\left\{\frac{4}{\sigma^2}(\rho_D - 1)[B_D^2(\rho_D - 2) + B_D b_D]\right\}. \end{aligned}$$

It is easily seen that

$$\begin{aligned} B_D^2(\rho_D - 2) + B_D b_D &= (b_D + \sigma \tilde{\eta}_D)^2(\rho_D - 2) + (b_D + \sigma \tilde{\eta}_D)b_D \\ &= b_D^2(\rho_D - 1) + 2\sigma \tilde{\eta}_D b_D(\rho_D - 3/2) + \sigma^2 \tilde{\eta}_D^2(\rho_D - 2), \end{aligned}$$

where $\{\tilde{\eta}_D\}_{D \in \mathbb{N}}$ is a sequence of independent standard Gaussian random variables. Therefore,

$$\mathbb{E}_0[L_\pi^2(Z_D) \mid B_D] = \exp\left\{\frac{4b_D^2}{\sigma^2}(1 - \rho_D)^2\right\} \exp\left\{\frac{8}{\sigma}b_D \tilde{\eta}_D(\rho_D - 1)\left(\rho_D - \frac{3}{2}\right) + 4\tilde{\eta}_D^2(\rho_D - 2)\right\}.$$

Since $\rho_D \in]1, 2[$, then $4\tilde{\eta}_D^2(\rho_D - 2) < 0$ and, hence,

$$\mathbb{E}_0[L_\pi^2(Z_D) | B_D] \leq \exp\left\{\frac{4b_D^2}{\sigma^2}(1 - \rho_D)^2 + \frac{8}{\sigma}b_D\tilde{\eta}_D(\rho_D - 1)\left(\rho_D - \frac{3}{2}\right)\right\}.$$

Using (A.44) with

$$\lambda_1 = 0, \quad \lambda_2 = \frac{8}{\sigma}b_D(\rho_D - 1)\left(\rho_D - \frac{3}{2}\right),$$

we get

$$\begin{aligned} \mathbb{E}_0[L_\pi^2(Z_D)] &= G_D^{-1}(C_0, C_1)\mathbb{E}(\mathbb{E}_0[L_\pi^2(Z_D) | B_D]) \\ &\leq \mathbb{E}\left(\exp\left\{\frac{4b_D^2}{\sigma^2}(1 - \rho_D)^2 + \frac{8}{\sigma}b_D\tilde{\eta}_D(\rho_D - 1)\left(\rho_D - \frac{3}{2}\right)\right\}\right) \\ &= \exp\left\{\frac{4b_D^2}{\sigma^2}(1 - \rho_D)^2\left[1 + \left(\rho_D - \frac{3}{2}\right)^2\right]\right\} \\ &\leq \exp\left\{\frac{5b_D^2}{\sigma^2}(1 - \rho_D)^2\right\}. \end{aligned} \tag{A.45}$$

Using (A.43) and (A.45), we eventually obtain

$$\begin{aligned} \beta_{\alpha,\sigma,b}(\Theta(r_{\sigma,D}), \mathcal{B}(b)) &\geq 1 - \alpha - \frac{1}{2}\left(\exp\left\{\frac{5b_D^2}{\sigma^2}(1 - \rho_D)^2\right\} - 1\right)^{1/2} - (1 - G_D(C_0, C_1)) \\ &\geq \beta, \end{aligned}$$

as soon as,

$$\frac{5b_D^2}{\sigma^2}(1 - \rho_D)^2 \leq \ln(1 + (1 - \alpha - \beta)^2) \quad \text{and} \quad G(C_0, C_1) \geq 1 - \frac{1}{2}(1 - \alpha - \beta).$$

In particular,

$$\frac{5b_D^2}{\sigma^2}(1 - \rho_D)^2 \leq \ln(1 + (1 - \alpha - \beta)^2) \quad \Leftrightarrow \quad |\theta_D - \theta_{D,0}| \leq C_{\alpha,\beta}\sigma|\theta_D|b_D^{-1},$$

where

$$C_{\alpha,\beta} = \sqrt{\frac{\ln(1 + (1 - \alpha - \beta)^2)}{5}}. \tag{A.46}$$

Choice of θ : The sequence $\theta = (\theta_j)_{j \in \mathbb{N}}$ is chosen as follows

$$\theta_j = \begin{cases} 0 & \text{if } j \neq D, \\ a_D^{-1}/2 & \text{if } j = D. \end{cases}$$

It can be easily seen that $\theta \in \mathcal{E}_a$.

Choice of θ_0 : The sequence $\theta_0 = (\theta_{j,0})_{j \in \mathbb{N}}$ is chosen as follows

$$\theta_{j,0} = \begin{cases} 0 & \text{if } j \neq D, \\ a_D^{-1}/2 + C_{\alpha,\beta}\sigma a_D^{-1}b_D^{-1}/2 & \text{if } j = D. \end{cases}$$

Note that $\theta_0 \in \mathcal{E}_\alpha$ as soon as

$$C_{\alpha,\beta}\sigma b_D^{-1} \leq 1. \quad (\text{A.47})$$

Indeed, using the standard inequality $(x+y)^2 \leq 2(x^2+y^2)$, for $x, y \in \mathbb{R}$, we immediately get

$$\sum_{j \in \mathbb{N}} a_j^2 \theta_{j,0}^2 = a_D^2 \theta_{D,0}^2 \leq a_D^2 \left(2 \frac{a_D^{-2}}{4} + 2C_{\alpha,\beta}^2 \frac{\sigma^2 a_D^{-2}}{b_D^2} \right) \leq \frac{1}{2} + C_{\alpha,\beta}^2 \frac{\sigma^2}{2b_D^2} \leq 1, \quad (\text{A.48})$$

as soon as (A.47) is satisfied. Furthermore, as soon as (A.47) is satisfied, it is easily seen that $\theta - \theta_0 \in \mathcal{E}_\alpha$.

Moreover, for the specific choices of θ and θ_0 given above, it is immediately seen that

$$|\theta_D - \theta_{D,0}| = C_{\alpha,\beta}\sigma b_D^{-1} |\theta_D| \Leftrightarrow \|\theta - \theta_0\| = \frac{C_{\alpha,\beta}}{2} \sigma b_D^{-1} a_D^{-1}.$$

In other words, we have proved that for all $D \in \mathbb{N}$ satisfying (A.47) then

$$\beta_{0,\sigma,b}(\Theta(r_{\sigma,D}), \mathcal{B}(b)) > \beta, \quad \text{where } r_{\sigma,D} = \frac{C_{\alpha,\beta}}{2} \sigma b_D^{-1} a_D^{-1},$$

for any given $\beta \in]0, 1 - \alpha[$. This implies that, for every $\rho > 0$, $\beta_{0,\sigma,b}(\Theta(\rho), \mathcal{B}(b)) \geq \beta$ as soon as

$$\rho \leq \frac{C_{\alpha,\beta}}{2} \sigma b_D^{-1} a_D^{-1} \quad \text{for some } D \in \mathbb{N} : C_{\alpha,\beta}\sigma b_D^{-1} \leq 1,$$

which holds, as soon as

$$\rho \leq \frac{C_{\alpha,\beta}}{2} \sigma b_D^{-1} a_D^{-1} \quad \text{for some } 1 \leq D \leq M_2,$$

on noting that

$$M_2 := \sup \left\{ D \in \mathbb{N} : C_{\alpha,\beta}\sigma b_D^{-1} \leq 1 \text{ and } G_D(C_0, C_1) \geq 1 - \frac{1}{2}(1 - \alpha - \beta) \right\}.$$

In particular,

$$\beta_{0,\sigma,b}(\Theta(\rho), \mathcal{B}(b)) > \beta \quad \text{for all } \rho \leq \frac{C_{\alpha,\beta}}{2} \sigma \max_{1 \leq D \leq M_2} [b_D^{-1} a_D^{-1}].$$

Hence,

$$\tilde{r}_{0,\sigma} \geq \frac{C_{\alpha,\beta}}{2} \sigma \max_{1 \leq D \leq M_2} [b_D^{-1} a_D^{-1}].$$

This completes the proof of the proposition.

A.4.2. Proof of Theorem 4.1

The proof is splitted in two parts. We first show that $\tilde{r}_{\varepsilon,\sigma} \geq \tilde{r}_{\varepsilon,0}$ and then show that $\tilde{r}_{\varepsilon,\sigma} \geq \tilde{r}_{0,\sigma}$.

Consider observations $Y = (Y_j)_{j \in \mathbb{N}}$ from the GSM (4.8). Introduce the following goodness-of-fit testing algorithm:

- Generate a sequence $\tilde{X} = (\tilde{X}_j)_{j \in \mathbb{N}}$ according to the GSM

$$\tilde{X}_j = b_j + \sigma \tilde{\eta}_j, \quad j \in \mathbb{N}, \quad (\text{A.49})$$

where $\tilde{\eta} = (\tilde{\eta}_j)_{j \in \mathbb{N}}$ is a sequence of independent standard Gaussian random variables (that is also independent of the sequence $\xi = (\xi_j)_{j \in \mathbb{N}}$). (Note that the GSM (A.49) is an independent copy of the second equation in the GSM (1.1).)

- Let $\tilde{\Psi}_\alpha := \tilde{\Psi}_\alpha(Y, \tilde{X})$ be a given (non-randomized) α -level goodness-of-fit testing procedure based on observations $(Y, \tilde{X}) = (Y_j, \tilde{X}_j)_{j \in \mathbb{N}}$ from the GSMs (4.8) and (A.49).
- Define the randomized test $\Psi_\alpha := \Psi_\alpha(Y)^2$ as

$$\Psi_\alpha(Y) := \mathbb{E}[\tilde{\Psi}_\alpha | Y],$$

where $\mathbb{E}[\cdot]$ refers to expectation with respect to the independent standard Gaussian sequence $\tilde{\eta}$.

In particular, for every $\varepsilon > 0$ and $\sigma > 0$, the randomized test Ψ_α is an α -level test. Indeed,

$$\begin{aligned} \alpha_\varepsilon(\Psi_\alpha) &= \mathbb{E}_{\theta_0, b}[\Psi_\alpha] \\ &= \mathbb{E}_{\theta_0, b}[\mathbb{E}[\tilde{\Psi}_\alpha | Y]] \\ &= \mathbb{E}_{\theta_0, b}[\tilde{\Psi}_\alpha] \\ &= \mathbb{P}_{\theta_0, b}(\tilde{\Psi}_\alpha = 1) = \alpha, \end{aligned} \tag{A.50}$$

since $\tilde{\Psi}_\alpha$ is an α -level test.

Let $\theta \in l^2(\mathbb{N})$ and $\theta - \theta_0 \in \mathcal{E}_a$ be fixed. Then, the associated second kind error probability satisfies

$$\begin{aligned} \mathbb{E}_{\theta, b}(1 - \Psi_\alpha(Y)) &= \mathbb{E}_{\theta, b}(1 - \mathbb{E}[\tilde{\Psi}_\alpha | Y]) \\ &= \mathbb{E}_{\theta, b}(1 - \tilde{\Psi}_\alpha) \\ &= \mathbb{P}_{\theta, b}(\tilde{\Psi}_\alpha = 0) \leq \beta, \end{aligned} \tag{A.51}$$

as soon as

$$\|\theta - \theta_0\| \geq r_{\varepsilon, \sigma}(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta).$$

This implies that for any α -level goodness-of-fit testing procedure $\tilde{\Psi}_\alpha$, based on observations (Y, \tilde{X}) from the GSMs (4.8)–(A.49), we can associate an α -level goodness-of-fit testing procedure Ψ_α , based on observations Y from the GSM (4.8), such that the separation radius of Ψ_α is smaller than the separation radius of $\tilde{\Psi}_\alpha$, i.e.,

$$r_{\varepsilon, 0}(\mathcal{E}_a, \Psi_\alpha, \beta) \leq r_{\varepsilon, \sigma}(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta).$$

Hence, it is immediately seen that, for any α -level goodness-of-fit testing procedure $\tilde{\Psi}_\alpha$, based on observations (Y, \tilde{X}) from the GSMs (4.8) and (A.49),

$$\begin{aligned} \tilde{r}_{\varepsilon, 0} &:= \inf_{\Psi_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} r_{\varepsilon, 0}(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta) \\ &\leq r_{\varepsilon, 0}(\mathcal{E}_a, \Psi_\alpha, \beta) \\ &\leq r_{\varepsilon, \sigma}(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta), \end{aligned} \tag{A.52}$$

implying that

$$\tilde{r}_{\varepsilon, 0} \leq \tilde{r}_{\varepsilon, \sigma}.$$

The proof of the assertion

$$\tilde{r}_{0, \sigma} \leq \tilde{r}_{\varepsilon, \sigma},$$

²A measurable function of the observation $Y = (Y_j)_{j \in \mathbb{N}}$ from the GSM (4.8) with values in the interval $[0, 1]$: for any given radius $\rho > 0$, the null hypothesis is rejected with probability $\Psi_\alpha(Y)$ and it is not rejected with probability $1 - \Psi_\alpha(Y)$. In this case, $\alpha_\varepsilon(\Psi_\alpha) := \mathbb{E}_{\theta_0, b}(\Psi_\alpha(Y))$ and $\beta_\varepsilon(\Theta_a(\rho), \Psi_\alpha) := \sup_{\theta_0 \in \mathcal{E}_a, \theta - \theta_0 \in \Theta_a(\rho)} \mathbb{E}_{\theta, b}(1 - \Psi_\alpha(Y))$.

follows similarly, along the lines of the proof of the previous assertion, and it is therefore omitted. This completes the proof of (4.12).

Finally, (4.13) follows immediately from (4.12), taking into account (4.3) and (4.10). This completes the proof of the theorem.

A.5. Lower bounds: Specific cases

For the sake of convenience, we give the proof of each item (i)–(iv) in Theorem 4.2 in different sections.

A.5.1. Case (i): Mildly ill-posed problems with ordinary smooth functions

We assume that (A.28) holds true, i.e.,

$$b_j \sim j^{-t}, \quad t > 0, \quad \text{and} \quad a_j \sim j^s, \quad s > 0, \quad j \in \mathbb{N}.$$

Proposition A.5. Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.28). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim \varepsilon^{\frac{4s}{2s+2t+1/2}} \vee \sigma^{2(\frac{s}{t} \wedge 1)}. \tag{A.53}$$

Proof. For the second term in (4.13), it is known that (see [18,21]),

$$\sup_{D \in \mathbb{N}} \left[c_{\alpha, \beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \sim \varepsilon^{\frac{4s}{2s+2t+1/2}}.$$

Consider now the first term in (4.13). If $s > t$, then the sequence $\{b_j^{-1} a_j^{-1}\}_{j \in \mathbb{N}}$ is non-increasing and, thus,

$$\frac{C_{\alpha, \beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \sim \sigma^2.$$

On the other hand, if $s \leq t$, then the sequence $\{b_j^{-1} a_j^{-1}\}_{j \in \mathbb{N}}$ is non-decreasing. Hence, thanks to (4.4),

$$\sigma^2 \sim b_{M_2}^2 \Leftrightarrow M_2 \sim \sigma^{-1/t},$$

and, thus,

$$\frac{C_{\alpha, \beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \sim \sigma^2 b_{M_2}^{-2} a_{M_2}^{-2} \sim a_{M_2}^{-2} \sim \sigma^{\frac{2s}{t}}.$$

Combining the above terms, we arrive at (A.53). This completes the proof of the proposition. □

A.5.2. Case (ii): Mildly ill-posed problems with super smooth functions

We assume that (A.32) holds true, i.e.,

$$b_j \sim j^{-t}, \quad t > 0, \quad \text{and} \quad a_j \sim \exp\{js\}, \quad s > 0, \quad j \in \mathbb{N}.$$

Proposition A.6. Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.32). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim \varepsilon^2 [\ln(1/\varepsilon)]^{(2t+\frac{1}{2})} \vee \sigma^2. \tag{A.54}$$

Proof. For the second term in (4.13), it is known that (see [18,21]),

$$\sup_{D \in \mathbb{N}} \left[c_{\alpha, \beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \sim \varepsilon^2 [\ln(1/\varepsilon)]^{(2t + \frac{1}{2})}.$$

Consider now the first term in (4.13). Then, the sequence $\{b_j^{-1} a_j^{-1}\}_{j \in \mathbb{N}}$ is non-increasing for each $s, t > 0$, and, thus,

$$\frac{C_{\alpha, \beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \sim \sigma^2.$$

Combining the above terms, we arrive at (A.54). This completes the proof of the proposition. \square

A.5.3. Case (iii): Severely ill-posed problems with ordinary smooth functions

We assume that (A.34) holds true, i.e.,

$$b_j \sim \exp\{-jt\}, \quad t > 0, \quad \text{and} \quad a_j \sim j^s, \quad s > 0, \quad j \in \mathbb{N}.$$

Proposition A.7. Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.34). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim [\ln(1/\varepsilon)]^{-2s} \vee [\ln(1/\sigma)]^{-2s}. \quad (\text{A.55})$$

Proof. For the second term in (4.13), it is known that (see [18,21]),

$$\sup_{D \in \mathbb{N}} \left[c_{\alpha, \beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \sim [\ln(1/\varepsilon)]^{-2s}.$$

Consider now the first term in (4.13). Then, the sequence $\{b_j^{-1} a_j^{-1}\}_{j \in \mathbb{N}}$ is non-decreasing. Hence, thanks to (4.4),

$$\sigma^2 \sim b_{M_2}^2 \quad \Leftrightarrow \quad M_2 \sim \frac{1}{t} \ln(1/\sigma)$$

and, thus,

$$\frac{C_{\alpha, \beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \sim \sigma^2 b_{M_2}^{-2} a_{M_2}^{-2} \sim a_{M_2}^{-2} \sim [\ln(1/\sigma)]^{-2s}.$$

Combining the above terms, we arrive at (A.55). This completes the proof of the proposition. \square

A.5.4. Case (iv): Severely ill-posed problems with super smooth functions

We assume that (A.36) holds true, i.e.,

$$b_j \sim \exp\{-jt\}, \quad t > 0, \quad \text{and} \quad a_j \sim \exp\{js\}, \quad s > 0, \quad j \in \mathbb{N}.$$

Proposition A.8. Assume that the sequences $b = (b_j)_{j \in \mathbb{N}}$ and $a = (a_j)_{j \in \mathbb{N}}$ are given by (A.36). Then, there exists $\varepsilon_0, \sigma_0 \in]0, 1[$ such that, for all $0 < \varepsilon \leq \varepsilon_0$ and $0 < \sigma \leq \sigma_0$, the minimax separation radius $\tilde{r}_{\varepsilon, \sigma}$ satisfies

$$\tilde{r}_{\varepsilon, \sigma}^2 \gtrsim \varepsilon^{\frac{2s}{s+t}} \vee \sigma^{2(\frac{s}{t} \wedge 1)}. \quad (\text{A.56})$$

Proof. For the second term in (4.13), it is known that (see [18,21]),

$$\sup_{D \in \mathbb{N}} \left[c_{\alpha, \beta} \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \sim \varepsilon^{\frac{2s}{s+t}}.$$

Consider now the first term in (4.13). If $s > t$, then the sequence $\{b_j^{-1} a_j^{-1}\}_{j \in \mathbb{N}}$ is non-increasing and, thus,

$$\frac{C_{\alpha, \beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \sim \sigma^2.$$

On the other hand, if $s \leq t$, then the sequence $\{b_j^{-1} a_j^{-1}\}_{j \in \mathbb{N}}$ is non-decreasing. Hence, thanks to (4.4),

$$\sigma^2 \sim b_{M_2}^2 \Leftrightarrow M_2 \sim \frac{1}{t} \ln(1/\sigma),$$

and, thus,

$$\frac{C_{\alpha, \beta}^2}{4} \sigma^2 \max_{1 \leq D \leq M_2} [b_D^{-2} a_D^{-2}] \sim \sigma^2 b_{M_2}^{-2} a_{M_2}^{-2} \sim a_{M_2}^{-2} \sim \sigma^{\frac{2s}{t}}.$$

Combining the above terms, we arrive at (A.53). This completes the proof of the proposition. \square

Acknowledgements

We would like to thank the referees and the associate editor for useful comments and suggestions.

References

- [1] Y. Baraud. Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** (5) (2002) 577–606. [MR1935648](#)
- [2] J. Bigot. Fréchet means of curves for signal averaging and application to ECG data analysis. *Ann. Appl. Stat.* **7** (4) (2013) 2384–2401. [MR3161727](#)
- [3] J. Bigot and S. Gadat. A deconvolution approach to estimation of a common shape in a shifted curves model. *Ann. Statist.* **38** (4) (2010) 2422–2464. [MR2676894](#)
- [4] L. Birgé. An alternative point of view on Lepski’s method. In *State of the Art in Probability and Statistics (Leiden, 1999)* 113–133. *IMS Lecture Notes Monogr. Ser.* **36**. Inst. Math. Statist, Beachwood, OH, 2001. [MR1836557](#)
- [5] N. Bissantz, G. Claeskens, H. Holzmann and A. Munk. Testing for lack of fit in inverse regression – with applications to biophotonic imaging. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** (1) (2009) 25–48. [MR2655522](#)
- [6] C. Butucea. Goodness-of-fit testing and quadratic functional estimation from indirect observations. *Ann. Statist.* **35** (5) (2007) 1907–1930. [MR2363957](#)
- [7] L. Cavalier. Inverse problems in statistics. In *Inverse Problems and High-Dimensional Estimation* 3–96. *Lect. Notes Stat. Proc.* **203**. Springer, Heidelberg, 2011. [MR2868199](#)
- [8] L. Cavalier and Y. Golubev. Risk hull method and regularization by projections of ill-posed inverse problems. *Ann. Statist.* **34** (2006) 1653–1677. [MR2283712](#)
- [9] L. Cavalier, G. K. Golubev, D. Picard and A. B. Tsybakov. Oracle inequalities for inverse problems. *Ann. Statist.* **30** (2002) 843–874. [MR1922543](#)
- [10] L. Cavalier and N. W. Hengartner. Adaptive estimation for inverse problems with noisy operators. *Inverse Probl.* **21** (4) (2005) 1345–1361. [MR2158113](#)
- [11] L. Cavalier and A. B. Tsybakov. Penalized blockwise Stein’s method, monotone oracles and sharp adaptive estimation. *Math. Methods Statist.* **3** (2001) 247–282. [MR1867161](#)
- [12] L. Cavalier and A. B. Tsybakov. Sharp adaptation for inverse problems with random noise. *Probab. Theory Related Fields* **123** (2002) 323–354. [MR1918537](#)
- [13] S. Delattre, M. Hoffmann, D. Picard and T. Vareschi. Blockwise SVD with error in the operator and application to blind deconvolution. *Electron. J. Stat.* **6** (2012) 2274–2308. [MR3020263](#)
- [14] H. W. Engl, M. Hanke and A. Neubauer. *Regularization of Inverse Problems. Mathematics and Its Applications* **375**. Kluwer Academic Publishers Group, Dordrecht, 1996. [MR1408680](#)

- [15] M. Hoffmann and M. Reiss. Nonlinear estimation for linear inverse problems with error in the operator. *Ann. Statist.* **36** (1) (2008) 310–336. [MR2387973](#)
- [16] J. Horowitz. Testing a parametric model against a nonparametric alternative with identification through instrumental variable. *Econometrica* **74** (2006) 521–538. [MR2207399](#)
- [17] Y. I. Ingster, T. Sapatinas and I. A. Suslina. Minimax nonparametric testing in a problem related to the Radon transform. *Math. Methods Statist.* **20** (4) (2011) 347–364. [MR2886641](#)
- [18] Y. I. Ingster, T. Sapatinas and I. A. Suslina. Minimax signal detection in ill-posed inverse problems. *Ann. Statist.* **40** (2012) 1524–1549. [MR3015034](#)
- [19] Y. I. Ingster and I. A. Suslina. *Nonparametric Goodness-of-Fit Testing Under Gaussian Models. Lecture Notes in Statistics* **169**. Springer-Verlag, New York, 2003. [MR1991446](#)
- [20] J. Johannes and M. Schwarz. Adaptive Gaussian inverse regression with partially unknown operator. *Comm. Statist. Theory Methods* **42** (7) (2013) 1343–1362. [MR3031285](#)
- [21] B. Laurent, J.-M. Loubes and C. Marteau. Non asymptotic minimax rates of testing in signal detection with heterogeneous variances. *Electron. J. Stat.* **6** (2012) 91–122. [MR2879673](#)
- [22] B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** (5) (2000) 1302–1338. [MR1805785](#)
- [23] C. Marteau and T. Sapatinas. A unified treatment for non-asymptotic and asymptotic approaches to minimax signal detection. *Stat. Surv.* **9** (2015) 253–297. [MR3452237](#)
- [24] A. Meister. *Deconvolution Problems in Nonparametric Statistics*. Springer-Verlag, New York, 2009. [MR2768576](#)
- [25] V. Nourrit, B. Vohnsen and P. Artal. Blind deconvolution for high-resolution confocal scanning laser ophthalmoscopy. *J. Opt. A, Pure Appl. Opt.* **7** (2005) 585–592.