Minimax hypothesis testing for curve registration

Olivier Collier

▶ To cite this version:

Olivier Collier. Minimax hypothesis testing for curve registration. Electronic journal of statistics , Shaker Heights, OH : Institute of Mathematical Statistics, 2012, 6, pp.1129-1154. <hal-00619808v2>

HAL Id: hal-00619808
https://hal.archives-ouvertes.fr/hal-00619808v2

Submitted on 3 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Electronic Journal of Statistics
ISSN: 1935-7524

Minimax hypothesis testing for curve registration

Olivier Collier

IMAGINE, LIGM, Université Paris Est, Ecole des Ponts ParisTech, France and ENSAE, CREST

Abstract: This paper is concerned with the problem of goodness-of-fit for curve registration, and more precisely for the shifted curve model, whose application field reaches from computer vision and road traffic prediction to medicine. We give bounds for the asymptotic minimax separation rate, when the functions in the alternative lie in Sobolev balls and the separation from the null hypothesis is measured by the l_2-norm. We use the generalized likelihood ratio to build a nonadaptive procedure depending on a tuning parameter, which we choose in an optimal way according to the smoothness of the ambient space. Then, a Bonferroni procedure is applied to give an adaptive test over a range of Sobolev balls. Both achieve the asymptotic minimax separation rates, up to possible logarithmic factors.

Keywords and phrases: Adaptive testing, composite null hypothesis, generalized maximum likelihood, minimax hypothesis testing.

Received January 2012.

Introduction

Curve registration

Our concern is the statistical problem of curve registration, which appears naturally in a large number of applications, when the available data consist of a set of noisy, distorted signals that possess a common structure or pattern. This pattern constitutes the essential information that we want to dig out from the observations. However, the deformations of the signals are generally nonlinear and relatively complex, which complicates the statistical task. Fortunately it is relevant in some cases to assume that the signals only differ from each other by a horizontal shift: we call this modeling the shifted curve model. For instance, it was successfully adopted for the interpretation of the ElectroCardioGramms: each deflection is considered as a repetition of the same signal starting at a random time. Isserles et al. [28] proposed an estimator of the common pattern. Interestingly, the assumptions on the deformations are in practice violated due to the baseline wandering, a periodic vertical perturbation of the potential, but the estimation of the structural pattern performs well yet.

By contrast, SIFT descriptors (*cf.* Lowe [31]) in computer vision are an example where the specification of the deformations is essential: selected keypoints of an image are assigned with descriptors including a histogram of the local gradient. If the image is rotated, the histogram of each keypoint is simply shifted by the angle of the rotation. To match the keypoints of the two images, it is then sufficient to test the adequation of their histograms with the shifted curve model. So, testing the model is sometimes the main concern, and even when estimation matters, the adequation of the model may have to be tested, as the estimation techniques depend on the structure of the deformations.

We refer to the papers Bigot and Gadat [5], Bigot, Gadat, and Loubes [6], Bigot, Gamboa, and Vimond [7], Castillo and Loubes [10], Dalalyan, Golubev, and Tsybakov [14], Dalalyan [13] and Gamboa, Loubes, and Maza [19] for results on the estimation of different features of the curve registration model. The present work builds on Collier and Dalalyan [11], where a comprehensive overview can be found.

Shifted curve model

This paper deals with the shifted curve model, which we will state in a Gaussian sequence form, but which originally relates on two 2π-periodic functions f and $f^\#$ in $\mathbb{L}_2$. Expanding these functions in the complex Fourier basis, we get

$$f(t) = \sum_{j=-\infty}^{+\infty} c_j(f) e^{\mathrm{i}jt} \text{ and } f^\#(t) = \sum_{j=-\infty}^{+\infty} c_j(f^\#) e^{\mathrm{i}jt} \text{ for } t \in [0, 2\pi],$$

where $c_j(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-\mathrm{i}jt}\,dt$ and $c_j(f^\#) = \frac{1}{2\pi}\int_0^{2\pi} f^\#(t)e^{-\mathrm{i}jt}\,dt$.

With this notation, if f and $f^\#$ only differ from each other by a shift, then the Fourier coefficients verify $c_j(f^\#) = e^{\mathrm{i}j\tau}c_j(f)$, for some real τ in $[0, 2\pi]$ and all non-zero integers j. Hence, if we introduce the pseudo-distance d such that

$$d^2((c_1, \ldots), (c_1^\#, \ldots)) \triangleq \inf_\tau \sum_{j=1}^{+\infty} |c_j - e^{-\mathrm{i}j\tau}c_j^\#|^2, \tag{1}$$

and given that $c_j(f) = \overline{c_{-j}(f)}$ for every integer j, testing that f was shifted from $f^\#$ amounts to testing if $d(\boldsymbol{c}, \boldsymbol{c}^\#) = 0$ where $\boldsymbol{c} = (c_1(f), c_2(f), \ldots)$ and $\boldsymbol{c}^\# = (c_1(f^\#), c_2(f^\#), \ldots)$.

Now, if we assume that the observations are given by the white noise model

$$dY(t) = f(t)\,dt + \sigma\,dW(t) \text{ and } dY^\#(t) = f^\#(t)\,dt + \sigma\,dW^\#(t),$$

where $\sigma > 0$ and $W, W^\#$ are independent Wiener processes, we can state our model in a more convenient Gaussian sequence form:

$$\begin{cases} Y_j = c_j + \sigma\xi_j \\ Y_j^\# = c_j^\# + \sigma\xi_j^\# \end{cases}, \qquad j = 1, 2, \ldots, \tag{2}$$

where

- $\{\xi_j, \xi_j^\#; j = 1, 2, \ldots\}$ is a family of independent complex random variables, whose real and imaginary parts are independent standard Gaussian variables,
- σ is assumed to be known.

Our problem amounts to testing H_0 against H_1 with

$$\begin{cases} H_0: \ d(\boldsymbol{c}, \boldsymbol{c}^{\#}) = 0, \\ H_1: \ d(\boldsymbol{c}, \boldsymbol{c}^{\#}) \geq C\rho_\sigma, \end{cases} \tag{3}$$

where C is a positive constant and ρ_σ is a sequence of positive real numbers. For reasons that we shall explain later, we assume that $\boldsymbol{c}$ and $\boldsymbol{c}^{\#}$ belong under the alternative to a Sobolev ball

$$\mathcal{F}_{s,L} \triangleq \Big\{ \boldsymbol{u} = (u_1, u_2, \ldots) : \|\boldsymbol{u}^{(s)}\|_2^2 \triangleq \sum_{j=1}^{\infty} j^{2s} |u_j|^2 \leq L^2 \Big\}, \tag{4}$$

with $s > 0$. With this notation, we denote Θ_0 and Θ_1 the parameter sets corresponding to the hypotheses H_0 and H_1, $\boldsymbol{Y}$ and $\boldsymbol{Y}^{\#}$ the sequences $(Y_1, Y_2, \ldots)$ and $(Y_1^{\#}, Y_2^{\#}, \ldots)$, and we call $\mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^{\#}}$ the probability engendered by $(\boldsymbol{Y}, \boldsymbol{Y}^{\#})$ when the parameters are $\boldsymbol{c}$ and $\boldsymbol{c}^{\#}$.

A detailed discussion of the model is deferred to Section 4, but before that, we point out that our choice of the Gaussian sequence model is not restrictive, since this model is equivalent in Le Cam's sense to many other models, including Gaussian white noise, density estimation (*cf.* Nussbaum [32]), nonparametric regression (*cf.* Brown and Low [8], in the case of random design in Reiß [34], in the case of nonGaussian noise in Grama and Nussbaum [22] and Grama and Nussbaum [23]), ergodic diffusion (*cf.* Dalalyan and Reiß [12]). On the other hand, the Gaussian noise is accepted in computer vision as a good approximation of the Poisson noise, that is more natural in this context.

Minimax testing

A randomized test in our model is a random variable taking values in $[0, 1]$ and measurable with respect to the σ-algebra engendered by $(\boldsymbol{Y}, \boldsymbol{Y}^{\#})$. In practice, the user simulates an independent random variable with a Bernoulli distribution of parameter the value of the test, which was computed from the data $(\boldsymbol{Y}, \boldsymbol{Y}^{\#})$. The null hypothesis is accepted, respectively rejected, when the result of the simulation is 0 or 1. We say that a test is nonrandomized when it only takes the values 0 or 1.

To measure the performance of a test ψ, we choose the minimax point of view, in which the errors of first and second kind are defined by

$$\begin{cases} \alpha(\psi, \Theta_0) = \sup_{\Theta_0} \mathbf{E}_{\boldsymbol{c},\boldsymbol{c}^{\#}}\big(\psi\big), \\ \beta(\psi, \Theta_1) = \sup_{\Theta_1} \mathbf{E}_{\boldsymbol{c},\boldsymbol{c}^{\#}}\big(1 - \psi\big). \end{cases} \tag{5}$$

Note that in the nonrandomized case, $\alpha(\psi, \Theta_0) = \sup_{\Theta_0} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^{\#}}(\psi = 1)$ and $\beta(\psi, \Theta_1) = \sup_{\Theta_1} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^{\#}}(\psi = 0)$.

We say that consistent testing in the asymptotic minimax sense is possible if for all $\alpha, \beta > 0$, there exists a test ψ_σ such that

$$\begin{cases} \overline{\lim\limits_{\sigma \to 0}}\, \alpha(\psi_\sigma, \Theta_0) \leq \alpha, \\ \overline{\lim\limits_{\sigma \to 0}}\, \beta(\psi_\sigma, \Theta_1) \leq \beta. \end{cases} \tag{6}$$

The distance between the null and the alternative hypotheses, $C\rho_\sigma$, determines the existence of such tests. Indeed, if $C\rho_\sigma$ is too small, no testing procedure is asymptotically better than a blind guess, for which $\alpha(\psi, \Theta_0) + \beta(\psi, \Theta_1) = 1$. For a fixed pair α, β, we call ρ_σ^* the asymptotic minimax separation rate if there are two positive constants C_* and C^* such that consistent testing is impossible for $\rho_\sigma = \rho_\sigma^*$ and $C < C_*$, and possible for $\rho_\sigma = \rho_\sigma^*$ and $C > C^*$. The best constants C_* and C^* satisfying these conditions are called exact separation constants. Conventionally, one applies the informal minimal writing length rule to avoid nonuniqueness of the minimax separation rate and of these constants. Moreover, a test which is consistent when $\rho_\sigma = \rho_\sigma^*$ and for some $C > 0$ is called asymptotically minimax rate optimal.

There is a vast literature on the subject of minimax testing: minimax separation rates were investigated in many models, including the Gaussian white noise model, the regression model, the Gaussian sequence model and the probability density model, for the greater part in signal detection, *i.e.,* testing the hypothesis "$f \equiv 0$" against the alternative "$\|f\| \geq C\rho_\sigma$". We present a selective overview of the papers that are the most relevant in the context of this work.

Starting from Ingster [25], Ermakov [15] and Ermakov [16], where the minimax separation rate and the exact separation constants were obtained when the functions in the alternative lie in ellipsoids and the separation from 0 is measured by the l_2-norm, various cases were considered: l_p-bodies as well as Sobolev, Hölder and Besov classes. We refer to Ingster and Suslina [27] and Ingster [26] for a survey. The cases when the functions in the alternative set lie in Sobolev or Hölder classes and the separation from 0 is measured by the sup-norm or by their values at a fixed point were studied in Lepski and Tsybakov [30]. Finally, the case of the $\mathbb{L}_p$-norm with $p < 2$ in Besov classes was considered in Lepski and Spokoiny [29].

Now, all the previously cited results are asymptotic, in the sense that the noise level σ (in the white noise model) tends to 0. But from a practical point of view, it may be interesting to look at the problem from a nonasymptotic point of view. In the regression and Gaussian sequence models, Baraud [1] derived nonasymptotic minimax separation rates when the functions in the alternative lie in l_p-bodies $(0 < p \leq 2)$ and the separation from 0 is measured by the l_2-norm. Baraud, Huet, and Laurent [2, 3] proposed procedures for testing linear or convex hypotheses in the regression model, and Fromont and Lévy-Leduc [18] inspected the improvement implied by a further hypothesis on the periodicity of the signal in the periodic Sobolev balls.

Composite null hypothesis testing

Up to here, we have reviewed results dealing mainly with a simple null hypothesis, namely in the case of signal detection: "$f \equiv 0$". In contrast, the testing problem in the shifted curve model deals with a composite null hypothesis. Here, we give a brief overview of the papers presenting hypothesis testing problems with composite null hypotheses.

The series of papers Baraud [1], Baraud, Huet, and Laurent [2, 3] tackled the case of a nonparametric null hypothesis, but their assumptions are not applicable in our set-up, since our null hypothesis as defined in (3) is neither linear nor convex. On the other hand, the test of a parametric model against a nonparametric one was studied in a substantial number of papers (*cf.* Horowitz and Spokoiny [24] and references therein), but only in Horowitz and Spokoiny [24] from a minimax point of view. The minimax separation rate that they obtained is the same as with a simple null hypothesis. This is due to the strong assumptions made on the behaviour of the estimator of the parameter characterizing the model under H_0.

On a related note, Gayraud and Pouet [20, 21] treated a more general composite null hypothesis in the regression model, that is mainly characterized by its entropy. In fact, the set of functions in the null hypothesis can grow with the sample size, and so be nonparametric. Their rate is the same as in the case of a simple hypothesis. Finally, Butucea and Tribouley [9] also considers the case of a nonparametric null hypothesis, since H_0 is "$f = g$", where f and g are two density functions.

Adaptive testing

A limitation of the minimax approach is that the optimal tests depend on the smoothness class. This is not convenient from a practical point of view, because the choice of the smoothness seems to be unnatural and arbitrary. To obtain handier procedures, we need an adaptive definition for hypothesis testing.

Prior to testing, some sets of smoothness parameters s, L must be chosen, over which adaptation is performed. Typically, these sets are taken as compact intervals $[s_1, s_2]$, $[L_1, L_2]$. To each couple of smoothness parameters (s, L), we associate the smoothness set $\mathcal{F}_{s,L}$, and we write $\Theta_0^{s,L}$ and $\Theta_1^{s,L}$ the corresponding null and alternative hypotheses. Note that, in our problem, $\Theta_0^{s,L} \equiv \Theta_0$ is independent of the smoothness parameters, and that $\Theta_1^{s,L}$ depends on (s, L), not only because $\boldsymbol{c}$ and $\boldsymbol{c}^{\#}$ are in $\mathcal{F}_{s,L}$, but also since ρ_σ is allowed to be a function of s: as a matter of fact, $\Theta_1^{s,L}$ depends on the choice of the radius $C\rho_\sigma(s)$. The easiest way to achieve adaptation is to use the test corresponding to the most constraining smoothness (s_1, L_2), but this entails a significant loss of efficiency if the tested parameters are in fact smoother.

Thus, we prefer a more economical approach and we will say that consistent adaptive testing is possible uniformly over $s \in [s_1, s_2]$ and $L \in [L_1, L_2]$, if for all $\alpha, \beta > 0$, there is a test ψ_σ depending only on $s_1, s_2, L_1, L_2, \alpha$ and β such that

$$\begin{cases} \overline{\lim}_{\sigma \to 0} \alpha(\psi_\sigma, \Theta_0) \leq \alpha, \\ \overline{\lim}_{\sigma \to 0} \sup_{s,L} \beta(\psi_\sigma, \Theta_1^{s,L}) \leq \beta. \end{cases} \tag{7}$$

However, adaptive testing is not always possible without loss of efficiency, *i.e.,* taking $\rho_\sigma(s) = \rho_\sigma^*(s)$ for each s. That is why it was suggested in Spokoiny [37] to replace σ by σd_σ in the expression of $\rho_\sigma^*(s)$, where d_σ is a sequence of positive real numbers, which can be seen as a necessary payment regarding the intensity of the noise to achieve adaptivity.

Now, we say that $\rho_{\sigma d_\sigma}^*(s), s \in [s_1, s_2]$ is the adaptive asymptotic minimax separation rate if there are two positive constants C_* and C^* such that adaptive consistent testing is impossible for $\rho_\sigma(s) = \rho_{\sigma d_\sigma}^*(s)$ and $C < C_*$, and possible for $\rho_\sigma(s) = \rho_{\sigma d_\sigma}^*(s)$ and $C > C^*$.

Spokoiny [37] proved that the optimal asymptotic factor is $(\log\log\sigma^{-1})^{1/4}$, for signal detection in Besov balls. Gayraud and Pouet [21] extended this result for Hölder classes in the regression model.

Fan, Zhang, and Zhang [17] provided a generic tool to construct minimax and adaptive minimax tests: the generalized maximum likelihood, that we also use in the present work to build our procedures both in the nonadaptive and adaptive contexts.

Our contribution

The problem considered in the present work is qualitatively different from the aforementioned works on the minimax separation rate, since our null hypothesis is not only composite but also semiparametric. Furthermore, it seems that the finite-dimensional parameter cannot be uniformly consistently estimated, which contrasts with the situation of Horowitz and Spokoiny [24].

Nevertheless, we propose a testing procedure which is consistent when the separation rate is of order $(\sigma^2\sqrt{\log\sigma^{-1}})^{2s/4s+1}$. This rate is then proven to be minimax, up to a possible logarithmic factor. Indeed, no testing procedure is consistent for a separation rate smaller than $\sigma^{4s/4s+1}$, which is the rate of signal detection in the Gaussian sequence model when the signal to be detected belongs to a Sobolev ball and the separation from 0 is measured by the l_2-norm.

Further, an adaptive test is proposed to circumvent the limitations of the nonadaptive approach. This test is minimax rate optimal, up to a possible logarithmic factor, uniformly over a family of Sobolev balls.

Finally, there is a gap between our lower and upper bounds for the asymptotic minimax separation rate. It could be argued that the lower bound is suboptimal, and that the minimax separation rate for the shifted curve model does contain our logarithmic factor. Indeed, the problem of testing the goodness-of-fit of the shifted curve model can be regarded as an adaptation to the unknown shift parameter. As a matter of fact, if adaptation to the unknown smoothness typically entails a loglog-factor, other types of adaptation can bring simple logarithmic ones: it is proved in Lepski and Tsybakov [30] that the asymptotic minimax separation rate for signal detection when the signal to be detected belongs to a Sobolev or Hölder ball and the separation from 0 is measured by the sup-norm is $(\sigma^2\sqrt{\log\sigma^{-1}})^{s/2s+1}$, while it is $\sigma^{2s/2s+1}$ when the separation from 0 is measured by the value of the signal at a fixed point. The logarithmic factor can be interpreted as a payment for the adaptation of the problem of testing at one point when this point is unknown. Furthermore, note that the same logarithmic factor appears in Fromont and Lévy-Leduc [18], where upper bounds on the minimax separation rate are established in the problem of periodic signal detection with unknown period.

Organization of the paper

The rest of this paper is organized as follows: a nonadaptive procedure is proposed in Section 1, and adjusted in Section 2 to obtain an adaptive test. We also state their minimax performances, which Section 3 indicates to be at least nearly optimal in the minimax sense. The theorems are proved in Sections 5 to 7, and the lemmas used in their proofs are presented in Section 8. The model is discussed in Section 4.

1. Nonadaptive testing procedure

Here, we build a test which will be proven later to be minimax, up to a possible logarithmic factor. Indeed, the procedure achieves the rate $(\sigma^2\sqrt{\log\sigma^{-1}})^{2s/4s+1}$.

Our proposal, which carries on the work presented in Collier and Dalalyan [11], is based on standardized versions $\lambda_\sigma(N)$ of estimators of $d(\boldsymbol{c},\boldsymbol{c}^{\#})$:

$$\left\{\begin{array}{ll} \lambda_\sigma(N) = & \frac{1}{4\sigma^2\sqrt{N}}\min_\tau\left[\sum_{j=1}^N\left|Y_j - e^{-\mathrm{i}j\tau}Y_j^{\#}\right|^2\right]-\sqrt{N}, \\ \psi_\sigma(N,q) = & \mathbb{1}_{\{\lambda_\sigma(N)>q\}}, \end{array}\right. \tag{8}$$

for $N\in\mathbb{N}^*$ and $q\in\mathbb{R}$. Put into words, the test $\psi_\sigma(N,q)$ rejects the null hypothesis when the statistic $\lambda_\sigma(N)$ exceeds the threshold q and accepts it otherwise. The following theorem establishes the minimax properties of this testing procedure for a proper choice of the tuning parameters.

Theorem 1. *Set*

$$\left\{\begin{array}{l} \Theta_0 = \Big\{(\boldsymbol{c},\boldsymbol{c}^{\#})\in l_2\times l_2 \,|\, d(\boldsymbol{c},\boldsymbol{c}^{\#})=0\Big\}, \\ \Theta_1 = \Big\{(\boldsymbol{c},\boldsymbol{c}^{\#})\in \mathcal{F}_{s,L}\times\mathcal{F}_{s,L} \,|\, d(\boldsymbol{c},\boldsymbol{c}^{\#})\geq C\rho_\sigma\Big\}, \end{array}\right. \tag{9}$$

with s and L are positive real numbers, $\rho_\sigma = \big(\sigma^2\sqrt{\log\sigma^{-1}}\big)^{\frac{2s}{4s+1}}$ and $C^2 > 4L^2c_{s,L}^{-2s}+\sqrt{\frac{256\,c_{s,L}}{4s+1}}$, $c_{s,L}=(4sL^2\sqrt{4s+1})^{2/4s+1}$. Denote ψ_σ the test $\psi_\sigma(N,q)$ defined in (8) with $N=N_\sigma(s,L)=[c_{s,L}\,\rho_\sigma^{-1/s}]$ and $q=q_\alpha$, the quantile of order $1-\alpha$ of the standard Gaussian distribution. Then

$$\overline{\lim_{\sigma\to 0}}\,\alpha(\psi_\sigma,\Theta_0)\leq\alpha\,, \tag{10}$$

$$\lim_{\sigma\to 0}\beta(\psi_\sigma,\Theta_1)=0\,. \tag{11}$$

Remark 1. In the rest of this section and in the proof, we skip the dependence of $N_\sigma(s,L)$ in s and L when no confusion is possible.

The proof of this result is given in Section 5. Let us now develop a brief heuristic describing how one could have guessed the optimal value of ρ_σ.

Heuristic for the performance of the nonadaptive procedure

Our proof will show that, under H_0, $\lambda_\sigma(N_\sigma)$ is bounded from above in probability. Thus, we decide to reject the null hypothesis when $\lambda_\sigma(N_\sigma)$ is larger than a constant to be chosen properly.

On the other hand, we inspect the behaviour of the statistic under the alternative hypothesis and give a condition on ρ_σ under which the test statistic is orders of magnitude larger than a constant, so that the procedure can have the desired power.

We derive the lower bound

$$\lambda_\sigma(N_\sigma) \geq \frac{1}{4\sqrt{N_\sigma}\sigma^2} \min_\tau \sum_{j=1}^{N_\sigma} |c_j - e^{-\mathrm{i}j\tau} c_j^\#|^2 - \Big| \sum_{j=1}^{N_\sigma} \frac{|\xi_j|^2 + |\xi_j^\#|^2 - 4}{4\sqrt{N_\sigma}} \Big| \tag{12}$$
$$- \frac{1}{2\sqrt{N_\sigma}} \max_\tau \Big| \sum_{j=1}^{N_\sigma} \mathrm{Re}\left(e^{\mathrm{i}j\tau} \xi_j \overline{\xi_j^\#}\right) \Big| + \text{ negligible terms.}$$

The proof will establish that the second term is bounded in probability, while the third, that we call perturbative, is of order $\sqrt{\log N_\sigma}$. The first term, up to a $4\sqrt{N_\sigma}\sigma^2$ factor, is an approximation of the square of the pseudo-distance $d(\boldsymbol{c}, \boldsymbol{c}^\#)$. Since $\boldsymbol{c}$ and $\boldsymbol{c}^\#$ lie in $\mathcal{F}_{s,L}$, the remainder of the sum can be bounded from above, up to a constant factor, by N_σ^{-2s}. In a nutshell, we get the heuristical lower bound

$$\lambda_\sigma(N_\sigma) \geq Cste \cdot \Big(\frac{d^2(\boldsymbol{c}, \boldsymbol{c}^\#) - Cste \cdot N_\sigma^{-2s}}{\sqrt{N_\sigma}\sigma^2} - O_P(\sqrt{\log N_\sigma}) \Big).$$

Consequently, the alternative is detected as soon as

$$\rho_\sigma^2 \gg \max\left(\sigma^2 \sqrt{N_\sigma}, N_\sigma^{-2s}, \sigma^2 \sqrt{N_\sigma \log N_\sigma}\right) \sim \left(\sigma^2 \sqrt{\log \sigma^{-1}}\right)^{\frac{4s}{4s+1}}.$$

Heuristic for the constant C

We may now ask how small the constant C can be without making our testing procedure inefficient. This constant is only optimized for our test, and we do not claim it to be optimal in the minimax sense.

The previous optimization shows that the test achieves its best rate when N_σ is of the order of $\rho_\sigma^{*\,-1/s}$. Now, denoting $N_\sigma = [c\rho_\sigma^{*\,-1/s}]$, a similar heuristic can give an optimized constant C in the definition of Θ_1. Indeed, Lemma 6 gives the more precise lower bound $(C^2 - 4L^2c^{-2s})\rho_\sigma^{*\,2}$ for the sum in the first term, and we will prove the exact order of magnitude of the third to be $\sqrt{\frac{256\,c}{4s+1}\log N_\sigma}$. Thus

$$\lambda_\sigma(N_\sigma) \geq \left(C^2 - 4L^2 c^{-2s} - \sqrt{\frac{256\,c}{4s+1}}\right)\sqrt{\log N_\sigma}.$$

and this leads to a minimization problem determining the choice of c (*cf.* Theorem 1).

2. Adaptive testing procedure

The procedure given in the previous section possesses asymptotic minimax optimality properties thanks to an appropriate choice of the tuning parameter N_σ, but the practician needs to determine values of s and L to implement the test. As it seems arbitrary and nonintuitive to make assumptions on the smoothness of the signals, it is necessary to design testing procedures independent of s and L that are nearly as good, in the minimax sense, as the procedure proposed in the previous section.

In this section, we only assume that an interval $[s_1, s_2]$ is available such that $c, c^\# \in \mathcal{F}_{s,L}$ for some $s \in [s_1, s_2]$ and $L \in [0, +\infty[$. We propose a testing procedure depending on s_1 and s_2 but independent of s and L, that achieves the same rate of separation, *i.e.*, $\left(\sigma^2\sqrt{\log\sigma^{-1}}\right)^{2s/4s+1}$, as the test based on the precise knowledge of s and L. Furthermore, this rate is achieved uniformly over the Sobolev classes $\mathcal{F}_{s,L}$ with $s \in [s_1, s_2]$ and L belonging to any compact interval included in $\mathbb{R}^+$.

Here is the idea of its construction. The nonadaptive testing procedure proposed above depends on s only via the tuning parameter $N_\sigma(s, L)$. In the followings, we will change the definition of $N_\sigma(s, L)$ to avoid the dependence on L and we will write $N_\sigma(s)$. Using a Bonferroni procedure like in Gayraud and Pouet [21] or Horowitz and Spokoiny [24], we consider the maximum of these tests for several values of $N_\sigma(s)$, more precisely, we consider tests of the form $\tilde\psi_\sigma(q) = \max_{N\in\mathcal{N}} \psi_\sigma(N, q)$. For this kind of test, the next proposition gives bounds for the first and second type errors:

Proposition 1. *Let $\mathcal{N}$ be a set of positive integers and denote $\tilde\psi_\sigma(q)$ the test $\max_{N\in\mathcal{N}} \psi_\sigma(N, q)$, where ψ_σ is defined in 8, then*

$$\begin{cases} \alpha(\tilde\psi_\sigma(q), \Theta_0) \le \sum_{N\in\mathcal{N}} \alpha(\psi_\sigma(N, q), \Theta_0) \\ \beta(\tilde\psi_\sigma(q), \Theta_1^{s,L}) \le \min_{N\in\mathcal{N}} \beta(\psi_\sigma(N, q), \Theta_1^{s,L}). \end{cases}$$

Consequently, the set $\mathcal{N}$ has to be as small as possible (to control the first kind error), but rich enough to approximate the set of all $N_\sigma(s)$ for $s \in [s_1, s_2]$. We will show in the proof that each $N \in \mathcal{N}$ brings adaptation over all Sobolev balls of regularity s such that there is a S such that $N = N_\sigma(S)$ and $S \le s \le S + 1/\log\sigma^{-1}$. Hence, we introduce the following notation leading to a proper choice of $\mathcal{N}$.

For every $s_2 > s_1 > 0$, define

$$\begin{cases} \Sigma(s_1, s_2) = \left\{ s_1 + \frac{j}{\log\sigma^{-1}} \,|\, j \ge 0, s_1 + \frac{j}{\log\sigma^{-1}} \le s_2 \right\}, \\ \mathcal{N}(s_1, s_2) = \left\{ N_\sigma(s) = \left[\rho^*_\sigma(s)^{-1/s}\right] \,|\, s \in \Sigma(s_1, s_2) \right\}. \end{cases} \tag{13}$$

Theorem 2. *Set*

$$\begin{cases} \Theta_0 = \left\{ (\boldsymbol{c}, \boldsymbol{c}^\#) \in l_2 \times l_2 \,|\, d(\boldsymbol{c}, \boldsymbol{c}^\#) = 0 \right\}, \\ \Theta_1^{s,L} = \left\{ (\boldsymbol{c}, \boldsymbol{c}^\#) \in \mathcal{F}_{s,L} \times \mathcal{F}_{s,L} \,|\, d(\boldsymbol{c}, \boldsymbol{c}^\#) \ge \rho_\sigma(s) \right\}, \end{cases} \tag{14}$$

with $C > 0$, $\rho_\sigma(s) = C\rho^*_\sigma(s)$, $\rho^*_\sigma(s) = \left(\sigma^2\sqrt{\log\sigma^{-1}}\right)^{\frac{2s}{4s+1}}$.

Consider the test $\tilde{\psi}_\sigma = \max_{N \in \mathcal{N}(\sigma_1, \sigma_2)} \psi_\sigma\big(N, \sqrt{2 \log\log \sigma^{-1}}\big)$, where ψ_σ is defined in (8). Then, for the interval $[s_1, s_2]$ used in the construction of the test $\tilde{\psi}_\sigma$ and for any interval $[L_1, L_2]$ included in $\mathbb{R}_*^+$, there is a constant C such that

$$\lim_{\sigma \to 0} \alpha\big(\tilde{\psi}_\sigma, \Theta_0\big) = 0, \tag{15}$$

$$\lim_{\sigma \to 0} \sup_{[L_1, L_2]} \sup_{[s_1, s_2]} \beta\big(\tilde{\psi}_\sigma, \Theta_1^{s,L}\big) = 0, \tag{16}$$

Remark 2. In the statement of this theorem, one observes that the constants L_1 and L_2 are not used in the definition of the test, while L was, in the definition of the nonadaptive procedure. Indeed, we optimized the separation constant C and gave an expression depending on L, while this optimization was not our matter in the second theorem.

Remark 3. The theorem claims that there exists a value of C for which the first and second type errors can be controlled. From the proof of the theorem, we see that it is sufficient that such a constant satisfies

$$\begin{cases} C^2 - 4L_2^2 e^{\frac{8}{(4s_1+1)^2}} - \frac{C}{2} > 0 \\ C > \frac{64}{\sqrt{4s_1+1}} \end{cases},$$

which is verified when $C > \max\big(\frac{64}{\sqrt{4s_1+1}}, \frac{1}{4} + (\frac{1}{16} + 4L_2^2 e^{\frac{8}{(4s_1+1)^2}})^{1/2}\big)$.

Heuristic for the performance of the adaptive procedure

Here we explain why our adaptive procedure achieves the same rate as the nonadaptive one. The heuristic of the previous section roughly holds, with this difference that $\max_N \lambda_\sigma(N)$ is of loglog-order under the null hypothesis. But this term is negligible in view of the perturbative term, so that the performances of the test do not deteriorate in the adaptive problem.

3. Lower bound for the minimax rate

After stating the performance of our tests, we prove in this section that they are at least nearly rate optimal. Indeed, we are able to establish a lower bound for our model, by proving that the detection of a signal lying in a Sobolev ball when the separation from 0 is measured by the l_2-norm (*cf.* (17) for a precise definition) is simpler than ours, in the sense that every lower bound result for this model is adaptable for our purpose.

Let us first introduce the classical signal detection problem, for which the minimax separation rate, and even the exact separation constants, are known:

$$\begin{cases} Y_j = c_j + \sigma \xi_j, \quad j = 1, 2, \ldots, \\ \Theta_0^{\mathbf{class}} = \{0\}, \\ \Theta_1^{\mathbf{class}} = \Big\{ \boldsymbol{c} \in \mathcal{F}_{s,L} \,\Big|\, \|\boldsymbol{c}\|_2 \geq C\rho_\sigma \Big\}. \end{cases} \tag{17}$$

For this model, we define the errors of first and second kind of a test $\psi^{\mathbf{class}}$ by

$$\begin{cases} \alpha^{\mathbf{class}}(\psi^{\mathbf{class}}, \Theta_0^{\mathbf{class}}) = \sup_{\Theta_0^{\mathbf{class}}} \mathbf{E}_{\boldsymbol{c}}\big(\psi^{\mathbf{class}}\big), \\ \beta^{\mathbf{class}}(\psi^{\mathbf{class}}, \Theta_1^{\mathbf{class}}) = \sup_{\Theta_1^{\mathbf{class}}} \mathbf{E}_{\boldsymbol{c}}\big(1 - \psi^{\mathbf{class}}\big), \end{cases} \tag{18}$$

where we denote $\mathbf{P}_{\boldsymbol{c}}$ the probability engendered by $\boldsymbol{Y} = (Y_1, Y_2, \ldots)$ when $(c_1, c_2, \ldots) = \boldsymbol{c}$.

Theorem 3. *Given the two models exposed in (2) and (17), we have*

$$\inf_{\psi_\alpha} \beta(\psi_\alpha, \Theta_1) \geq \inf_{\psi_\alpha^{\mathbf{class}}} \beta^{\mathbf{class}}(\psi_\alpha^{\mathbf{class}}, \Theta_1^{\mathbf{class}}) \,, \tag{19}$$

where the infima are taken over all tests of level α respectively for our model and for the classical one.

Thus, our model can benefit from every lower bound result on model (17). We choose to exploit the nonasymptotic results presented in Baraud [1], Proposition 3. The following theorem shows that the asymptotic minimax separation rate for our problem is not smaller than $\sigma^{4s/4s+1}$.

Corollary. *Let α and β be in $]0,1]$. Define $\eta = 2(1-\alpha-\beta)$, $\mathcal{L} = \log(1+\eta^2)$ and $\rho^2 = \sup_{d\geq 1}\big[\sqrt{2\mathcal{L}d}\sigma^2 \wedge L^2 d^{-2s}\big]$. Then*

$$\rho_\sigma \leq \rho \quad \Rightarrow \quad \inf_{\psi^\alpha} \beta\Big(\psi^\alpha, \Theta_1\Big) \geq \beta, \tag{20}$$

where the infimum is taken over all tests of level α for the shifted curve model.

Remark 4. We can approximate ρ by computing

$$\sup_{x\in\mathbb{R}^+} \big[\sqrt{2\mathcal{L}x}\sigma^2 \wedge L^2 x^{-2s}\big] = L^{\frac{1}{4s+1}} \big(\sigma^2\sqrt{2\mathcal{L}}\big)^{\frac{2s}{4s+1}}.$$

Remark 5. Our proof shows that every lower bound result for adaptive testing could be used for our purpose as well, for instance Gayraud and Pouet [21].

4. Discussion

Model

The choice of our model was inspired by practical considerations, and we intend to apply it to a problem in computer vision: that of keypoint matching as briefly discussed in Collier and Dalalyan [11]. Accordingly, it is necessary to justify the realism of model (2).

Variance

Although the theoretical analysis of this paper is carried out for the Gaussian sequence model, the procedure we propose admits a simple counterpart in the regression model, at least in the case of deterministic equidistant design. According to the theory on the asymptotic equivalence, our results hold true for this model as well, provided that $s > 1/2$ (*cf.* Rohde [36]). However, in the model of regression, it is not realistic to assume that the variance of noise is known in advance.

Nevertheless, one can compute a consistent estimator of the variance (*cf.* Rice [35]) and plug this estimator in the testing procedure. In an analogous setup, it is proved in Gayraud and Pouet [20] for example, that this plug-in strategy preserves the rate-optimality of the testing procedure. We believe that a similar result can be deduced in our set-up as well.

Symmetry of the model

In our modelization, the two parts corresponding in the Gaussian white noise model to two different functions are treated symmetrically: the same model, with the same variance and the same noise, applies to both. But, in applications, the signals that we want to match with each other are thought to have the same nature. In addition, it seems that it is not meaningful to consider the case when the regularities of the Sobolev balls are different for the signals: under H_0, the regularity has to be the same.

Besides, one could want to normalize both equations to get the same variance for both sides. But, this would also change the functions, which would not only differ from each other by a shift, but also by a dilatation. Therefore, the application of our methodology to this case is not straightforward. However, a detailed inspection reveals that our results carry over to the case when we replace σ by $\max(\sigma, \sigma^{\#})$.

Weighted estimator

In Collier and Dalalyan [11], another estimator of $d^2(\boldsymbol{c}, \boldsymbol{c}^{\#})$ is used, stemming from a penalization of the log-likelihood ratio. This could be adapted in our context by considering the test statistic

$$\lambda_\sigma^{\mathbf{w}} = \frac{1}{4\sigma^2\sqrt{N_\sigma}} \min_\tau \left[\sum_{j=1}^{+\infty} w_j \big|Y_j - e^{-\mathrm{i}j\tau} Y_j^{\#}\big|^2\right] - \|\mathbf{w}\|_2, \tag{21}$$

where $\mathbf{w} = (w_1, w_2, \ldots)$ is a sequence of real numbers in $[0,1]$ depending on σ. Under some conditions on $\mathbf{w}$, our study would undergo only few modifications, and only the optimal constants would be changed. For simplicity sake, we chose not to consider the weighted estimator.

From classical signal detection to shift testing

A first guess to try solving our problem could be to use an estimator $\hat{\tau}$ of the shift and to apply the classical signal detection methods to the sequence

$(Y_j - e^{\mathrm{i}j\hat{\tau}} Y_j^{\#})$. But this approach fails, since it is not possible to get any consistent estimator of the shift. Indeed, for example, the shift may not be identifiable. Consequently, the study of the perturbative term (*cf.* first heuristic after Theorem 1) is unavoidable, in order to take into account every possible shift. We think that this uncertainty entails a price, *i.e.,* a supplementary factor in the minimax separation rate.

Future research

Our model is only a simple version of the curve registration problem. In further work, we could study what happens when the signals are shifted and dilated by considering the pseudo-distance

$$\tilde{d}^2(\boldsymbol{c}, \boldsymbol{c}^{\#}) = \inf_{\tau, a} \sum_{j=1}^{+\infty} |c_j - a\, e^{\mathrm{i}j\tau} c_j^{\#}|^2.$$

Once again, the problem is whether it is possible to estimate the dilatation parameter consistently.

5. Proof of Theorem 1

5.1. First kind error

Here, we prove that the asymptotic first kind error of the test ψ_σ does not exceed the prescribed level α. To this end, denote τ^* a real number such that, under H_0, $\forall j \geq 1$, $c_j^{\#} = e^{\mathrm{i}j\tau^*} c_j$. We skip the dependence of τ^* on $\boldsymbol{c}$ and $\boldsymbol{c}^{\#}$. Using the inequality

$$\min_\tau \sum_{j=1}^{N_\sigma} \left|Y_j - e^{-\mathrm{i}j\tau} Y_j^{\#}\right|^2 \leq \sum_{j=1}^{N_\sigma} \left|Y_j - e^{-\mathrm{i}j\tau^*} Y_j^{\#}\right|^2 = \sigma^2 \sum_{j=1}^{N_\sigma} \left|\xi_j - e^{-\mathrm{i}j\tau^*} \xi_j^{\#}\right|^2,$$

we get

$$\begin{aligned}
\alpha(\psi_\sigma, \Theta_0) &= \sup_{\Theta_0} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^{\#}} \Big(\frac{1}{4\sigma^2\sqrt{N_\sigma}} \min_\tau \sum_{j=1}^{N_\sigma} \left|Y_j - e^{-\mathrm{i}j\tau} Y_j^{\#}\right|^2 - \sqrt{N_\sigma} > q_\alpha \Big) \\
&\leq \mathbf{P}\Big(\frac{1}{4\sqrt{N_\sigma}} \sum_{j=1}^{N_\sigma} \left(\eta_j^2 + \tilde{\eta}_j^2 - 4\right) > q_\alpha \Big), \\
&\quad \text{where } \eta_j = \mathrm{Re}(\xi_j - e^{-\mathrm{i}j\tau^*}\xi_j^{\#}), \tilde{\eta}_j = \mathrm{Im}(\xi_j - e^{-\mathrm{i}j\tau^*}\xi_j^{\#}) \overset{iid}{\sim} \mathcal{N}(0,2).
\end{aligned}$$

Finally, using Berry-Esseen's inequality (*cf.* Theorem 5), we get

$$\alpha(\psi_\sigma, \Theta_0) \leq \alpha + \frac{1}{\sqrt{2\pi N_\sigma}}\,,$$

and this gives the desired asymptotic level.

5.2. Second kind error

It remains to study the second kind error of the test, and to show that it tends to 0. Our proof is based on the heuristic given earlier in Section 1: we decompose $\lambda_\sigma(N_\sigma)$ into several terms, and make use of their respective orders of magnitude. The decomposition gives

$$\begin{aligned}
&4\sigma^2\sqrt{N_\sigma}\lambda_\sigma(N_\sigma)\\
&\geq \min_\tau \Big\{ \sum_{j=1}^{N_\sigma} |c_j - e^{-\mathrm{i}j\tau} c_j^\#|^2 + 2\sigma \sum_{j=1}^{N_\sigma} \mathrm{Re}\left((c_j - e^{-\mathrm{i}j\tau}c_j^\#)\overline{(\xi_j - e^{-\mathrm{i}j\tau}\xi_j^\#)}\right) \Big\} \qquad (22)\\
&- \sigma^2\sqrt{N_\sigma}\Big|\sum_{j=1}^{N_\sigma} \frac{|\xi_j|^2 + |\xi_j^\#|^2 - 4}{\sqrt{N_\sigma}}\Big| - 2\sigma^2 \max_\tau \Big|\sum_{j=1}^{N_\sigma} \mathrm{Re}\left(e^{\mathrm{i}j\tau}\xi_j\overline{\xi_j^\#}\right)\Big|.
\end{aligned}$$

For simplicity sake, we introduce some notation:

$$\begin{cases}
D_\sigma(\boldsymbol{c}, \boldsymbol{c}^\#) = \min_\tau \Big\{ \sum_{j=1}^{N_\sigma} |c_j - e^{-\mathrm{i}j\tau}c_j^\#|^2 \\
+ 2\sigma \sum_{j=1}^{N_\sigma} \mathrm{Re}\left((c_j - e^{-\mathrm{i}j\tau}c_j^\#)\overline{(\xi_j - e^{-\mathrm{i}j\tau}\xi_j^\#)}\right)\Big\},\\
A_\sigma = \Big|\sum_{j=1}^{N_\sigma} \frac{|\xi_j|^2+|\xi_j^\#|^2-4}{\sqrt{N_\sigma}}\Big|,\\
B_\sigma = \max_\tau \Big|\sum_{j=1}^{N_\sigma} \mathrm{Re}\left(e^{\mathrm{i}j\tau}\xi_j\overline{\xi_j^\#}\right)\Big|,
\end{cases}$$

which, combined with (22), leads to:

$$\beta(\psi_\sigma, \Theta_1) \leq \sup_{\Theta_1} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(D_\sigma(\boldsymbol{c}, \boldsymbol{c}^\#) - \sigma^2\sqrt{N_\sigma}A_\sigma - 2\sigma^2 B_\sigma \leq 4q_\alpha\sigma^2\sqrt{N_\sigma}\Big).$$

In addition to $c_{s,L}$, introduced in the definition of N_σ, we will need the constant c' and ϵ, defined as

$$\begin{cases}
c' = \sqrt{\frac{256\, c_{s,L}}{4s+1}},\\
\epsilon = \frac{1}{2}\left(C^2 - 4L^2 c_{s,L}^{-2s} - \sqrt{\frac{256\, c_{s,L}}{4s+1}}\right).
\end{cases}$$

Separating the different terms to study them independently, we write

$$\begin{aligned}
\beta(\psi_\sigma, \Theta_1) &\leq \sup_{\Theta_1} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(D_\sigma(\boldsymbol{c}, \boldsymbol{c}^\#) \leq (c' + \epsilon + \frac{4q_\alpha\sqrt{c_{s,L}}}{\sqrt{\log\sigma^{-1}}})\rho_\sigma^2\Big)\\
&+ \mathbf{P}\Big(\sigma^2\sqrt{N_\sigma}A_\sigma > \epsilon\rho_\sigma^2\Big) + \mathbf{P}\Big(2\sigma^2 B_\sigma > c'\rho_\sigma^2\Big).
\end{aligned}$$

- Let us first study $\sup_{\Theta_1} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\left(D_\sigma(\boldsymbol{c}, \boldsymbol{c}^\#) \leq (c' + \epsilon + \frac{4q_\alpha\sqrt{c_{s,L}}}{\sqrt{\log\sigma^{-1}}})\rho_\sigma^2\right)$, which contains the dominant term when ρ_σ is too large.

Denoting $\delta = \sqrt{C^2 - 4L^2 c_{s,L}^{-2s}}$, Lemma 1 allows to apply Lemma 2 with $x_0 = \delta\rho_\sigma$ and $M = (c' + \epsilon + \frac{4q_\alpha\sqrt{c_{s,L}}}{\sqrt{\log\sigma^{-1}}})\rho_\sigma^2$. The choice of the parameters yields for σ small enough

$$(\frac{\delta}{4} - \frac{c'+\epsilon}{4\delta} - \frac{q_\alpha\sqrt{c_{s,L}}}{\delta\sqrt{\log\sigma^{-1}}})\rho_\sigma > 0\,,$$

so that the second part of Lemma 2 holds:

$$\begin{aligned}
&\sup_{\Theta_1} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(D_\sigma(\boldsymbol{c},\boldsymbol{c}^\#) \le (c' + \epsilon + \frac{4q_\alpha\sqrt{c_{s,L}}}{\sqrt{\log\sigma^{-1}}})\rho_\sigma^2\Big)\\
&\le 2\Big(1 + \delta^{-1}L\,\rho_\sigma^{-1}\max\{1, N_\sigma^{1-s}\}\Big)\\
&\times\Big[\exp\Big\{-\Big(\delta^2 - c' - \epsilon - \frac{4q_\alpha\sqrt{c_{s,L}}}{\sqrt{\log\sigma^{-1}}}\Big)^2\frac{\rho_\sigma^2}{32\delta^2\sigma^2}\Big\} + \exp\Big\{-\frac{\rho_\sigma^2\delta^2}{8\sigma^2}\Big\}\Big]\\
&\overset{\sigma\to 0}{\longrightarrow} 0,\ \text{ since } \rho_\sigma/\sigma \to 0 \text{ as } \sigma \to 0.
\end{aligned}$$

- Let us now turn to $\mathbf{P}\big(\sigma^2\sqrt{N_\sigma}A_\sigma > \epsilon\rho_\sigma^2\big)$. Prior to using Berry-Esseen's inequality (*cf.* Theorem 5), we derive $\frac{\epsilon\rho_\sigma^2}{4\sigma^2\sqrt{N_\sigma}} \ge \frac{\epsilon}{4\sqrt{c_{s,L}}}\sqrt{\log\sigma^{-1}}$, so that, putting $x = \frac{\epsilon}{4\sqrt{c_{s,L}}}\sqrt{\log\sigma^{-1}}$ into the formula of the theorem and using the bound $1 - \Phi(x) \le \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}$ for every positive x,

$$\mathbf{P}\Big(\sigma^2\sqrt{N_\sigma}A_\sigma > \epsilon\rho_\sigma^2\Big) \le \sqrt{\frac{2}{\pi N_\sigma}} + \sqrt{\frac{32c_{s,L}}{\pi\epsilon^2}}\frac{\sigma^{\frac{\epsilon^2}{32c}}}{\sqrt{\log\sigma^{-1}}} \to 0.$$

- Finally, it remains to control $\mathbf{P}\big(2\sigma^2B_\sigma > c'\rho_\sigma^2\big)$. We apply Lemma 3:

$$\begin{aligned}
\mathbf{P}\Big(2\sigma^2B_\sigma > c'\rho_\sigma^2\Big) &\le 2c(\log\sigma^{-1})^{\frac{-1}{4s+1}}\sigma^{\frac{c'^2}{64c} - \frac{4}{4s+1}} + e^{-N_\sigma/2}\\
&\le 2c(\log\sigma^{-1})^{\frac{-1}{4s+1}} + e^{-N_\sigma/2} \to 0.
\end{aligned}$$

6. Proof of Theorem 2

6.1. Proposition 1

Let $\mathcal{N}$ be a set of positive integers and denote $\tilde\psi_\sigma(q) = \max_{N\in\mathcal{N}}\psi_\sigma(N,q)$, where ψ_σ is defined in 8.

- Concerning the first kind error:

$$\alpha(\tilde\psi_\sigma(q),\Theta_0) = \sup_{(\boldsymbol{c},\boldsymbol{c}^\#)\in\Theta_0}\mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(\max_{N\in\mathcal{N}}\min_\tau\sum_{j=1}^N|Y_j - e^{-\mathrm{i}j\tau}Y_j^\#|^2 > q\Big)$$

$$\leq \sum_{N\in\mathcal{N}} \sup_{(\boldsymbol{c},\boldsymbol{c}^\#)\in\Theta_0} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(\min_\tau \sum_{j=1}^N |Y_j - e^{-\mathrm{i}j\tau}Y_j^\#|^2 > q\Big)$$

$$= \sum_{N\in\mathcal{N}} \alpha(\psi_\sigma(N,q),\Theta_0).$$

- Concerning the second kind error:

$$\beta(\tilde{\psi}_\sigma(q),\Theta_1^{s,L}) = \sup_{\Theta_1^{s,L}} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(\max_{N\in\mathcal{N}} \min_\tau \sum_{j=1}^N |Y_j - e^{-\mathrm{i}j\tau}Y_j^\#|^2 \leq q\Big)$$

$$\leq \sup_{\Theta_1^{s,L}} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(\min_\tau \sum_{j=1}^N |Y_j - e^{-\mathrm{i}j\tau}Y_j^\#|^2 \leq q\Big), \forall N\in\mathcal{N},$$

$$\leq \min_{N\in\mathcal{N}} \beta(\psi_\sigma(N,q),\Theta_1^{s,L}).$$

6.2. First kind error

Here, we prove that the first kind error of the test $\tilde{\psi}_\sigma$ converges to 0. To this end, denote τ^* a real number such that, under H_0, $\forall j\geq 1$, $c_j^\# = e^{\mathrm{i}j\tau^*}c_j$. We skip the dependence of τ^* on $\boldsymbol{c}$ and $\boldsymbol{c}^\#$. Using the inequality

$$\min_\tau \sum_{j=1}^{N_\sigma} \big|Y_j - e^{-\mathrm{i}j\tau}Y_j^\#\big|^2 \leq \sum_{j=1}^{N_\sigma} \big|Y_j - e^{-\mathrm{i}j\tau^*}Y_j^\#\big|^2 = \sigma^2 \sum_{j=1}^{N_\sigma} \big|\xi_j - e^{-\mathrm{i}j\tau^*}\xi_j^\#\big|^2,$$

we get

$$\alpha\big(\tilde{\psi}_\sigma,\Theta_0\big) \leq \sum_{N\in\mathcal{N}(s_1,s_2)} \mathbf{P}\bigg(\frac{1}{4\sqrt{N}}\sum_{j=1}^N(\eta_j^2+\tilde{\eta}_j^2-4) > \sqrt{2\,\log\log\sigma^{-1}}\bigg),$$

where $\eta_j = \mathrm{Re}(\xi_j - e^{-\mathrm{i}j\tau^*}\xi_j^\#), \tilde{\eta}_j = \mathrm{Im}(\xi_j - e^{-\mathrm{i}j\tau^*}\xi_j^\#) \overset{iid}{\sim} \mathcal{N}(0,2)$.

Thus, using Berry-Esseen's inequality (*cf.* Theorem 5 with $x=\sqrt{2\;\log\log\sigma^{-1}}$) and the bound $1-\Phi(x)\leq \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}$ for every positive x,

$$\alpha\big(\tilde{\psi}_\sigma,\Theta_0\big) \leq \sum_{N\in\mathcal{N}(s_1,s_2)} \Big\{\frac{1}{\sqrt{2\pi N}} + \frac{\exp(-\log\log\sigma^{-1})}{\sqrt{4\pi\;\log\log\sigma^{-1}}}\Big\}$$

$$\leq \frac{1}{\sqrt{2\pi}}\frac{\mathbf{Card}\,\mathcal{N}(s_1,s_2)}{\sqrt{N_\sigma(s_2)}} + \frac{1}{\sqrt{4\pi}}\frac{\mathbf{Card}\,\mathcal{N}(s_1,s_2)}{\log\sigma^{-1}\sqrt{\log\log\sigma^{-1}}}.$$

As $\mathbf{Card}\,\mathcal{N}(s_1,s_2) = 1+\big[\,(s_2-s_1)\log\sigma^{-1}\big]$ is of logarithmic order, this implies that $\alpha\big(\tilde{\psi}_\sigma,\Theta_0\big)\to 0$.

6.3. Second kind error

Finally, we study the second kind error and prove that it converges to 0.

For $s \in [s_1, s_2]$, define $S = \max\{t \in \Sigma(s_1, s_2) \mid t \le s\}$, where we omit the dependence of S in s for simplicity sake. Note that $0 \le s - S \le \frac{1}{\log \sigma^{-1}}$. S is an approximation of s which will be sufficient for our purpose according to Lemma 6.

We introduce the notation

$$\begin{cases} D_\sigma^s(\boldsymbol{c}, \boldsymbol{c}^\#) = \min_\tau \Big\{ \sum_{j=1}^{N_\sigma(s)} |c_j - e^{-\mathrm{i}j\tau} c_j^\#|^2 \\ + 2\sigma \sum_{j=1}^{N_\sigma(s)} \mathrm{Re}\left((c_j - e^{-\mathrm{i}j\tau} c_j^\#)(\overline{\xi_j - e^{-\mathrm{i}j\tau}\xi_j^\#})\right)\Big\}, \\ A_\sigma^s = \Big| \sum_{j=1}^{N_\sigma(s)} \frac{|\xi_j|^2 + |\xi_j^\#|^2 - 4}{\sqrt{N_\sigma(s)}} \Big|, \\ B_\sigma^s = \max_\tau \Big| \sum_{j=1}^{N_\sigma(s)} \mathrm{Re}\left(e^{\mathrm{i}j\tau}\xi_j \overline{\xi_j^\#}\right)\Big|. \end{cases}$$

and computations similar to those of the previous section yield

$$\begin{aligned} &\sup_{[L_1,L_2]} \sup_{[s_1,s_2]} \beta(\tilde{\psi}_\sigma, \Theta_1^{s,L}) \\ &\le \sup_{s,L} \sup_{\Theta_1^{s,L}} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(D_\sigma^S(\boldsymbol{c}, \boldsymbol{c}^\#) \le \sigma^2 \sqrt{32\, N_\sigma(S)\, \log\log \sigma^{-1}} + \frac{C}{2}\rho_\sigma^2(S)\Big) \\ &+ \sum_{s\in\Sigma} \mathbf{P}\Big(\sigma^2 \sqrt{N_\sigma(s)} A_\sigma^s > \frac{C}{4}\rho_\sigma^2(s)\Big) + \sum_{s\in\Sigma} \mathbf{P}\Big(2\sigma^2 B_\sigma^s > \frac{C}{4}\rho_\sigma^2(s)\Big). \end{aligned}$$

- Let us study $\sup_{s,L} \sup_{\Theta_1^{s,L}} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\big(D_\sigma^S(\boldsymbol{c}, \boldsymbol{c}^\#) \le \sigma^2 \sqrt{32\, N_\sigma(S)\, \log\log\sigma^{-1}} + \frac{C}{2}\rho_\sigma^2(S)\big)$.
  Lemma 6 implies

  $$\big(N_\sigma(S) + 1\big)^{-2s} \le \rho_\sigma^*(S)^2 \le e^{\frac{8}{(4s_1+1)^2}}\, \rho_\sigma^*(s)^2,$$

  so that, denoting $\delta^2 = C^2 - 4L^2 e^{\frac{8}{(4s_1+1)^2}}$, Lemma 1 allows to apply Lemma 2 with $x_0 = \delta\rho_\sigma^*(s)$ and $M = \sigma^2\sqrt{32\, N_\sigma(S) \log\log\sigma^{-1}} + \frac{C}{2}\rho_\sigma^2(s)$. On the other hand, the choice of δ entails that for C large and σ small enough

  $$\forall\, s \in [s_1, s_2], \quad \big(\frac{\delta}{4} - \frac{C}{8\delta}\big)\rho_\sigma^*(s) - \frac{\sigma^2\sqrt{2\, N_\sigma(S)\log\log\sigma^{-1}}}{\delta\rho_\sigma^*(s)} > 0.$$

  Hence, applying the second part of Lemma 5, we get an inequality where the right-hand side converges to 0 as σ tends to 0:

  $$\begin{aligned} &\sup_s \sup_{\Theta_1^{s,L}} \mathbf{P}_{\boldsymbol{c},\boldsymbol{c}^\#}\Big(D_\sigma^S(\boldsymbol{c}, \boldsymbol{c}^\#) \le \sigma^2\sqrt{32\, N_\sigma(S)\, \log\log\sigma^{-1}} + \frac{C}{2}\rho_\sigma^2(S)\Big) \\ &\le 2\Big(1 + \delta^{-1} L\, \rho_\sigma(s_2)^{-1} \max\{1, N_\sigma(s_1)^{1-s_1}\}\Big) \end{aligned}$$

$$\times \Bigg[\exp\Big\{ - \Big((\delta^2 - \frac{C}{2})\rho_\sigma^2(s_1) - \sqrt{32\, N_\sigma(s_1) \log\log \sigma^{-1}} \Big)^2 / 32\delta^2 \rho_\sigma^2(s_1)\sigma^2 \Big\}$$
$$+ \exp\Big\{ - \frac{\rho_\sigma^2(s_2)\delta^2}{8\sigma^2} \Big\} \Bigg].$$

- Consider the second term. Berry-Esseen's theorem (*cf.* Theorem 5) implies the following inequality, where the right-hand side converges to 0 as σ tends to 0:

$$\sum_{s\in\Sigma} \mathbf{P}\Big(\sigma^2 \sqrt{N_\sigma(s)} A_\sigma^s > \frac{C}{4}\rho_\sigma^2(s) \Big)$$
$$\leq \mathbf{Card}\, \mathcal{N}(s_1, s_2) \cdot \Bigg[\sqrt{\frac{2}{\pi N_\sigma(s_2)}} + \sqrt{\frac{128}{\pi C}} \frac{\sigma^{\frac{C}{128}}}{\sqrt{\log \sigma^{-1}}} \Bigg].$$

- Let us turn to the third term. We apply Lemma 3 and get an inequality where once again the right-hand side converges to 0 as σ tends to 0:

$$\sum_{s\in\Sigma} \mathbf{P}\Big(2\sigma^2 B_\sigma^s > \frac{C}{4}\rho_\sigma^2(s) \Big)$$
$$\leq \mathbf{Card}\, \mathcal{N}(s_1, s_2) \cdot \Big[2(\log \sigma^{-1})^{\frac{-1}{4s_2+1}} \sigma^{\frac{C^2}{1024} - \frac{4}{4s_1+1}} + e^{-N_\sigma/2} \Big].$$

7. Proof of Theorem 3

Consider a randomized test ψ in the shifted curve model. We will define a corresponding test in the classical model with smaller first and second kind errors, and it is sufficient to establish the result.

First note that there is a measurable function f with respect to the σ-algebra engendered by the sequences $\boldsymbol{Y}$ and $\boldsymbol{Y}^\#$ and with values in $[0,1]$ such that $\psi = f(\boldsymbol{Y}, \boldsymbol{Y}^\#)$. Denoting $\boldsymbol{\epsilon}$ a sequence of i.i.d random variables $\mathcal{N}(0, \sigma^2)$ independent from $\boldsymbol{Y}$, we define $\psi^{\mathbf{class}} = \mathbf{E}_{\boldsymbol{\epsilon}}\big(f(\boldsymbol{Y}, \boldsymbol{\epsilon}) | \boldsymbol{Y} \big)$, where $\mathbf{E}_{\boldsymbol{\epsilon}}$ is the integration with respect to the probability engendered by $\boldsymbol{\epsilon}$. $\psi^{\mathbf{class}}$ is $\sigma(\boldsymbol{Y})$-measurable and thus constitutes a test for the classical model.

This testing procedure can be interpreted as a test in the shifted curve model when $\boldsymbol{c}^\# = 0$. Indeed, $d(\boldsymbol{c}, \boldsymbol{c}^\#) = \|\boldsymbol{c}\|_2$ when $\boldsymbol{c}^\# = 0$, so that $\Theta_0^{\mathbf{class}} \times 0 \subseteq \Theta_0$ and $\Theta_1^{\mathbf{class}} \times 0 \subseteq \Theta_1$. By Tonelli-Fubini's theorem, $\psi^{\mathbf{class}}$ satisfies

$$\begin{aligned} \alpha^{\mathbf{class}}(\psi^{\mathbf{class}}, \Theta_0^{\mathbf{class}}) &= \sup_{\Theta_0^{\mathbf{class}}} \mathbf{E}_{\boldsymbol{c}}\big(\psi^{\mathbf{class}}\big) \\ &= \sup_{\Theta_0^{\mathbf{class}}} \mathbf{E}_{\boldsymbol{c},0}\big(f(\boldsymbol{Y}, \boldsymbol{Y}^\#) \big) \\ &\leq \alpha(\psi, \Theta_0). \end{aligned}$$

A similar inequality holds concerning the second kind error.

8. Lemmas

Lemma 1. *Let $\boldsymbol{c} = (c_1, c_2, \dots)$ and $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \tilde{c}_2, \dots)$ in $\mathcal{F}_{s,L}$, with $s > 0$, be such that $d(\boldsymbol{c}, \tilde{\boldsymbol{c}}) \geq C\rho$, and let $N + 1 \geq c\rho^{-1/s}$. Then*

$$\min_{\tau} \sum_{j=1}^{N} |c_j - e^{-ij\tau}\tilde{c}_j|^2 \geq (C^2 - 4L^2 c^{-2s})\rho^2.$$

Proof of Lemma 1. Since both $\boldsymbol{c}$ and $\tilde{\boldsymbol{c}}$ belong to the Sobolev ball, it holds that

$$\begin{aligned}
\sum_{j>N} |c_j - e^{-ij\tau}\tilde{c}_j|^2 &\leq \sum_{j>N} \big(2|c_j|^2 + 2|\tilde{c}_j|^2\big) \\
&\leq 2(N+1)^{-2s} \sum_{j>N} j^{2s}\big(|c_j|^2 + |\tilde{c}_j|^2\big) \\
&\leq 4L^2(N+1)^{-2s}.
\end{aligned}$$

Consequently, taking into account that $\sum_{j=1}^{\infty} |c_j - e^{-ij\tau}\tilde{c}_j|^2 \geq d^2(c, \tilde{c}) \geq C^2\rho^2$, we get

$$\begin{aligned}
\sum_{j=1}^{N} |c_j - e^{-ij\tau}\tilde{c}_j|^2 &= \sum_{j=1}^{\infty} |c_j - e^{-ij\tau}\tilde{c}_j|^2 - \sum_{j>N} |c_j - e^{-ij\tau}\tilde{c}_j|^2 \\
&\geq C^2\rho^2 - 4L^2(N+1)^{-2s},
\end{aligned}$$

and the result follows in view of $N + 1 \geq c\rho^{-1/s}$. $\square$

Lemma 2. *Let N be some positive integer, let ξ_j, $\tilde{\xi}_j$, $j = 1, \dots, N$ be independent complex valued random variables such that their real and imaginary parts are independent standard Gaussian variables, and let $\boldsymbol{c} = (c_1, \dots, c_N)$, $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \dots, \tilde{c}_N)$ be complex vectors. Denote $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N)$, $\tilde{\boldsymbol{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_N)$ and*

$$\begin{cases}
D_{\sigma,N}(\boldsymbol{c}, \tilde{\boldsymbol{c}}) = \min_{\tau} \Big\{ \sum_{j=1}^{N} |c_j - e^{-ij\tau}\tilde{c}_j|^2 \\
+2\sigma \sum_{j=1}^{N} \mathrm{Re}\,\big((c_j - e^{-ij\tau}\tilde{c}_j)\overline{(\xi_j - e^{-ij\tau}\tilde{\xi}_j)}\big)\Big\}, \\
d_{N,\tau}(\boldsymbol{c}, \tilde{\boldsymbol{c}}) = \sqrt{\sum_{j=1}^{N} \big|c_j - e^{-ij\tau}\tilde{c}_j\big|^2}, \\
u_N(\boldsymbol{\xi}, \boldsymbol{c}, \tilde{\boldsymbol{c}}) = \sup_{\tau} \Big| \sum_{j=1}^{N} \frac{\mathrm{Re}\,\big[\xi_j \overline{(c_j - e^{-ij\tau}\tilde{c}_j)}\big]}{d_{N,\tau}(\boldsymbol{c}, \tilde{\boldsymbol{c}})} \Big|.
\end{cases}$$

Assume that $x_0 \leq \min_{\tau} d_{N,\tau}(\boldsymbol{c}, \tilde{\boldsymbol{c}})$, then

$$\begin{aligned}
\forall M \in \mathbb{R}, \quad &\mathbf{P}\Big(D_{\sigma,N}(\boldsymbol{c}, \tilde{\boldsymbol{c}}) \leq M\Big) \\
&\leq 2\,\mathbf{P}\Big(\sigma u_N(\boldsymbol{\xi}, \boldsymbol{c}, \tilde{\boldsymbol{c}}) \geq \frac{x_0}{4} - \frac{M}{4x_0}\Big) + 2\,\mathbf{P}\Big(\frac{x_0}{2} < \sigma u_N(\boldsymbol{\xi}, \boldsymbol{c}, \tilde{\boldsymbol{c}})\Big).
\end{aligned}$$

Assume further that $\boldsymbol{c}$ *and* $\tilde{\boldsymbol{c}}$ *are in* $\mathcal{F}_{s,L}$ *and that* $\frac{x_0}{4} - \frac{M}{4x_0} > 0$*, then combining the last result with Lemma 5,*

$$\mathbf{P}\Big(D_{\sigma,N}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \le M\Big) \le 2\Big(1 + x_0^{-1} L \max\{1, N^{1-s}\}\Big) \times \Big(\exp\big\{-(x_0^2 - M)^2/32x_0^2\sigma^2\big\} + \exp\big\{-x_0^2/8\sigma^2\big\}\Big).$$

Proof of Lemma 2.

$$\begin{aligned}
&\sum_{j=1}^N \big|c_j - e^{-\mathrm{i}j\tau}\tilde{c}_j\big|^2 + 2\sigma \sum_{j=1}^N \mathrm{Re}\,\Big((c_j - e^{-\mathrm{i}j\tau}\tilde{c}_j)(\overline{\xi_j - e^{-\mathrm{i}j\tau}\tilde{\xi}_j})\Big)\\
&= d_{N,\tau}^2(\boldsymbol{c},\tilde{\boldsymbol{c}}) + 2\sigma d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \sum_{j=1}^N \frac{\mathrm{Re}\,\big[\xi_j(\overline{c_j - e^{-\mathrm{i}j\tau}\tilde{c}_j})\big]}{d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}})}\\
&+ 2\sigma d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \sum_{j=1}^N \frac{\mathrm{Re}\,\big[\tilde{\xi}_j(\overline{e^{\mathrm{i}j\tau}c_j - \tilde{c}_j})\big]}{d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}})}\\
&\ge d_{N,\tau}^2(\boldsymbol{c},\tilde{\boldsymbol{c}}) - 2\sigma d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \sup_\tau \Big|\sum_{j=1}^N \frac{\mathrm{Re}\,\big[\xi_j(\overline{c_j - e^{-\mathrm{i}j\tau}\tilde{c}_j})\big]}{d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}})}\Big|\\
&- 2\sigma d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \sup_\tau \Big|\sum_{j=1}^N \frac{\mathrm{Re}\,\big[\tilde{\xi}_j(\overline{e^{\mathrm{i}j\tau}c_j - \tilde{c}_j})\big]}{d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}})}\Big|.
\end{aligned}$$

With the notation $u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}}) = \sup_\tau \Big|\sum_{j=1}^N \frac{\mathrm{Re}\,\big[\xi_j(\overline{c_j - e^{-\mathrm{i}j\tau}\tilde{c}_j})\big]}{d_{N,\tau}(\boldsymbol{c},\tilde{\boldsymbol{c}})}\Big|$, we obtain

$$D_{\sigma,N}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \ge \min_{x \ge x_0}(x^2 - ax),$$

with $a = 2\sigma u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}}) + 2\sigma u_N(\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{c}},\boldsymbol{c})$. Now, using the fact that $\min_{x\ge x_0}(x^2 - ax)$ is reached at the point x_0 if $x_0 \ge \frac{a}{2}$, we get

$$\begin{aligned}
\mathbf{P}\Big(D_{\sigma,N}(\boldsymbol{c},\tilde{\boldsymbol{c}}) \le M\Big) &\le \mathbf{P}\Big(x_0^2 - 2x_0\sigma u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}}) - 2x_0\sigma u_N(\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{c}},\boldsymbol{c}) \le M\Big)\\
&+ \mathbf{P}\Big(x_0 < \sigma u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}}) + \sigma u_N(\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{c}},\boldsymbol{c}))\Big)\\
&\le 2\,\mathbf{P}\Big(\sigma u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}}) \ge \frac{x_0}{4} - \frac{M}{4x_0}\Big)\\
&+ 2\,\mathbf{P}\Big(\frac{x_0}{2} < \sigma u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}})\Big),
\end{aligned}$$

since $u_N(\boldsymbol{\xi},\boldsymbol{c},\tilde{\boldsymbol{c}})$ and $u_N(\tilde{\boldsymbol{\xi}},\tilde{\boldsymbol{c}},\boldsymbol{c})$ have the same distribution. ∎

Lemma 3. *Let $\xi_j, \tilde{\xi}_j$ be independent complex valued random variables such that their real and imaginary parts are independent standard Gaussian variables, let c, s and σ be some positive real numbers. Denote*

$$\begin{cases} \rho_\sigma = (\sigma^2\sqrt{\log\sigma^{-1}})^{\frac{2s}{4s+1}}, \\ N_\sigma = [c\rho_\sigma^{-1/s}], \\ B_\sigma = \max_\tau \Big|\sum_{j=1}^{N_\sigma} \mathrm{Re}\left(e^{ij\tau}\xi_j\tilde{\xi}_j\right)\Big|. \end{cases}$$

Then, for σ small enough and for every positive c',

$$\mathbf{P}\Big(2\sigma^2 B_\sigma > c'\rho_\sigma^2\Big) \leq 2c(\log\sigma^{-1})^{\frac{-1}{4s+1}}\sigma^{\frac{c'^2}{64c}-\frac{4}{4s+1}} + e^{-N_\sigma/2}.$$

Proof of Lemma 3. Applying Lemma 4, we state that, for σ small enough,

$$\mathbf{P}\Big(B_\sigma > 4x\sqrt{N_\sigma\log(\sigma^{-1})}\Big) \leq 2c(\log\sigma^{-1})^{\frac{-1}{4s+1}}\sigma^{x^2-\frac{4}{4s+1}} + e^{-N_\sigma/2},$$

from which follows that

$$\mathbf{P}\Big(B_\sigma > 4x\rho_\sigma^{-1/2s}\sqrt{c\log(\sigma^{-1})}\Big) \leq 2c(\log\sigma^{-1})^{\frac{-1}{4s+1}}\sigma^{x^2-\frac{4}{4s+1}} + e^{-N_\sigma/2}.$$

We conclude, observing that $4x\rho_\sigma^{-1/2s}\sqrt{c\log(\sigma^{-1})} = \frac{8x\rho_\sigma^2\sqrt{c}}{2\sigma^2}$. □

Lemma 4. *Let N be some positive integer and let ξ_j, $\tilde{\xi}_j$, $j = 1,\dots,N$, be independent complex valued random variables such that their real and imaginary parts are independent standard Gaussian variables. Let $\boldsymbol{u} = (u_1,\dots,u_N)$ be a vector of real numbers. Denote $S(t) = \sum_{j=1}^N u_j \,\mathrm{Re}\left(e^{ijt}\xi_j\tilde{\xi}_j\right)$ for every t in $[0,2\pi]$ and $\|S\|_\infty = \sup_{t\in[0,2\pi]}|S(t)|$. Then*

$$\forall x,y>0,\quad \mathbf{P}\Big(\|S\|_\infty > \sqrt{2}x\big(\|\boldsymbol{u}\|_2 + y\|\boldsymbol{u}\|_\infty\big)\Big) \leq (N+1)e^{-x^2/2} + e^{-y^2/2}.$$

Proof of Lemma 4. We refer to Collier and Dalalyan [11], Lemma 3, for a proof of this lemma. □

Lemma 5. *Let $\boldsymbol{c} = (c_1, c_2, \dots)$ and $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \tilde{c}_2, \dots)$ in $\mathcal{F}_{s,L}$ with $s > 0$ and let N be an integer. Denoting $\eta_j, \tilde{\eta}_j \overset{iid}{\sim} \mathcal{N}(0,1)$, we define*

$$S(t) = \sum_{j=1}^N \frac{\eta_j\,\mathrm{Re}(c_j - e^{-ijt}\tilde{c}_j) + \tilde{\eta}_j\,\mathrm{Im}(c_j - e^{-ijt}\tilde{c}_j)}{\sqrt{\sum_{j=1}^N\left|c_j - e^{-ijt}\tilde{c}_j\right|^2}}$$

for every t in $[0,2\pi]$. Then

$$\mathbf{P}\Big(\|S\|_\infty \geq x\Big) \leq \Big(\frac{L\cdot\max\{1, N^{1-s}\}}{\sqrt{\min_\tau\sum_{j=1}^N|c_j - e^{-ij\tau}\tilde{c}_j|^2}} + 1\Big)e^{-\frac{x^2}{2}}.$$

First recall Berman's formula, that we will need in the proof.

Theorem 4 (Berman [4]). *Let N be a positive integer, $a<b$ some real numbers and g_j, $j=1,\dots,N$ be continuously differentiable functions on $[a,b]$ satisfying $\sum_{j=1}^N g_j(t)^2=1$ for all $t\in\mathbb{R}$ and η_j, $j=1,\dots,N$, some independent standard Gaussian variables. Then*

$$\mathbf{P}\left(\sup_{[a,b]}\sum_{j=1}^N g_j(t)\eta_j\geq x\right)\leq \frac{I}{2\pi}e^{-\frac{x^2}{2}}+\int_x^\infty\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt$$

with

$$I=\int_a^b\left[\sum_{j=1}^N g_j'(t)^2\right]^{1/2}dt.$$

Proof of Lemma 5. Denote

$$\begin{cases} f_j(t)=\frac{\mathrm{Re}(c_j-e^{-\mathrm{i}jt}\tilde{c}_j)}{\sqrt{\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}},\\ g_j(t)=\frac{\mathrm{Im}(c_j-e^{-\mathrm{i}jt}\tilde{c}_j)}{\sqrt{\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}}.\end{cases}$$

We compute the derivatives of these functions:

$$\begin{aligned}f_j'(t)&=\frac{-\mathrm{Im}(je^{-\mathrm{i}jt}\tilde{c}_j)}{\sqrt{\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}}\\&\quad+\frac{\mathrm{Re}(c_j-e^{-\mathrm{i}jt}\tilde{c}_j)}{\left(\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2\right)^{\frac{3}{2}}}\sum_{k=1}^N\mathrm{Im}(k\overline{c}_k\tilde{c}_ke^{-\mathrm{i}kt})\\ \text{and } g_j'(t)&=\frac{\mathrm{Re}(je^{-\mathrm{i}jt}\tilde{c}_j)}{\sqrt{\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}}\\&\quad+\frac{\mathrm{Im}(c_j-e^{-\mathrm{i}jt}\tilde{c}_j)}{\left(\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2\right)^{\frac{3}{2}}}\sum_{k=1}^N\mathrm{Im}(k\overline{c}_k\tilde{c}_ke^{-\mathrm{i}kt}),\end{aligned}$$

whence

$$\begin{aligned}\sum_{j=1}^N\left(f_j'(t)^2+g_j'(t)^2\right)&=\frac{\sum_{j=1}^N j^2|\tilde{c}_j|^2}{\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}-\left(\frac{\sum_{k=1}^N\mathrm{Im}(k\overline{c}_k\tilde{c}_ke^{-\mathrm{i}kt})}{\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}\right)^2\\&\leq\frac{L^2\max\{1,N^{2-2s}\}}{\min_t\sum_{k=1}^N|c_k-e^{-\mathrm{i}kt}\tilde{c}_k|^2}\end{aligned}$$

The conclusion follows from Berman's formula. □

Lemma 6. *Let σ be a positive real number and s,S in $[s_1,s_2]\subseteq\mathbb{R}_*^+$ be such that $0\leq s-S\leq\frac{1}{\log\sigma^{-1}}$. Denote $\rho_\sigma^*(s)=\left(\sigma^2\sqrt{\log\sigma^{-1}}\right)^{\frac{2s}{4s+1}}$, then, for σ small enough,*

$$\frac{\rho_\sigma^*(S)}{\rho_\sigma^*(s)}\leq e^{\frac{4}{(4s_1+1)^2}}.$$

Proof of Lemma 6. By the definition of $\rho^*_\sigma(s)$, we have

$$\frac{\rho^*_\sigma(S)}{\rho^*_\sigma(s)} = \Big(\sigma^2\sqrt{\log(\sigma^{-1})}\Big)^{\frac{2(S-s)}{(4s+1)(4S+1)}},$$

which, when σ is so small that $\sigma^2\sqrt{\log\sigma^{-1}} \le 1$, leads, with the hypothesis on s and S,

$$\frac{\rho^*_\sigma(S)}{\rho^*_\sigma(s)} \le \Big(\sigma^2\sqrt{\log(\sigma^{-1})}\Big)^{\frac{-2}{(4s_1+1)^2\log\sigma^{-1}}}.$$

Then, we compute

$$\begin{aligned}
&\Big(\sigma^2\sqrt{\log(\sigma^{-1})}\Big)^{\frac{-2}{(4s_1+1)^2\log\sigma^{-1}}}\\
&= \exp\Big\{\frac{-2}{(4s_1+1)^2\log\sigma^{-1}}(2\log\sigma + \frac{1}{2}\log\log\sigma^{-1})\Big\}\\
&= \exp\Big\{\frac{4}{(4s_1+1)^2}(1 - \frac{\log\log\sigma^{-1}}{4\log\sigma^{-1}})\Big\}\\
&\le e^{\frac{4}{(4s_1+1)^2}},
\end{aligned}$$

and this concludes the proof. $\square$

Finally, we recall here Berry-Esseen's inequality, in a simpler version than Theorem 5.4 of Petrov [33].

Theorem 5 (Berry-Esseen's inequality)**.** *Let N be a positive integer and some random variables $X_1,\dots,X_N \overset{iid}{\sim} X$ and such that $\mathbf{E}(X) = 0$, $\mathbf{V}ar(X) = \gamma^2$, $\mathbf{E}|X|^3 = m^3 < +\infty$. Denote $F_N(x) = \mathbf{P}\big(\frac{1}{\sqrt{N}\gamma}\sum_{j=1}^N X_j < x\big)$ and Φ the distribution function of the standard Gaussian variable. Then*

$$\sup_x |F_N(x) - \Phi(x)| \le \frac{Am^3}{\gamma^3}\frac{1}{\sqrt{N}},$$

for an absolute constant number A. Moreover, in the case when $X = Y^2 - 1$ and Y has a centered Gaussian distribution, and using the majoration $A \le \frac{1}{2}$,

$$\sup_x |F_N(x) - \Phi(x)| \le \frac{1}{\sqrt{2\pi N}}.$$

References

[1] Y. Baraud. Non-asymptotic minimax rates of testing in signal detection. *Bernoulli*, 8(5):577–606, 2002. MR1935648

[2] Y. Baraud, S. Huet, and B. Laurent. Adaptive tests of linear hypotheses by model selection. *Ann. Statist.*, 31(1):225–251, 2003. MR1962505

[3] Y. Baraud, S. Huet, and B. Laurent. Testing convex hypotheses on the mean of a Gaussian vector. Application to testing qualitative hypotheses on a regression function. *Ann. Statist.*, 33(1):214–257, 2005. MR2157802

[4] S. M. Berman. Sojourns and extremes of a stochastic process defined as a random linear combination of arbitrary functions. *Comm. Statist. Stochastic Models*, 4(1): 1–43, 1988. MR0938574

[5] J. Bigot and S. Gadat. A deconvolution approach to estimation of a common shape in a shifted curves model. *Ann. Statist.*, 38(4):2422–2464, 2010. MR2676894

[6] J. Bigot, S. Gadat, and J.-M. Loubes. Statistical M-estimation and consistency in large deformable models for image warping. *J. Math. Imaging Vis.*, 34(3):270–290, 2009. MR2515449

[7] J. Bigot, F. Gamboa, and M. Vimond. Estimation of translation, rotation, and scaling between noisy images using the Fourier-Mellin transform. *SIAM J. Imaging Sci.*, 2(2):614–645, 2009. MR2519925

[8] L. D. Brown and M. G. Low. Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.*, 24(6):2384–2398, 1996. MR1425958

[9] C. Butucea and K. Tribouley. Nonparametric homogeneity tests. *Journal of statistical planning and inference*, 136 (3):597–639, 2006. MR2181971

[10] I. Castillo and J.-M. Loubes. Estimation of the distribution of random shifts deformation. *Math. Meth. Statist.*, 18(1):21–42, 2009. MR2508947

[11] O. Collier and A. S. Dalalyan. Wilks' phenomenon and penalized likelihood-ratio test for nonparametric curve registration. *Journal of Machine Learning Research - Proceedings Track*, 22:264–272, 2012.

[12] A. Dalalyan and M. Reiss. Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields*, 134(2): 248–282, 2006. MR2222384

[13] A. S. Dalalyan. Stein shrinkage and second-order efficiency for semiparametric estimation of the shift. *Math. Methods Statist.*, 16(1):42–62, 2007. MR2319470

[14] A. S. Dalalyan, G. K. Golubev, and A. B. Tsybakov. Penalized maximum likelihood and semiparametric second-order efficiency. *Ann. Statist.*, 34(1):169–201, 2006. MR2275239

[15] M. S. Ermakov. Minimax detection of a signal in Gaussian white noise. *Teor. Veroyatnost. i Primenen.*, 35(4): 704–715, 1990. MR1090496

[16] M. S. Ermakov. Asymptotically minimax criteria for testing complex nonparametric hypotheses. *Problems Inform. Transmission*, 33:184–196, 1996. MR1441739

[17] J. Fan, C. Zhang, and J. Zhang. Generalized likelihood ratio statistics and wilks phenomenon. *Ann. Statist.*, 29(1):153–193, 2001. MR1833962

[18] M. Fromont and C. Lévy-Leduc. Adaptive tests for periodic signal detection with applications to laser vibrometry. *ESAIM P. S.*, 10:46–75 (electronic), 2006. MR2197102

[19] F. Gamboa, J.-M. Loubes, and E. Maza. Semi-parametric estimation of shifts. *Electron. J. Statist.*, 1:616–640, 2007. MR2369028

[20] G. Gayraud and C. Pouet. Minimax testing composite null hypotheses in the discrete regression scheme. *Math. Methods Statist.*, 10(4):375–

394 (2002), 2001. Meeting on Mathematical Statistics (Marseille, 2000). MR1887339

[21] G. Gayraud and C. Pouet. Adaptive minimax testing in the discrete regression scheme. *Probab. Theory Relat. Fields*, 133(4): 531–558, 2005. MR2197113

[22] I. Grama and M. Nussbaum. Asymptotic equivalence for nonparametric generalized linear models. *Probability Theory and Related Fields*, 111 (2):167–214, 1998. MR1633574

[23] I. Grama and M. Nussbaum. Asymptotic equivalence for nonparametric regression. *Math. Methods Statist.*, 11(1):1–36, 2002. MR1900972

[24] J. L. Horowitz and V. G. Spokoiny. An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica*, 69(3):599–631, 2001. MR1828537

[25] Y. I. Ingster. Minimax nonparametric detection of signals in white Gaussian noise. *Problems of Information Transmission*, 18:130–140, 1982. MR0689340

[26] Y. I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. I,II,III. *Math. Methods Statist.*, 2(2):85–114, 1993. MR1257978

[27] Y.I. Ingster and I.A. Suslina. *Nonparametric goodness-of-fit testing under Gaussian models*. Springer Verlag, 2003. MR1991446

[28] U. Isserles, Y. Ritov, and T. Trigano. Semiparametric curve alignment and shift density estimation for biological data. *IEEE Transactions on Signal Processing, in press*, 2011.

[29] O. V. Lepski and V. G. Spokoiny. Minimax nonparametric hypothesis testing: the case of an inhomogeneous alternative. *Bernoulli*, 5(2):333–358, 1999. MR1681702

[30] O. V. Lepski and A. B. Tsybakov. Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. *Probability Theory and Related Fields*, 117 (1):17–48, 2000. MR1759508

[31] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60 (2):91–110, 2004.

[32] M. Nussbaum. Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.*, 24(6):2399–2430, 1996. MR1425959

[33] V. V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1995. Sequences of independent random variables, Oxford Science Publications. MR1353441

[34] M. Reiss. Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.*, 36(4):1957–1982, 2008. MR2435461

[35] J. Rice. Bandwidth choice for nonparametric regression. *The Annals of Statistics*, 12(4): 1215–1230, 1984. MR0760684

[36] A. Rohde. On the asymptotic equivalence and rate of convergence of nonparametric regression and gaussian white noise. *Statistics & Deci-*

sions/International mathematical journal for stochastic methods and models, 22(3/2004):235–243, 2004. MR2125610

[37] V. G. Spokoiny. Adaptive hypothesis testing using wavelets. *Ann. Statist.*, 24(6):2477–2498, 1996. MR1425962