Adaptive tests for periodic signal detection with applications to laser vibrometry

By Magalie FROMONT and Céline LÉVY-LEDUC*

Université Paris-Sud

Abstract

Initially motivated by a practical issue in target detection via laser vibrometry, we are interested in the problem of periodic signal detection in a Gaussian fixed design regression framework. Assuming that the signal belongs to some periodic Sobolev ball and that the variance of the noise is known, we first consider the problem from a minimax point of view: we evaluate the so-called *minimax separation rate* which corresponds to the minimal l_2-distance between the signal and zero so that the detection is possible with prescribed probabilities of error. Then, we propose a testing procedure which is available when the variance of the noise is unknown and which does not use any prior information about the smoothness degree or the period of the signal. We prove that it is adaptive in the sense that it achieves, up to a possible logarithmic factor, the *minimax separation rate* over various periodic Sobolev balls simultaneously. The originality of our approach as compared to related works on the topic of signal detection is that our testing procedure is sensitive to the periodicity assumption on the signal. A simulation study is performed in order to evaluate the effect of this prior assumption on the power of the test. We do observe the gains that we could expect from the theory. At last, we turn to the application to target detection by laser vibrometry that we had in view.

AMS Classification: Primary 62G10, 62G08; Secondary 62G20.
Key words and phrases: Periodic signal detection, adaptive test, minimax separation rates, nonparametric regression.

1 Introduction

We consider the following periodic fixed design regression model:

$$Y_j = f\left(\frac{j}{n}\right) + \sigma\varepsilon_j \text{ for } j \in \{1, \ldots, n\}, \tag{1}$$

where f is some unknown real valued periodic function called the signal, the ε_j's are independent standard Gaussian random variables and σ is a positive real number. We deal with the

*supported by Thalès Optronique (France)

problem of testing the null hypothesis (H_0) : "$f = 0$" against the alternative (H_1) : "$f \neq 0$", focusing in particular on the impact of the periodicity assumption.

This problem appears in various practical situations. For instance, it can be seen as a first step to target identification in laser vibrometry. Indeed, after emission of a continuous wave by a laser vibrometer, reflection on an illuminated target and demodulation, we can suppose that the received signal is of the form (1). Testing "$f = 0$" against "$f \neq 0$" amounts in this case to detecting the presence of any target.

From a theoretical point of view, when σ is known, the performances of a level α test of (H_0) can be evaluated in terms of uniform separation rates with respect to some prescribed norm $||.||$ over classes of smooth functions. Given $\delta \in]0,1[$, a class of functions $\mathcal{S}$ and a level α test Φ_α with values in $\{0,1\}$ rejecting (H_0) when $\Phi_\alpha = 1$, the uniform separation rate $\rho(\Phi_\alpha, \mathcal{S}, \delta, \sigma)$ of Φ_α over the class $\mathcal{S}$ is defined as the smallest number ρ such that the test guarantees a power at least equal to $(1-\delta)$ for all alternatives f in $\mathcal{S}$ at a distance ρ from 0. More precisely, if $\mathbb{P}_f$ denotes the distribution of the observation $Y = (Y_1, \ldots, Y_n)$ defined by (1),

$$\begin{aligned} \rho(\Phi_\alpha, \mathcal{S}, \delta, \sigma) &= \inf\left\{\rho > 0, \inf_{f \in \mathcal{S}, ||f|| \geq \rho} \mathbb{P}_f[\Phi_\alpha = 1] \geq 1 - \delta\right\} \\ &= \inf\left\{\rho > 0, \sup_{f \in \mathcal{S}, ||f|| \geq \rho} \mathbb{P}_f[\Phi_\alpha = 0] < \delta\right\}. \end{aligned}$$

We equip $[0,1]$ with the measure μ_n given by $\mu_n = n^{-1}\sum_{j=1}^n \delta_{j/n}$.

The first purpose of this paper is to evaluate the quantity

$$\tilde{\rho}(\mathcal{S}, \alpha, \delta, \sigma) = \inf_{\Phi_\alpha} \rho(\Phi_\alpha, \mathcal{S}, \delta, \sigma), \tag{2}$$

where the infimum is taken over all level α tests Φ_α, $\mathcal{S}$ being some periodic Sobolev ball with a known smoothness parameter and $||.||$ the usual norm of $\mathbb{L}_2([0,1], \mu_n)$. This quantity introduced by Baraud (2002) as the (α, δ)-minimax rate of testing over $\mathcal{S}$ or the minimax separation rate over $\mathcal{S}$ is the non-asymptotic version of the (asymptotic) minimax rate of testing usually considered. The key reference for the computation of minimax rates of testing in various statistical models is the series of papers due to Ingster (1982) and (1993). In the Gaussian white noise model, which is originally considered to study the problem of signal detection, Ingster establishes the minimax rates of testing with respect to the $\mathbb{L}_r-$norm $(1 \leq r \leq +\infty)$ over some classes of alternatives like Sobolev, Hölder balls in $\mathbb{L}_p$ ($p \geq r$ if $r \leq 2$, $p = r$ if $r > 2$) or ellipsoids in l_r. In particular, for Sobolev or Hölder balls with smoothness degree s, he obtains a rate with respect to the $\mathbb{L}_2-$norm of order $n^{-2s/(4s+1)}$. Other kinds of alternatives such as Besov balls $B_{s,p,q}(R)$ with $p \in]0,2[$ for the $\mathbb{L}_2-$norm are considered by Lepski and Spokoiny (1999). Furthermore, Lepski and Tsybakov (2000) extend Ingster's study by finding the exact separation constants over Hölder and Sobolev classes for the $\mathbb{L}_\infty-$norm. Baraud (2002) focuses on the Gaussian sequence model and gives the minimax separation rates with respect to the l_2-norm over some ellipsoids in l_p (with $p \in]0,2]$) and some Besov bodies. His results, which allow (as explained by Spokoiny (1996)) to recover some of the results obtained in the Gaussian white noise model, can be distinguished by their non-asymptotic characteristic. As for the fixed design regression model, Härdle and Kneip (1999) state a first minimax result over some particular Sobolev classes. Gayraud and Pouet (2001) consider the more general problem of testing composite null hypotheses such as "f belongs to some parametric family". They give the minimax rates of testing for this problem with respect to the $\mathbb{L}_2-$norm over some Hölder classes and prove that these rates are of the same order as those obtained for simple hypotheses in the Gaussian white noise model (Ingster (1993)). The particularity of our work lies in the fact that we exhibit how the minimax separation rates in a periodic fixed design regression model depend on the period of the signal.

The second purpose of this paper is to propose a testing procedure which does not use any prior information on the smoothness of the signal f and which is available when the variance of the noise and the period of f are unknown. The problem of finding adaptive (assumption free) tests of "$f = 0$" is studied from a minimax point of view by Spokoiny (1996) and (1998) in the Gaussian white noise model over some Besov balls and by Baraud (2002) in the Gaussian sequence model when the variance is known over some ellipsoids. They prove that adaptation is impossible without some loss of efficiency within an extra $\log\log n$ factor and they propose adaptive testing procedures which are rate optimal over several classes of alternatives simultaneously. In the fixed design regression model without any periodicity assumption, various testing procedures using no prior assumption on the smoothness of the signal f have been proposed. Horowitz and Spokoiny (2001) use kernel smoothing methods to construct a procedure for the problem of testing that the signal belongs to some parametric family of functions. This test achieves a rate of order $(\sqrt{\log\log n}/n)^{2s/(4s+1)}$ for the normalized l_2-norm over Hölder, Sobolev and Besov classes of alternatives with smoothness parameter s. Besides, some methods of model selection by minimization of a penalized criterion allow to construct other assumption free testing procedures. The test proposed by Eubank and Hart (1992) is built from a criterion connected with Mallows' C_p. It achieves the parametric rate of testing over directional alternatives. The test developed by Baraud, Huet and Laurent (2003) is based on a criterion related to the one used by Laurent and Massart (2000) to estimate quadratic functionals in the regression framework. It consists in a multiple testing procedure which can be described as follows. Introduce some suitable collection of linear subspaces $\{S_m,\ m \in \mathcal{M}\}$ of $\mathbb{R}^n$ based on orthonormal functions of $\mathbb{L}_2([0,1])$ such that for each m, $S_m \neq \{0\}$ and $S_m \neq \mathbb{R}^n$, and some collection $\{\alpha_m,\ m \in \mathcal{M}\}$ of numbers in $]0,1[$. Consider, for each m in $\mathcal{M}$, the Fisher test with level α_m of the null hypothesis "$(f(1/n),\ldots,f(1))^T = 0$" against the alternative "$(f(1/n),\ldots,f(1))^T \in S_m$". One thus obtains a collection of Fisher tests and one decides to reject the null hypothesis if one of the tests of the collection does. From a theoretical point of view, Baraud, Huet and Laurent (2003) evaluate the quality of their test by the means of uniform separation rates over some Hölder balls with respect to the normalized l_2-norm. They obtain some results similar to Horowitz and Spokoiny's ones, except that they look at the problem from a non-asymptotic angle. From a practical point of view, they illustrate their results by a simulation study. In the fixed design regression model with a periodic regression function, we can at least distinguish two kinds of signal detection procedures: some rely on a likelihood ratio test and others use Bayesian techniques. The ones based on likelihood ratio allow to test the null hypothesis (H_0) against (H_1) when the periodic regression function is a trigonometric polynomial with a not necessarily known degree. Such procedures are developed in the books of Brockwell and Davis (1991) p.337 *et seq.* or Quinn and Hannan (2001) p.73 *et seq.* They consist in rejecting the null hypothesis if the periodogram of the observations contains a value substantially larger than the average value, recalling that the periodogram of $Y = (Y_1,\ldots,Y_n)$ is defined by $I_Y(q) = n^{-1}|\sum_{j=1}^n e^{-2i\pi jq/n}Y_j|^2$ for $q \in \{1,\ldots,n-1\}$. As for the Bayesian methods, they are detailed in the paper by Gregory and Loredo (1992). The relevance of these different tests is essentially justified by simulation studies.

Our approach consists in using the above procedure due to Baraud, Huet and Laurent (2003) with a collection of models $\{S_m,\ m \in \mathcal{M}\}$ which fits the periodicity assumption on the signal. We evaluate the uniform separation rates of the obtained test over some periodic Sobolev balls with respect to the usual norm of $\mathbb{L}_2([0,1],\mu_n)$. In particular, we point up in Section 3 that they are of the same order, up to a logarithmic factor, as the minimax separation rates that we establish in Section 2. An experimental comparison between the powers of the considered test and the one originally investigated by Baraud, Huet and Laurent (2003) (which does not take any periodicity assumption into account) is presented in Section 4. In Section 5, we finally focus on the application to target detection by laser vibrometry which has initially motivated this work. The proofs of the theorems and the propositions stated in the paper are detailed in Section 6.

2 Minimax separation rates over periodic Sobolev balls

Throughout this section, we assume that σ is known. Given $k \in \{1,\ldots,n\}$, $s \in \mathbb{N}^*$ and $R > 0$, we consider the periodic Sobolev ball:

$$\mathcal{S}_k(s,R) = \left\{ f \in \mathcal{C}^s([0,1]),\ f \text{ is periodic with period } k/n,\ \|f^{(s)}\|_{2,k/n} \leq R \right\}, \tag{3}$$

where $\mathcal{C}^s([0,1])$ denotes the set of functions f from $[0,1]$ with a continuous $s-$th order derivative denoted by $f^{(s)}$, and for $\theta > 0$, for all $g \in \mathbb{L}_2([0,1],dx)$,

$$\|g\|_{2,\theta}^2 = \theta^{-1} \int_0^\theta g^2(x)dx.$$

Let us introduce the usual norm $\|.\|_n$ of $\mathbb{L}_2([0,1],\mu_n)$ given, for all $g \in \mathbb{L}_2([0,1],\mu_n)$, by

$$\|g\|_n^2 = \frac{1}{n}\sum_{j=1}^n g^2\left(\frac{j}{n}\right).$$

Fixing some levels $\alpha \in]0,1[$ and $\delta \in]0,1-\alpha[$, our aim in this section is to evaluate the minimax separation rate over $\mathcal{S}_k(s,R)$ defined by

$$\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma) = \inf_{\Phi_\alpha} \inf \left\{ \rho > 0,\ \sup_{f \in \mathcal{S}_k(s,R), \|f\|_n \geq \rho} \mathbb{P}_f[\Phi_\alpha = 0] < \delta \right\},$$

where the infimum is taken over all level α tests Φ_α rejecting "$f=0$" when $\Phi_\alpha = 1$.

2.1 Lower bound

Following a general idea due to Ingster, we derive in this section a lower bound for $\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma)$.

For all $\mathcal{B} \subset \mathbb{L}_2([0,1],\mu_n)$, we define

$$\beta(\mathcal{B}) = \inf_{\Phi_\alpha} \sup_{f \in \mathcal{B}} \mathbb{P}_f(\Phi_\alpha = 0), \tag{4}$$

which allows us to state the first result of the paper.

Theorem 1. *Given some levels $\alpha \in]0,1[$ and $\delta \in]0, 1-\alpha[$, there exist some absolute constant $\kappa \in]0,1]$ and some positive constant $c(s)$ such that for every integer k in $[\kappa^{-1}, n/2]$, setting*

$$\rho_k^2(s,R) = c(s) \sup_{D\in\{1,\ldots,[\kappa k]\}} \left(\frac{(1-\alpha-\delta)\sqrt{D}\sigma^2}{n} \wedge R^2 D^{-2s}\left(\frac{k}{n}\right)^{2s}\right),$$

one has, with β defined by (4),

$$\beta(\{f \in \mathcal{S}_k(s,R), ||f||_n^2 \geq \rho_k^2(s,R)\}) \geq \delta.$$

This theorem allows us to obtain the following explicit lower bound for $\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma)$.

Corollary 1. *Given some levels $\alpha \in]0,1[$ and $\delta \in]0, 1-\alpha[$, there exist some absolute constant $\kappa \in]0,1]$ and some positive constant $c(s,\alpha,\delta)$ such that for every integer k in $[\kappa^{-1}, n/2]$,*

$$\tilde{\rho}^2(\mathcal{S}_k(s,R),\alpha,\delta,\sigma) \geq c(s,\alpha,\delta)\left(R^{\frac{2}{1+4s}}\left(\frac{k}{n}\right)^{\frac{2s}{4s+1}}\left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}} \wedge \frac{\sqrt{k}\sigma^2}{n} \wedge R^2\left(\frac{k}{n}\right)^{2s}\right).$$

In particular, for all $k \in [\kappa^{-1}, n/2]$, if

$$\frac{\sigma}{\sqrt{n}}\left(\frac{k}{n}\right)^{-s} \leq R \leq \sigma n^{s-1/4}\left(\frac{k}{n}\right)^{1/4},$$

then

$$\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma) \geq c'(s,\alpha,\delta) R^{\frac{1}{1+4s}}\left(\frac{k}{n}\right)^{\frac{s}{4s+1}}\left(\frac{\sigma^2}{n}\right)^{\frac{2s}{4s+1}}.$$

Proof: It is easy to see that the result

$$\beta(\{f \in \mathcal{S}_k(s,R), ||f||_n^2 \geq \rho_k^2(s,R)\}) \geq \delta$$

of Theorem 1 leads to the inequality:

$$\tilde{\rho}^2(\mathcal{S}_k(s,R),\alpha,\delta,\sigma) \geq \rho_k^2(s,R).$$

Then, we only need to compute an explicit lower bound for $\rho_k^2(s,R)$.
Let us prove that for $\kappa^{-1} \leq k \leq n/2$,

$$\sup_{D\in\{1,\ldots,[\kappa k]\}}\left(R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} \wedge \frac{\sqrt{D}}{n}\sigma^2\right) \geq \frac{1}{\sqrt{2}} R^{\frac{2}{1+4s}}\left(\frac{k}{n}\right)^{\frac{2s}{1+4s}}\left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}} \wedge \frac{\sqrt{[\kappa k]}\sigma^2}{n} \wedge R^2\left(\frac{k}{n}\right)^{2s}.$$

Remark that $R^2(k/n)^{2s}D^{-2s} \leq \sqrt{D}\sigma^2/n$ if and only if $D \geq D^*$, where D^* is the smallest integer at least equal to $\left((k/n)^{2s} nR^2/\sigma^2\right)^{2/(1+4s)}$, that is

$$D^* = \inf\left\{ D \in \mathbb{N},\ D \geq \left(\left(\frac{k}{n}\right)^{2s}\frac{nR^2}{\sigma^2}\right)^{\frac{2}{1+4s}}\right\}.$$

We can distinguish three cases.
If $D^* = 1$, for all D in $\mathcal{D}_k = \{1,\ldots,[\kappa k]\}$, $D \geq D^*$, which implies that $R^2(k/n)^{2s}D^{-2s} \leq \sqrt{D}\sigma^2/n$. Hence,

$$\sup_{D\in\mathcal{D}_k}\left(R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} \wedge \frac{\sqrt{D}}{n}\sigma^2\right) = R^2\left(\frac{k}{n}\right)^{2s}.$$

If $D^* > [\kappa k]$, for all D in $\mathcal{D}_k$, $\sqrt{D}\sigma^2/n < R^2(k/n)^{2s}D^{-2s}$, and

$$\sup_{D\in\mathcal{D}_k}\left(R^2\left(\frac{k}{n}\right)^{2s}D^{-2s}\wedge\frac{\sqrt{D}}{n}\sigma^2\right)=\frac{\sqrt{[\kappa k]}\sigma^2}{n}.$$

Finally, let us see what happens when $2\le D^*\le[\kappa k]$. Since in this case $(D^*-1)\in\mathcal{D}_k$, we have that

$$\sup_{D\in\mathcal{D}_k}\left(R^2\left(\frac{k}{n}\right)^{2s}D^{-2s}\wedge\frac{\sqrt{D}}{n}\sigma^2\right)\ge R^2\left(\frac{k}{n}\right)^{2s}(D^*-1)^{-2s}\wedge\frac{\sqrt{D^*-1}}{n}\sigma^2.$$

So, from the definition of D^* and the fact that $D^*\ge 2$ leads to $D^*-1\ge D^*/2$, we derive:

$$\sup_{D\in\mathcal{D}_k}\left(R^2\left(\frac{k}{n}\right)^{2s}D^{-2s}\wedge\frac{\sqrt{D}}{n}\sigma^2\right)\ge\frac{\sqrt{D^*}}{\sqrt{2}n}\sigma^2\ge\frac{1}{\sqrt{2}}R^{\frac{2}{1+4s}}\left(\frac{k}{n}\right)^{\frac{2s}{1+4s}}\left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}},$$

which concludes the proof.

2.2 Upper bound

We now aim to establish the optimality of the above lower bound, provided that the radius R of the periodic Sobolev ball $\mathcal{S}_k(s,R)$ belongs to some specified interval. For all $\theta>0$, we introduce the Fourier basis on $[0,\theta]$ defined as follows:

$$\left\{\begin{array}{rcl}\varphi_{\theta,1}(x)&=&1,\\ \varphi_{\theta,2p}(x)&=&\sqrt{2}\cos\left(2\pi px/\theta\right),\forall p\ge 1,\\ \varphi_{\theta,2p+1}(x)&=&\sqrt{2}\sin\left(2\pi px/\theta\right),\forall p\ge 1.\end{array}\right.\tag{5}$$

For all $D\in\mathbb{N}^*$, we consider the linear subspace S_D of $\mathbb{R}^n$ spanned by the vectors:

$$\left\{\left(\varphi_{k/n,l}(1/n),\ldots,\varphi_{k/n,l}(1)\right)^T,\ l=1,\ldots,D\right\}.\tag{6}$$

Let us denote by Π_{S_D} the orthogonal projection onto S_D with respect to the Euclidean norm of $\mathbb{R}^n$ and introduce the norm $||.||_{\mathbb{R}^n}$ given by:

$$||V||^2_{\mathbb{R}^n}=\frac{1}{n}\sum_{j=1}^n V_j^2\text{ for all }V=(V_1,\ldots,V_n)^T\in\mathbb{R}^n.\tag{7}$$

Proposition 1. *Let $\alpha\in]0,1[$ and $\delta\in]0,1-\alpha[$. Define*

$$D_k^*=k\wedge\inf\{D\in\mathbb{N},\ D\ge((k/n)^{2s}nR^2/\sigma^2)^{2/(1+4s)}\},$$

and denote by D_k the dimension of the linear space $S_{D_k^*}$. Let

$$\Phi^*_\alpha=\mathbb{1}_{\left\{n||\Pi_{S_{D_k^*}}(Y)||^2_{\mathbb{R}^n}>t_{D_k,\alpha}\sigma^2\right\}},$$

where $t_{D_k,\alpha}$ is the $(1-\alpha)$ quantile of the χ^2 distribution with D_k degrees of freedom. There exists some positive constant $c(\alpha,\delta)$ such that the test Φ^*_α satisfies:

$$\mathbb{P}_0(\Phi^*_\alpha=1)\le\alpha,$$

and

$$\mathbb{P}_f(\Phi^*_\alpha = 0) \leq \delta$$

for all $f \in \mathcal{S}_k(s,R)$ such that:

$$\|f\|_n^2 \geq c(\alpha,\delta)\left(R^{\frac{2}{1+4s}}\left(\frac{k}{n}\right)^{\frac{2s}{4s+1}}\left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}} + \frac{R^2 k}{n^{2s}} + \frac{\sigma^2}{n}\right).$$

In particular, if

$$\frac{\sigma}{\sqrt{n}}\left(\frac{k}{n}\right)^{-s} \leq R \leq \sigma n^{\frac{8s^2-6s-1}{8s}}\left(\frac{k}{n}\right)^{-\frac{2s+1}{8s}}, \tag{8}$$

then

$$\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma) \leq c'(\alpha,\delta) R^{\frac{1}{1+4s}}\left(\frac{k}{n}\right)^{\frac{s}{4s+1}}\left(\frac{\sigma^2}{n}\right)^{\frac{2s}{4s+1}}.$$

Comments:
Assuming that k belongs to $[\kappa^{-1}, n/2]$ and that the radius R belongs to the range given by (8), we derive from Corollary 1 and Proposition 1 that the minimax separation rate $\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma)$ for the considered problem over $\mathcal{S}_k(s,R)$ is of order

$$R^{\frac{1}{1+4s}}\left(\frac{k}{n}\right)^{\frac{s}{4s+1}}\left(\frac{\sigma^2}{n}\right)^{\frac{2s}{4s+1}}.$$

Introducing the larger class $\mathcal{S}'_k(s,R) = \{f,\ \ f \text{ is periodic with period } k/n \text{ and } f(x) = \sum_{l\geq 1}\beta_l(f)\varphi_{k/n,l}(x) \text{ with } (2\pi)^{2s}\sum_{l\geq 2}[l/2]^{2s}\beta_l^2(f) \leq R^2(k/n)^{2s}\}$, one can easily see that the same result holds for $\tilde{\rho}(\mathcal{S}'_k(s,R),\alpha,\delta,\sigma)$. Such a result is in fact comparable to the minimax separation rate that one can obtain in the usual Gaussian white noise model (see Appendix for further details).

3 Adaptive testing procedures with unknown variance

We turn here to the problem of finding adaptive tests of (H_0) "$f = 0$" against (H_1) "$f \neq 0$" in $\mathbb{L}_2([0,1],\mu_n)$ when the variance σ^2 of the Y_j's given by (1) is unknown. We consider the following test of "$(f(1/n),\ldots,f(1))^T = 0$" against " $(f(1/n),\ldots,f(1))^T \neq 0$" proposed by Baraud, Huet and Laurent (2003). Let us fix some α in $]0,1[$. Given some finite collection of linear subspaces $\{S_m, m \in \mathcal{M}\}$ of $\mathbb{R}^n$, let

$$T_\alpha = \sup_{m\in\mathcal{M}}\left\{\frac{(n-D_m)\|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m\|Y-\Pi_m(Y)\|^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m,n-D_m}(u_\alpha)\right\} \tag{9}$$

where:

- $\|.\|_{\mathbb{R}^n}$ is the norm defined by (7),
- $\overline{F}_{D_m,n-D_m}(u)$ is the probability for a Fisher with D_m and $n-D_m$ degrees of freedom variable to be larger than u,
- Π_m is the orthogonal projection onto S_m with respect to the Euclidean norm,

- D_m is the dimension of S_m,
- $u_\alpha = \sup \left\{ u \in]0,1[,\ \mathbb{P}_0 \left(\sup_{m \in \mathcal{M}} \left\{ \frac{(n-D_m)||\Pi_m(Y)||^2_{\mathbb{R}^n}}{D_m||Y-\Pi_m(Y)||^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m,n-D_m}(u) \right\} > 0 \right) \leq \alpha \right\}$.

The null hypothesis "$(f(1/n),\ldots,f(1))^T = 0$" is rejected when the test statistic T_α is positive.

We propose to use this testing procedure by taking particular collections of linear subspaces $\{S_m, m \in \mathcal{M}\}$ which fit the periodicity assumption on the signal. Our aim in the two following sections is to evaluate the uniform separation rates of the consequent tests over some Sobolev balls. To do this, we need a fundamental result due to Baraud, Huet and Laurent that we recall here. Let $\bar{f} = (f(1/n),\ldots,f(1))^T$ and let d_n denote the distance associated with the norm $||.||_{\mathbb{R}^n}$.

Theorem 2 (Baraud, Huet, Laurent). *Given some levels α and δ in $]0,1[$, let T_α be the test statistic defined by (9). Assume that $n \geq 2$ and that the collection $\{S_m,\ m \in \mathcal{M}\}$ satisfies:*

$$(\mathcal{H}_\mathcal{M}) : \forall m \in \mathcal{M}, \left\{ \begin{array}{l} u_\alpha \geq e^{\left(\frac{D_m - n}{10}\right)}, \\ \delta \geq 2e^{\left(\frac{D_m - n}{21}\right)}. \end{array} \right.$$

Then,

$$\mathbb{P}_f(T_\alpha > 0) \geq 1 - \delta$$

for all f satisfying

$$||f||^2_n \ \geq \ \inf_{m \in \mathcal{M}} \left\{ \begin{array}{c} \left[1 + 10\left(\frac{D_m + \log(1/u_\alpha)}{n - D_m}\right)\right] d^2_n(\overline{f}, S_m) \\ + \frac{\sigma^2}{n}\left[5\left(1 + \sqrt{\frac{D_m}{n-D_m}}\right)\sqrt{D_m \log\left(\frac{2}{u_\alpha \delta}\right)} + \frac{25}{2}\left(1 + 2\frac{D_m}{n-D_m}\right)\log\left(\frac{2}{u_\alpha\delta}\right)\right] \end{array} \right\}.$$

The proof of this theorem is based on some deviations inequalities for noncentral χ^2 variables due to Birgé (2000) combined with some exponential inequalities for Fisher variables.

3.1 Testing procedure when the period of the signal is known

Throughout this section, we assume that $n \geq 3$ and that f is periodic with period k/n for k in $\{3,\ldots,n\}$. We recall that $\{\varphi_{k/n,l}, l \geq 1\}$ denotes the Fourier basis on $[0,k/n]$. We take

$$\mathcal{M} = \left\{1, 2, 2^2, \ldots, 2^{[\log_2(k/2)]}\right\}, \tag{10}$$

and for D in $\mathcal{M}$, S_D as the linear subspace of $\mathbb{R}^n$ spanned by the vectors:

$$\left\{ \left(\varphi_{k/n,l}(1/n), \ldots, \varphi_{k/n,l}(1)\right)^T,\ l = 1,\ldots,D \right\}. \tag{11}$$

The following proposition gives upper bounds for the uniform separation rates of the corresponding testing procedure over the periodic Sobolev balls $\mathcal{S}_k(s,R)$ defined in Section 2.

Proposition 2. *Let $n \geq 3$ and $k \in \{3,\ldots,n\}$. Given some levels α and $\delta \in]0,1[$, let T_α be the test statistic defined by (9) with $\mathcal{M}$ and $\{S_m, m \in \mathcal{M}\}$ chosen as in (10) and (11). Assume that n is large enough so that $\alpha \geq e^{-\frac{n}{20}} \log n/\log 2$ and $\delta \geq 2e^{-\frac{n}{42}}$. For all $s \in \mathbb{N}^*$,* *there exist some positive constants $c(\alpha,\delta)$ and $c(s,\alpha,\delta)$ such that for all $R>0$, if f belongs to $\mathcal{S}_k(s,R)$ and satisfies:*

$$||f||_n^2 \quad \geq \quad c(\alpha,\delta)\left\{R^{\frac{2}{4s+1}}\left(\frac{k}{n}\right)^{\frac{2s}{4s+1}}\left(\frac{\sqrt{\log\log k}\sigma^2}{n}\right)^{\frac{4s}{4s+1}}+\frac{\log\log k}{n}\sigma^2\right\}+c(s,\alpha,\delta)\frac{R^2k}{n^{2s}},$$

then

$$\mathbb{P}_f(T_\alpha>0)\geq 1-\delta.$$

In particular, if

$$\frac{\sigma}{\sqrt{n}}\left(\frac{k}{n}\right)^{-s}(\log\log k)^{s+1/2}\leq R\leq \sigma n^{\frac{8s^2-6s-1}{8s}}\left(\frac{k}{n}\right)^{-\frac{2s+1}{8s}}(\log\log k)^{1/4}, \tag{12}$$

then

$$\rho(\mathbb{1}_{\{T_\alpha>0\}},\mathcal{S}_k(s,R),\delta,\sigma)\leq c'(s,\alpha,\delta)R^{\frac{1}{1+4s}}\left(\frac{k}{n}\right)^{\frac{s}{4s+1}}\left(\frac{\sqrt{\log\log k}\sigma^2}{n}\right)^{\frac{2s}{4s+1}}.$$

Comments:
According to the results of Section 2, this means that the testing procedure is rate optimal, up to a possible $\log\log k$ factor, over all the Sobolev balls $\mathcal{S}_k(s,R)$ such that (12) holds simultaneously. In view of the results due to Spokoiny (1996) in the Gaussian white noise model for Besov balls and Baraud (2002) in the Gaussian sequence model for families of nested ellipsoids, we can rightfully think that this loss of efficiency is unavoidable when we deal with an adaptive procedure. However, we have obtained no result which would confirm it.

3.2 Testing procedure when the period of the signal is unknown

We now consider the more realistic case where the period of the signal is unknown. Let us recall that $\{\varphi_{\theta,l}, l\geq 1\}$ denotes the Fourier basis on $[0,\theta]$. We consider the set

$$\mathcal{M}=\left\{(k,D), k\in\{2,\ldots,n\}, D\in\left\{1,2,\ldots,2^{[\log_2(k/2)]}\right\}\right\}. \tag{13}$$

For $m=(k,D)$ in $\mathcal{M}$, we define S_m as the linear space spanned by

$$\left\{\left(\varphi_{k/n,l}(1/n),\ldots,\varphi_{k/n,l}(1)\right)^T,\ l=1,\ldots,D\right\}. \tag{14}$$

In the following result, we establish upper bounds for the uniform separation rates of the corresponding procedure over some Sobolev balls precised below.

Proposition 3. *Let $n\geq 3$. Given some levels α and $\delta\in]0,1[$, let T_α be the test statistic defined by (9) with $\mathcal{M}$ and $\{S_m, m\in\mathcal{M}\}$ chosen as in (13) and (14). Introduce for $s\in\mathbb{N}^*$, $R>0$ and $\tau_1\in[2/n,1]$,*

$$\mathcal{S}(s,R,\tau_1)=\left\{f\in\mathcal{C}^s([0,1]),\ f \text{ is periodic with period } \tau(f)\in[2/n,\tau_1],\ \|f^{(s)}\|_{2,\tau(f)}\leq R\right\}.$$

Assume that n is large enough so that $\alpha\geq e^{-\frac{n}{20}}\log(n!)/\log 2$ and $\delta\geq 2e^{-\frac{n}{42}}$. For all $s\in\mathbb{N}^*$, there exists some positive constant $c(s,\alpha,\delta)$ such that for all $R>0$, $\tau_1\in[2/n,1]$, if f belongs to $\mathcal{S}(s,R,\tau_1)$ and satisfies:

$$||f||_n^2 \quad\geq\quad c(s,\alpha,\delta)\left\{R^{\frac{2}{4s+1}}\tau_1^{\frac{2s}{4s+1}}\left(\frac{\sqrt{\log(n)}\sigma^2}{n}\right)^{\frac{4s}{4s+1}}+\frac{\log(n)}{n}\sigma^2+\left(\frac{R^2\tau_1^{2s}}{n\tau_1^2}\wedge R^2\tau_1^{2s}\right)\log n\right\},$$

then

$$\mathbb{P}_f(T_\alpha > 0) \geq 1 - \delta.$$

In particular, if

$$\frac{\sigma}{\sqrt{n}}\tau_1^{-s}(\log n)^{s+1/2} \leq R \leq \frac{\sigma}{\sqrt{n}}\tau_1^{-s}(n\tau_1^2)^{\frac{4s+1}{8s}}(\log n)^{-\frac{2s+1}{8s}}, \tag{15}$$

$$\rho(\mathbb{1}_{\{T_\alpha>0\}}, \mathcal{S}(s, R, \tau_1), \delta, \sigma) \leq c'(s, \alpha, \delta) R^{\frac{1}{1+4s}} \tau_1^{\frac{s}{4s+1}} \left(\frac{\sqrt{\log n}\sigma^2}{n}\right)^{\frac{2s}{4s+1}}.$$

Comments:

The upper bound for the uniform separation rate of the test over $\mathcal{S}(s, R, \tau_1)$ when R and τ_1 satisfy (15) is similar to the one obtained when the period of f is known to be equal to k/n (see Proposition 2), but with a loss of efficiency of order $\log n$ instead of $\log\log k$. This is technically due to the fact that we choose a collection $\{S_m,\ m \in \mathcal{M}\}$ which is rich enough to ensure that it contains, for any function f with period $\tau(f)$ in $[2/n, 1]$, some subspace S_m close to f enough. Our choice for $\{S_m,\ m \in \mathcal{M}\}$ essentially relies on the properties of orthogonality of the family $\{(\varphi_{k/n,l}(1/n), \ldots, \varphi_{k/n,l}(k/n)),\ 1 \leq l \leq k\}$ with respect to the Euclidean norm of $\mathbb{R}^k$. However, we do not know if the consequent extra $\log n$ factor is optimal or not.

4 Simulation study

We evaluate here the performances of the proposed test from a practical point of view. In particular, we aim to highlight the improvement carried by the periodicity assumption. In the definition (9) of the test statistic T_α, we have to choose the set $\mathcal{M}$ and the collection $\{S_m, m \in \mathcal{M}\}$ of linear subspaces onto which Y is projected. For numerical reasons which are explained below, we do not only consider periods of the form k/n: we propose to take

$$\mathcal{M} = \{(q, D), q \in \{1, \ldots, Q\}, D \in \{2d+1, 0 \leq d \leq d_0\}\},$$

where Q and d_0 are some fixed integers with $2d_0Q < n$, and for $m = (q, D) \in \mathcal{M}$, S_m as the linear space spanned by the vectors

$$\left\{\left(\varphi_{q^{-1},l}(1/n), \ldots, \varphi_{q^{-1},l}(1)\right)^T,\ l = 1, \ldots, D\right\}.$$

With the same notations as in Section 3, we set

$$T_{per} = \sup_{m\in\mathcal{M}} \left\{\frac{(n - D_m)\|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m\|Y - \Pi_m(Y)\|^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m,n-D_m}(u_\alpha)\right\},$$

where

$$u_\alpha = \sup\left\{u \in]0,1[,\ \mathbb{P}_0\left(\sup_{m\in\mathcal{M}}\left\{\frac{(n - D_m)\|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m\|Y - \Pi_m(Y)\|^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m,n-D_m}(u)\right\} > 0\right) \leq \alpha\right\}.$$

Without taking the periodicity assumption into account, the corresponding statistic is

$$T_{nper} = \sup_{m\in\mathcal{M}'} \left\{\frac{(n - D_m)\|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m\|Y - \Pi_m(Y)\|^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m,n-D_m}(u'_\alpha)\right\},$$

where $\mathcal{M}' = \{2d+1, 0 \le d \le d_0\}$, for $m = D$ in $\mathcal{M}'$, S_m is the linear space spanned by the vectors

$$\left\{ (\varphi_{1,l}(1/n), \ldots, \varphi_{1,l}(1))^T ,\ l = 1, \ldots, D \right\},$$

and

$$u'_\alpha = \sup \left\{ u \in]0,1[,\ \mathbb{P}_0 \left(\sup_{m \in \mathcal{M}'} \left\{ \frac{(n - D_m) \|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m \|Y - \Pi_m(Y)\|^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m, n-D_m}(u) \right\} > 0 \right) \le \alpha \right\}.$$

To implement the testing procedures based on T_{per} and T_{nper}, the main point is to compute $||\Pi_{(q,D)}(Y)||^2_{\mathbb{R}^n}$ for every (q, D) in $\mathcal{M}$. Considering

$$A(q,D) = (A_{j,l}(q,D))_{1 \le j \le n, 1 \le l \le D} = \left(\varphi_{q^{-1},l}(j/n) \right)_{1 \le j \le n, 1 \le l \le D},$$

we can see by straightforward calculations that when q is an integer such that $2d_0 q < n$, $A^T(q,D)A(q,D) = nI_D$ (I_D denotes the identity matrix with dimension D). Minimizing the least squares criterion $||Y - A(q,D)C||^2_{\mathbb{R}^n}$ for C in $\mathbb{R}^n$ then leads to $\Pi_{(q,D)}(Y) = n^{-1} A(q,D) A^T(q,D) Y$ and consequently,

$$||\Pi_{(q,D)}(Y)||^2_{\mathbb{R}^n} = \frac{1}{n} \|A^T(q,D)Y\|^2_{\mathbb{R}^n} = \frac{1}{n^2} \sum_{l=1}^{D} \left(\sum_{j=1}^{n} \varphi_{q^{-1},l}(j/n) Y_j \right)^2 .$$

Finally, if $D = 2d+1$, it is easy to see that:

$$||\Pi_{(q,D)}(Y)||^2_{\mathbb{R}^n} = \frac{1}{n^2} \sum_{l=-d}^{d} \left| \sum_{j=1}^{n} e^{-2i\pi l j q/n} Y_j \right|^2 .$$

Such an expression is interesting for two reasons: first, it corresponds to a cumulated periodogram, the periodogram being a quantity frequently used in periodic signal detection, secondly, the values of $\{\sum_{j=1}^n e^{-2i\pi jm/n} Y_j,\ m \in \{1, \ldots, n-1\}\}$ can be computed by the means of a Fast Fourier Transform, which reduces our computation time.

For our simulations, we take a level $\alpha = 5\%$ and a number of observations $n = 100$. We choose d_0 and Q both equal to 7. The alternatives that we consider are the following ones:

$$\begin{aligned}
&f_a(x) = a \sin(12\pi x), \\
&g_a(x) = a(\sin(16\pi x) + \sin(8\pi x)), \\
&h_a(x) = a \left(\cos\left(\tfrac{16\pi}{15} \cos(12\pi x)\right) + \cos\left(\tfrac{8\pi}{15} \cos(4\pi x)\right) \right), \\
&j_a(x) = a \cos\left(\tfrac{16\pi}{15} \cos(8\pi x)\right).
\end{aligned}$$

We present in Table 1 the estimated powers for the tests based on T_{per} and T_{nper} under various alternatives. These powers are evaluated by 5000 simulations.

Alternatives f_a		
a	T_{per}	T_{nper}
1	1	1
0.75	0.97	0.89
0.5	0.61	0.42
0.25	0.14	0.10
Alternatives g_a		
a	T_{per}	T_{nper}
1	1	0.97
0.5	0.89	0.43
0.25	0.20	0.12

Alternatives h_a		
a	T_{per}	T_{nper}
1	1	0.23
0.5	0.56	0.09
0.25	0.11	0.06
Alternatives j_a		
a	T_{per}	T_{nper}
0.5	0.38	0.05
0.25	0.08	0.03

Table 1: Estimated powers for the tests based on T_{per} and T_{nper}

Comments:
As expected from our theoretical study, the above experimental results show that taking the periodicity assumption into account in the testing procedure improves the power of the test. This is all the more true if the number of significant Fourier coefficients (or harmonics) in the expansion of the periodic signal is great, as we can see for alternatives h_a and j_a.

5 Applications to laser vibrometry

In this section, we are interested in the problem of target detection which can be seen as a preliminary step to long range target identification by laser vibrometry (see Lavielle and Lévy-Leduc (2003) and Prenat (2001) for further details). More precisely, we show how the above method of testing can find some applications to this problem.

After emission of a continuous laser wave, reflection of it on an illuminated object, reception and demodulation, we receive a complex type signal of the form:

$$Y_j = f(j/n) + \sigma(\varepsilon_{1,j} + i\,\varepsilon_{2,j}) \text{ for } j \in \{1,\dots,n\}, \tag{16}$$

where $i^2 = -1$, f is a complex valued function, the $\varepsilon_{1,j}$'s and the $\varepsilon_{2,j}$'s are independent standard Gaussian random variables. Testing the null hypothesis "$f = 0$" against "$f \neq 0$" in this context then means testing the presence of any target.

When the target consists in a great number of reflectors (or vibrators) vibrating in a sinusoidal way, the signal f has the following expression (see Prenat (2001)):

$$f(t) = \sum_{m=1}^{M} a_m \exp\left[\frac{4i\pi\gamma_m}{\lambda}\cos(2\pi F_s t)\right] \tag{17}$$

where M may be large ($M \approx 200$).

Figure 1(a) displays the imaginary part of $f(j/n)$ for $j \in \{1,\dots,9000\}$ when $n = 2^{18}$, $M = 1$, $a_1 = 0.06$, $\gamma_1 = 35 \times 10^{-6}$, $\lambda = 1.5 \times 10^{-6}$ and $F_s = 48$Hz. Figure 1(b) displays the periodogram of $(f(j/n), j \in \{1,\dots,n\})$ defined by

$$I_f(q) = \frac{1}{n}\left|\sum_{j=1}^{n} e^{-\frac{2i\pi qj}{n}} f\left(\frac{j}{n}\right)\right|^2 \quad \text{for } q \in \{1,\dots,n-1\}.$$

A zoom of this periodogram is presented in Figure 1(c).

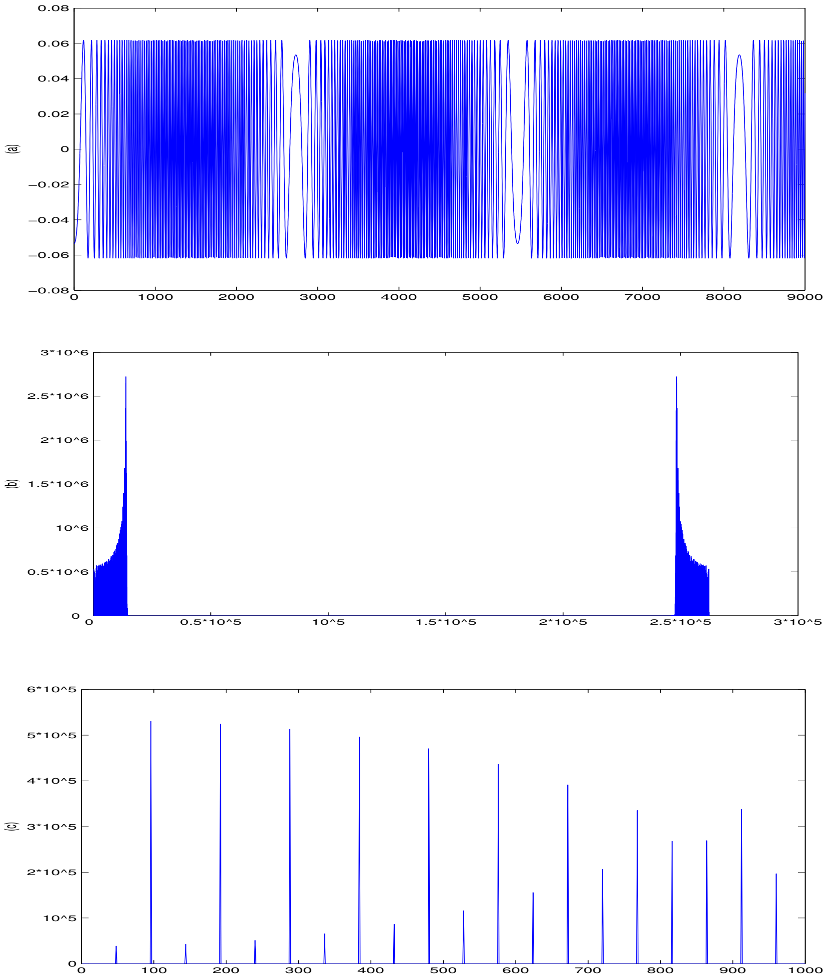

Figure 1: A synthetic signal f. (a) the imaginary part of f, (b) the periodogram of f, (c) zoom of (b).

We can remark in this example that all the information is contained in a frequency band $[0, F_{Max}]$, where $F_{Max} = 4\pi\gamma_1 F_s/\lambda \approx 330 \times 48 = 15840$Hz. This means that the signal f has about 330 significant Fourier coefficients (or harmonics).

Following Prenat (2001), when we add some noise ε_j with variance $2\sigma^2$ to $f(j/n)$, we characterize the signal to noise ratio by the so-called $msnr$ (in dB) where

$$msnr = 10 \log \left(n \sum_{m=1}^{M} a_m^2 / 2\sigma^2 \right).$$

The $msnr$ corresponds in fact to the ratio between the mean power of the signal and the mean power of the noise in a frequency band equal to the inverse of the observation time.

Figure 2(a) displays the imaginary part of the Y_j's satisfying (16) when f is the above deterministic signal and $msnr = 27$dB for $j \in \{1, \ldots, 9000\}$. In Figure 2(b) and Figure 2(c), one can see the corresponding periodogram defined by

$$I_Y(q) = \frac{1}{n}\left|\sum_{j=1}^{n} e^{-\frac{2i\pi qj}{n}} Y_j\right|^2, \text{ for } q \in \{1, \ldots, n-1\},$$

and a zoom of it. We can notice in this example that, because of the high level of noise, the original signal and its harmonics cannot be visually detected.

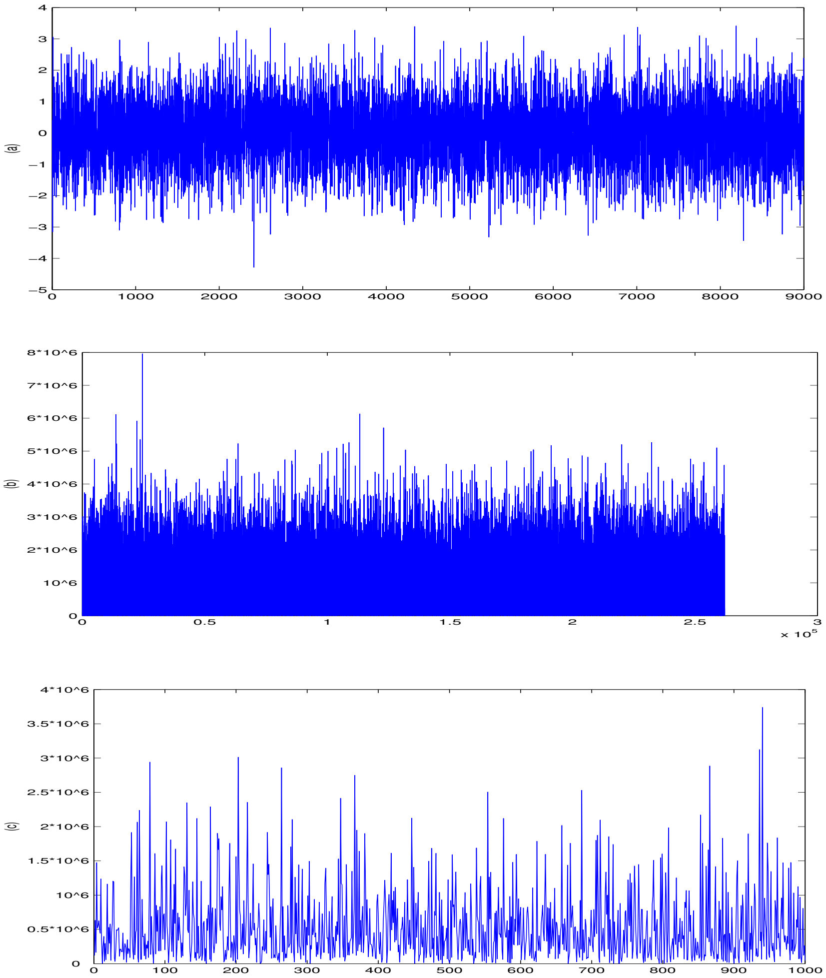

Figure 2: The observed signal $f + \varepsilon$. (a) the imaginary part of $f + \varepsilon$, (b) the periodogram of $f + \varepsilon$, (c) zoom of (b).

Observing some Y_j's following the model (16) amounts in fact to observing two real data sets given by the imaginary and real parts of the Y_i's, which both satisfy (1). We can therefore use the procedure described in Section 3, at the price of slight changes in the definition of the test statistic T_α. More precisely, we replace the norm $\|.\|_{\mathbb{R}^n}$ by $\|.\|_{\mathbb{C}^n}$, where $\|V\|^2_{\mathbb{C}^n} = n^{-1}\sum_{j=1}^{n}|V_j|^2$ for all $V = (V_1, \ldots, V_n)^T \in \mathbb{C}^n$, and we choose the collection $\{S_m, m \in \mathcal{M}\}$ as follows:

$$\mathcal{M} = \left\{(q, D),\ q \in \{1, \ldots, Q\},\ D \in \{2^d,\ d \in \{1, \ldots, d_0\}\}\right\},$$

and for each (q,D) in $\mathcal{M}$, $S_{(q,D)}$ is the linear space spanned by the vectors

$$\left\{\left(e^{2i\pi ql/n},\dots,e^{2i\pi ql}\right)^T,\ l=-D,\dots,D\right\}.$$

We consider an alternative which corresponds to the signal (17) where $M=1$, $F_s=48$Hz, $\gamma_1=35\times10^{-6}$, $\lambda=1.5\times10^{-6}$, with a noise such that $\sigma^2=1$. The computation of the test statistic is conducted in the same way as in Section 4: with very similar calculations, we actually obtain that when $2^{d_0+1}Q<n$,

$$||\Pi_{(q,D)}(Y)||^2_{\mathbb{C}^n}=\frac{1}{n^2}\sum_{l=-D}^{D}\left|\sum_{j=1}^{n}e^{-2i\pi ljq/n}Y_j\right|^2.$$

For a number of observations $n=2^{18}$, we have chosen $Q=50$ and $d_0=10$. We take a level of significance $\alpha=5\%$. The quantiles $\overline{F}^{-1}_{D_m,n-D_m}(u_\alpha)$ are estimated by 40000 simulations. We use 20000 simulations for the estimation of $\overline{F}^{-1}_{D_m,n-D_m}(u)$ for u varying on a regular grid of $]0,\alpha[$ with mesh 1/20000 and 20000 for the estimation of the probabilities

$$\mathbb{P}_0\left(\sup_{m\in\mathcal{M}}\left\{\frac{(n-D_m)||\Pi_m(Y)||^2_{\mathbb{R}^n}}{D_m||Y-\Pi_m(Y)||^2_{\mathbb{R}^n}}-\overline{F}^{-1}_{D_m,n-D_m}(u)\right\}>0\right).$$

Table 2 presents the estimated powers for various $msnr$, namely $20,21,\dots,30$ dB.

$msnr$	20	21	22	23	24	25	26	27	28	29	30
power	0.05	0.07	0.10	0.13	0.18	0.29	0.45	0.67	0.85	0.97	0.99

Table 2: Estimated powers of the test for various $msnr$

Comments:
Since the minimal $msnr$ needed to perform target detection is directly linked to the power that the laser has to emit, a $msnr$ about 30 dB is reasonable for the application that we have in view. The results summarized in Table 2 show that, at such a signal to noise ratio, our test is much powerful. The classical tests used in periodic signal detection, such as those developed in the books of Brockwell and Davis (1991) or Quinn and Hannan (2001), essentially use a parametric approach. They seem to be adapted to the detection of periodic signals having a small number of harmonics such as sinusoids or trigonometric polynomials with small degrees, which is not the case of the signals of our interest. Therefore, we can think that they would not be effective in our context. This is furthermore confirmed by an experimental study: by implementing the test proposed by Brockwell and Davis (1991), we found for a $msnr$ of 30 dB a power of 0.06.

6 Proofs

6.1 Proof of Theorem 1

The idea of the proof is to construct, for each Sobolev ball $\mathcal{S}_k(s,R)$, a sphere Ω of maximal radius such that $\beta(\Omega)\geq\delta$ and $\Omega\subset\mathcal{S}_k(s,R)$. Let $k\in\{1,\dots,n\}$, $s\in\mathbb{N}^*$ and $R>0$. As

Ingster (1993) suggests, we introduce a compactly supported and smooth enough function ϕ: let ϕ be some function in $\mathcal{C}^{\infty}(\mathbb{R})$ supported by $]0,1[$ and positive on its support (we can take for example the function defined on $]0,1[$ by $\phi(x) = \exp\left(-1/x^2 - 1/(x-1)^2\right)$). For all $D \in \mathbb{N}^*$, we consider

$$\widetilde{S_D} = \left\{ f,\ f(x) = \sum_{l=1}^{D} \beta_l \phi_D \left(\frac{nDx}{k} - l + 1 \right), \beta_l \in \mathbb{R} \right\},$$

where ϕ_D is the function in $\mathcal{C}^{\infty}(\mathbb{R})$ which is periodic with period D and which agrees with ϕ on $]0,D]$.
Let

$$\kappa_1 = 1 \wedge \frac{\int_0^1 \phi^2(x)dx}{16\|\phi\phi'\|_\infty} \text{ and } \kappa_2 = \frac{1}{\sqrt{6}} \left(\frac{\int_0^1 \phi^2(x)dx}{4\|\phi\phi'\|_\infty + \int_0^1 \phi^2(x)dx} \right). \tag{18}$$

The following paragraph is devoted to a study of the spheres of the form $\{f \in \widetilde{S_D}, \|f\|_n = \rho\}$.

- **A lower bound for $\tilde{\rho}(\widetilde{S_D}, \alpha, \delta, \sigma)$**

Proposition 4. *Assume that $\kappa_1^{-1} \leq k \leq n$. For all $D \in \{1, \ldots, [\kappa_1 k]\}$ and for all $\rho^2 \leq \kappa_2(1-\alpha-\delta)\sqrt{D}\sigma^2/n$,*

$$\beta\left(\left\{f \in \widetilde{S_D},\ \|f\|_n = \rho\right\}\right) \geq \delta.$$

Comments:
This proposition gives a result in terms of minimax separation rates. Indeed, it is easy to see that it implies the lower bound:

$$\tilde{\rho}^2(\widetilde{S_D}, \alpha, \delta, \sigma) \geq \kappa_2(1-\alpha-\delta)\sqrt{D}\sigma^2/n.$$

Proof: By homogeneity, we can only consider the case where $\sigma = 1$. Let $D \in \{1, \ldots, [\kappa_1 k]\}$ and for $\rho > 0$

$$\widetilde{S_D}[\rho] = \left\{f \in \widetilde{S_D},\ \|f\|_n = \rho\right\}.$$

Recall that $\mathbb{P}_f$ denotes the distribution of the observation $Y = (Y_1, \ldots, Y_n)$ defined by (1). The proof of Proposition 4 is based on the following classical bound on the error of tests : in the notation of Baraud (2002), if

- μ_ρ is a probability measure on $\widetilde{S_D}[\rho]$,
- $P_{\mu_\rho} = \int \mathbb{P}_f d\mu_\rho(f)$,
- $L_{\mu_\rho}(y) = \frac{dP_{\mu_\rho}}{d\mathbb{P}_0}(y)$,
- $\mathbb{E}_0$ is the expectation with respect to $\mathbb{P}_0$,

then

$$\beta\left(\widetilde{S_D}[\rho]\right) \geq 1 - \alpha - \frac{1}{2}\left(\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] - 1\right)^{1/2}.$$

The main difficulty then consists in finding some μ_ρ such that for all $\rho^2 \leq \kappa_2(1-\alpha-\delta)\sqrt{D}\sigma^2/n$,

$$\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] \leq 1 + 4(1-\alpha-\delta)^2.$$

Let $\xi = (\xi_1, \ldots, \xi_D)$ be a sequence of Rademacher random variables that is a sequence of i.i.d. random variables taking the values +1 or -1 with probability 1/2.
Let $f_\xi(x) = \lambda \sum_{l=1}^{D} \xi_l \phi_D\left(nDx/k - l + 1\right)$ where λ is chosen such that $||f_\xi||_n = \rho$. By the properties of the support and the period of ϕ_D:

$$
\begin{aligned}
||f_\xi||_n^2 &= \frac{\lambda^2}{n}[n/k] \sum_{j=1}^{k} \left(\sum_{l=1}^{D} \xi_l \phi_D \left(\frac{Dj}{k} - l + 1 \right) \right)^2 + \frac{\lambda^2}{n} \sum_{j=1}^{n-[n/k]k} \left(\sum_{l=1}^{D} \xi_l \phi_D \left(\frac{Dj}{k} - l + 1 \right) \right)^2 \\
&= \frac{\lambda^2}{n}[n/k] \sum_{l=1}^{D} \sum_{j \in A_{l,D}} \phi_D^2 \left(\frac{Dj}{k} - l + 1 \right) + \frac{\lambda^2}{n} \sum_{l=1}^{D} \sum_{j \in A'_{l,D}} \phi_D^2 \left(\frac{Dj}{k} - l + 1 \right) \\
&= \frac{\lambda^2}{n} \sum_{l=1}^{D} \left([n/k] \sum_{j \in A_{l,D}} \phi_D^2 \left(\frac{Dj}{k} - l + 1 \right) + \sum_{j \in A'_{l,D}} \phi_D^2 \left(\frac{Dj}{k} - l + 1 \right) \right),
\end{aligned}
$$

where $A_{l,D} = \left\{ j \in \left] \frac{(l-1)k}{D}, \frac{lk}{D} \right[\right\}$ and $A'_{l,D} = A_{l,D} \cap \{1, \ldots, n - [n/k]k\}$.
Taking

$$
\lambda^2 = n\rho^2 \left[\sum_{l=1}^{D} \left([n/k] \sum_{j \in A_{l,D}} \phi_D^2 \left(\frac{Dj}{k} - l + 1 \right) + \sum_{j \in A'_{l,D}} \phi_D^2 \left(\frac{Dj}{k} - l + 1 \right) \right) \right]^{-1} \tag{19}
$$

implies that $f_\xi \in \widetilde{S_D}[\rho]$, which allows us to take μ_ρ as the distribution of f_ξ. With such a choice for μ_ρ, we have

$$
\begin{aligned}
L_{\mu_\rho}(y) &= \mathbb{E}_\xi \left(\exp \left\{ -\frac{1}{2} \sum_{j=1}^{n} \left[y_j - \lambda \sum_{l=1}^{D} \xi_l \phi_D \left(\frac{Dj}{k} - l + 1 \right) \right]^2 + \frac{1}{2} \sum_{j=1}^{n} y_j^2 \right\} \right) \\
&= e^{-\frac{n\rho^2}{2}} \mathbb{E}_\xi \left(\exp \left\{ \lambda \sum_{j=1}^{n} y_j \sum_{l=1}^{D} \xi_l \phi_D \left(\frac{Dj}{k} - l + 1 \right) \right\} \right) \\
&= e^{-\frac{n\rho^2}{2}} \mathbb{E}_\xi \left(\exp \left\{ \begin{array}{l} \lambda \sum_{l=1}^{D} \xi_l \sum_{m=0}^{[n/k]-1} \sum_{j=mk+1}^{mk+k} y_j \phi_D \left(\frac{Dj}{k} - l + 1 \right) \\ + \lambda \sum_{l=1}^{D} \xi_l \sum_{j=[n/k]k+1}^{n} y_j \phi_D \left(\frac{Dj}{k} - l + 1 \right) \end{array} \right\} \right).
\end{aligned}
$$

From the properties of the support of ϕ_D, we get

$$
L_{\mu_\rho}(y) = e^{-\frac{n\rho^2}{2}} \mathbb{E}_\xi \left(\exp \left\{ \begin{array}{l} \lambda \sum_{l=1}^{D} \xi_l \sum_{m=0}^{[n/k]-1} \sum_{j \in A_{m,l,D}} y_j \phi_D \left(\frac{Dj}{k} - l + 1 \right) \\ + \lambda \sum_{l=1}^{D} \xi_l \sum_{j \in \{[n/k]k+1, \ldots, n\} \cap A_{[n/k],l,D}} y_j \phi_D \left(\frac{Dj}{k} - l + 1 \right) \end{array} \right\} \right),
$$

where

$$
A_{m,l,D} = \left\{ j \in \left] mk + \frac{(l-1)k}{D}, mk + \frac{lk}{D} \right[\right\},
$$

and by independence of the ξ_l's, we have

$$\mathbb{E}_0(L^2_{\mu_\rho}(Y)) = e^{-n\rho^2}\,\mathbb{E}_0\left[\prod_{l=1}^{D}\cosh^2\left(\begin{array}{c}\lambda\sum\limits_{m=0}^{[n/k]-1}\sum\limits_{j\in A_{m,l,D}} Y_j\phi_D\left(\frac{Dj}{k}-l+1\right)\\ +\lambda\sum\limits_{j\in\{[n/k]k+1,\dots,n\}\cap A_{[n/k],l,D}} Y_j\phi_D\left(\frac{Dj}{k}-l+1\right)\end{array}\right)\right].$$

Let

$$B_l = \lambda\sum_{m=0}^{[n/k]-1}\sum_{j\in A_{m,l,D}} Y_j\phi_D\left(\frac{Dj}{k}-l+1\right)+\lambda\sum_{j\in\{[n/k]k+1,\dots,n\}\cap A_{[n/k],l,D}} Y_j\phi_D\left(\frac{Dj}{k}-l+1\right).$$

Under the null hypothesis, the Y_j's are independent Gaussian random variables with mean 0 and variance 1. Hence, the B_l's are independent Gaussian random variables with mean 0 and variance $\lambda^2\Sigma_l$ where

$$\Sigma_l = \sum_{m=0}^{[n/k]-1}\sum_{j\in A_{m,l,D}} \phi_D^2\left(\frac{Dj}{k}-l+1\right)+\sum_{j\in\{[n/k]k+1,\dots,n\}\cap A_{[n/k],l,D}} \phi_D^2\left(\frac{Dj}{k}-l+1\right).$$

By periodicity of ϕ_D, we have that

$$\Sigma_l = [n/k]\sum_{j\in A_{l,D}}\phi_D^2\left(\frac{Dj}{k}-l+1\right)+\sum_{j\in A'_{l,D}}\phi_D^2\left(\frac{Dj}{k}-l+1\right).$$

This implies via (19) that

$$\lambda^2 = n\rho^2\left(\sum_{l=1}^{D}\Sigma_l\right)^{-1},$$

and we get

$$\begin{aligned}
\mathbb{E}_0(L^2_{\mu_\rho}(Y)) &= e^{-n\rho^2}\prod_{l=1}^{D}e^{\lambda^2\Sigma_l}\cosh(\lambda^2\Sigma_l)\\
&\le e^{-n\rho^2}\exp\left(\lambda^2\sum_{l=1}^{D}\Sigma_l\right)\exp\left(\frac{\lambda^4}{2}\sum_{l=1}^{D}\Sigma_l^2\right)\\
&\le \exp\left(\frac{\lambda^4}{2}\sum_{l=1}^{D}\Sigma_l^2\right)\le\exp\left(\frac{n^2\rho^4}{2}\frac{\sum_{l=1}^{D}\Sigma_l^2}{\left(\sum_{l=1}^{D}\Sigma_l\right)^2}\right). \qquad (20)
\end{aligned}$$

Let us turn to an evaluation of Σ_l, which is the object of the following Lemma.

Lemma 1. *Let $1\le D\le k\le n$ and let for $l\in\{1,\dots,D\}$, $A_{l,D}=\left\{j\in\left]\frac{(l-1)k}{D},\frac{lk}{D}\right[\right\}$, and $A'_{l,D}=A_{l,D}\cap\{1,\dots,n-[n/k]k\}$. The quantity Σ_l defined by*

$$\Sigma_l = [n/k]\sum_{j\in A_{l,D}}\phi_D^2\left(\frac{Dj}{k}-l+1\right)+\sum_{j\in A'_{l,D}}\phi_D^2\left(\frac{Dj}{k}-l+1\right)$$

satisfies the two following inequalities:

$$\Sigma_l \le 2[n/k]\frac{k}{D}\left(4\|\phi\phi'\|_\infty + \int_0^1 \phi^2(x)dx\right),$$

and if $8\|\phi\phi'\|_\infty \le (k/2D)\int_0^1 \phi^2(x)dx$ *(that is if* $D \le \kappa_1 k$*),*

$$\Sigma_l \ge [n/k]\frac{k}{2D}\int_0^1 \phi^2(x)dx.$$

From Lemma 1, we deduce that for all $D \in \{1, \dots, [\kappa_1 k]\}$,

$$\sum_{l=1}^{D}\Sigma_l^2\left(\sum_{l=1}^{D}\Sigma_l\right)^{-2} \le \frac{8}{3\kappa_2^2 D}.$$

Hence, as soon as

$$\rho^2 \le \kappa_2(1-\alpha-\delta)\frac{\sqrt{D}}{n} \le \kappa_2\frac{\sqrt{3}}{2}\frac{\sqrt{\log(1+4(1-\alpha-\delta)^2)D}}{n},$$

from (20), we derive that

$$\mathbb{E}_0(L^2_{\mu_\rho}(Y)) \le 1 + 4(1-\alpha-\delta)^2,$$

and the result follows.

Proof of Lemma 1: To simplify the expressions, we set $x_{j,l} = Dj/k - l + 1$. Let us first give a lower bound for Σ_l. It is obvious that

$$\begin{aligned}
\Sigma_l \ge& [n/k]\sum_{j=[(l-1)k/D]}^{[lk/D]}\phi_D^2(x_{j,l}) - [n/k]\phi_D^2\left(x_{[(l-1)k/D],l}\right)\\
\ge& [n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]}^{[lk/D]}\left(\frac{D}{k}\phi_D^2(x_{j,l}) - \int_{x_{j,l}}^{x_{j+1,l}}\phi_D^2(x)dx\right)\\
&+ [n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]}^{[lk/D]}\int_{x_{j,l}}^{x_{j+1,l}}\phi_D^2(x)dx - [n/k]\phi_D^2\left(x_{[(l-1)k/D],l}\right).
\end{aligned}$$

Since $\phi_D(0) = 0$, this implies that

$$\begin{aligned}
\Sigma_l \quad \ge \quad & [n/k]\frac{k}{D}\int_0^1 \phi_D^2(x)dx - [n/k]\left(\phi_D^2(x_{[(l-1)k/D],l}) - \phi_D^2(0)\right)\\
& -[n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]}^{[lk/D]}\int_{x_{j,l}}^{x_{j+1,l}}\left|\phi_D^2(x_{j,l}) - \phi_D^2(x)\right|dx,
\end{aligned}$$

which allows us to use the Lipshitz properties of ϕ_D. We thus obtain

$$\begin{aligned}
\Sigma_l \quad \ge \quad & [n/k]\frac{k}{D}\int_0^1 \phi_D^2(x)dx - [n/k]2\|\phi_D\phi_D'\|_\infty|x_{[(l-1)k/D],l}| - 2\|\phi_D\phi_D'\|_\infty[n/k]\frac{k}{D}\left(\frac{k}{D}+2\right)\left(\frac{D}{k}\right)^2\\
\ge \quad & [n/k]\left(\frac{k}{D}\int_0^1 \phi_D^2(x)dx - 8\|\phi_D\phi_D'\|_\infty\right).
\end{aligned}$$

Since ϕ_D agrees with ϕ on $]0,1[$, $\Sigma_l \geq [n/k]\left(\frac{k}{D}\int_0^1 \phi^2(x)dx - 8\|\phi\phi'\|_\infty\right)$. So if $8\|\phi\phi'\|_\infty \leq (k/2D)\int_0^1 \phi^2(x)dx$,

$$\Sigma_l \geq [n/k]\frac{k}{2D}\int_0^1 \phi^2(x)dx.$$

The upper bound is obtained by similar arguments. Since $D \leq k \leq n$,

$$\begin{aligned}
\Sigma_l \leq & 2[n/k]\sum_{j=[(l-1)k/D]+1}^{[lk/D]} \phi_D^2(x_{j,l}) \\
\leq & 2[n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]+1}^{[lk/D]-1}\left(\frac{D}{k}\phi_D^2(x_{j,l}) - \int_{x_{j,l}}^{x_{j+1,l}}\phi_D^2(x)dx\right) \\
& + 2[n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]+1}^{[lk/D]-1}\int_{x_{j,l}}^{x_{j+1,l}}\phi_D^2(x)dx + 2[n/k]\phi_D^2(x_{[lk/D],l}).
\end{aligned}$$

The fact that $\phi_D(1)=0$ leads as above to

$$\begin{aligned}
\Sigma_l \leq & 2[n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]+1}^{[lk/D]-1}\int_{x_{j,l}}^{x_{j+1,l}}\left(\phi_D^2(x_{j,l}) - \phi_D^2(x)\right)dx \\
& + 2[n/k]\frac{k}{D}\int_0^1 \phi_D^2(x)dx + 2[n/k]\left(\phi_D^2(x_{[lk/D]}) - \phi_D^2(1)\right) \\
\leq & 4[n/k]\frac{k}{D}\sum_{j=[(l-1)k/D]+1}^{[lk/D]-1}\|\phi_D\phi_D'\|_\infty\left(\frac{D}{k}\right)^2 + 2[n/k]\frac{k}{D}\int_0^1\phi_D^2(x)dx + 4[n/k]\|\phi_D\phi_D'\|_\infty\frac{D}{k} \\
\leq & 2[n/k]\frac{k}{D}\left(4\|\phi\phi'\|_\infty + \int_0^1\phi^2(x)dx\right),
\end{aligned}$$

which completes the proof.

- **A lower bound for $\tilde{\rho}(\mathcal{S}_k(s,R),\alpha,\delta,\sigma)$**

Let k satisfy $\kappa_1^{-1} \leq k \leq n/2$. For all $D \in \{1,\ldots,[\kappa_1 k]\}$, we introduce

$$r_{D,k,s}^2 = \frac{1}{4}\frac{\int_0^1\phi^2(x)dx}{\int_0^1 {\phi^{(s)}}^2(x)dx}R^2D^{-2s}\left(\frac{k}{n}\right)^{2s} \wedge \kappa_2(1-\alpha-\delta)\frac{\sqrt{D}\sigma^2}{n},$$

and we aim here to obtain the lower bound:

$$\tilde{\rho}^2(\mathcal{S}_k(s,R),\alpha,\delta,\sigma) \geq \sup_{D\in\{1,\ldots,[\kappa_1 k]\}} r_{D,k,s}^2.$$

The fact that $r_{D,k,s}^2 \leq \kappa_2(1-\alpha-\delta)\sqrt{D}\sigma^2/n$ implies by Proposition 4 that

$$\beta\left(\{f\in\widetilde{S_D},\ \|f\|_n = r_{D,k,s}\}\right) \geq \delta.$$

Hence, we only need to prove that for every $D \in \{1, \ldots, [\kappa_1 k]\}$,

$$\{f \in \widetilde{S_D},\ \|f\|_n = r_{D,k,s}\} \subset \{f \in \mathcal{S}_k(s,R),\ \|f\|_n \geq r_{D,k,s}\},$$

to ensure that

$$\beta\left(\{f \in \mathcal{S}_k(s,R),\ \|f\|_n \geq r_{D,k,s}\}\right) \geq \delta,$$

for all $D \in \{1, \ldots, [\kappa_1 k]\}$. The expected lower bound for $\tilde{\rho}^2(\mathcal{S}_k(s,R), \alpha, \delta, \sigma)$ will then directly follow.

So, fixing $D \in \{1, \ldots, [\kappa_1 k]\}$, let us take f in $\widetilde{S_D}$ with

$$f(x) = \sum_{l=1}^{D} \beta_l \phi_D \left(\frac{Dnx}{k} - l + 1\right),$$

and prove that $||f||_n = r_{D,k,s}$ implies that f belongs to $\mathcal{S}_k(s,R)$. It is easy to see that f belongs to $\mathcal{C}^s([0,1])$ and that f is periodic with period k/n. Furthermore, by the properties of the support of ϕ_D,

$$\begin{aligned}
\frac{n}{k}\int_0^{k/n} f^{(s)^2}(x)dx &= \frac{n}{k}\int_0^{k/n} \left(\sum_{l=1}^{D} \beta_l^2 \left(\frac{Dn}{k}\right)^{2s} \phi_D^{(s)^2}\left(\frac{Dnx}{k} - l + 1\right)\right) dx \\
&= \left(\frac{Dn}{k}\right)^{2s} \frac{1}{D} \sum_{l=1}^{D} \beta_l^2 \int_0^1 \phi^{(s)^2}(y)dy. \qquad (21)
\end{aligned}$$

From the assumptions on the period and the support of ϕ_D, we can see that:

$$||f||_n^2 = \frac{1}{n}\sum_{l=1}^{D} \beta_l^2 \left(\left[\frac{n}{k}\right] \sum_{j \in A_{l,D}} \phi_D^2\left(\frac{Dj}{k} - l + 1\right) + \sum_{j \in A'_{l,D}} \phi_D^2\left(\frac{Dj}{k} - l + 1\right) \right),$$

with

$$A_{l,D} = \left\{ j \in \left]\frac{(l-1)k}{D}, \frac{lk}{D}\right[\right\}$$

and

$$A'_{l,D} = A_{l,D} \cap \{1, \ldots, n - [n/k]k\}.$$

Lemma 1 then leads to the inequality:

$$||f||_n^2 \geq \frac{1}{4D}\sum_{l=1}^{D} \beta_l^2 \int_0^1 \phi^2(x)dx.$$

Hence, if

$$||f||_n^2 = r_{D,k,s}^2 \leq \frac{1}{4}\frac{\int_0^1 \phi^2(x)dx}{\int_0^1 \phi^{(s)^2}(x)dx} R^2 \left(\frac{k}{n}\right)^{2s} D^{-2s}$$

we have that

$$\frac{1}{D}\sum_{l=1}^{D} \beta_l^2 \leq \frac{R^2 \left(\frac{k}{n}\right)^{2s} D^{-2s}}{\int_0^1 \phi^{(s)^2}(x)dx}.$$

From (21), we finally deduce that

$$\frac{n}{k}\int_0^{k/n} f^{(s)^2}(x)dx \leq R^2,$$

so, $f \in \mathcal{S}_k(s,R)$. This concludes the proof of Theorem 1.

6.2 Proof of Proposition 1

Let $\bar{f} = (f(1/n), \ldots, f(1))^T$. As in Theorem 2 from Baraud (2002), by using some deviations inequalities due to Birgé (2000) and Laurent and Massart (2000), since the dimension of $S_{D_k^*}$ is smaller than D_k^*, we can prove that for all f satisfying:

$$\|f\|_n^2 \geq \left\{ \|\bar{f} - \Pi_{S_{D_k^*}}(\bar{f})\|_{\mathbb{R}^n}^2 + 2(\sqrt{5}+4)\log\left(\frac{1}{\alpha\delta}\right)\frac{\sqrt{D_k^*}}{n}\sigma^2 \right\}, \tag{22}$$

then

$$\mathbb{P}_f(\Phi_\alpha^* = 0) \leq \delta.$$

The final result is obtained by computing an upper bound for $\|\bar{f} - \Pi_{S_{D_k^*}}(\bar{f})\|_{\mathbb{R}^n}^2$ when f belongs to $\mathcal{S}_k(s,R)$. This can be derived via the following lemma.

Lemma 2. *Let d_n be the distance associated with the norm $\|.\|_{\mathbb{R}^n}$ and for $D \in \mathbb{N}^*$, consider S_D defined by (6). Let f be some periodic function with period k/n such that for all $x \in [0,1]$, $f(x) = \sum_{l\geq 1}\beta_l(f)\varphi_{k/n,l}(x)$. If*

$$(2\pi)^{2s}\sum_{l\geq 2}[l/2]^{2s}\beta_l^2(f) \leq R^2\left(\frac{k}{n}\right)^{2s},$$

then, there exists some absolute constant $C > 0$ such that, setting $\bar{f} = (f(1/n), \ldots, f(1))^T$:

$$d_n^2\left(\bar{f}, S_D\right) \leq C\left(R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} + \frac{R^2 k}{n^{2s}}\right).$$

The proof of Lemma 2 is postponed to the end of the section.

Assume that f belongs to $\mathcal{S}_k(s,R)$. Then f satisfies the assumptions of Lemma 2 and we deduce from Lemma 2 and (22) that there exists some positive constant $c(\alpha,\delta)$ such that if f satisfies

$$\|f\|_n^2 \geq c(\alpha,\delta)\left\{R^2\left(\frac{k}{n}\right)^{2s} D_k^{*-2s} + \frac{R^2k}{n^{2s}} + \frac{\sqrt{D_k^*}}{n}\sigma^2\right\},$$

$$\mathbb{P}_f(\Phi_\alpha^* = 0) \leq \delta.$$

Recalling that $D_k^* = D^* \wedge k$ where D^* is the smallest integer at least equal to $(R^2(k/n)^{2s}n/\sigma^2)^{2/(1+4s)}$, we now distinguish three cases.
If $D^* = 1$, then $(R^2(k/n)^{2s}n/\sigma^2)^{2/(1+4s)} \leq 1$ and $D_k^* = 1$. Hence

$$R^2\left(\frac{k}{n}\right)^{2s} D_k^{*-2s} + \frac{R^2k}{n^{2s}} + \frac{\sqrt{D_k^*}}{n}\sigma^2 \leq 3\frac{\sigma^2}{n}.$$

If $D^* > k$, $(R^2(k/n)^{2s}n/\sigma^2)^{2/(1+4s)} \geq k$ and $D_k^* = k$. This implies that

$$R^2\left(\frac{k}{n}\right)^{2s} D_k^{*-2s} + \frac{R^2k}{n^{2s}} + \frac{\sqrt{D_k^*}}{n}\sigma^2 \leq 3\frac{R^2k}{n^{2s}}.$$

Finally, if $2 \leq D^* \leq k$, since $D^* - 1 < (R^2(k/n)^{2s}n/\sigma^2)^{2/(1+4s)} \leq D^*$ and $D^* \leq 2(D^* - 1)$, we obtain that

$$R^2\left(\frac{k}{n}\right)^{2s} D_k^{*-2s} + \frac{R^2 k}{n^{2s}} + \frac{\sqrt{D_k^*}}{n}\sigma^2 \leq (1+\sqrt{2})R^{\frac{2}{1+4s}}\left(\frac{k}{n}\right)^{\frac{2s}{1+4s}}\left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}} + \frac{R^2 k}{n^{2s}}.$$

Furthermore, assuming that (8) holds,

$$R^2\left(\frac{k}{n}\right)^{2s} D_k^{*-2s} + \frac{R^2 k}{n^{2s}} + \frac{\sqrt{D_k^*}}{n}\sigma^2 \leq (7+\sqrt{2})R^{\frac{2}{1+4s}}\left(\frac{k}{n}\right)^{\frac{2s}{1+4s}}\left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}},$$

which ends the proof of Proposition 1.

Proof of Lemma 2: We have that

$$\begin{aligned} d_n^2(\overline{f}, S_D) &\leq \frac{1}{n}\sum_{j=1}^{n}\left(f(j/n) - \sum_{l=1}^{D}\beta_l(f)\varphi_{k/n,l}(j/n)\right)^2 \\ &\leq \frac{1}{n}\left([n/k]+1\right)\sum_{j=1}^{k}\left(\sum_{l\geq D+1}\beta_l(f)\varphi_{k/n,l}(j/n)\right)^2. \end{aligned}$$

Hence,

$$\begin{aligned} d_n^2(\overline{f}, S_D) &\leq \frac{2}{n}\left([n/k]+1\right)\sum_{j=1}^{k}\left(\sum_{l=D+1}^{k\vee(D+1)}\beta_l(f)\varphi_{k/n,l}(j/n)\right)^2 \\ &\quad +\frac{2}{n}\left([n/k]+1\right)\sum_{j=1}^{k}\left(\sum_{l\geq(k+1)\vee(D+2)}\beta_l(f)\varphi_{k/n,l}(j/n)\right)^2. \end{aligned}$$

We can see by straightforward computations that for $l, l' \in \{1, \ldots, k\}, l \neq l'$:

$$\sum_{j=1}^{k}\varphi_{k/n,l}(j/n)\varphi_{k/n,l'}(j/n) = 0.$$

So, we get

$$d_n^2(\overline{f}, S_D) \leq \frac{4}{n}\left([n/k]+1\right)k\sum_{l\geq D+1}\beta_l^2(f) + \frac{4}{n}\left([n/k]+1\right)k\sum_{l\neq l'\geq(k+1)\vee(D+2)}|\beta_l(f)||\beta_{l'}(f)|.$$

Therefore,

$$d_n^2(\overline{f}, S_D) \leq 8\left(\sum_{l\geq D+1}\beta_l^2(f) + \left(\sum_{l\geq k+1}|\beta_l(f)|\right)^2\right).$$

On the one hand, since

$$\sum_{l\geq D+1}\beta_l^2(f)\left(\frac{l}{2}-\frac{1}{2}\right)^{2s} \leq (2\pi)^{-2s}R^2\left(\frac{k}{n}\right)^{2s},$$

then for $D \geq 1$,

$$\sum_{l \geq D+1} \beta_l^2(f) \leq (2\pi)^{-2s} R^2 \left(\frac{k}{n}\right)^{2s} \left(\frac{D}{2}\right)^{-2s}.$$

On the other hand, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
\left(\sum_{l \geq k+1} |\beta_l(f)|\right)^2 &\leq \left(\sum_{l \geq k+1} \beta_l^2(f)[l/2]^{2s}\right)\left(\sum_{l \geq k+1} \frac{1}{[l/2]^{2s}}\right) \\
&\leq (2\pi)^{-2s} R^2 \left(\frac{k}{n}\right)^{2s} \left(\sum_{l \geq k+1} \left(\frac{l}{2} - \frac{1}{2}\right)^{-2s}\right) \\
&\leq (2\pi)^{-2s} R^2 \left(\frac{k}{n}\right)^{2s} \left(\sum_{l \geq k+1} \left(\frac{l}{4}\right)^{-2s}\right).
\end{aligned}$$

Since s is supposed to be larger than 1,

$$\Big(\sum_{l \geq k+1} |\beta_l(f)|\Big)^2 \leq R^2 \left(\frac{k}{n}\right)^{2s} k^{1-2s} \leq R^2 k n^{-2s},$$

and finally,

$$d_n^2(\overline{f}, S_D) \leq C\left(R^2 \left(\frac{k}{n}\right)^{2s} D^{-2s} + \frac{R^2 k}{n^{2s}}\right).$$

6.3 Proof of Proposition 2

Since

$$\begin{aligned}
\mathbb{P}_0\Bigg(\sup_{m \in \mathcal{M}} \left\{\frac{(n-D_m)\|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m\|Y - \Pi_m(Y)\|^2_{\mathbb{R}^n}} - \overline{F}^{-1}_{D_m, n-D_m}\left(\frac{\alpha}{|\mathcal{M}|}\right)\right\} > 0\Bigg) \\
\leq \sum_{m \in \mathcal{M}} \mathbb{P}_0\left(\frac{(n-D_m)\|\Pi_m(Y)\|^2_{\mathbb{R}^n}}{D_m\|Y - \Pi_m(Y)\|^2_{\mathbb{R}^n}} > \overline{F}^{-1}_{D_m, n-D_m}\left(\frac{\alpha}{|\mathcal{M}|}\right)\right) \\
\leq \alpha,
\end{aligned}$$

by the definition of u_α, we have that $u_\alpha \geq \alpha/|\mathcal{M}|$. This implies that $u_\alpha \geq \alpha \log 2/\log k$ and since $k \leq n$, under the assumptions of the proposition, $(\mathcal{H}_\mathcal{M})$ is satisfied. Following Theorem 2, we then look for an upper bound for the quantity:

$$\inf_{m \in \mathcal{M}} \left\{ \begin{array}{c} \left[1 + 10\left(\frac{D_m + \log(1/u_\alpha)}{n - D_m}\right)\right] d_n^2(\overline{f}, S_m) \\ + \frac{\sigma^2}{n}\left[5\left(1 + \sqrt{\frac{D_m}{n-D_m}}\right)\sqrt{D_m \log\left(\frac{2}{\delta u_\alpha}\right)} + \frac{25}{2}\left(1 + 2\frac{D_m}{n-D_m}\right)\log\left(\frac{2}{u_\alpha \delta}\right)\right] \end{array} \right\}.$$

From the inequalities $u_\alpha \geq \alpha \log 2/\log k$ and $D_m \leq m \leq n/2$ for all $m \in \mathcal{M}$, together with $(\mathcal{H}_\mathcal{M})$, we derive that we only need to find an upper bound for

$$\inf_{D \in \mathcal{M}} \left\{c_1 d_n^2(\overline{f}, S_D) + c_2(\alpha, \delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\,\sigma^2\right\} + c_3(\alpha, \delta)\log\log k\frac{\sigma^2}{n}.$$

Assuming that $f \in \mathcal{S}_k(s,R)$, Lemma 2 gives:

$$d_n^2(\overline{f}, S_D) \le C\left(R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} + \frac{R^2 k}{n^{2s}}\right),$$

and we obtain

$$\begin{aligned}
&\inf_{D\in\mathcal{M}}\left\{c_1 d_n^2(\overline{f}, S_D) + c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\} + c_3(\alpha,\delta)\log\log k\frac{\sigma^2}{n} \\
&\le \inf_{D\in\mathcal{M}}\left\{C_1 R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} + c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\} + c_3(\alpha,\delta)\log\log k\frac{\sigma^2}{n} + C_1\frac{R^2 k}{n^{2s}}.
\end{aligned}$$

We can now compute the expected upper bound by noticing that

$$R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} \le \frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2 \text{ and } D\in\mathcal{M} \quad \text{if and only if} \quad D \ge D^* \text{ and } D \in \mathcal{M},$$

where D^* is the smallest dyadic integer at least equal to

$$R^{\frac{4}{4s+1}}\left(\frac{k}{n}\right)^{\frac{4s}{4s+1}} n^{\frac{2}{4s+1}}\left(\sqrt{\log\log k}\right)^{-\frac{2}{4s+1}}\sigma^{-\frac{4}{4s+1}},$$

that is

$$D^* = \inf\left\{2^J,\ J\in\mathbb{N},\ 2^J \ge \left(R^2(k/n)^{2s} n\sqrt{\log\log k}^{-1}\sigma^{-2}\right)^{2/(4s+1)}\right\}.$$

If $D^* \in \{2,\ldots,2^{[\log_2(k/2)]}\}$, then

$$\begin{aligned}
&\inf_{D\in\mathcal{M}}\left\{C_1 R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} + c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\} \\
&\qquad\le C_1 R^2\left(\frac{k}{n}\right)^{2s} D^{*-2s} + c_2(\alpha,\delta)\frac{\sqrt{D^*}}{n}\sqrt{\log\log k}\sigma^2 \\
&\qquad\le (C_1 + c_2(\alpha,\delta))\frac{\sqrt{D^*}}{n}\sqrt{\log\log k}\sigma^2.
\end{aligned}$$

Since $D^* \le 2\left(R^2(k/n)^{2s} n\sqrt{\log\log k}^{-1}\sigma^{-2}\right)^{\frac{2}{4s+1}}$, this leads to

$$\begin{aligned}
&\inf_{D\in\mathcal{M}}\left\{C_1 R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} + c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\} \\
&\qquad\le c(\alpha,\delta) R^{\frac{2}{4s+1}}\left(\frac{k}{n}\right)^{\frac{2s}{4s+1}}\left(\frac{\sqrt{\log\log k}\sigma^2}{n}\right)^{\frac{4s}{4s+1}}.
\end{aligned}$$

If $D^* = 1$, then for all $D \in \mathcal{M}$,

$$R^2\left(\frac{k}{n}\right)^{2s} D^{-2s} \le \frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2,$$

and

$$\inf_{D\in\mathcal{M}}\left\{C_1R^2\left(\frac{k}{n}\right)^{2s}D^{-2s}+c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\}\leq c'(\alpha,\delta)\sqrt{\log\log k}\frac{\sigma^2}{n}.$$

If $D^*>2^{[\log_2(k/2)]}$, then for all $D\in\mathcal{M}$,

$$\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2<R^2\left(\frac{k}{n}\right)^{2s}D^{-2s},$$

and

$$\begin{aligned}\inf_{D\in\mathcal{M}}\left\{C_1R^2\left(\frac{k}{n}\right)^{2s}D^{-2s}+c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\} &\leq c(\alpha,\delta)R^2\left(\frac{k}{n}\right)^{2s}\left(2^{[\log_2(k/2)]}\right)^{-2s}\\ &\leq 2^{4s}c(\alpha,\delta)\frac{R^2}{n^{2s}}.\end{aligned}$$

So, in all cases

$$\begin{aligned}&\inf_{D\in\mathcal{M}}\left\{C_1R^2\left(\frac{k}{n}\right)^{2s}D^{-2s}+c_2(\alpha,\delta)\frac{\sqrt{D}}{n}\sqrt{\log\log k}\sigma^2\right\}+c_3(\alpha,\delta)\log\log k\frac{\sigma^2}{n}+C_1\frac{R^2k}{n^{2s}}\\ &\quad\leq c(\alpha,\delta)R^{\frac{2}{4s+1}}\left(\frac{k}{n}\right)^{\frac{2s}{4s+1}}\left(\frac{\sqrt{\log\log k}\sigma^2}{n}\right)^{\frac{4s}{4s+1}}+c'(\alpha,\delta)\frac{\sqrt{\log\log k}\sigma^2}{n}\\ &\qquad+c(s,\alpha,\delta)\frac{R^2}{n^{2s}}+c_3(\alpha,\delta)\frac{\log\log k\sigma^2}{n}+C_1\frac{R^2k}{n^{2s}},\end{aligned}$$

which gives the result.

6.4 Proof of Proposition 3

As in the proof of Proposition 2, since $u_\alpha\geq\alpha/|\mathcal{M}|$, under the assumptions of Proposition 3, $(\mathcal{H}_\mathcal{M})$ is satisfied and we look for an upper bound for

$$\inf_{(k,D)\in\mathcal{M}}\left\{c_1d_n^2(\overline{f},S_{(k,D)})+c_2(\alpha,\delta)\sqrt{D\log n}\frac{\sigma^2}{n}\right\}+c_3(\alpha,\delta)\log n\frac{\sigma^2}{n},$$

when $f\in\mathcal{S}(s,R,\tau_1)$, for $s\in\mathbb{N}^*$, $R>0$, $\tau_1\in[2/n,1]$.

Assume that $f\in\mathcal{S}(s,R,\tau_1)$. Let $k(f)=[\tau(f)n]+1$ if $\tau(f)n\notin\mathbb{N}^*$, $\tau(f)n$ if $\tau(f)n\in\mathbb{N}^*$. Then $k(f)\in\{2,\ldots,n\}$ and

$$\begin{aligned}&\inf_{(k,D)\in\mathcal{M}}\left\{c_1d_n^2(\overline{f},S_{(k,D)})+c_2(\alpha,\delta)\sqrt{D\log n}\frac{\sigma^2}{n}\right\}\\ &\qquad\leq\inf_{D\in\{1,\ldots,2^{[\log_2(k(f)/2)]}\}}\left\{c_1d_n^2(\overline{f},S_{(k(f),D)})+c_2(\alpha,\delta)\sqrt{D\log n}\frac{\sigma^2}{n}\right\}.\end{aligned}$$

Denote by $\beta_l(f)$ the l-th Fourier coefficient of f. Recall that if $f\in\mathcal{S}(s,R,\tau_1)$, for all $x\in[0,1]$, $f(x)=\beta_1(f)+\sum_{l\geq2}\beta_l(f)\varphi_{\tau(f),l}(x)$ with

$$(2\pi)^{2s}\sum_{l\geq2}[l/2]^{2s}\beta_l^2(f)\leq R^2\tau(f)^{2s},\tag{23}$$

and $\sum_{l\geq 1}|\beta_l(f)| < +\infty$. By introducing $\tilde{f}$ such that $\tilde{f}(x) = \beta_1(f) + \sum_{l\geq 2}\beta_l(f)\varphi_{k(f)/n,l}(x)$, we have:

$$d_n^2\left(\overline{f}, S_{(k(f),D)}\right) \leq 2\left(d_n^2\left(\overline{f}, \overline{\tilde{f}}\right) + d_n^2\left(\overline{\tilde{f}}, S_{(k(f),D)}\right)\right),$$

where $\overline{\tilde{f}} = (\tilde{f}(1/n), \dots, \tilde{f}(1))^T$.

On the one hand, since $\tau(f) \leq k(f)/n$, (23) implies that $\tilde{f}$ satisfies the assumptions of Lemma 2 with $k = k(f)$, so:

$$\begin{aligned} d_n^2\left(\overline{\tilde{f}}, S_{(k(f),D)}\right) &\leq C\Big(R^2(k(f)/n)^{2s}D^{-2s} + R^2k(f)n^{-2s}\Big) \\ &\leq c(s)\Big(R^2\tau(f)^{2s}D^{-2s} + R^2\tau(f)n^{1-2s}\Big). \end{aligned}$$

On the other hand, using the Cauchy-Schwarz inequality,

$$\begin{aligned} d_n^2\left(\overline{f}, \overline{\tilde{f}}\right) &= \frac{1}{n}\sum_{j=1}^{n}\left(\sum_{l\geq 2}\beta_l(f)\left[\varphi_{\tau(f),l}(j/n) - \varphi_{k(f)/n,l}(j/n)\right]\right)^2 \\ &\leq \frac{8}{n}\sum_{j=1}^{n}\left(\sum_{l\geq 2}[l/2]^{2s}\beta_l^2(f)\right)\left(\sum_{l\geq 2}[l/2]^{-2s}\left|\sin\left(\frac{\pi j[l/2]}{n\tau(f)} - \frac{\pi j[l/2]}{k(f)}\right)\right|^2\right). \end{aligned}$$

Let $\gamma \in [0,1[$ that will be chosen later. For all $x \in \mathbb{R}$, $\sin^2(x) \leq |x|^\gamma$, which implies that:

$$\begin{aligned} d_n^2\left(\overline{f}, \overline{\tilde{f}}\right) &\leq \frac{c(s)}{n}R^2\tau(f)^{2s}\sum_{j=1}^{n}\left(\sum_{l\geq 2}j^\gamma[l/2]^{\gamma-2s}\frac{1}{(\tau(f)n)^{2\gamma}}\right) \\ &\leq c(s)\frac{R^2\tau(f)^{2s-2\gamma}}{n^{2\gamma+1}}\left(\sum_{j=1}^{n}j^\gamma\right)\left(\sum_{l\geq 2}[l/2]^{\gamma-2s}\right) \\ &\leq \frac{c'(s)}{2s-\gamma-1}\frac{R^2\tau(f)^{2s-2\gamma}}{n^\gamma}. \end{aligned}$$

Then, we have:

$$\begin{aligned} &\inf_{(k,D)\in\mathcal{M}}\left\{c_1d_n^2(\overline{f}, S_{(k,D)}) + c_2(\alpha,\delta)\sqrt{D\log n}\frac{\sigma^2}{n}\right\} + c_3(\alpha,\delta)\log n\frac{\sigma^2}{n} \\ &\quad\leq \inf_{D\in\{1,\dots,2^{[\log_2(k(f)/2)]}\}}\left\{c_1(s)R^2D^{-2s}\tau(f)^{2s} + c_2(\alpha,\delta)\sqrt{D\log n}\frac{\sigma^2}{n}\right\} \\ &\qquad\qquad + c_3(\alpha,\delta)\log n\frac{\sigma^2}{n} + \frac{c_4(s)}{2s-\gamma-1}\frac{R^2\tau(f)^{2s-2\gamma}}{n^\gamma} + c_1(s)\frac{R^2\tau(f)}{n^{2s-1}}. \end{aligned}$$

Note again that $R^2D^{-2s}\tau(f)^{2s} \leq \sqrt{D\log n}\sigma^2/n$ and D is dyadic if and only if $D \geq D^*(f)$, where $D^*(f)$ is the smallest dyadic integer at least equal to $\left(R^2\tau(f)^{2s}n\sigma^{-2}(\log n)^{-1/2}\right)^{2/(4s+1)}$, that is

$$D^*(f) = \inf\left\{2^J,\ J\in\mathbb{N},\ 2^J \geq \left(R^2\tau(f)^{2s}n\sqrt{\log n}^{-1}\sigma^{-2}\right)^{2/(4s+1)}\right\}.$$

If $D^*(f) \in \{2, \ldots, 2^{[\log_2(k(f)/2)]}\}$, then

$$D^*(f) \leq 2 \left(\frac{R^2 \tau(f)^{2s} n}{\sigma^2 \sqrt{\log n}} \right)^{\frac{2}{4s+1}},$$

and

$$\begin{aligned} \inf_{D \in \{1, \ldots, 2^{[\log_2(k(f)/2)]}\}} & \left\{ c_1(s) R^2 D^{-2s} \tau(f)^{2s} + c_2(\alpha, \delta) \sqrt{D \log n} \frac{\sigma^2}{n} \right\} \\ & \leq c_1(s) R^2 D^*(f)^{-2s} \tau(f)^{2s} + c_2(\alpha, \delta) \sqrt{D^*(f) \log n} \frac{\sigma^2}{n} \\ & \leq c(s, \alpha, \delta) \left(R^2 \tau(f)^{2s} \right)^{\frac{1}{4s+1}} \left(\sqrt{\log n} \frac{\sigma^2}{n} \right)^{\frac{4s}{4s+1}}. \end{aligned}$$

If $D^*(f) = 1$, then for all $D \in \{1, \ldots, 2^{[\log_2(k(f)/2)]}\}$, $R^2 D^{-2s} \tau(f)^{2s} \leq \sqrt{D \log n} \sigma^2 / n$ and

$$\begin{aligned} \inf_{D \in \{1, \ldots, 2^{[\log_2(k(f)/2)]}\}} & \left\{ c_1(s) R^2 D^{-2s} \tau(f)^{2s} + c_2(\alpha, \delta) \sqrt{D \log n} \frac{\sigma^2}{n} \right\} \\ & \leq c'(s, \alpha, \delta) \inf_{D \in \{1, \ldots, 2^{[\log_2(k(f)/2)]}\}} \left\{ \sqrt{D \log n} \frac{\sigma^2}{n} \right\} \\ & \leq c'(s, \alpha, \delta) \sqrt{\log n} \frac{\sigma^2}{n}. \end{aligned}$$

At last, if $D^*(f) > 2^{[\log_2(k(f)/2)]}$, for all $D \in \{1, \ldots, 2^{[\log_2(k(f)/2)]}\}$, $R^2 D^{-2s} \tau(f)^{2s} > \sqrt{D \log n} \sigma^2 / n$, which leads to:

$$\inf_{D \in \{1, \ldots, 2^{[\log_2(k(f)/2)]}\}} \left\{ c_1(s) R^2 D^{-2s} \tau(f)^{2s} + c_2(\alpha, \delta) \sqrt{D \log n} \frac{\sigma^2}{n} \right\} \leq c''(s, \alpha, \delta) R^2 n^{-2s}.$$

Finally,

$$\begin{aligned} \inf_{(k,D) \in \mathcal{M}} & \left\{ c_1 d_n^2(\overline{f}, S_{(k,D)}) + c_2(\alpha, \delta) \sqrt{D \log n} \frac{\sigma^2}{n} \right\} + c_3(\alpha, \delta) \log n \frac{\sigma^2}{n} \\ & \leq C(s, \alpha, \delta) \Bigg\{ \left(R^2 \tau_1^{2s} \right)^{\frac{1}{4s+1}} \left(\sqrt{\log n} \frac{\sigma^2}{n} \right)^{\frac{4s}{4s+1}} + \log n \frac{\sigma^2}{n} + \frac{R^2 \tau_1}{n^{2s-1}} \\ & \qquad + \frac{R^2}{n^{2s}} + \frac{1}{2s - \gamma - 1} \frac{R^2 \tau_1^{2s}}{(n \tau_1^2)^\gamma} \Bigg\}. \end{aligned}$$

Since $n^{-1} \leq \tau_1 \leq 1$, we can see that for all $\gamma \in [0, 1[$, for all $s \in \mathbb{N}^*$,

$$\frac{R^2}{n^{2s}} \leq \frac{R^2 \tau_1}{n^{2s-1}} \leq \frac{R^2 \tau_1^{2s}}{(n \tau_1^2)^\gamma}.$$

To prove the final result, we need to distinguish between two cases in order to optimize the choice of γ. If $n \tau_1^2 \leq 2$, by taking $\gamma = 0$, we obtain:

$$\frac{1}{2s - \gamma - 1} \frac{R^2 \tau_1^{2s}}{(n \tau_1^2)^\gamma} = \frac{1}{2s - 1} R^2 \tau_1^{2s} = \frac{1}{2s - 1} \left(2 \frac{R^2 \tau_1^{2s}}{n \tau_1^2} \wedge R^2 \tau_1^{2s} \right).$$

If $n\tau_1^2 > 2$, by taking $\gamma_n = 1 - \log 2/\log(n\tau_1^2)$, we get:

$$\frac{1}{2s-1-\gamma_n}\frac{R^2\tau_1^{2s}}{(n\tau_1^2)^{\gamma_n}} \leq \frac{2}{2s-1-\gamma_n}\frac{R^2\tau_1^{2s}}{n\tau_1^2} \leq \frac{2\log n}{\log 2}\frac{R^2\tau_1^{2s}}{n\tau_1^2} \leq \frac{2\log n}{\log 2}\left(2\frac{R^2\tau_1^{2s}}{n\tau_1^2} \wedge R^2\tau_1^{2s}\right),$$

which ends the proof.

7 Appendix: the Gaussian white noise model

In this section, we consider the following ideal Gaussian white noise model instead of our current regression framework:

$$dY(t) = f(t)dt + \frac{1}{\sqrt{n}}dW(t),\ t \in [0,1]. \tag{24}$$

Let τ such that $1/\tau$ is an integer. Assume that f belongs to $\mathbb{L}_2([0,1])$ and that f is periodic with period τ. By projecting the above model onto the Fourier basis $\{\varphi_{\tau,l},\ l \geq 1\}$ defined by (5) on $[0,1]$, we obtain the Gaussian sequence model:

$$Y_l = \beta_l(f) + \frac{\varepsilon_l}{\sqrt{n}},\ l \in \mathbb{N}^*, \tag{25}$$

where

$$Y_l = \int_0^1 \varphi_{\tau,l}(t)dY(t),\ \beta_l(f) = \int_0^1 \varphi_{\tau,l}(t)f(t)dt,\ \varepsilon_l = \int_0^1 \varphi_{\tau,l}(t)dW(t).$$

Testing the null hypothesis "$f = 0$" against "$f \neq 0$" in the model (24) is equivalent to testing "$\beta_l(f) = 0, \forall l \geq 1$" against "$\exists l \geq 1,\ \beta_l(f) \neq 0$" in the model (25).
Let $\mathcal{E}(\tau, R)$ be the ellipsoid defined by

$$\mathcal{E}(\tau,R) = \left\{(\beta_l(f),\ l \geq 1),\ \sum_{l\geq 2}\frac{\beta_l^2(f)}{([l/2]^{-s})^2} \leq (2\pi)^{-2s}R^2\tau^{2s}\right\}.$$

Considering the problem of testing "$\beta_l(f) = 0, \forall l \geq 1$" in the model (25), Baraud has comprehensively studied the minimax separation rate $\bar{\rho}$ over such an ellipsoid with respect to the classical norm of $l^2(\mathbb{N})$. From his results, we can deduce the following bounds:

$$\bar{\rho}^2\left(\mathcal{E}(\tau,R),\alpha,\delta\right) \geq \sup_{D\geq 1}\left\{\left(\frac{\sqrt{2\mathcal{L}(\alpha,\delta)D}}{n}\right) \wedge \left((2\pi)^{-2s}R^2\tau^{2s}[D/2]^{-2s}\right)\right\}$$

where $\mathcal{L}(\alpha,\delta) = \log\left(1 + 4(1-\alpha-\delta)^2\right)$, and if $R > \tau^{-s}n^{-1/2}$,

$$\bar{\rho}^2\left(\mathcal{E}(\tau,R),\alpha,\delta\right) \leq c(s,\alpha,\delta)\sup_{D\geq 1}\left\{\frac{\sqrt{D}}{n} \wedge R^2\tau^{2s}D^{-2s}\right\} \leq 2c(s,\alpha,\delta)R^{\frac{2}{4s+1}}\tau^{\frac{2s}{4s+1}}n^{-\frac{4s}{4s+1}}.$$

Since $\sum_{l\geq 1}\beta_l^2(f) = \tau^{-1}\int_0^\tau f^2(t)\mathrm{d}t = \|f\|_{2,\tau}^2$, we thus find lower bounds similar to the ones obtained in our Gaussian periodic fixed design regression framework over the classes $\mathcal{S}'_k(s,R) = \{f,\ f$ is periodic with period k/n and $f(x) = \sum_{l\geq 1}\beta_l(f)\varphi_{k/n,l}(x)$ with $(2\pi)^{2s}\sum_{l\geq 2}[l/2]^{2s}\beta_l^2(f) \leq R^2(k/n)^{2s}\}$ (see Theorem 1), except that the supremum over $D \geq 1$

is replaced by a supremum over $D \in \{1, \dots, [\kappa k]\}$. As for the upper bounds, we can see via Proposition 1, that if $R > (k/n)^{-s} n^{-1/2}$,

$$\tilde{\rho}^2 \left(\mathcal{S}'_k(s,R), \alpha, \delta, \sigma\right) \leq 2c(\alpha,\delta) \left(R^{\frac{2}{1+4s}} \left(\frac{k}{n}\right)^{\frac{2s}{4s+1}} \left(\frac{\sigma^2}{n}\right)^{\frac{4s}{4s+1}} + \frac{R^2 k}{n^{2s}} \right).$$

The losses that one can observe are essentially due to the fact that the Fourier family $\{\varphi_{k/n,l},\ l \geq 1\}$ is not orthogonal with respect to the norm $\|.\|_n$ that we consider.

References

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

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

Birgé, L. (2000). An alternative point of view on Lepski's method. *To appear in State of the Art in Probability and Statistics; Festschrift for W.R. van Zwet, Institute of Mathematical Statistics, Lecture Notes-Monograph Series*.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods* (Second ed.). Springer Series in Statistics. New York: Springer-Verlag.

Eubank, R. and Hart, J. (1992). Testing goodness-of-fit in regression via order selection criteria. *Ann. Stat. 20*(3), 1412–1425.

Gayraud, G. and Pouet, C. (2001). Minimax testing composite null hypotheses in the discrete regression scheme. *Math. Methods Stat. 10*(4), 375–394.

Gregory, P. and Loredo, T. (1992). A new method for the detection of a periodic signal of unknown shape and period. http://bayes.wustl.edu/gregory/gregory.html.

Härdle, W. and Kneip, A. (1999). Testing a regression model when we have smooth alternatives in mind. *Scand. J. Stat. 26*(2), 221–238.

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

Ingster, Y. (1982). Minimax nonparametric detection of signals in white Gaussian noise. *Probl. Inf. Transm. 18*, 130–140.

Ingster, Y. (1993). Asymptotically minimax testing for nonparametric alternatives I-II-III. *Math. Methods Statist. 2*, 85–114, 171–189, 249–268.

Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist. 28*(5), 1302–1338.

Lavielle, M. and Lévy-Leduc, C. (2003). A model selection approach for estimating the period of periodic functions with unknown shape and applications in laser vibrometry. Technical report, 2003-21 Université Paris 11.

Lepski, O. and Spokoiny, V. (1999). Minimax nonparametric hypothesis testing: The case of an inhomogeneous alternative. *Bernoulli 5*(2), 333–358.

Lepski, O. and Tsybakov, A. (2000). Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. *Probab. Theory Relat. Fields 117*(1), 17–48.

Prenat, M. (2001). Vibration modes and laser vibrometry performance in noise. In *Proceedings of the Physics in Signal and Image Processing (PSIP2001) conference, Marseille, France.*

Quinn, B. G. and Hannan, E. J. (2001). *The estimation and tracking of frequency.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press.

Spokoiny, V. (1996). Adaptive hypothesis testing using wavelets. *Ann. Stat. 24*(6), 2477–2498.

Spokoiny, V. (1998). Adaptive and spatially adaptive testing of a nonparametric hypothesis. *Math. Methods Stat. 7*(3), 245–273.

Magalie FROMONT
Laboratoire de Mathématiques
U.M.R C 8628
UNIVERSITÉ PARIS-SUD
Bât. 425
91405 ORSAY CEDEX
FRANCE
e-mail: magalie.fromont@math.u-psud.fr

Céline LÉVY-LEDUC
Laboratoire de Mathématiques
U.M.R C 8628
UNIVERSITÉ PARIS-SUD
Bât. 425
91405 ORSAY CEDEX
FRANCE
e-mail: celine.levy-leduc@math.u-psud.fr
THALÈS OPTRONIQUE
Rue Guynemer BP55 78283 Guyancourt Cédex
FRANCE
e-mail : celine.levy-leduc@fr.thalesgroup.com