Goodness-of-fit test for noisy directional data

Claire Lacour, Thanh Mai Pham Ngoc

Goodness-of-fit test for noisy directional data

Claire LACOUR * Thanh Mai PHAM NGOC †

March 8, 2012

Abstract

We consider spherical data X_i noised by a random rotation $\varepsilon_i \in \mathrm{SO}(3)$ so that only the sample $Z_i = \varepsilon_i X_i$, $i = 1, \dots, N$ is observed. We define a nonparametric test procedure to distinguish H_0 : "the density f of X_i is the uniform density f_0 on the sphere" and H_1 : "$\|f - f_0\|_2^2 \geq \mathcal{C}\psi_N^2$ and f is in a Sobolev space with smoothness s". For a noise density f_ε with smoothness index ν, we show that an adaptive procedure (i.e. s in not assumed to be known) cannot have a faster rate of separation than $\psi_N^{ad}(s) = (N/\sqrt{\log\log(N)})^{-2s/(2s+2\nu+1)}$ and we provide a procedure which reaches this rate. We also deal with the case of super smooth noise. We illustrate the theory by implementing our test procedure for various kinds of noise on SO(3) and show that it yields promising numerical results.

Keywords : Adaptive testing, spherical deconvolution, minimax hypothesis testing, nonparametric alternatives, spherical harmonics.
MSC 2010. Primary 62G10, secondary 62H11.

1 Introduction

We consider the spherical convolution model. We observe:

$$Z_i = \varepsilon_i X_i, \quad i = 1, \dots, N \tag{1}$$

where the ε_i are i.i.d. random variables of SO(3) the rotation group in $\mathbb{R}^3$ and the X_i's are i.i.d. random variables of $\mathbb{S}^2$, the unit sphere of $\mathbb{R}^3$. We suppose that X_i and ε_i are independent. We also assume that the distributions of Z_i and X_i are absolutely continuous with respect to the uniform measure on $\mathbb{S}^2$ and we set f_Z and f the densities of Z_i and X_i respectively. The distribution of ε_i is absolutely continuous with respect to the Haar measure on SO(3) and we will denote it f_ε. Then we have

$$f_Z = f_\varepsilon * f,$$

where $*$ denotes the convolution product which is defined below in (5).

Roughly speaking, the density spherical convolution model provides a setup where each genuine observation X_i is contaminated by a small random rotation. The aim of the present paper is to provide nonparametric adaptive minimax goodness-of-fit testing procedures on f from the

*Laboratoire de Mathématique, UMR 8628, Université Paris Sud, 91405 Orsay Cedex, France, Email: claire.lacour@math.u-psud.fr

†Laboratoire de Mathématique, UMR 8628, Université Paris Sud, 91405 Orsay Cedex, France Email: thanh.pham_ngoc@math.u-psud.fr

noisy observations Z_i. More precisely, let f_0 being the uniform density on $\mathbb{S}^2$, we consider the problem of testing the null hypothesis $f = f_0$ with alternatives expressed in $\mathbb{L}_2$ norm over Sobolev classes.

To the best of our knowledge, goodness-of-fit testing in spherical convolution model does not exist. Our results are strongly motivated by astrophysical interests. In astrophysics, a burning issue consists in understanding the behaviour of the so-called Ultra High Energy Cosmic Rays (UHECR). These latter are cosmic rays with an extreme kinetic energy (of the order of 10^{19} eV) and the rarest particles in the universe. The source of those most energetic particles remains a mystery and the stake lies in finding out their origins and which process produces them. Astrophysicists have at their disposal directional data which are measurements of the incoming directions of the UHECR on Earth.

Needless to say that finding out more about the law of probability of those incoming directions is crucial to gain an insight into the mechanisms generating the UHECR. Several hypotheses are made. A uniform density would suggest that the High Energy Cosmic Rays are generated by cosmological effects, such as the decay of relic particles from the Big Bang. On the contrary, if these UHECR are generated by astrophysical phenomena (such as acceleration into Active Galactic Nuclei (AGN)), then we should observe a density function which is highly non-uniform and tightly correlated with the the local distribution of extragalactic supermassive black holes at the center of nearby galaxies (AGN). First results seemed to favour a non-uniform density but as underlined by Faÿ et al. (2011), a more recent analysis based on 69 observations of UHECR softens this conclusion of anisotropy. To this prospect, these relevant considerations lead naturally to goodness-of-fit testing on the uniform density.

Faÿ et al. (2011) recently developed isotropy goodness-of-fit tests based on the so-called needlets for the direct case. Their study is focused on the practical aspect with simulations connected to realistic cosmic rays scenari. In our work, we consider the indirect case in which some noise occurs. Indeed, the observations of the events X_i are always most often perturbated by a secondary noise (ε_i) which leads to the convolution statistical model (1). Here, we put the emphasis on theoretical statements and completed them with promising simulations. In particular, we consider complex noise distributions on SO(3) such as the Gaussian and the Laplace ones which have never been implemented so far.

Considering goodness-of-fit testing in the spherical convolution model not only finds its interest in the above important applications, but it also fills a gap both in the noisy setup testing literature and the spherical convolution one. Indeed, convolution models have been extensively studied on the real line and more recently on the sphere. However, so far, only estimation has been treated in the spherical setup. For estimation problem, one is interested in recovering the underlying density f from noisy observations Z_i. The pioneer works of Healy et al. (1998), Kim and Koo (2002), Kim et al. (2004) introduced a minimax estimation procedure based on the Fourier basis of $\mathbb{L}_2(\mathbb{S}^2)$. Recently, Kerkyacharian et al. (2011) proposed an optimal and adaptive hard thresholding estimation procedure based on needlets.

Nonparametric goodness-of-fit testing has aroused a lot of interest. For minimax testing, we refer to the work of Ingster (1993) who settled down the theoretical basements. Spokoiny (1996) first established adaptive testing procedure based on wavelets over Besov bodies. Nonetheless, goodness-of-fit testing has mainly focused on the case of direct observations. Indeed, very few works have been devoted to the case of indirect observations and if so, only on $\mathbb{R}$. On this point, let us cite the pioneer works of Butucea (2007) who built minimax nonparametric goodness-of-fit testing for convolution models based on kernels methods, of Butucea et al. (2009) which made a

step forward by building an adaptive testing procedure in the noisy setup.

We would also like to bring to the reader's attention some interesting fact when encountering testing problems with indirect observations. Indeed, there is a natural pas de deux between the following approaches : to test $f = f_0$ or to test $f_\varepsilon * f = f_\varepsilon * f_0$. This question has been the object of the recent work of Laurent et al. (2011) and has been previously evoked by Butucea et al. (2009). In the case of the convolution model on the real line, Laurent, Loubes and Marteau prove that if a test procedure is minimax for testing problem : $H_0^D : f_\varepsilon * f = f_\varepsilon * f_0$ versus $H_1^D : f_\varepsilon * (f - f_0) \in \mathcal{F}_D$ where

$$\mathcal{F}_D = \{g \text{ with smoothness } s' \text{ and } \|g\|^2 \geq C' n^{-4s'/(4s'+1)}, \text{ with } s' = s + \nu\},$$

then it is minimax for $H_0^I : f = f_0$ versus $H_1^I : f - f_0 \in \mathcal{F}_I$ where

$$\mathcal{F}_I = \{f \text{ with smoothness } s \text{ and } \|f\|^2 \geq C n^{-4s/(4s+4\nu+1)}\}$$

but the reverse is not true (here n is the number of data and ν the smoothness index of the noise). This interesting conclusion (that we can conjecture true in our context also) does not make it any the less necessary to study the inverse problem here. Indeed, until the present work, the minimax rates was not known in the context of noisy spherical data. Moreover, when dealing with adaptive procedures, the link between the direct and inverse problems is not established yet.

In the present paper, the whole difficulty actually lies in the spherical geometry which complicates every steps that one encounters on $\mathbb{R}$. Indeed, the efficient test statistic of Butucea (2007) was built upon a deconvolution kernel estimator of the quadratic functional $\int (f - f_0)^2$. It is well-known that such an estimator is closely linked to the Fourier transform on $\mathbb{R}$. There exist kernel methods to treat density estimation for spherical data but only for direct observations (see Hall et al. (1987), and Bai et al. (1988)). Here in the spherical convolution context, Fourier analysis has a different behaviour and we resort to existing procedures to estimate the quadratic risk $\int (f - f_0)^2$. Those procedures (see Kim and Koo (2002)) are based on Fourier series which come down to projections. Consequently, the approach proves to be quite different than the one on the real line. The difficulty of testing in a spherical deconvolution model can be seen in the following way. If you use an orthogonal basis (ψ_k) to estimate the unknown function f, then using U-statistics requires that the "deconvoluted" basis ϕ_k (s.t. $\psi_k = f_\varepsilon * \phi_k$) is also (almost) orthogonal, which is delicate to realize. Thus one has to circumvent new problems linked to estimation of the quadratic functional, Fourier series, spherical context and convolution model setting. This explains why we choose to use spherical harmonics and their good properties in term of orthogonality.

The plan of the paper is as follows. In Section 2, we give a brief overview about harmonic analysis on SO(3) and $\mathbb{S}^2$ which will be necessary throughout the paper. In Section 3 we define the test hypotheses and the smoothness assumptions about the unknown density f and the noise ε_i. We also introduce the adaptive goodness-of-fit testing procedure. In Sections 4 and 5 we compute lower and upper bounds for testing rates for the ordinary smooth noise case. The super smooth noise case is treated in Section 6. Finally, we give a simulation study in Section 7. The proofs of the results are detailed in Section 8.

2 Some preliminaries about harmonic analysis on SO(3) and $\mathbb{S}^2$

We will provide a brief overview of Fourier analysis on SO(3) and $\mathbb{S}^2$. Most of the material can be found in expanded form in Vilenkin (1968), Talman (1968), Terras (1985), Kim and Koo (2002) and Healy et al. (1998).

Let D^l_{mn} for $-l \leq m,\ n \leq l,\ l = 0,\ 1, \ldots$ be the eigenfunctions of the Laplace Beltrami operator on SO(3), hence, $\sqrt{2l+1}D^l_{mn},\ -l \leq m,\ n \leq l,\ l = 0,\ 1, \ldots$ is a complete orthonormal basis for $\mathbb{L}_2(\mathrm{SO}(3))$ with respect to the probability Haar measure. Explicit formulae of the rotational harmonics D^l_{mn} in term of Euler angles exist but we do not need it here. Next, for $f \in \mathbb{L}_2(\mathrm{SO}(3))$, we define the rotational Fourier transform on SO(3) by the $(2l+1) \times (2l+1)$ matrices $f^{\star l}$ with entries

$$f^{\star l}_{mn} = \int_{\mathrm{SO}(3)} f(g) D^l_{mn}(g) dg,$$

where dg is the probability Haar measure on SO(3). The rotational inversion can be obtained by

$$f(g) = \sum_l \sum_{-l \leq m,\ n \leq l} f^{\star l}_{mn} (2l+1) \overline{D^l_{mn}(g)}. \tag{2}$$

(2) is to be understood in $\mathbb{L}_2$-sense although with additional smoothness conditions, it can hold pointwise.

A parallel spherical Fourier analysis is available on $\mathbb{S}^2$. Any point on $\mathbb{S}^2$ can be represented by

$$\omega = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta)^t,$$

with $\phi \in [0, 2\pi),\ \theta \in [0, \pi)$. We also define the functions:

$$Y^l_m(\omega) = Y^l_m(\theta, \phi) = \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P^l_m(\cos\theta) e^{im\phi},$$

for $-l \leq m\ \leq l,\ l = 0,\ 1, \ldots,\ \phi \in [0, 2\pi),\ \theta \in [0, \pi)$ and where P^l_m are the associated Legendre functions. The functions Y^l_m obey

$$Y^l_{-m}(\theta, \phi) = (-1)^m \overline{Y^l_m(\theta, \phi)}. \tag{3}$$

Let $\mathbb{L}_2(\mathbb{S}^2)$ denote the space of square integrable functions on $\mathbb{S}^2$, that is, the set of measurable functions f on $\mathbb{S}^2$ for which

$$\|f\|_2 = \left(\int_{\mathbb{S}^2} |f(x)|^2 dx \right)^{\frac{1}{2}} < \infty,$$

where dx is the Lebesgue measure on the sphere $\mathbb{S}^2$. It is well-known that $\mathbb{L}_2(\mathbb{S}^2)$ is a Hilbert space with the inner product

$$\langle f, g \rangle_{\mathbb{L}_2} = \int_{\mathbb{S}^2} f(x) \overline{g(x)} dx, \quad f, g \in \mathbb{L}_2(\mathbb{S}^2).$$

The set $\{Y^l_m,\ -l \leq m\ \leq l,\ l = 0,\ 1, \ldots\}$ is forming an orthonormal basis of $\mathbb{L}_2(\mathbb{S}^2)$, generally referred to as the spherical harmonic basis. Again, as above, for $f \in \mathbb{L}_2(\mathbb{S}^2)$, we define the spherical Fourier transform on $\mathbb{S}^2$ by

$$f_m^{\star l} = \int_{\mathbb{S}^2} f(x)\overline{Y_m^l(x)}dx. \tag{4}$$

We think of (4) as the vector entries of the $(2l+1)$ vector

$$f^{\star l} = [f_m^{\star l}]_{-l\leq m\leq l},\ l = 0,\ 1, \ldots$$

The spherical inversion can be obtained by

$$f(\omega) = \sum_l \sum_{-l\leq m\leq l} f_m^{\star l} Y_m^l(\omega).$$

The bases detailed above are important because they realize a singular value decomposition of the convolution operator created by our model. In effect, we define for $f_\varepsilon \in \mathbb{L}_2(\mathrm{SO}(3)),\ f \in \mathbb{L}_2(\mathbb{S}^2)$ the convolution by the following formula:

$$f_\varepsilon * f(\omega) = \int_{\mathrm{SO}(3)} f_\varepsilon(u) f(u^{-1}\omega)du \tag{5}$$

and we have for all $-l \leq m\ \leq l,\ l = 0,\ 1, \ldots,$

$$(f_\varepsilon * f)_m^{\star l} = \sum_{n=-l}^{l} f_{\varepsilon,mn}^{\star l} f_n^{\star l} = (f_\varepsilon^{\star l} f^{\star l})_m. \tag{6}$$

We shall recall some basic facts which will be useful throughout the paper. Let $\mathbb{H}_l$ the vector space spanned by $\{Y_m^l = -l \leq m \leq l\}$ for each $l = 0, 1, \ldots$. Any element $h \in \mathbb{H}_l$ can be written as $h = \sum_{m=-l}^{l} h_m^{\star l} Y_m^l$ and thanks to Parseval equality we have $\|h\|_2^2 = \sum_{m=-l}^{l} |h_m^{\star l}|^2$. Now according to (6) we have

$$f_\varepsilon^{\star l} : \mathbb{H}_l \to \mathbb{H}_l \text{ defined by } f_\varepsilon^{\star l} h = \sum_{m=-l}^{l} \left(\sum_{n=-l}^{l} (f_\varepsilon^{\star l})_{mn} h_n^{\star l} \right) Y_m^l.$$

We finally get the operator inequality

$$\|f_\varepsilon^{\star l} h\|_2 \leq \|f_\varepsilon^{\star l}\|_{op} \|h\|_2, \quad \text{where} \quad \|f_\varepsilon^{\star l}\|_{op} = \sup_{h\neq 0, h\in\mathbb{H}_l} \frac{\|f_\varepsilon^{\star l} h\|_2}{\|h\|_2}.$$

3 Model and assumptions

We would like to present our results in terms of Sobolev classes (see e.g. Healy et al. (1998) for a definition on the sphere). On the space $C^\infty(\mathbb{S}^2)$ of infinitely continuous differentiable functions on $\mathbb{S}^2$, consider the so-called Sobolev norm $\|\|_{W_s}$ of order s defined in the following way. For any function $f = \sum_{lm} f_m^{\star l} Y_m^l$ let

$$\|f\|_{W_s}^2 = \sum_{l\geq 0} \sum_{m=-l}^{l} (1 + l(l+1))^s |f_m^{\star l}|^2. \tag{7}$$

We denote by $W_s(\mathbb{S}^2)$ the vector space completion of $C^\infty(\mathbb{S}^2)$ with respect to the Sobolev norm (7) of order s. For some fixed constant $R > 0$, let $W_s(\mathbb{S}^2, R)$ denote the smoothness class of functions $f \in W_s(\mathbb{S}^2)$ which satisfy

$$\|f\|_{W_s} \leq R. \tag{8}$$

For the uniform density of probability on the sphere namely $f_0 = (4\pi)^{-1}\mathbb{1}_{\mathbb{S}^2}$, we want to test the hypothesis

$$H_0 : \quad f = f_0,$$

from observations $Z_1, \ldots, Z_N$ given by model (1). We consider the alternative

$$H_1(s, R, \mathcal{C}\psi_N) : f \in W_s(\mathbb{S}^2, R) \text{ and } \|f - f_0\|_2^2 \geq \mathcal{C}\psi_N$$

where $\mathcal{C}$ is a constant and ψ_N is the testing rate.

We will say that the distribution of ε is ordinary smooth of order ν if the rotational Fourier transform of f_ε satisfies the following assumption.

Assumption 1. *For all $l \geq 0$, the matrix $f_\varepsilon^{\star l}$ is invertible and there exist positive constants d_0, d_1, ν such that*

$$\|f_{\varepsilon^{-1}}^{\star l}\|_{op} \leq d_0^{-1} l^\nu \quad \text{and} \quad \|f_\varepsilon^{\star l}\|_{op} \leq d_1 l^{-\nu},$$

where we have denoted the matrix $(f_\varepsilon^{\star l})^{-1}$ by $f_{\varepsilon^{-1}}^{\star l}$.

Recall that we assume that f_ε is known, consequently d_0 and ν are also considered known. Some examples satisfying this assumption are given in Section 7.

In order to build a test statistic, as usual, we first have to construct an unbiased estimator of the quadratic functional $\int_{\mathbb{S}^2}(f - f_0)^2 = \|f - f_0\|_2^2$. To do so, we remark that thanks to Parseval equality:

$$\int_{\mathbb{S}^2}(f - f_0)^2 = \sum_{l \geq 0}\sum_{m=-l}^{l} |f_m^{\star l} - {f_0}_m^{\star l}|^2 = \sum_{l \geq 1}\sum_{m=-l}^{l} |f_m^{\star l}|^2,$$

the last equality coming from the fact that $(f_0)_m^{\star l} \neq 0$ only for $(l, m) = (0, 0)$. Since $f^{\star l} = f_{\varepsilon^{-1}}^{\star l} f_Z^{\star l}$ for $l = 0, 1, \ldots$, we can write under Assumption 1

$$f_m^{\star l} = \sum_{n=-l}^{l} (f_{\varepsilon^{-1}}^{\star l})_{mn} (f_Z^{\star l})_n.$$

A natural estimator of $f_m^{\star l}$ is given by

$$\hat{f}_m^{\star l} = \frac{1}{N}\sum_{i=1}^{N}\sum_{n=-l}^{l} (f_{\varepsilon^{-1}}^{\star l})_{mn} \overline{Y_n^l(Z_i)}.$$

If we denote by $\Phi_{lm}(x) = \sum_{n=-l}^{l} (f_{\varepsilon^{-1}}^{\star l})_{mn} \overline{Y_n^l}(x)$ then

$$\hat{f}_m^{\star l} = \frac{1}{N}\sum_{i=1}^{N} \Phi_{lm}(Z_i).$$

Consequently, we can derive an unbiased estimator T_{lm} of $|f_m^{\star l}|^2$

$$T_{lm} = \frac{2}{N(N-1)} \sum_{i_1<i_2} \Phi_{lm}(Z_{i_1})\overline{\Phi_{lm}(Z_{i_2})},$$

and finally an estimator of $\|f - f_0\|_2^2$

$$T_L = \sum_{l=1}^{L} \sum_{m=-l}^{l} \frac{2}{N(N-1)} \sum_{i_1<i_2} \Phi_{lm}(Z_{i_1})\overline{\Phi_{lm}(Z_{i_2})}.$$

We can now define a test procedure

$$\Delta = \begin{cases} 1 & \text{if } |T_L| > t^2 \\ 0 & \text{otherwise} \end{cases}$$

for a threshold t^2 to be suitably chosen. The choice of L is crucial too, ans this point will be solved in Sections 5 & 6.

As one may have noticed, the noise smoothness hypothesis and hence the test procedure only rely on the Fourier transform of the noise density f_ε. Consequently, we do not need the existence of the density f_ε but only its Fourier transform.

4 Lower bound for testing rate

It is known that the rate of separation in the case of direct observations in dimension two is $N^{-4s/4s+2}$. Let us see how it is modified by the presence of a noise with smoothness ν.

Theorem 1. *Assume that $R > 1/\sqrt{2\pi}$. Let $s \geq 1$ and $\psi_N = N^{-2s/(2s+2\nu+1)}$. Let $\eta \in (0,1)$. If $\mathcal{C} \leq KR^2$ where K is a constant only depending on d_0, d_1, ν, s, η, then*

$$\liminf_{N\to\infty} \inf_{\Delta_N} \left\{ \mathbb{P}_{f_0}(\Delta_N = 1) + \sup_{f \in H_1(s,R,\mathcal{C}\psi_N)} P_f(\Delta_N = 0) \right\} \geq \eta$$

where the infimum is taken over all test procedures Δ_N based on the observations $Z_1, \ldots, Z_N$.

This means that testing with a faster rate than $\psi_N = N^{-2s/(2s+2\nu+1)}$ is impossible. If the distance beetween f_0 and the alternative is smaller than $\psi_N = N^{-2s/(2s+2\nu+1)}$, the sum of the error of the two kinds is close to 1. Nevertheless, it requires the knowledge of the smoothness index s. That is why we want to build on a so-called adaptive test procedure which does not depend on s. But we prove in the next statement that we have to face a phenomenon of "lack of adaptability" for our problem, i.e. it is not possible to test adaptively with the same rate. Indeed, in the context where s is unknown and belongs to some set $\mathcal{S}$, there is not any universal test with small error for each $s \in \mathcal{S}$. The price to pay for adaptivity is an extra factor $\sqrt{\log\log N}$ in the separation rate.

Theorem 2. *Assume that $R > 1/\sqrt{2\pi}$. For all $s \geq 1$, let $\psi_N^{ad}(s) = (N/\sqrt{\log\log(N)})^{-2s/(2s+2\nu+1)}$. Let $\eta \in (0,1)$. If $\mathcal{C} \leq KR^2$ where K is a constant only depending on d_0, d_1, ν, s, η, then, if $\mathcal{S}$ is a set containing an interval,*

$$\liminf_{N\to\infty} \inf_{\Delta_N} \left\{ \mathbb{P}_{f_0}(\Delta_N = 1) + \sup_{s\in\mathcal{S}} \sup_{f\in H_1(s,R,\mathcal{C}\psi_N^{ad}(s))} P_f(\Delta_N = 0) \right\} \geq \eta$$

where the infimum is taken over all test procedures Δ_N based on the observations $Z_1, \ldots, Z_N$. Moreover any rate faster than ψ_N^{ad} will also lead to a lower-bounded error.

5 Upper bound for testing rate

In order to construct an adaptive procedure of testing, we shall use the following exponential inequality.

Lemma 1. *There exists K_0, K_1 such that, for all sequence u_N,*

$$\mathbb{P}_0(|T_L| > L^{2\nu+1}u_N/N) \leq K_1 \exp(-K_0 u_N^2)$$

provided that $u_N L^{-1}$, $LN^{-2}u_N^8$ and $LN^{-1}u_N^3$ are bounded.

Actually the term $L^{2\nu+1}/N$ is the order of the variance of T_L under H_0. We denote $\lceil x \rceil$ the smallest integer larger than or equal to x.

Theorem 3. *Assume $s \geq 1$ and $\psi_N = (N/\sqrt{\log\log N})^{-2s/(2s+2\nu+1)}$. We consider the set $\mathcal{L} = \{2^{j_0}, \ldots, 2^{j_m}\}$ where $j_0 = \lceil \log_2(\log\log N) \rceil$, $j_m = \lceil \log_2(N(\log\log N)^{-3/2}) \rceil$ and the adaptive test statistic*

$$D_N = \mathbb{1}_{\{\max_{L\in\mathcal{L}}(|T_L|/t_L^2) > \sqrt{2/K_0}\}}$$

with $t_L^2 = L^{2\nu+1}\sqrt{\log\log N}/N$. Let $0 < \eta < 1$. Then, if $\mathcal{C} > \sqrt{2K_0^{-1}} + R^2$,

$$\limsup_{N\to\infty} \left\{ \mathbb{P}_{f_0}(D_N = 1) + \sup_{f\in H_1(s,R,\mathcal{C}\psi_N)} \mathbb{P}_f(D_N = 0) \right\} \leq \eta.$$

Note that the direct case (without noise) is included in this result, taking $\varepsilon = Id$, $f_\varepsilon^{\star l} = Id$, $\nu = 0$. In this case, the separation rate is $(N/\sqrt{\log\log N})^{-2s/(2s+1)}$. To our knowledge, even in this simpler case, this result was not established yet.

6 Super smooth noise

In this section, we deal with the case of a super smooth noise. This kind of noise is of interest since it includes the Gaussian distribution. We will say that the distribution of ε is super smooth of order ν if the rotational Fourier transform of f_ε satisfies

Assumption 2. *For all $l \geq 0$, the matrix $f_\varepsilon^{\star l}$ is invertible and there exist reals $\nu_1 \leq \nu_0$, and positive constants d_0, d_1, δ, β such that*

$$\|f_{\varepsilon^{-1}}^{\star l}\|_{op} \leq d_0^{-1} l^{-\nu_0} \exp(l^\beta/\delta) \quad \text{and} \quad \|f_\varepsilon^{\star l}\|_{op} \leq d_1 l^{\nu_1} \exp(-l^\beta/\delta).$$

In this case, we present a similar test statistic but with a different threshold t_L. Moreover it is sufficient to consider only one L^* instead of a maximum.

Theorem 4. *Let $\psi_N = (\log N)^{-2s/\beta}$ and $K_0 > 0$. We consider $L^* = \left\lfloor (\delta \log(N)/8)^{1/\beta} \right\rfloor$ and the test statistic*

$$D_N = \mathbb{1}_{\{|T_{L^*}|/t^2_{L^*} > K_0\}}$$

with $t_L^2 = L^{-2\nu_0+1} \exp(2L^\beta/\delta)/N$. Let $0 < \eta < 1$. Then, if $\mathcal{C} > K_0 + R^2(\delta/16)^{-2s/\beta}$,

$$\limsup_{N\to\infty} \left\{ \mathbb{P}_{f_0}(D_N = 1) + \sup_{f \in H_1(s,R,\mathcal{C}\psi_n)} \mathbb{P}_f(D_N = 0) \right\} \leq \eta.$$

We observe that in this case the separation rate is very slow $\psi_N = (\log N)^{-2s/\beta}$. However, this rate is reached without any knowledge on the smoothness of f. Moreover, we prove that this is the optimal rate:

Theorem 5. *Assume that $R > 1/\sqrt{2\pi}$. Let $s \geq 1/2$ and $\psi_N = (\log N)^{-2s/\beta}$. Let $\eta \in (0,1)$. If $\mathcal{C} \leq KR^2$ where K is a constant only depending on $d_0, d_1, \nu_0, \beta, \delta, s, \eta$, then*

$$\liminf_{N\to\infty} \inf_{\Delta_N} \left\{ \mathbb{P}_{f_0}(\Delta_N = 1) + \sup_{f \in H_1(s,R,\mathcal{C}\psi_N(s))} P_f(\Delta_N = 0) \right\} \geq \eta$$

where the infimum is taken over all test procedures Δ_N based on the observations $Z_1, \ldots, Z_N$.

The deterioration of the rate in the case of a super smooth noise is a well-known phenomenon in convolution models (see e.g. Fan (1991)).

7 Simulations

In this section we implement our test procedure for two kinds of noise. First we will present the noises on SO(3) we consider, then we explain the protocol we are going to follow and eventually we give the performances of our goodness-of-fit test in term of power.

In our theoretical statements, we talked about ordinary and super smooth noises on the group SO(3), but what does it mean in practice? Does it exist concrete examples of random matrices which could be generated according to densities which meet those smoothness assumptions? The answer is yes. We will particularly highlight two cases, the Rotational Laplace and the Gaussian densities on SO(3) (for further details see Kim and Koo (2002)). To the best of our knowledge, they have never been implemented in practice. The first one is an ordinary smooth density and the second one a super smooth one.

As explained in section 3, the noise smoothness can be characterized by the decay of its rotational Fourier transform. The Rotational Laplace distribution is the rotational analogue of the well-known Euclidean Laplace distribution (known also as double exponential distribution). It has been discussed in depth in Healy et al. (1998). Its expanded form in term of rotational harmonics is the following

$$f_\varepsilon = \sum_{l\geq 0} \sum_{m=-l}^{l} (1 + \sigma^2 l(l+1))^{-1} (2l+1) \overline{D^l_{mm}}, \tag{9}$$

for some $\sigma^2 > 0$ which is a variance parameter. Hence we have

$$f^{\star l}_{\varepsilon,mn} = (1 + \sigma^2 l(l+1))^{-1}\delta_{mn},$$

for $l = 0, 1, \ldots$ and where $\delta_{mn} = 1$ if $m = n$ and is 0 otherwise. The Laplace distribution is ordinary smooth with a smoothness index $\nu = 2$.

Let us present now the Gaussian distribution. The distribution can be written as follows (see Kim and Koo (2002))

$$f_\varepsilon = \sum_{l\geq 0}\sum_{m=-l}^{l} \exp(-\sigma^2 l(l+1)/2)(2l+1)\overline{D^l_{mm}}, \tag{10}$$

for $\sigma > 0$. We have

$$f^{\star l}_{\varepsilon,mn} = \exp(-\sigma^2 l(l+1)/2)\delta_{mn}.$$

This is an example of a super smooth distribution with $\delta = 2/\sigma^2$ and $\beta = 2$ following the terminology in Section 6.

Here is our testing procedure steps.

1. Under H_0 generate N observations uniformly.
2. With these N uniform observations, compute by 1000 Monte Carlo runs the 10% quantile of the statistics $\max_{L\in\mathcal{L}}(|T_L|/t_L^2)$ defined in Theorem 3. We will denote it q_{10}.

At this stage it is possible to apply the test to any noised data. Here we study the performance of the test with the following next steps.

3. Generate N random rotation matrices following the Laplace or the Gaussian distribution and apply them to N observations generated according to the alternative density g in H_1 (see Figure 1).
4. Compute the test power given by $\mathbb{P}_g(\max_{L\in\mathcal{L}}(|T_L|/t_L^2) > q_{10})$ by 200 Monte Carlo runs.

We point out that the numerical procedure is notably fast all the more so as we are in dimension 2.

The alternative density g is represented in spherical coordinates on Figure 1. It has two modes, one very sharp and the other one oblate.

For more clarity about how the noise matrices are involved in the problem, we represent 500 observations generated under the alternative density g with and without noise, see Figure 2. To plot the figure, we consider the Laplace distribution noise with a variance parameter $\sigma^2 = 0.5$. It turned out that applying random rotations tends to spread uniformly the observations which makes the separation between the null and the alternative hypothesis a difficult task.

We would also like to make a remark about how to generate random matrices according to the Laplace or the Gaussian distribution. After rewriting carefully their density expressions in terms of rotational harmonics given by (9) and (10), it turned out that $f_\varepsilon(u)$ only depends on the angle of the rotation u, say θ. Then the simulation of a rotation following f_ε amounts to pick at random an axis and perform a rotation about this axis by an angle following the law $f_\varepsilon(\theta)(1-\cos(\theta))/\pi$.

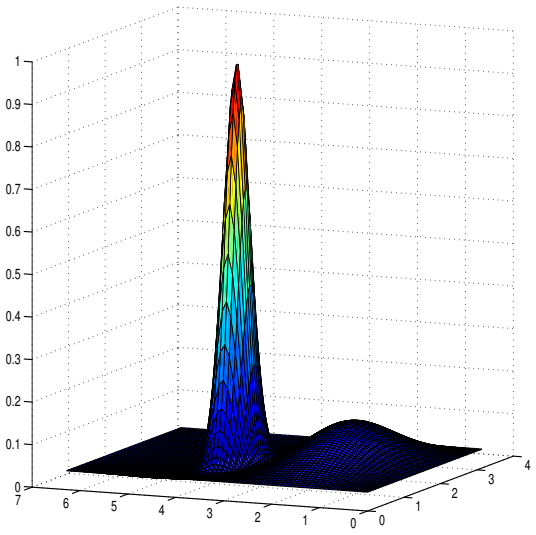

Figure 1: The alternative density

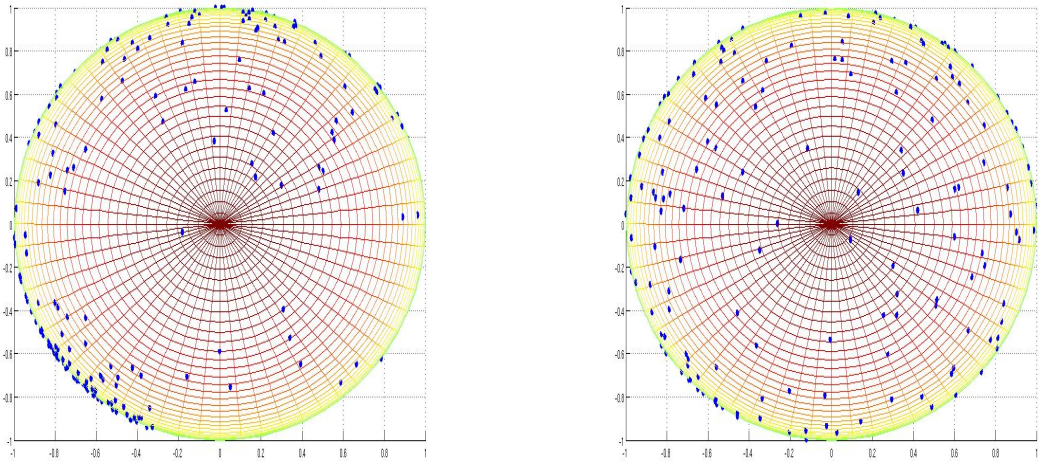

Figure 2: Observations under H_1 without and with noise

For more conciseness, we do not present the empirical level of the test which is whatever the noise, the number of observations, between 6% and 12%. Indeed, the procedure is devised in such a way that we have a level around 10%. Let us focus now on the test power which is more informative. We computed it for several number of observations going from 50 only to 500 and give our results in Table 1. We point out that we deliberately considered the cases of very few observations such as 50 and 100 because these scenari are realistic in astrophysics when collecting UHECR observations. Our test procedure performs pretty well with excellent results if we have 500 observations. For reasonable values of the variance parameters such as 0.1, the procedure yields reasonable results even for 50 or 100 observations only.

Table 1: Test Power

Noise distribution	$N = 50$	$N = 100$	$N = 250$	$N = 500$
Laplace, $\sigma^2 = 0.1$	0.59	0.93	1	1
$\sigma^2 = 0.25$	0.37	0.68	0.99	1
$\sigma^2 = 0.5$	0.20	0.30	0.63	0.96
Gaussian, $\sigma^2 = 0.1$	0.42	0.81	0.99	1
$\sigma^2 = 0.25$	0.30	0.35	0.77	0.98
$\sigma^2 = 0.5$	0.16	0.33	0.50	0.83

8 Proofs

8.1 Proof of Theorem 1

As usual in the proofs of lower bounds, we build a set of functions quite far from f_0 in term of $\mathbb{L}_2$ norm, but whose distance between the resulting models is small. More precisely, let $\gamma = \gamma(N)$ and $L = L(N)$, respectively a scale factor and a resolution level to be specified later. For all $l < L$ and $m \in \{-l, \dots, l\}$, we define φ_{lm} the function such that

$$f_\varepsilon * \varphi_{lm} = Y_m^l.$$

Here Y_m^l denotes the real form of the spherical harmonic, that we denote as the complex form for the sake of simplicity. Using the real form ensures that φ_{lm} is real. The existence of such a function is ensured by the assumption of invertibility of matrices $f_\varepsilon^{\star l}$ and we can write $\varphi_{lm} = \sum_{n=-l}^{l} (f_{\varepsilon^{-1}}^{\star l})_{nm} Y_n^l$. Now, for θ_{lm}, $l < L$, $m \in \{-l, \dots, l\}$, independent random variables with distribution $\mathbb{P}(\theta_{lm} = \pm\gamma) = 1/2$, we introduce

$$f_\theta = f_0 + \sum_{l=L/2}^{L-1} \sum_{m=-l}^{l} \theta_{lm} \varphi_{lm}.$$

In the sequel we show that, for good choices of γ and L,

- f_θ belongs to $W_s(\mathbb{S}^2, R)$,
- f_θ is a density function,
- $\|f_\theta - f_0\|^2 \geq \mathcal{C}\psi_N$,
- $\chi^2(P_\theta, P_{f_0}) \leq (1-\eta)^2$ for N large enough.

Then, if T is the maximum likelihood test which is equal to 1 if $\prod_{i=1}^N f_\theta(X_i)/f_0(X_i) \geq 1$ and 0 otherwise,

$$P_{f_0}(T=1) + P_\theta(T=0) \geq \int \min(dP_\theta, dP_{f_0}) \geq 1 - \sqrt{\chi^2(dP_\theta, dP_{f_0})} \geq 1 - (1-\eta) \tag{11}$$

using section 2.4.1 in Tsybakov (2009). Thus, for N large enough,

$$P_{f_0}(T=1) + \sup_{f \in H_1(s,R,\mathcal{C}\psi_N)} P_f(T=0) \geq \eta$$

and the result is proved.

•*Belonging to the Sobolev space:*
We compute

$$\begin{aligned}
\sum_{ln}(1+l(l+1))^s|<f_\theta,Y_n^l>|^2 &= |<f_\theta,Y_0^0>|^2+\sum_{l=L/2}^{L-1}\sum_n(1+l(l+1))^s|<f_\theta,Y_n^l>|^2\\
&\leq |<f_0,Y_0^0>|^2+\sum_{l=L/2}^{L-1}(1+l(l+1))^s\sum_n|\sum_{m=-l}^{l}\theta_{lm}(f_{\varepsilon^{-1}}^{\star l})_{nm}|^2\\
&\leq \frac{1}{4\pi}+\sum_{l=L/2}^{L-1}(1+l(l+1))^s\|f_{\varepsilon^{-1}}^{\star l}\|_{op}^2\sum_{m=-l}^{l}|\theta_{lm}|^2\\
&\leq \frac{1}{4\pi}+d_0^{-2}\gamma^2\sum_{l=L/2}^{L-1}(1+l(l+1))^s l^{2\nu}(2l+1)\\
&\leq \frac{R^2}{2}+C_1(d_0,s,\nu)\gamma^2L^{2s+2\nu+2}.
\end{aligned}$$

Hence, belonging to the Sobolev ball imposes that

$$\gamma^2L^{2(s+\nu+1)}\leq\frac{R^2}{2C_1(d_0,s,\nu)}.$$

Then it is sufficient to choose $\gamma^2=c_1L^{-2(\nu+s+1)}$ with $c_1\leq R^2/(2C_1(d_0,s,\nu))$.

•*Density:*
Since, for $l>0$,

$$\frac{1}{\sqrt{4\pi}}\int\varphi_{lm}=(\varphi_{lm})_0^{\star 0}=0$$

we obtain $\int f_\theta=1$. Let us now show that $f_\theta\geq 0$. We first use the Cauchy-Schwarz inequality

$$|f_\theta(x)-f_0(x)|^2 \leq \sum_{l=L/2}^{L-1}\sum_n|\langle f_\theta-f_0,Y_n^l\rangle|^2\sum_{l=L/2}^{L-1}\sum_n|Y_n^l(x)|^2$$

Next, since spherical harmonics have the property $\sum_n|Y_n^l|^2\leq(2l+1)/(4\pi)$,

$$\begin{aligned}
|f_\theta(x)-f_0(x)|^2 &\leq \sum_{l=L/2}^{L-1}\sum_n\left|\sum_{m=-l}^{l}\theta_{lm}(f_{\varepsilon^{-1}}^{\star l})_{nm}\right|^2\sum_{l=L/2}^{L-1}\frac{2l+1}{4\pi}\\
&\leq \frac{3}{8\pi}\sum_{l=L/2}^{L-1}\|f_{\varepsilon^{-1}}^{\star l}\|_{op}^2\sum_{m=-l}^{l}|\theta_{lm}|^2L^2\\
&\leq \frac{9d_0^{-2}}{8\pi}\gamma^2\sum_{l=L/2}^{L-1}l^{2\nu+1}L^2\leq C_2^2\gamma^2L^{2\nu+4}
\end{aligned}$$

with $C_2^2 = 9d_0^{-2}/(8\pi(2\nu+2))$. Thus, replacing γ by its value,

$$f_\theta \geq f_0 - \|f_\theta - f_0\|_\infty \geq \frac{1}{4\pi} - C_2\gamma L^{\nu+2} \geq \frac{1}{4\pi} - C_2\sqrt{c_1}L^{1-s}.$$

Now, since $s \geq 1$, f_θ is a density as soon as

$$\frac{1}{4\pi} - C_2\sqrt{c_1} \geq 0 \Leftrightarrow c_1 \leq \frac{\nu+1}{9\pi d_0^{-2}}.$$

•*Separation rate:*
We denote $p_\theta = f_\varepsilon * f_\theta = f_0 + \sum_{l=L/2}^{L-1}\sum_m \theta_{lm}Y_m^l$. Then

$$\begin{aligned}\|p_\theta - f_0\|^2 &= \|f_\varepsilon * (f_\theta - f_0)\|^2 = \sum_{l=L/2}^{L-1}\sum_{m=-l}^{l} |(f_\varepsilon^{\star l}(f_\theta - f_0)^{\star l})_m|^2 \\ &\leq \sum_{l=L/2}^{L-1} \|f_\varepsilon^{\star l}\|_{op}^2 \sum_{m=-l}^{l} |(f_\theta - f_0)_m^{\star l}|^2 \leq d_1^2(L/2)^{-2\nu}\|f_\theta - f_0\|^2.\end{aligned}$$

Moreover

$$\|p_\theta - f_0\|^2 = \sum_{l=L/2}^{L-1}\sum_m |\theta_{lm}|^2 = \gamma^2 \sum_{l=L/2}^{L-1}(2l+1) \geq C\gamma^2 L^2.$$

Finally $\|f_\theta - f_0\|^2 \geq C_3(d_1,\nu)\gamma^2 L^{2\nu+2} = C_3(d_1,\nu)c_1 L^{-2s}$. Now we choose

$$L = \left\lfloor N^{1/(2s+2\nu+1)} \right\rfloor$$

where $\lfloor x \rfloor$ denotes the largest integer which is smaller than or equal to x. Thus $\|f_\theta - f_0\|^2 \geq C_3(d_1,\nu)c_1\psi_N \geq \mathcal{C}\psi_N$ as soon as $\mathcal{C} \leq C_3(d_1,\nu)c_1$.

•*Chi-square divergence:*
We denote μ the measure defined by $d\mu(\theta) = \prod_{lm}(\delta_1 + \delta_{-1})/2$. We want to show that

$$\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}} - 1\right)^2\right) \leq (1-\eta)^2$$

where

$$\frac{dP_\mu}{dP_{f_0}} = \int \prod_{i=1}^N \frac{p_\theta(Z_i)}{p_0(Z_i)}\mu(d\theta) = \mathbb{E}_\mu \prod_{i=1}^N 4\pi p_\theta(Z_i).$$

First, note that $\mathbb{E}_{f_0}(4\pi p_\theta(Z_1)) = 1 + \sum_{lm} 4\pi\theta_{lm}\int Y_m^l = 1$. Then, using Fubini and the independence of the Z_i, $\mathbb{E}_{f_0}(\frac{dP_\mu}{dP_{f_0}}) = \mathbb{E}_\mu \prod_{i=1}^n \mathbb{E}_{f_0}(4\pi p_\theta(Z_i)) = 1$. So it is sufficient to prove that

$$\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) \leq 1 + (1-\eta)^2.$$

Using Fubini,

$$\begin{aligned}\frac{1}{(4\pi)^{2N}}\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) &= \mathbb{E}_{f_0}\left(\mathbb{E}_{\mu\times\mu}\prod_{i=1}^N p_\theta(Z_i)\prod_{i=1}^N p_{\theta'}(Z_i)\right) = \mathbb{E}_{f_0}\left(\mathbb{E}_{\mu\times\mu}\prod_{i=1}^N p_\theta(Z_i)p_{\theta'}(Z_i)\right)\\ &= \mathbb{E}_{\mu\times\mu}\left(\mathbb{E}_{f_0}\prod_{i=1}^N p_\theta(Z_i)p_{\theta'}(Z_i)\right)\end{aligned}$$

where $d(\mu\times\mu)((\theta,\theta')) = d\mu(\theta)d\mu(\theta')$. Now, the Z_i being independent, we write

$$\mathbb{E}_{f_0}\left(\prod_{i=1}^N p_\theta(Z_i)p_{\theta'}(Z_i)\right) = \prod_{i=1}^N \mathbb{E}_{f_0}(p_\theta(Z_i)p_{\theta'}(Z_i)) = \left(\int p_\theta p_{\theta'}\frac{1}{4\pi}\right)^N. \tag{12}$$

Using the definition of p_θ and the orthogonality of the Y_m^l, we obtain

$$\int p_\theta p_{\theta'} = \frac{1}{4\pi} + \sum_{l=L/2}^{L-1}\sum_{m=-l}^{l}\theta_{lm}\theta'_{lm}. \tag{13}$$

Combining (12) and (13), and then inequality $1+a\le e^a$, gives

$$\begin{aligned}\mathbb{E}_{f_0}\left(\prod_{i=1}^N p_\theta(Z_i)p_{\theta'}(Z_i)\right) &= \left(\frac{1}{(4\pi)^2} + \frac{1}{4\pi}\sum_{lm}\theta_{lm}\theta'_{lm}\right)^N \le \frac{1}{(4\pi)^{2N}}\exp(N4\pi\sum_{lm}\theta_{lm}\theta'_{lm})\\ &\le \frac{1}{(4\pi)^{2N}}\prod_{lm}\exp(N4\pi\theta_{lm}\theta'_{lm})\end{aligned}$$

so that

$$\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) \le \mathbb{E}_{\mu\times\mu}\prod_{lm}\exp(N4\pi\theta_{lm}\theta'_{lm}).$$

Using the distribution of (θ,θ'), we obtain

$$\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) \le \prod_{lm}\frac{1}{2}\exp(N\gamma^2 4\pi) + \frac{1}{2}\exp(-N\gamma^2 4\pi) \le \prod_{lm}\cosh(N\gamma^2 4\pi).$$

Now, using $\cosh(2x) = 1+2\sinh^2(x)$ and inequality $1+a\le e^a$,

$$\begin{aligned}\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) &\le \prod_{lm}(1+2\sinh^2(N\gamma^2 2\pi)) \le \prod_{lm}\exp(2\sinh^2(N\gamma^2 2\pi))\\ &\le \exp(2\sum_{lm}\sinh^2(N\gamma^2 2\pi)).\end{aligned}$$

Since $N\gamma^2\to 0$ and $\sinh(x) = x+o(x)$, there exists $C_4>0$ such that, for N large enough,

$$\sinh^2(N\gamma^2 2\pi) \le C_4 N^2\gamma^4.$$

Then

$$\sum_{lm} \sinh^2(N\gamma^2 2\pi) \leq C_4 N^2\gamma^4 \sum_{lm} 1 \leq C_5 N^2\gamma^4 L^2.$$

That yields, for N large enough,

$$\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) \leq \exp(2C_5 N^2\gamma^4 L^2) \leq \exp(2C_5 c_1^2 N^2 L^{-4s-4\nu-2}).$$

But remember that $N^2 < (L+1)^{4s+4\nu+2}$ so that $\mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) \leq \exp(C_6(s,\nu)c_1^2) \leq 1+(1-\eta)^2$ for a good choice of c_1.

8.2 Proof of Theorem 2

We follow the same proof as the one of Theorem 1 but with this time a random L. Let $k_N = \lfloor (\log N)^\delta \rfloor$. We choose k_N elements of $\mathcal{S}$: $s_1 < \cdots < s_{k_N}$ such that $s_{j+1} - s_j \geq \log(N)^{-1}$. Then it is possible to find $J_1 > \cdots > J_{k_N} > 1$ such that for all $1 \leq j \leq k_N$, $N/\sqrt{\log(k_N)} = c_2 2^{J_j(2\nu+2s_j+1)}$. We also define $\gamma_j^2 = c_1 2^{-2J_j(\nu+s_j+1)}$. We consider hypothesis functions

$$f_\theta = f_0 + \sum_L \sum_{l=L/2}^{L-1} \sum_{m=-l}^{l} \theta_{Llm}\varphi_{lm}$$

and we take a prior of the form $\mu = {k_N}^{-1}\sum_{j=1}^{k_N}\mu_j$. Then θ is randomly chosen such that $\mu_j(\theta_{lm} = \pm\gamma_j) = 1/2$ if $L = 2^{J_j}, 2^{J_j-1} \leq l < 2^{J_j}$, $-l \leq m \leq l$ and $\mu_j(\theta_{lm} = 0) = 1$ otherwise. This means that L is fixed equal to 2^{J_j} with probability $1/k_N$ and random densities with respect to the measures μ_j have the following form

$$f_\theta = f_0 + \sum_{l=L/2}^{L-1} \sum_{m=-l}^{l} \theta_{lm}\varphi_{lm}$$

where $\mu_j(\theta_{lm} = \pm\gamma_j) = 1/2$.

Given the proof of Theorem 1, we easily verify that μ_j-a.s. $f_\theta \in H_1(s_j, R, \mathcal{C}\psi_N^{ad}(s_j))$ if c_1 is chosen small enough. Now, since $\sup_{s\in\mathcal{S}} \sup_{f\in H_1(s,R,\mathcal{C}\psi_N^{ad}(s))} P_f(\Delta_N = 0) \geq P_\mu(\Delta_N = 0)$ and according to (11), it is sufficient to bound the χ^2-divergence. So we will show that

$$\limsup_N \mathbb{E}_{f_0}\left(\left(\frac{dP_\mu}{dP_{f_0}}\right)^2\right) \leq 1 + (1-\eta)^2$$

which comes back to

$$\limsup_N \frac{1}{k_N^2}\sum_{p,q=1}^{k_N} \mathbb{E}_{f_0}\left(\frac{dP_{\mu_p}}{dP_{f_0}}\frac{dP_{\mu_q}}{dP_{f_0}}\right) \leq 1 + (1-\eta)^2.$$

Using Fubini's Theorem and independence of the Z_i's,

$$\mathbb{E}_{f_0}\left(\frac{dP_{\mu_p}}{dP_{f_0}}\frac{dP_{\mu_q}}{dP_{f_0}}\right) = \mathbb{E}_{\mu_p\times\mu_q}\left(\mathbb{E}_{f_0}\prod_{i=1}^{N} p_\theta(Z_i)p_{\theta'}(Z_i)\right) = \mathbb{E}_{\mu_p\times\mu_q}\left(\left(\int 4\pi p_\theta p_{\theta'}\right)^N\right).$$

Denoting $a_l(L_1,L_2) = 4\pi \mathbb{1}_{L_1/2\le l<L_1}\mathbb{1}_{L_2/2\le l<L_2}$ we can write

$$\int 4\pi p_\theta p_{\theta'} = 1 + \sum_{L_1,L_2}\sum_{lm} \theta_{Llm}\theta'_{Llm}a_l(L_1,L_2).$$

Thus

$$\left(4\pi\int p_\theta p_{\theta'}\right)^N \le \exp\left(N\sum_{L_1L_2lm}\theta_{Llm}\theta'_{Llm}a_l(L_1,L_2)\right) \le \prod_{L_1L_2lm}\exp(N\theta_{Llm}\theta'_{Llm}a_l(L_1,L_2)).$$

Using the distribution of (θ,θ'), we obtain

$$\begin{aligned}\mathbb{E}_{f_0}\left(\frac{dP_{\mu_p}}{dP_{f_0}}\frac{dP_{\mu_q}}{dP_{f_0}}\right) &\le \prod_{lm}\frac{1}{2}\exp(N\gamma_p\gamma_q a_l(2^{J_p},2^{J_q})) + \frac{1}{2}\exp(-N\gamma_p\gamma_q a_l(2^{J_p},2^{J_q}))\\ &\le \prod_{lm}\cosh(N\gamma_p\gamma_q a_l(2^{J_p},2^{J_q})).\end{aligned}$$

Now, using $\cosh(2x) = 1+2\sinh^2(x)$, inequality $1+a\le e^a$ and $\sinh(x) = x+o(x)$

$$\mathbb{E}_{f_0}\left(\frac{dP_{\mu_p}}{dP_{f_0}}\frac{dP_{\mu_q}}{dP_{f_0}}\right) \le \exp(2\sum_{lm}\sinh^2(N\gamma_p\gamma_q a_l(2^{J_p},2^{J_q})/2)) \le \exp(C_1N^2\gamma_p^2\gamma_q^2\sum_{lm}|a_l(2^{J_p},2^{J_q})|^2).$$

We observe that $a_l(2^{J_p},2^{J_q}) = 0$ as soon as $J_p\neq J_q$. That yields

$$\mathbb{E}_{f_0}\left(\frac{dP_{\mu_p}}{dP_{f_0}}\frac{dP_{\mu_q}}{dP_{f_0}}\right) \le \exp(C_2N^2\gamma_p^4 2^{2J_p}\mathbb{1}_{p=q}).$$

Then

$$\frac{1}{k_N^2}\sum_{p,q=1}^{k_N}\mathbb{E}_{f_0}\left(\frac{dP_{\mu_p}}{dP_{f_0}}\frac{dP_{\mu_q}}{dP_{f_0}}\right) \le \frac{1}{k_N^2}\sum_{p=1}^{k_N}\exp(C_2N^2\gamma_p^4 2^{2J_p}) \le \frac{1}{k_N}\exp(C_2c_1^2c_2^2\log(k_N)) \le k_N^{C_2c_1^2c_2^2-1}$$

which is bounded by $1+(1-\eta)^2$ for N large enough if we choose c_1 small enough.

8.3 Proof of Lemma 1

We recall the result from Giné et al. (2000).

Lemma 2. *Let u a bounded canonical kernel, completely degenerate of the i.i.d. variables $Z_1,\dots,Z_N$. There exist universal constants $K_1,K_2>0$ such that, for all $x>0$,*

$$\mathbb{P}\left(\left|\sum_{1\le i_1\neq i_2\le N}u(Z_{i_1},Z_{i_2})\right|\ge x\right) \le K_1\exp\left(-K_2\min\left(\frac{x^2}{C^2},\frac{x}{D},\frac{x^{2/3}}{B^{2/3}},\frac{x^{1/2}}{A^{1/2}}\right)\right)$$

where A,B,C,D are defined by

$$A = \|u\|_\infty, \qquad B^2 = N\|E(|u|^2(Z,.)\|_\infty, \qquad C^2 = N^2\mathbb{E}[|u|^2(Z_1,Z_2)]$$

and

$$D = N\sup\left\{\left|\mathbb{E}[u(Z_1,Z_2)|u_1(Z_1)u_2(Z_2)]\right|, \mathbb{E}[u_1^2(Z)]\le 1, \mathbb{E}[u_2^2(Z)]\le 1\right\}.$$

We apply this Lemma to the kernel

$$u(x,y) = \sum_{l=1}^{L} \sum_{m=-l}^{l} \Phi_{lm}(x)\overline{\Phi_{lm}(y)}.$$

which is degenerate for Z_i under H_0. As one may have noticed, we stated the lemma above with a kernel u taking complex values. Normally, the result of Giné et al. (2000) was stated for real valued kernel. But their result can be extended to complex valued kernel by simply separating the real and imaginary parts as shown below. Indeed if we denote u_R and u_I the real and imaginary part of u it follows that

$$\mathbb{P}\left(\left|\sum_{1\le i_1\neq i_2\le N} u(Z_{i_1},Z_{i_2})\right| \ge x\right) = \mathbb{P}\left(\left|\sum_{1\le i_1\neq i_2\le N} u_R(Z_{i_1},Z_{i_2})\right|^2 + \left|\sum_{1\le i_1\neq i_2\le N} u_I(Z_{i_1},Z_{i_2})\right|^2 \ge x^2\right)$$
$$\le \mathbb{P}\left(\left|\sum_{1\le i_1\neq i_2\le N} u_R(Z_{i_1},Z_{i_2})\right|^2 \ge \frac{x^2}{2}\right) + \mathbb{P}\left(\left|\sum_{1\le i_1\neq i_2\le N} u_I(Z_{i_1},Z_{i_2})\right|^2 \ge \frac{x^2}{2}\right).$$

Hence, it amounts to a real valued problem. We only deal with the real part since exactly the same arguments remain true for the imaginary part. Let us show now that the bounds A, B, C, D of Lemma 2 hold for the real part $\mathbb{P}\left(\left|\sum_{1\le i_1\neq i_2\le N} u_R(Z_{i_1},Z_{i_2})\right|^2 \ge x^2/2\right)$. Because $u_R^2 + u_I^2 = |u|^2$, we have $u_R \le |u_R| \le |u|$. Then

$$\begin{aligned} \|u_R\|_\infty \le \|u\|_\infty &\le A \\ N\|\mathbb{E}((u_R)^2(Z,.))\|_\infty \le N\|\mathbb{E}(|u|^2(Z,.))\|_\infty &\le B^2 \\ N^2\mathbb{E}[(u_R)^2(Z_1,Z_2)] \le N^2\mathbb{E}[|u|^2(Z_1,Z_2)] &\le C^2. \end{aligned}$$

As for the last term D, since u_1 and u_2 are real valued

$$|\mathbb{E}(u(Z_1,Z_2)u_1(Z_1)u_2(Z_2))|^2 = |\mathbb{E}(u_R(Z_1,Z_2)u_1(Z_1)u_2(Z_2))|^2 + |\mathbb{E}(u_I(Z_1,Z_2)u_1(Z_1)u_2(Z_2))|^2$$

it entails that

$$N\sup\{\mathbb{E}(u_R(Z_1,Z_2)u_1(Z_1)u_2(Z_2))\} \le N\sup\{|\mathbb{E}(u(Z_1,Z_2)u_1(Z_1)u_2(Z_2))|\} \le D$$

which concludes the justification our lemma.

Let us compute now the four bounds, A, B, C, D.

▷ *Computation of A*
Denoting by $Y^l(x)$ the vector $(Y_m^l(x))_{-l\le m\le l}$ and using algebraic properties of the spherical harmonics,

$$\sum_{m=-l}^{l} |\Phi_{lm}(x)|^2 = \sum_m |(f_{\varepsilon^{-1}}^{\star l} Y^l(x))_m|^2 \le \|f_{\varepsilon^{-1}}^{\star l}\|_{op}^2 \sum_m |Y_m^l(x)|^2 \le d_0^{-2} l^{2\nu} \frac{2l+1}{4\pi}. \qquad (14)$$

We deduce that for all $x, y \in \mathbb{S}^2$,

$$|u(x,y)| \leq \sum_{l=1}^{L} \left(\sum_{m=-l}^{l} |\Phi_{lm}(x)|^2 \sum_{m=-l}^{l} |\Phi_{lm}(y)|^2 \right)^{1/2} \leq \sum_{l=1}^{L} \pi^{-1} d_0^{-2} l^{2\nu+1}$$

so that $A \leq (\pi^{-1} d_0^{-2}/(2\nu+2))(L+1)^{2\nu+2}$.

▷ *Computation of C*

We state the following Lemma, which allows to control the order of the variance of the test statistic.

Lemma 3. *Under Assumption 1, denoting* $c_3 = 3d_0^{-4} 2^{4\nu+2}/(4\nu+2)$

$$\sum_{l_1,l_2=1}^{L} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} |\mathbb{E}_{f_0}(\Phi_{l_1 m_1}(Z)\overline{\Phi_{l_2 m_2}(Z)})|^2 \leq c_3 L^{4\nu+2}$$

and $$\sum_{l_1,l_2=1}^{L} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} |\mathbb{E}_{f_0}(\Phi_{l_1 m_1}(Z)\Phi_{l_2 m_2}(Z))|^2 \leq c_3 L^{4\nu+2}.$$

Proof. Under H_0, the Z_i are uniformly distributed on the sphere. Then

$$\begin{aligned}
\mathbb{E}_{f_0}(\Phi_{l_1 m_1}\overline{\Phi_{l_2 m_2}(Z)}) &= \int \Phi_{l_1 m_1}\overline{\Phi_{l_2 m_2}(z)}dz \\
&= \sum_{n_1,n_2} (f_{\varepsilon^{-1}}^{\star l_1})_{m_1 n_1} (\overline{f_{\varepsilon^{-1}}^{\star l_2}})_{m_2 n_2} \int \overline{Y_{n_1}^{l_1}(z)} Y_{n_2}^{l_2}(z)dz \\
&= \sum_{n} (f_{\varepsilon^{-1}}^{\star l_1})_{m_1,n} (\overline{f_{\varepsilon^{-1}}^{\star l_1}})_{m_2,n} \mathbb{1}_{l_1=l_2}.
\end{aligned}$$

But, for any matrices $A = (a_{mn})_{-l\leq m\leq l, -l\leq n\leq l}$, $B = (b_{mn})_{-l\leq m\leq l, -l\leq n\leq l}$

$$\sum_{m_1} |\sum_{n} a_{m_1 n} b_{m_2 n}|^2 \leq \|A\|_{op}^2 \sum_{n} |b_{m_2 n}|^2 \leq \|A\|_{op}^2 \|B^T\|_{op}^2 = \|A\|_{op}^2 \|B\|_{op}^2$$

Then

$$\sum_{m_1,m_2} |\mathbb{E}_{f_0}(\Phi_{l_1 m_1}\overline{\Phi_{l_2 m_2}(Z)})|^2 \leq \|f_{\varepsilon^{-1}}^{\star l_1}\|_{op}^4 (2l_1+1)\mathbb{1}_{l_1=l_2} \tag{15}$$

and

$$\sum_{l_1,l_2=1}^{L} \sum_{m_1,m_2} |\mathbb{E}_{f_0}(\Phi_{l_1 m_1}\overline{\Phi_{l_2 m_2}(Z)})|^2 \leq \sum_{l_1=1}^{L} \|f_{\varepsilon^{-1}}^{\star l_1}\|_{op}^4 (2l_1+1) \leq 3d_0^{-4} \sum_{l_1=1}^{L} l_1^{4\nu+1} \leq \frac{3d_0^{-4}}{4\nu+2}(L+1)^{4\nu+2}.$$

For the second term, we can write, using (3)

$$\mathbb{E}_{f_0}(\Phi_{l_1 m_1}\Phi_{l_2 m_2}(Z)) = \sum_{n} (f_{\varepsilon^{-1}}^{\star l_1})_{m_1,n} (f_{\varepsilon^{-1}}^{\star l_1})_{m_2,-n} (-1)^n \mathbb{1}_{l_1=l_2}.$$

Then it is sufficient to apply the same method with matrix B such that $b_{mn} = (-1)^n (f_{\varepsilon^{-1}}^{\star l_1})_{m,-n} = (-1)^n a_{m,-n}$. The conclusion results from equality $\|B\|_{op} = \|A\|_{op}$. □

Lemma 3 gives $C^2 \leq 3d_0^{-4}/(4\nu+2)N^2(L+1)^{4\nu+2}$.

▷ *Computation of B*
Let $x \in \mathbb{S}^2$. We can write

$$\mathbb{E}_{f_0}[|u(Z,x)|^2] = \sum_{l_1,l_2} \sum_{m_1,m_2} \mathbb{E}_{f_0}[\Phi_{l_1 m_1} \overline{\Phi_{l_2 m_2}(Z)}] \Phi_{l_2 m_2}(x) \overline{\Phi_{l_1 m_1}(x)}.$$

But we have seen previously than $\mathbb{E}_{f_0}[\Phi_{l_1 m_1} \overline{\Phi_{l_2 m_2}(Z)}]$ vanishes when $l_1 \neq l_2$. Then, using Cauchy-Schwarz inequality, we compute

$$\mathbb{E}_{f_0}[|u(Z,x)|^2] \leq \sum_l \left(\sum_{m_1,m_2} |\mathbb{E}_{f_0}[|\Phi_{lm_1} \overline{\Phi_{lm_2}(Z)}]|^2 \sum_{m_1,m_2} |\Phi_{lm_1}(x)\Phi_{lm_2}(x)|^2 \right)^{1/2}.$$

Now, we use previous computations (15) and (14) to state

$$\begin{aligned}
\mathbb{E}_{f_0}[u^2(Z,x)] &\leq \sum_l \left(\|f_{\varepsilon^{-1}}^{\star l_1}\|_{op}^4 (2l+1) \right)^{1/2} \sum_m |\Phi_{lm}(x)|^2 \\
&\leq \sum_l \left(3d_0^{-4} l^{4\nu+1}\right)^{1/2} d_0^{-2} l^{2\nu} \frac{2l+1}{4\pi} \\
&\leq \sum_{l=1}^{L} \frac{\sqrt{3}}{\pi} d_0^{-4} l^{4\nu+3/2} \leq \frac{\sqrt{3} d_0^{-4}}{\pi(4\nu+5/2)} (L+1)^{4\nu+5/2}.
\end{aligned}$$

Thus $B^2 \leq /\sqrt{3}\pi^{-1} d_0^{-4}/(4\nu+5/2)N(L+1)^{4\nu+5/2}$.

▷ *Computation of D*
Let us first compute $\mathbb{E}_{f_0}(\Phi_{lm}(Z_1)u_1(Z_1))$ under H_0. We denote by U_1^l the vector of the Fourier coefficients of u_1 with harmonic order l: $U_1^l = (< u_1, Y_n^l >)_{-l \leq n \leq l}$.

$$\begin{aligned}
\mathbb{E}_{f_0}(\Phi_{lm}(Z_1)u_1(Z_1)) &= \int \Phi_{lm}(x) u_1(x) dx = \sum_{n=-l}^{l} (f_{\varepsilon^{-1}}^{\star l})_{mn} \int \overline{Y_n^l(x)} u_1(x) dx \\
&= \sum_{n=-l}^{l} (f_{\varepsilon^{-1}}^{\star l})_{mn} < u_1, Y_n^l > = (f_{\varepsilon^{-1}}^{\star l} U_1^l)_m.
\end{aligned}$$

Then

$$\sum_{m=-l}^{l} |\mathbb{E}_{f_0}(\Phi_{lm}(Z_1)u_1(Z_1))|^2 = \|f_{\varepsilon^{-1}}^{\star l} U_1^l\|^2 \leq d_0^{-2} l^{2\nu} \|U_1^l\|^2.$$

But, using Parseval's equality

$$\sum_{l \geq 0} \|U_1^l\|^2 = \sum_{l \geq 0} \sum_{n=-l}^{l} | < u_1, Y_n^l > |^2 = \int u_1^2(x) dx$$

so that, under H_0, $\sum_l \|U_1^l\|^2 \leq \mathbb{E}_{f_0}(u_1^2(Z_1))$. In the same way we can prove

$$\sum_{m=-l}^{l} |\mathbb{E}_{f_0}(\overline{\Phi_{lm}(Z_2)}u_2(Z_2))|^2 \leq d_0^{-2} l^{2\nu} \|U_2^l\|^2$$

with $\sum_l \|U_2^l\|^2 \leq \mathbb{E}_{f_0}(u_2^2(Z_1))$. Then, using repeatedly Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbb{E}_{f_0}(u(Z_1,Z_2)u_1(Z_1)u_2(Z_2)) &= \sum_{l=1}^{L}\sum_{m=-l}^{l} \mathbb{E}_{f_0}(\Phi_{lm}(Z_1)u_1(Z_1))\mathbb{E}_{f_0}(\overline{\Phi_{lm}(Z_2)}u_2(Z_2)) \\
&\leq \sum_{l=1}^{L}\left(\sum_{m=-l}^{l} |\mathbb{E}_{f_0}(\Phi_{lm}(Z_1)u_1(Z_1))|^2 \sum_{m=-l}^{l} |\mathbb{E}_{f_0}(\overline{\Phi_{lm}(Z_2)}u_2(Z_2))|^2\right)^{1/2} \\
&\leq \sum_{l=1}^{L} d_0^{-2} l^{2\nu} \|U_1^l\| \|U_2^l\| \leq d_0^{-2} L^{2\nu} \left(\sum_{l=1}^{L} \|U_1^l\|^2 \sum_{l=1}^{L} \|U_2^l\|^2\right)^{1/2} \\
&\leq d_0^{-2} L^{2\nu} \mathbb{E}_{f_0}^{1/2}(u_1^2(Z_1)) \mathbb{E}_{f_0}^{1/2}(u_2^2(Z_2)).
\end{aligned}$$

Thus $D \leq d_0^{-2} N L^{2\nu}$.

Conclusion
Now, using Lemma 2 with $x = N(N-1)t/2$, we obtain

$$P\left(|T_L| \geq t\right) \leq K_1 \exp\left(-K_3 \min\left(\frac{N^2t^2}{L^{4\nu+2}}, \frac{Nt}{L^{2\nu}}, \frac{Nt^{2/3}}{L^{4\nu/3+5/6}}, \frac{Nt^{1/2}}{L^{\nu+1}}\right)\right)$$

where K_3 only depends on d_0 and ν. Then

$$\begin{aligned}
P\left(|T_L| \geq L^{2\nu+1}u_N/N\right) &\leq K_1 \exp\left(-K_3 \min\left(u_N^2, u_N L, N^{1/3}u_N^{2/3}L^{-1/6}, N^{1/2}u_N^{1/2}L^{-1/2}\right)\right) \\
&\leq K_1 \exp(-K_0 u_N^2)
\end{aligned}$$

provided that $u_N = O(L)$, $L = O(N^2 u_N^{-8})$ and $L = O(N u_N^{-3})$.

8.4 Proof of Theorem 3

First, using Lemma 1

$$\begin{aligned}
\mathbb{P}_{f_0}(D_N = 1) &\leq \sum_{L\in\mathcal{L}} \mathbb{P}_{f_0}\left(|T_L| > \sqrt{2K_0^{-1}}t_L^2\right) \leq \sum_{L\in\mathcal{L}} \mathbb{P}_{f_0}\left(|T_L| > L^{2\nu+1}\sqrt{2K_0^{-1}\log\log N}/N\right) \\
&\leq \sum_{L\in\mathcal{L}} K_1 \exp(-K_0(2K_0^{-1}\log\log N)) = K_1 \sum_{L\in\mathcal{L}} \exp(-2\log\log N) \\
&\leq K_2 |\mathcal{L}| (\log(N))^{-2} = O(\log(N)^{-1}) = o(1)
\end{aligned}$$

since $|\mathcal{L}| = O(\log(N))$. Then, for N large enough, $\mathbb{P}_{f_0}(D_N = 1) \leq \eta/2$.

Now let $f \in H_1(s, R, \mathcal{C}\psi_N)$. Then

$$\mathbb{P}_f(D_N = 0) = \mathbb{P}_f\left(\forall L \in \mathcal{L},\ |T_L| \leq \sqrt{2K_0^{-1}}t_L^2\right) \leq \mathbb{P}_f\left(|T_{L^*}| \leq \sqrt{2K_0^{-1}}t_{L^*}^2\right)$$

with $L^* = 2^{j*}$ and $j* = \lfloor \log_2[(N/\sqrt{\log\log N})^{1/(2s+2\nu+1)}] \rfloor$. Remark that for N large enough, $4\log\log N \leq (N/\sqrt{\log\log N})^{1/(2s+2\nu+1)} \leq N/(\log\log N)^{3/2}$, so that $j_0 \leq j* \leq j_m$ and L^* belongs to $\mathcal{L}$. Note also that with this choice $t^2_{L^*} \leq \psi_N$ and $L^{*-2s} \leq 2^{2s}\psi_N$. Using triangle inequality we have that

$$\mathbb{P}_f\left(|T_{L^*}| \leq \sqrt{2K_0^{-1}}t^2_{L^*}\right) \leq \mathbb{P}_f\left(|T_{L^*} - \mathbb{E}_f(T_{L^*})| \geq \|f - f_0\|_2^2 - \sqrt{2K_0^{-1}}t^2_{L^*} - B_f(T_{L^*})\right) \quad (16)$$

where $B_f(T_L) = \|f - f_0\|_2^2 - \mathbb{E}_f(T_L)$. If f is in the Sobolev ball $W_s(\mathbb{S}^2, R)$, it directly follows from the definition of $W_s(\mathbb{S}^2, R)$ (8) that

$$B_f(T_L) = \sum_{l>L}\sum_m |f^{\star l}_m|^2 \leq R^2 L^{-2s} \leq R^2 2^{2s}\psi_N.$$

We set $C_1 = \sqrt{2K_0^{-1}} + R^2 2^{2s}$ and $C_2 = 1 - C_1/\mathcal{C} > 0$. Using the definition of H_1

$$\psi_N \leq \mathcal{C}^{-1}\|f - f_0\|_2^2.$$

Markov inequality yields the following upperbound for the expression (16)

$$\mathbb{P}_f(D_N = 0) \leq \frac{\mathrm{Var}_f(T_{L^*})}{C_2^2\|f - f_0\|^4}. \quad (17)$$

Let us now state the following Lemma which evaluate the variance of the estimator T_L.

Lemma 4. *If Assumption 1 is verified,*

$$\mathrm{Var}_f(T_L) \leq c_4\left(\frac{L^{4\nu+2}}{N^2} + \frac{\|f - f_0\|^2 L^{4\nu+4}}{N^2} + \frac{\|f - f_0\|^2 L^{2\nu+1}}{N} + \frac{\|f - f_0\|^3 L^{2\nu+2}}{N} + \frac{\|f - f_0\|^4}{N}\right)$$

where c_4 only depends on d_0 and ν.

Proof. We have

$$\mathrm{Var}_f(T_L) = \mathbb{E}((T_L - \mathbb{E}(T_L))(\overline{T_L - \mathbb{E}(T_L)}).$$

Simple calculations entail that

$$\mathrm{Var}_f(T_L) = -\sum_{l_1,m_1}\sum_{l_2,m_2} |f^{\star l_1}_{m_1}|^2 |f^{\star l_2}_{m_2}|^2$$
$$+\frac{4}{(N(N-1))^2}\left[\sum_{l_1,l_2=1}^{L}\sum_{m_1=-l_1}^{l_1}\sum_{m_2=-l_2}^{l_2}\mathbb{E}\left(\sum_{i_1<i_2}\Phi_{l_1m_1}(Z_{i_1})\overline{\Phi_{l_1m_1}(Z_{i_2})}\sum_{i_3<i_4}\overline{\Phi_{l_2m_2}(Z_{i_3})}\Phi_{l_2m_2}(Z_{i_4})\right)\right].$$

The term

$$\mathbb{E}\left(\sum_{i_1<i_2}\Phi_{l_1m_1}(Z_{i_1})\overline{\Phi_{l_1m_1}(Z_{i_2})}\sum_{i_3<i_4}\overline{\Phi_{l_2m_2}(Z_{i_3})}\Phi_{l_2m_2}(Z_{i_4})\right)$$

is bounded by

$$\sum_{i_1<i_2}\sum_{i_3<i_4}\mathbb{E}(\Phi_{l_1m_1}(Z_{i_1})\overline{\Phi_{l_1m_1}(Z_{i_2})\Phi_{l_2m_2}(Z_{i_3})}\Phi_{l_2m_2}(Z_{i_4}))$$
$$=\sum_{i_1<i_2}\sum_{i_3<i_4}\Big[|f^{\star l_1}_{m_1}|^2|f^{\star l_2}_{m_2}|^2\mathbb{1}_{i_1\neq i_2\neq i_3\neq i_4}+|\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2\mathbb{1}_{i_1=i_3,i_2=i_4}$$
$$+|\mathbb{E}(\Phi_{l_1m_1}(Z)\Phi_{l_2m_2}(Z))|^2\mathbb{1}_{i_1=i_4,i_2=i_3}+\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})\overline{f^{\star l_1}_{m_1}}f^{\star l_2}_{m_2}\mathbb{1}_{i_1=i_3,i_2\neq i_4}$$
$$+\mathbb{E}(\overline{\Phi_{l_1m_1}(Z)}\Phi_{l_2m_2}(Z))f^{\star l_1}_{m_1}\overline{f^{\star l_2}_{m_2}}\mathbb{1}_{i_1\neq i_3,i_2=i_4}+\mathbb{E}(\Phi_{l_1m_1}(Z)\Phi_{l_2m_2}(Z))\overline{f^{\star l_1}_{m_1}f^{\star l_2}_{m_2}}\mathbb{1}_{i_1=i_4,i_2\neq i_3}$$
$$+\mathbb{E}(\overline{\Phi_{l_1m_1}(Z)\Phi_{l_2m_2}(Z)})f^{\star l_1}_{m_1}f^{\star l_2}_{m_2}\mathbb{1}_{i_1\neq i_4,i_2=i_3})\Big].$$

Eventually we get that

$$\begin{aligned}\mathrm{Var}_f(T_L) &= \sum_{l_1,l_2=1}^{L}\sum_{m_1m_2}\Bigg[\left(\frac{(N-2)(N-3)}{N(N-1)}-1\right)|f^{\star l_1}_{m_1}|^2|f^{\star l_2}_{m_2}|^2 \\ &+\frac{1}{N(N-1)}\left(|\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2+|\mathbb{E}(\Phi_{l_1m_1}(Z)\Phi_{l_2m_2}(Z))|^2\right) \\ &+\frac{2(N-2)}{N(N-1)}\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})\overline{f^{\star l_1}_{m_1}}f^{\star l_2}_{m_2} \\ &+\frac{2(N-2)}{N(N-1)}\Re\left(\mathbb{E}(\Phi_{l_1m_1}(Z)\Phi_{l_2m_2}(Z))\overline{f^{\star l_1}_{m_1}f^{\star l_2}_{m_2}}\right)\Bigg]\end{aligned} \tag{18}$$

where $\Re(x)$ denotes the real part of x. We shall now upperbound each term that appears in the expression (18) above.

▷ First term. Since $\sum_{l=1}^{L}\sum_m|f^{\star l}_m|^2\le\|f-f_0\|^2$, we obtain

$$\sum_{l_1,l_2=1}^{L}\sum_{m_1m_2}\left(\frac{(N-2)(N-3)}{N(N-1)}-1\right)|f^{\star l_1}_{m_1}|^2|f^{\star l_2}_{m_2}|^2\le\frac{\|f-f_0\|_2^4}{N}.$$

▷ Second term. Firstly

$$\begin{aligned}|\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2 &= \left|\int\Phi_{l_1m_1}\overline{\Phi_{l_2m_2}}f_0+\int\Phi_{l_1m_1}\overline{\Phi_{l_2m_2}}(f_Z-f_0)\right|^2 \\ &\le 2|\mathbb{E}_{f_0}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2+2\left|\int\Phi_{l_1m_1}\overline{\Phi_{l_2m_2}}(f_Z-f_0)\right|^2 \\ &\le 2|\mathbb{E}_{f_0}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2+2\|\Phi_{l_1m_1}\overline{\Phi_{l_2m_2}}\|_2^2\|f_Z-f_0\|_2^2.\end{aligned}$$

We can remark that, under Assumption 1,

$$\|f_Z-f_0\|_2^2=\sum_{lm}|(f_Z-f_0)^{\star l}_m|^2\le\sum_l\|f^{\star l}_\varepsilon\|^2_{op}\sum_m|(f-f_0)^{\star l}_m|^2\le\|f-f_0\|_2^2$$

since $\|f^{\star l}_\varepsilon\|_{op}\le1$ for all l. Now let us show that there exists $C_1>0$ such that

$$\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}\|\Phi_{l_1m_1}\overline{\Phi_{l_2m_2}}\|_2^2\le C_1L^{4\nu+4}.$$

We deduce from (14) that

$$\begin{aligned}\sum_{m_1}\int|\Phi_{l_1m_1}(x)|^2|\Phi_{l_2m_2}(x)|^2dx &\leq \frac{3d_0^{-2}}{4\pi}l_1^{2\nu+1}\int|\Phi_{l_2m_2}|^2 \leq \frac{3d_0^{-2}}{4\pi}l_1^{2\nu+1}\sum_m|(f^{*l_2}_{\varepsilon^{-1}})_{m_2m}|^2\\ &\leq \frac{3d_0^{-4}}{4\pi}l_1^{2\nu+1}l_2^{2\nu}.\end{aligned}$$

Then

$$\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}\|\Phi_{l_1m_1}\overline{\Phi_{l_2m_2}}\|_2^2 \leq \frac{3d_0^{-4}}{4\pi}\sum_{l_1,l_2=1}^{L}\sum_{m_2}l_1^{2\nu+1}l_2^{2\nu} \leq C_1L^{4\nu+4}$$

and, using Lemma 3,

$$\begin{aligned}\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}|\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2 &\leq 2\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}|\mathbb{E}_{f_0}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2\\ &\quad +2C_1L^{4\nu+4}\|f-f_0\|_2^2\\ &\leq C_2(L^{4\nu+2}+L^{4\nu+4}\|f-f_0\|_2^2).\end{aligned}$$

In the same way

$$\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}|\mathbb{E}(\Phi_{l_1m_1}(Z)\Phi_{l_2m_2}(Z))|^2 \leq C_2(L^{4\nu+2}+L^{4\nu+4}\|f-f_0\|_2^2).$$

Thus, the second term is bounded by a constant times $L^{4\nu+2}/N^2 + L^{4\nu+4}\|f-f_0\|_2^2/N^2$.

▷ Third term. Using Cauchy Schwarz inequality we get

$$\begin{aligned}\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}f^{\star l_1}_{m_1}\overline{f^{\star l_2}_{m_2}}\mathbb{E}(\Phi_{lm_1}(Z)\overline{\Phi_{lm_2}(Z)}) &\leq \left(\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}|f^{\star l_1}_{m_1}|^2|f^{\star l_2}_{m_2}|^2\right)^{1/2}\\ &\quad \left(\sum_{l_1,l_2=1}^{L}\sum_{m_1,m_2}|\mathbb{E}(\Phi_{l_1m_1}(Z)\overline{\Phi_{l_2m_2}(Z)})|^2\right)^{1/2}\\ &\leq \sqrt{C_2}\ \ \|f-f_0\|_2^2(L^{2\nu+1}+L^{2\nu+2}\|f-f_0\|_2).\end{aligned}$$

The third term is of order $\|f-f_0\|_2^2L^{2\nu+1}/N + \|f-f_0\|_2^3L^{2\nu+2}/N$.

▷ Fourth term. We bound the fourth term in the same way as the third.

Finally we have the bound for $\mathrm{Var}_f(T_L)$

$$\frac{L^{4\nu+2}}{N^2}+\frac{\|f-f_0\|_2^2L^{4\nu+4}}{N^2}+\frac{\|f-f_0\|_2^2L^{2\nu+1}}{N}+\frac{\|f-f_0\|_2^3L^{2\nu+2}}{N}+\frac{\|f-f_0\|_2^4}{N}.$$

□

This gives

$$\frac{\mathrm{Var}_f(T_{L^*})}{\|f-f_0\|^4} \leq c_4\left(\frac{L^{*4\nu+2}}{N^2\|f-f_0\|^4}+\frac{L^{*4\nu+4}}{N^2\|f-f_0\|^2}+\frac{L^{*2\nu+1}}{N\|f-f_0\|^2}+\frac{L^{*2\nu+2}}{N\|f-f_0\|}+\frac{1}{N}\right).$$

Besides, as $\|f-f_0\|_2^2 \geq \mathcal{C}2^{-2s}L^{*-2s}$ and $N \geq L^{*2s+2\nu+1}\sqrt{\log\log N}$, we get an upperbound for (17) in terms of L^*

$$C_3\left(\frac{1}{C^2\log\log N}+\frac{L^{*2-2s}}{C\log\log N}+\frac{1}{C\sqrt{\log\log N}}+\frac{L^{*1-s}}{\sqrt{C\log\log N}}+\frac{1}{N}\right).$$

Since $s \geq 1$, all these terms tend to zero when N goes to infinity. In conclusion for N large enough, $\mathbb{P}_f(D_N=0) \leq \eta/2$.

8.5 Proof of Theorem 4

This proof follows the same line as the one of Theorem 3. We first give an adaptation of Lemma 4 in order to control the variance of T_L:

$$\begin{aligned}\mathrm{Var}_f(T_L) \leq C_0\Bigg(&\frac{L^{-4\nu_0+2-\beta}}{N^2}e^{4L^\beta/\delta}+\frac{\|f-f_0\|^2L^{-4\nu_0+4-2\beta}}{N^2}e^{4L^\beta/\delta}\\&+\frac{\|f-f_0\|^2L^{-2\nu_0+1-\beta/2}}{N}e^{2L^\beta/\delta}+\frac{\|f-f_0\|^3L^{-2\nu_0+2-\beta}}{N}e^{2L^\beta/\delta}+\frac{\|f-f_0\|^4}{N}\Bigg).\end{aligned} \tag{19}$$

This result is obtained with standard integrals evaluation which give for any real α,

$$\sum_{l=1}^{L} l^\alpha e^{l^\beta/\delta} \leq C\int_1^{L+1} x^\alpha e^{x^\beta/\delta}dx \leq C'L^{\alpha+1-\beta}e^{L^\beta/\delta} \tag{20}$$

(for L large enough if $\alpha<0$). Now, we evaluate the first type error. Using that $\mathbb{E}_{f_0}(T_{L^*})=0$, we write

$$\begin{aligned}\mathbb{P}_{f_0}(D_N=1) &= \mathbb{P}_{f_0}\left(|T_{L^*}|>K_0t_{L^*}^2\right) \leq K_0^{-2}t_{L^*}^{-4}\mathrm{Var}_{f_0}(T_{L^*})\\ &\leq K_0^{-2}C_0t_{L^*}^{-4}L^{*-4\nu_0+2-\beta}\exp\left(4L^{*\beta}/\delta\right)N^{-2} \leq K_0^{-2}C_0L^{*-\beta} \leq \eta/2\end{aligned}$$

for N large enough. To bound the error of the second kind, let $f \in H_1(s,R,\mathcal{C}\psi_N)$. We have

$$\mathbb{P}_f(D_N=0) \leq \mathbb{P}_f\left(|T_{L^*}| \leq K_0t_{L^*}^2\right) \leq \mathbb{P}_f\left(|T_{L^*}-\mathbb{E}_f(T_{L^*})| \geq \|f-f_0\|_2^2-K_0t_{L^*}^2-B_f(T_{L^*})\right).$$

The definition of L^* implies that, for N large enough

$$\left(\frac{\delta}{16}\log(N)\right)^{1/\beta} \leq L^* \leq \left(\frac{\delta}{8}\log(N)\right)^{1/\beta}.$$

That ensures that $L^{*-2s} \leq (\delta/16)^{-2s/\beta}\psi_N$ and $t_{L^*}^2 \leq (\delta\log N/8)^{(-2\nu_0+1)/\beta}N^{-3/4} \leq \psi_N$ for N large enough. We set $C_1=K_0+R^2(\delta/16)^{-2s/\beta}$ and $C_2=1-C_1/\mathcal{C}$ (which is positive if $\mathcal{C}$ large enough). Markov inequality yields

$$\mathbb{P}_f(D_N=0) \leq \frac{\mathrm{Var}_f(T_{L^*})}{C_2^2\|f-f_0\|^4}. \tag{21}$$

Using (19), we bound

$$\begin{aligned}\frac{\mathrm{Var}_f(T_{L^*})}{\|f-f_0\|^4} \;\leq\; & C_0\left(\frac{L^{*-4\nu_0+2-\beta}}{N^2\|f-f_0\|^4}e^{(4L^{*\beta}/\delta)}+\frac{L^{*-4\nu_0+4-2\beta}}{N^2\|f-f_0\|^2}e^{(4L^{*\beta}/\delta)}\right.\\ & \left.+\frac{L^{*-2\nu_0+1-\beta/2}}{N\|f-f_0\|^2}e^{(2L^{*\beta}/\delta)}+\frac{L^{*-2\nu_0+2-\beta}}{N\|f-f_0\|}e^{(2L^{*\beta}/\delta)}+\frac{1}{N}\right).\end{aligned}$$

Besides, as $\|f-f_0\|_2^2 \geq C_3 L^{*-2s}$, we get the following upperbound

$$\begin{aligned}C_4\left(\frac{(\log N)^{(-4\nu_0+2-\beta+4s)/\beta}N^{1/2}}{N^2}+\frac{(\log N)^{(-4\nu_0+4-2\beta+2s)/\beta}N^{1/2}}{N^2}\right.\\ \left.+\frac{(\log N)^{(-2\nu_0+1-\beta/2+2s)/\beta}N^{1/4}}{N}+\frac{(\log N)^{(-2\nu_0+2-\beta+s)/\beta}N^{1/4}}{N}+\frac{1}{N}\right).\end{aligned}$$

and all these terms tend to zero when N goes to infinity.

8.6 Proof of Theorem 5

The proof is analogous to the proof of Theorem 1, with hypothesis functions

$$f_\theta = f_0 + \sum_{m=-L}^{L}\theta_{Lm}\varphi_{Lm}, \qquad \mathbb{P}(\theta_{Lm}=\pm\gamma)=1/2,$$

where

$$\gamma^2 = c_1\exp(-2L^\beta/\delta)L^{-2s+2\nu_0-1}$$

and

$$L = \left\lfloor (2\delta\log(N))^{1/\beta}\right\rfloor.$$

This choice of L ensures that, for N large enough,

$$(\delta\log(N))^{1/\beta} \leq L \leq (2\delta\log(N))^{1/\beta}.$$

Aknowledgements

We would like to thank Erwan Le Pennec and Richard Nickl for interesting discussions and suggestions.

References

Bai, Z. D., Radhakrishna Rao, C., and Zhao, L. C. (1988). Kernel estimators of density function of directional data. *J. Multivariate Anal.*, 27(1):24–39.

Butucea, C. (2007). Goodness-of-fit testing and quadratic functional estimation from indirect observations. *Ann. Statist.*, 35(5):1907–1930.

Butucea, C., Matias, C., and Pouet, C. (2009). Adaptive goodness-of-fit testing from indirect observations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(2):352–372.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.*, 19(3):1257–1272.

Faÿ, G., Delabrouille, J., Kerkyacharian, G., and Picard, D. (2011). Testing the isotropy of high energy cosmic rays using spherical needlets. *arXiv:1107.5658v1*.

Giné, E., Latała, R., and Zinn, J. (2000). Exponential and moment inequalities for U-statistics. In *High dimensional probability, II (Seattle, WA, 1999)*, volume 47 of *Progr. Probab.*, pages 13–38. Birkhäuser Boston, Boston, MA.

Hall, P., Watson, G. S., and Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika*, 74(4):751–762.

Healy, Jr., D. M., Hendriks, H., and Kim, P. T. (1998). Spherical deconvolution. *J. Multivariate Anal.*, 67(1):1–22.

Ingster, Y. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I, II, III. *Math. Methods Statist.*, 2(1):85–114.

Kerkyacharian, G., Pham Ngoc, T. M., and Picard, D. (2011). Localized spherical deconvolution. *Ann. Statist.*, 39(2):1042–1068.

Kim, P. T. and Koo, J.-Y. (2002). Optimal spherical deconvolution. *J. Multivariate Anal.*, 80(1):21–42.

Kim, P. T., Koo, J.-Y., and Park, H. J. (2004). Sharp minimaxity and spherical deconvolution for super-smooth error distributions. *J. Multivariate Anal.*, 90(2):384–392.

Laurent, B., Loubes, J.-M., and Marteau, C. (2011). Testing inverse problems: a direct or an indirect problem? *J. Statist. Plann. Inference*, 141(5):1849–1861.

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

Talman, J. D. (1968). *Special functions: A group theoretic approach.* Based on lectures by Eugene P. Wigner. With an introduction by Eugene P. Wigner. W. A. Benjamin, Inc., New York-Amsterdam.

Terras, A. (1985). *Harmonic analysis on symmetric spaces and applications. I.* Springer-Verlag, New York.

Tsybakov, A. B. (2009). *Introduction to nonparametric estimation.* Springer Series in Statistics. Springer, New York. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.

Vilenkin, N. J. (1968). *Special functions and the theory of group representations.* Translated from the Russian by V. N. Singh. Translations of Mathematical Monographs, Vol. 22. American Mathematical Society, Providence, R. I.