Electronic Journal of Statistics
Vol. 0 (2022) 1–34
ISSN: 1935-7524
DOI: 10.1214/154957804100000000

Deconvolution of spherical data corrupted with unknown noise

Jérémie Capitao-Miniconi and Elisabeth Gassiat

Universite Paris-Saclay, CNRS
Laboratoire de Mathematiques d'Orsay
91405 Orsay Cedex, France
e-mail: Jeremie.capitao-miniconi@universite-paris-saclay.fr
; Elisabeth.gassiat@unversite-paris-saclay.fr

Abstract: We consider the deconvolution problem for densities supported on a $(d-1)$-dimensional sphere with unknown center and unknown radius, in the situation where the distribution of the noise is unknown and without any other observations. We propose estimators of the radius, of the center, and of the density of the signal on the sphere that are proved consistent without further information. The estimator of the radius is proved to have almost parametric convergence rate for any dimension d. When $d = 2$, the estimator of the density is proved to achieve the same rate of convergence over Sobolev regularity classes of densities as when the noise distribution is known.

Keywords and phrases: deconvolution, spherical data, unknown noise.

Contents

1. Introduction

In this paper, we study the deconvolution problem of random data on a sphere corrupted by independent additive noise. The observations are

$$Y_i = X_i + \varepsilon_i,\ i = 1, \dots, n \tag{1}$$

where $(X_i)_{i\geq 1}$ (the signal) is a sequence of independent identically distributed (i.i.d.) random variables taking values on a $(d-1)$-dimensional sphere (for $d \geq 2$) with unknown center $C^\star$ and unknown radius $R^\star$, $(\varepsilon_i)_{i\geq 1}$ (the noise) is a sequence of i.i.d. random variables in $\mathbb{R}^d$ independent of the signal and with totally unknown distribution. The distribution of the signal is also unknown, it is only known that it is spherically supported. To solve the deconvolution problem and estimate the structural parameters $C^\star$ and $R^\star$, the only assumption we shall put on the noise is that its d coordinates are independently distributed. Just notice that in model (1), the observed data may be outside the sphere. This is different from the model studied in [12], [11], where the observed noisy data remain on the sphere.

The statistical estimation of the center and of the radius of the sphere is of interest in various applications such as object tracking, robotics, pattern recognition, see for instance [5], [6], [16], [19], among others, see also [3] and references therein. For example, in target tracking, one aim is to recover the shape of the target from multiple measurements of the extent of the target at each time, one possible shape being a circle. In biometrics, the aim is to identify circular irises in an image. Several methods have been proposed to estimate the center and the radius of a sphere, based on least squares, maximum likelihood, see [13] for a recent likelihood based algorithm, most of them modeling the noise distribution with a Gaussian distribution.

The deconvolution problem of the distribution of the signal when the radius and the center are known is studied for circular signals (that is when $d = 2$) in [10]. The author proves that the minimax rate of convergence of the estimator over a wide collection of smoothness classes of the density of the signal on the circle does not depend on the (known) noise distribution, for a variety of different noise distributions, contrasting with the situation where the signal has a density with respect to Lebesgue over the whole space.

Recently, it has been proved in [7] that deconvolution with unknown noise distribution is possible for multivariate signals, as soon as the signal can be decomposed in two components that satisfy a mild dependence assumption, that its distribution has light enough tails, and without any assumption on the noise distribution except that its two corresponding components are independently distributed. The authors of [7] then consider the situation where the probability of X_1 has a density with respect to Lebesgue measure, and they prove that not knowing the noise distribution does not deteriorate the estimation rate of the density on Sobolev regularity classes for compactly supported signals.

Here, the probability distribution of the signal is singular with respect to Lebesgue measure on $\mathbb{R}^d$ and their convergence results do not apply. However, we prove that the general conditions they propose under which deconvolution with unknown noise is possible are satisfied for spherical signals, this is our first main identifiability result Theorem 2. The main contribution of our work is then to exhibit estimators that achieve remarkable properties:

- We propose estimators of the radius, the center, and the distribution of the signal, which are proved consistent whatever the noise distribution, see Proposition 2.
- Under the mild assumption that the noise has finite variance, we prove that the radius of the sphere can be estimated at almost parametric rate with totally unknown noise distribution, see Theorem 3.
- When $d = 2$, that is for circular signals, we prove that the center can be estimated at almost parametric rate and that the density of the signal distribution on the circle can be estimated at the same rate as when the distribution of the noise is known on some Sobolev regularity classes, with a rate which is minimax as proved in [10], see Theorem 4 and Theorem 5.

In Section 2, we first recall general results of [7] and we prove in Proposition 1 a strengthened version of the local L^2-consistency of the general estimator of the characteristic function of the signal that will be a basic stone for all our convergence rates theorems. We then state our identifiability theorem, give the definition of the estimators and prove their consistency. Section 3 studies the rates of convergence of our estimators, and in section 4 we study the situation where radius and center of the sphere together with the noise distribution are unknown, though the distribution of the angles of the random signal is known. Simulations illustrating our findings are given in Section 5. We discuss possible further work and related questions in Section 6. Proofs of propositions and lemmas are detailed in Section 7.

2. Identifiability and estimation method

In this section, we prove that model (1) is identifiable with no more assumptions. We then explain the estimation method and define the estimators which will be studied in Section 3.

2.1. Preliminaries: deconvolution with unknown noise

We first recall general results in [7]. Then, we prove a proposition which will be used to obtain the nearly parametric rate of our estimators of the radius and the center. In [7], the authors consider the situation where the observations $Y_i \in \mathbb{R}^d$ come from the model

$$\begin{pmatrix} Y^{(1)} \\ Y^{(2)} \end{pmatrix} = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} + \begin{pmatrix} \varepsilon^{(1)} \\ \varepsilon^{(2)} \end{pmatrix}$$

in which $Y^{(1)} \in \mathbb{R}^{d_1}$, $d_1 \geq 1$, and $Y^{(2)} \in \mathbb{R}^{d_2}$, $d_2 \geq 1$, with $d_1 + d_2 = d$, and where $\varepsilon^{(1)}$ is independent of $\varepsilon^{(2)}$. They prove identifiability under very mild assumptions on the signal distribution. The first one is about the tail of its distribution.

A(ρ) There exists $\rho < 2$, $a > 0$ and $b > 0$ such that for all $\lambda \in \mathbb{R}^d$, $E\left[\exp\left(\lambda^\top X\right)\right] \leq a\exp\left(b\|\lambda\|_2^\rho\right)$.

Here, $X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}$ and $\|\cdot\|_2$ is the Euclidian norm.

Under A(ρ), the characteristic function of the signal can be extended into a multivariate analytic function denoted by

$$\begin{array}{rcl} \Phi : \mathbb{C}^{d_1} \times \mathbb{C}^{d_2} & \longrightarrow & \mathbb{C} \\ (z_1, z_2) & \longmapsto & E\left[\exp\left(iz_1^\top X^{(1)} + iz_2^\top X^{(2)}\right)\right]. \end{array}$$

The second assumption is a mild dependence assumption (see the discussion after Theorem 2.1 in [7]).

A(dep) For any $z_0 \in \mathbb{C}^{d_1}$, $z \mapsto \Phi(z_0, z)$ is not the null function and for any $z_0 \in \mathbb{C}^{d_2}$, $z \mapsto \Phi(z, z_0)$ is not the null function.

Obviously, if no centering constraint is put on the signal or on the noise, it is possible to translate the signal by a fixed vector $m \in \mathbb{R}^d$ and the noise by $-m$ without changing the observation. The model can thus be identifiable only up to translation.

Theorem 1 (from [7])**.** *If the distribution of the signal satisfies A(ρ) and A(dep), then the distribution of the signal and the distribution of the noise can be recovered from the distribution of the observation, up to translation.*

An important step of the identifiabiity proof is to prove that, since the characteristic function of the distribution of the signal is a multivariate analytic function, it is enough to recover it in a neighborhood of 0. Thus, investigation of the characteristic functions outside a neigborhood of 0 is not needed, and decays of the characteristic function of the noise will have no impact on the convergence rates. In fact, as proved in Theorem 4 below when $d = 2$, the rate of convergence of the estimator of the density of the signal on the sphere will not depend on the unknown distribution of the noise.

The first step in the estimation procedure is the estimation of the characteristic function of the signal by a method inspired by the proof of the identifiability theorem. For any $S > 0$, let $\Upsilon_{\rho,S}$ be the subset of multivariate analytic functions from $\mathbb{C}^d$ to $\mathbb{C}$ defined as follows.

$$\Upsilon_{\rho,S} = \Bigg\{ \phi \text{ analytic s.t. } \forall z \in \mathbb{R}^d, \overline{\phi(z)} = \phi(-z), \phi(0) = 1$$
$$\text{and } \forall j \in \mathbb{N}^d \setminus \{0\}, \left| \frac{\partial^j \phi(0)}{\prod_{a=1}^d j_a!} \right| \le \frac{S^{\|j\|_1}}{\|j\|_1^{\|j\|_1/\rho}} \Bigg\}$$

where $\|j\|_1 = \sum_{a=1}^d j_a$. For all Φ satisfying A(ρ), there exists S such that $\Phi \in \Upsilon_{\rho,S}$ (Lemma 3.1 in [7]). Let $\Phi_{\varepsilon^{(p)}}$ be the characteristic function of $\varepsilon^{(p)}$, $p = 1, 2$, and define for all $\phi \in \Upsilon_{\rho,S}$ and any $\nu > 0$,

$$M(\phi; \nu|\Phi) = \int_{B_\nu^{d_1} \times B_\nu^{d_2}} |\phi(t_1, t_2)\Phi(t_1, 0)\Phi(0, t_2) - \Phi(t_1, t_2)\phi(t_1, 0)\phi(0, t_2)|^2$$
$$|\Phi_{\varepsilon^{(1)}}(t_1)\Phi_{\varepsilon^{(2)}}(t_2)|^2 dt_1 dt_2,$$

where $B_\nu = [-\nu, \nu]$. It is proved in [7] that if $\phi \in \Upsilon_{\rho,S}$ satisfies A(dep), $M(\phi; \nu|\Phi) = 0$ for a fixed ν if and only if $\phi = \Phi$ (up to translation). The estimator of the characteristic function of the signal can then be defined as a minimizer of the empirical estimator of M.
Fix some $\nu_{\text{est}} > 0$. Let $\mathcal{H}$ be a subset of functions from $\mathbb{R}^d$ to $\mathbb{C}^d$ such that all elements of $\mathcal{H}$ satisfy A(dep) and which is closed in $L^2(B^d_{\nu_{\text{est}}})$. Define $\widehat{\phi}_n$ as a (up to $1/n$) measurable minimizer of the functional M_n over $\Upsilon_{\rho,S} \cap \mathcal{H}$, where M_n is defined as

$$M_n(\phi) = \int_{B^{d_1}_{\nu_{\text{est}}} \times B^{d_2}_{\nu_{\text{est}}}} |\phi(t_1, t_2)\tilde{\phi}_n(t_1, 0)\tilde{\phi}_n(0, t_2) - \tilde{\phi}_n(t_1, t_2)\phi(t_1, 0)\phi(0, t_2)|^2 dt_1 dt_2,$$

where for all $(t_1, t_2) \in \mathbb{C}^{d_1} \times \mathbb{C}^{d_2}$,

$$\tilde{\phi}_n(t_1, t_2) = \frac{1}{n} \sum_{\ell=1}^n \exp\left\{ i t_1^\top Y_\ell^{(1)} + i t_2^\top Y_\ell^{(2)} \right\}.$$

It appears that, for any $\nu > 0$, $\widehat{\phi}_n$ is a consistent estimator of Φ in $L^2([-\nu, \nu]^d)$ at almost parametric rate. The constants c_1,c_2 and c_3 in Proposition 1 will depend on the signal through ρ and S, and on the noise through its second moment and the following quantity:

$$c_\nu = \inf\{|\Phi_{\varepsilon^{(1)}}(t)|,\ t \in B_\nu^{d_1}\} \wedge \inf\{|\Phi_{\varepsilon^{(2)}}(t)|,\ t \in B_\nu^{d_2}\}. \tag{2}$$

For any noise distribution, for small enough ν, c_ν is a positive real number. Moreover, notice that for fixed $\rho > 0$, for $S \le S'$, we have $\Upsilon_{\rho,S} \subset \Upsilon_{\rho,S'}$, that is, for S large enough, we can find $\nu \in [(d + 4/3)e/S, \nu_{\text{est}}]$ such that $c_\nu > 0$. We prove the following.

Proposition 1. *Assume $\Phi \in \Upsilon_{\rho,S} \cap \mathcal{H}$ and ε_1 has finite variance. Fix some $\nu \in [(d + 4/3)e/S, \nu_{est}]$ such that $c_\nu > 0$. For all $\delta \in (0, 1)$, there exist positive* *constants* c_1, c_2, c_3 *which depend on* δ, ν_{est}, ν, c_ν,ρ, S, $\mathcal{H}$, d *and* $E(\|Y_1\|^2)$ *such that for all* $x \geq 1$ *and* $n \geq (1 \vee xc_1)/c_2$, *with probability at least* $1 - e^{-x}$,

$$\int_{B_\nu^d} |\widehat{\phi}_n(t) - \Phi(t)|^2 dt \leq c_3 \left(\frac{x}{n^{1-\delta}} \vee \frac{x^2}{n^{2-2\delta}} \right).$$

Proposition 1 improves on Proposition A.3 in [7] and is proved in Section 7.1.

2.2. *Identifiability theorem*

For any $Z \in \mathbb{R}^d$, denote $Z^{(1)}, \ldots, Z^{(d)}$ its d coordinates. We shall parametrize a vector on a sphere through angles. For any $u \in [0,1]^{d-1}$, define $S(u)$ on the unit d-dimensional sphere as

$$S(u) = \begin{pmatrix} \cos(2\pi u^{(1)}) \\ \sin(2\pi u^{(1)}) \cos(\pi u^{(2)}) \\ \sin(2\pi u^{(1)}) \sin(\pi u^{(2)}) \cos(\pi u^{(3)}) \\ \vdots \\ \sin(2\pi u^{(1)}) \sin(\pi u^{(2)}) \cdots \sin(\pi u^{(d-2)}) \cos(\pi u^{(d-1)}) \\ \sin(2\pi u^{(1)}) \sin(\pi u^{(2)}) \cdots \sin(\pi u^{(d-2)}) \sin(\pi u^{(d-1)}) \end{pmatrix}.$$

Then for a sequence $(U_i)_{i\geq 1}$ of i.i.d random vectors taking values in $[0,1]^{d-1}$, we have for all $i \geq 1$,

$$X_i = C^\star + R^\star S(U_i), \tag{3}$$

with $C^\star$ the center of the sphere and $R^\star$ its radius.

We shall also make the following assumptions.

(H1) The coordinates $\varepsilon_1^{(1)}, \ldots, \varepsilon_1^{(d)}$ of the noise are independently distributed. We denote $\mathbb{Q}^\star = \otimes_{j=1}^d \mathbb{Q}_j^\star$ the distribution of ε_1, with $\mathbb{Q}_j^\star$ the distribution of $\varepsilon_1^{(j)}$, $j = 1, \ldots, d$.

(H2) The distribution of U_1 has a density $f^\star$ with respect to Lebesgue measure on $[0,1]^{d-1}$. When $d > 2$, we assume $f^\star$ positive on $(0, \zeta^\star) \times (0,1)^{d-2}$ for some $\zeta^\star > 0$.

We shall sometimes call $f^\star$ *exploration density of the angles* or *exploration density*. For any $\mathbb{Q} = \otimes_{j=1}^d \mathbb{Q}_j$, with $\mathbb{Q}_j$, $j = 1, \ldots, d$, probability distributions on $\mathbb{R}$, any probability density f on $[0,1]^{d-1}$, any $C \in \mathbb{R}^d$ and any $R \geq 0$, let $\mathbb{P}_{C,R,f,\mathbb{Q}}$ be the distribution of Y_1 when X_1 lies on the sphere with center C, radius R, and U_1 has density f.

Theorem 2. *Assume (H1) and (H2). For any* $\mathbb{Q} = \otimes_{j=1}^d \mathbb{Q}_j$, *any probability density* f *on* $[0,1]^{d-1}$, *any* $C \in \mathbb{R}^d$ *and any* $R \geq 0$, $\mathbb{P}_{C,R,f,\mathbb{Q}} = \mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}$ *if and only if* $R = R^\star$, $f = f^\star$, *and there exists* $m \in \mathbb{R}^d$ *such that* $C = C^\star + m$ *and* $\mathbb{Q}(\cdot) = \mathbb{Q}^\star(\cdot + m)$. *If moreover* $\mathbb{Q}$ *and* $\mathbb{Q}^\star$ *have finite first moment and are centered distributions, then* $m = 0$, *that is* $C = C^\star$ *and* $\mathbb{Q} = \mathbb{Q}^\star$.

Proof of Theorem 2.
We shall apply Theorem 1. For any probability density f on $[0,1]^{d-1}$, $C \in \mathbb{R}^d$ and $R \geq 0$, A(ρ) holds with $\rho = 1$ and with the constants $a = 1$ and $b = \|C\|_2 + R$. To verify A(dep), since all coordinates of the noise are independently distributed, we first choose a decomposition of the signal in two components. We define $\widetilde{X}^{(1)} = X_1^{(1)}$, and $\widetilde{X}^{(2)} = (X_1^{(2)}, \ldots, X_1^{(d)})^T$, and we prove in Section 7.2 the following Lemma from which A(dep) follows.

Lemma 1. *Assume (H2). Then for all $z \in \mathbb{C}$, $E\left[\exp\left(iz\widetilde{X}^{(1)}\right)|\widetilde{X}^{(2)}\right]$ is not $P_{\widetilde{X}^{(2)}}$-a.s. the null random variable, and for all $z \in \mathbb{C}^{d-1}$, $E\left[\exp\left(iz^T\widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]$ is not $P_{\widetilde{X}^{(1)}}$-a.s. the null random variable. Here, $P_{\widetilde{X}^{(p)}}$ denotes the distribution of $\widetilde{X}^{(p)}$, $p = 1, 2$.*

Then, translation of the spherical signal does not change the radius of the sphere and the exploration density of the angles on the sphere, but only the centering of the sphere and correspondingly the distribution of the noise. Applying Theorem 1 leads then to the conclusion of Theorem 2.

The proof of Lemma 1 proceeds by computing explicitly the conditional expectation, and then to give an argument why it can not be the null random variable. The argument for $d = 2$ does not apply to $d > 2$, in which case we use another argument needing the positivity of $f^\star$ near the origin. Since the choice of the positive first coordinate to define the angles and the density is arbitrary, the proof still holds under the assumption that the density is positive near the point of the sphere at the intersection with one of the $2d$ axis directions.

2.3. Estimation method and consistency

We shall apply the method described in Section 2.1 to estimate $R^\star$ and $f^\star$. For any positive real number R and any probability density f on $[0,1]^{d-1}$, define $\Psi_{f,R}$ the characteristic function of the random variable with distribution on the centered sphere with radius R, and exploration density of the angles f, that is, for all $t \in \mathbb{R}^d$,

$$\Psi_{f,R}(t) = \int_{(0,1)^{d-1}} \exp\left\{iRt^T S(u)\right\} f(u)du. \tag{4}$$

We shall consider functions $\Psi_{f,R}$ for any function f on $(0,1)^{d-1}$ (not only probability densities) as defined by (4). Notice that $\Psi_{f,R}$ can be extended to $\mathbb{C}^d$.
Since all components of ϵ_1 are independent, we have to make a choice of d_1 and d_2 for the definition of M_n and M, thus, in the following, we assume to have $d_1 = 1$ and $d_2 = d - 1$, as in the proof of Theorem 2. For any $\nu > 0$, define

$$\begin{aligned} M(f,R) = \int_{B_\nu \times B_\nu^{d-1}} & |\Psi_{f,R}(t_1,t_2)\Psi_{f^\star,R^\star}(t_1,0)\Psi_{f^\star,R^\star}(0,t_2) \\ & -\Psi_{f^\star,R^\star}(t_1,t_2)\Psi_{f,R}(t_1,0)\Psi_{f,R}(0,t_2)|^2 |\Phi_\epsilon(t_1,t_2)|^2 dt_1 dt_2. \end{aligned}$$

The parameter ν does not appear in the notation of M and can be chosen as needed.
Fix some $\nu_{\text{est}} > 0$, and define

$$M_n(f,R) = \int_{B_{\nu_{\text{est}}} \times B^{d-1}_{\nu_{\text{est}}}} |\Psi_{f,R}(t_1,t_2)\tilde{\psi}_n(t_1,0)\tilde{\psi}_n(0,t_2) \\ - \tilde{\psi}_n(t_1,t_2)\Psi_{f,R}(t_1,0)\Psi_{f,R}(0,t_2)|^2 dt_1 dt_2,$$

with

$$\tilde{\psi}_n(t_1,t_2) = \frac{1}{n}\sum_{\ell=1}^{n} \exp\left\{it_1\tilde{Y}_\ell^{(1)} + it_2^\top \tilde{Y}_\ell^{(2)}\right\}.$$

where for all ℓ, $\tilde{Y}_\ell^{(1)} = Y_\ell^{(1)}$, and $\tilde{Y}_\ell^{(2)} = (Y_\ell^{(2)}, \dots, Y_\ell^{(d)})^T$.
We need to fix the compact subset on which we minimize M_n. We choose $\mathcal{F}$ a compact subset of $\mathbb{L}^2[(0,1)^{d-1}]$ such that for all $f \in \mathcal{F}$, $\int_{(0,1)^{d-1}} f(u)du = 1$. For example, $\mathcal{F}$ can be the intersection of a regularity class such as a Sobolev ball with the closed subset of functions f such that $\int_{(0,1)^{d-1}} f(u)du = 1$, see Section 3.2 in the case $d = 2$. We also choose real numbers $R_{\min}$ and $R_{\max}$ such that $0 < R_{\min} < R_{\max} < +\infty$. Since we shall study minimax rates in Section 3, we shall fix later $\mathcal{F}$ to include all Sobolev classes of interest in that paper. Then we define $(\widehat{f}, \widehat{R})$ as any measurable random variable such that

$$M_n\left(\widehat{f}, \widehat{R}\right) \le \inf_{(f,R)\in\mathcal{F}\times[R_{\min};R_{\max}]} M_n(f,R) + \frac{1}{n}. \tag{5}$$

Notice that we do not constrain functions in $\mathcal{F}$ to be non negative, that is we do not constrain $\widehat{f}$ to be a probability density. Using Proposition 1 we get the following corollary which will be the basic stone to obtain estimation rates of our estimators. For any $\nu > 0$, define $c_\nu^\star$ as in (2) with $\mathbb{Q} = \mathbb{Q}^\star$.

Corollary 1. *Assume $f^\star \in \mathcal{F}$, $R^\star \in (R_{\min}; R_{\max})$ and ε_1 has finite variance. For all $\nu \in (0, \nu_{est}]$ such that $c_\nu^\star > 0$, for all $\delta \in (0,1)$, there exist a positive constants c_1, c_2, c_3 which depend on δ, ν, $c_\nu^\star$, d and $E(\|Y\|^2)$ such that for all $x \ge 1$ and $n \ge (1 \vee xc_1)/c_2$, with probability at least $1 - e^{-x}$,*

$$\int_{[-\nu,\nu]^d} |\Psi_{\widehat{f},\widehat{R}}(t) - \Psi_{f^\star,R^\star}(t)|^2 dt \le c_3\left(\frac{x}{n^{1-\delta}} \vee \frac{x^2}{n^{2-2\delta}}\right).$$

We insist on the fact that the quantity $c_\nu^\star$ is unknown, and that its knowledge is not needed to construct the estimators and to get asymptotic rates, since there always exists a small enough ν such that $c_\nu^\star > 0$.
For any $\nu > 0$, $c(\nu) > 0$, $E > 0$, define $\mathcal{Q}^{(d)}(\nu, c(\nu), E)$ the set of distributions $\mathbb{Q} = \otimes_{j=1}^d \mathbb{Q}_j$ on $\mathbb{R}^d$ such that $c_\nu \ge c(\nu)$ and $\int_{\mathbb{R}^d} \|x\|^2 d\mathbb{Q}(x) \le E$. The following corollary gives an upper bound of the maximum risk for the integrated square loss, showing convergence at rate $n^{1-\delta}$ for any positive δ.

Corollary 2. *For all $\nu \in (0, \nu_{est}]$, $c(\nu) > 0$, $E > 0$ and $\delta \in (0,1)$, there exists a positive constant C which only depends on δ, ν, $c(\nu)$, d, $E(\|Y\|^2)$, $R_{\min}$ and*

$R_{\max}$ such that for n large enough, uniformly for $f \in \mathcal{F}, R \in [R_{\min}; R_{\max}], \mathbb{Q} \in \mathcal{Q}^{(d)}(\nu, c(\nu), E), C \in \mathbb{R}^2$,

$$\mathbb{E}_{C,R,f,\mathbb{Q}}\left[\int_{[-\nu,\nu]^d} |\Psi_{\widehat{f},\widehat{R}}(t) - \Psi_{f,R}(t)|^2 dt\right] \leq \frac{C}{n^{1-\delta}}.$$

When $\mathbb{Q}^\star$ has a finite first moment and is centered, we can estimate the center of the sphere. We define

$$\widehat{C} = \frac{1}{n}\sum_{i=1}^{n} Y_i - \widehat{R}\int_{(0,1)^{d-1}} S(u)\widehat{f}(u)du. \tag{6}$$

The estimators of the radius and of the exploration density can be proved to be consistent by applying M-estimator general results. Then consistency of the estimator of the radius follows. We give a detailed proof of the following proposition in Section 7.3. Here, it is not needed that the noise has finite variance.

Proposition 2. *Assume $f^\star \in \mathcal{F}$ and $R^\star \in (R_{\min}; R_{\max})$. Then $\widehat{R} = R^\star + o_{\mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}}(1)$ and $\int_{(0,1)^{d-1}}(\widehat{f}(u) - f^\star(u))^2 du = o_{\mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}}(1)$. If moreover $\mathbb{Q}^\star$ has finite first moment and is a centered distribution, then also $\widehat{C} = C^\star + o_{\mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}}(1)$.*

Comments on the practical computation of the estimator. In practice, computing the minimum over the infinite-dimensional set defined in (5) requires to introduce a truncation parameter. In other words, instead of minimizing M_n over all elements (f, R) of $\mathcal{F} \times [R_{\min}, R_{\max}]$, we would minimize it over all $(T_m f, R)$, where m is a truncation parameter and $T_m f$ is the truncated Fourier expansion of f (also defined in (11)). This truncation has no impact on the result proved in Theorem 3, Theorem 4 and Theorem 5 i.e. on the convergence rates derived in this paper, as long as this truncation parameter is chosen sufficiently large with respect to $\log n/\log\log n$ to obtain the rates for the estimation of $\Psi_{f^\star,R^\star}$ (see the end of Section 3.2 of [7]).

3. Convergence rates of the estimators

In this section, we prove that the estimator of the radius has almost parametric rate of convergence, whatever the dimension d of the sphere. We then get rates of convergence for the estimator of the exploration density and of the center in the case $d = 2$ that is for circular signals. Our estimator of the exploration density achieves the minimax rate on Sobolev regularity classes and the estimator of the center can be proved to have almost parametric rate.

3.1. The estimator of the radius

Our first main result is the fact that, without any knowledge of the noise distribution and of the exploration density, the radius of the sphere can be recovered at almost parametric rate.

Theorem 3. *Assume $f^\star \in \mathcal{F}$ and $R^\star \in (R_{\min}; R_{\max})$. Assume also that ε_1 has finite variance. For all $\nu \in (0, \nu_{est}]$ such that $c^\star_\nu > 0$, for all $\delta \in (0,1)$, there exist positive constants c_1, c_2, c_3 which depend on δ, ν, $c^\star_\nu$, d and $E(\|Y\|^2)$ such that for all $x \geq 1$ and $n \geq (1 \vee xc_1)/c_2$, with probability at least $1 - e^{-x}$,*

$$|\widehat{R} - R^\star|^2 \leq \frac{c_3}{R_{\min}^2 (1 - \frac{(\nu R_{\max})^2}{2d+8})^2} \left(\frac{x}{n^{1-\delta}} \vee \frac{x^2}{n^{2-2\delta}} \right).$$

Proof. We denote by Δ the Laplacian operator in the Cartesian coordinate system,

$$\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2},$$

and for all $k \geq 1$, $\Delta^k = \Delta^{k-1} \circ \Delta$ with Δ^0 the identity operator.
Notice that $\Psi_{f,R}$ is an eigenfunction of the Laplacian, with eigenvalue R^2,

$$\Delta \Psi_{f,R}(x) + R^2 \ \Psi_{f,R}(x) = 0,$$

so that for all $k \geq 1$,

$$\Delta^k \Psi_{f,R}(x) + (-1)^{k-1} R^{2k} \ \Psi_{f,R}(x) = 0.$$

Define $\mathbb{D}_d(0,\nu) = \{x \in \mathbb{R}^d, \ \|x\|_2 \leq \nu\}$ the d-dimensional disk centered at the origin and with radius ν, Γ the gamma function and λ_d the d-dimensional Lebesgue measure. Then, according to [15], for all $\nu > 0$ and for all multivariate analytic function ψ on $\mathbb{C}^d$,

$$\begin{aligned}
\frac{1}{\lambda_d(\mathbb{D}_d(0,\nu))} \int_{\mathbb{D}_d(0,\nu)} \psi(x) dx &= \sum_{k=0}^{\infty} \frac{\Delta^k \psi(0)}{2^k k! \prod_{j=1}^{k} (d+2j)} \nu^{2k} \\
&= \Gamma\left(\frac{d}{2} + 1\right) \sum_{k=0}^{\infty} \frac{\Delta^k \psi(0)}{2^{2k} k! \Gamma(\frac{d}{2} + k + 1)} \nu^{2k}.
\end{aligned}$$

Applying this equality to $\Psi_{\widehat{f},\widehat{R}}$ and to $\Psi_{f^\star,R^\star}$ we get that

$$\frac{1}{(\sqrt{\pi}\nu)^d} \int_{\mathbb{D}_d(0,\nu)} \left(\Psi_{\widehat{f},\widehat{R}}(x) - \Psi_{f^\star,R^\star}(x) \right) dx = \sum_{k=0}^{\infty} (-1)^k \frac{\widehat{R}^{2k} - (R^\star)^{2k}}{2^{2k} k! \Gamma(\frac{d}{2} + k + 1)} \nu^{2k}, \tag{7}$$

since $\lambda^d(\mathbb{D}_d(0,\nu)) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)} \nu^d$.
Let $J_{d/2}$ be the Bessel function of order $(d/2)$. We collect in Section 8 results on Bessel functions that will be useful in our analysis. Using identity (I) in Section 8 we get that for all $x \in \mathbb{R}$,

$$\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^{2k} k! \Gamma(\frac{d}{2} + k + 1)} = 2^{d/2} \frac{J_{d/2}(x)}{x^{d/2}}, \tag{8}$$

so that using (7), (8), Cauchy-Schwarz inequality and the fact that $\mathbb{D}_d(0,\nu) \subset B_\nu^d$ we obtain

$$2^d \left| \frac{J_{d/2}(\nu \widehat{R})}{(\nu \widehat{R})^{d/2}} - \frac{J_{d/2}(\nu R^\star)}{(\nu R^\star)^{d/2}} \right|^2 \leq \frac{1}{(\sqrt{\pi}\nu)^d \Gamma(\frac{d}{2}+1)} \int_{B_\nu^d} \left(\Psi_{\widehat{f},\widehat{R}}(x) - \Psi_{f^\star,R^\star}(x) \right)^2 dx. \tag{9}$$

Let H be the function defined by

$$\forall x \neq 0,\ H(x) = \frac{J_{d/2}(x)}{x^{d/2}},\ H(0) = \frac{1}{2^{d/2}\Gamma(\frac{d}{2}+1)}.$$

Then using (8), H has infinitely many derivatives so that there exists $\widetilde{R} \in (R^\star, \widehat{R})$ such that

$$H(\nu \widehat{R}) - H(\nu R^\star) = \nu(\widehat{R} - R^\star) H'(\nu \widetilde{R}).$$

Computation of the derivative and (IV) in Section 8 gives

$$H'(x) = \frac{x(J_{d/2-1}(x) - J_{d/2+1}(x)) - dJ_{d/2}(x)}{2x^{d/2+1}},$$

and using (V) in Section 8 we get

$$H'(x) = -\frac{J_{d/2+1}(x)}{x^{d/2}}.$$

Using lemma 3 in Section 8, we get that

$$J_{d/2+1}(\nu \widetilde{R}) \geq \frac{(\nu \widetilde{R})^{d/2+1}}{2^{d/2+1}\Gamma(\frac{d}{2}+2)} \left(1 - \frac{(\nu \widetilde{R})^2}{2d+8} \right).$$

Since $\widetilde{R} \in (R_{\min}, R_{\max})$, we deduce that for any $\nu \in (0, 1/R_{\max})$,

$$\begin{aligned} |H'(\nu \widetilde{R})| = \left| \frac{J_{d/2}(\nu \widetilde{R})}{(\nu \widetilde{R})^{d/2}} \right| &\geq \frac{\nu \widetilde{R}}{2^{d/2+1}\Gamma(\frac{d}{2}+2)} \left| 1 - \frac{(\nu \widetilde{R})^2}{2d+8} \right| \\ &\geq \frac{\nu R_{\min}}{2^{d/2+1}\Gamma(\frac{d}{2}+2)} \left(1 - \frac{(\nu R_{\max})^2}{2d+8} \right) > 0, \end{aligned}$$

so that

$$|\widehat{R} - R^\star| \leq \frac{2^{d/2+1}\Gamma(\frac{d}{2}+2)}{\nu^2 R_{\min}(1 - \frac{(\nu R_{\max})^2}{2d+8})} |H(\nu \widehat{R}) - H(\nu R^\star)|.$$

Using (9) we get

$$|\widehat{R} - R^\star|^2 \leq \frac{4\Gamma(\frac{d}{2}+2)^2}{\nu^4 (\sqrt{\pi}\nu)^d R_{\min}^2 (1 - \frac{(\nu R_{\max})^2}{2d+8})^2} \int_{B_\nu^d} \left(\Psi_{\widehat{f},\widehat{R}}(x) - \Psi_{f^\star,R^\star}(x) \right)^2 dx.$$

The end of the proof follows from Corollary 1. □

The following corollary gives an upper bound for the rate of convergence of the maximum risk.

Corollary 3. *For all $\nu \in (0, \nu_{est}]$, $c(\nu) > 0$, $E > 0$ and $\delta \in (0,1)$, there exists a positive constant C which only depends on δ, ν, $c(\nu)$, d, $E(\|Y\|^2)$, $R_{\min}$ and $R_{\max}$ such that for n large enough,*

$$\sup_{f\in\mathcal{F}, R\in[R_{\min};R_{\max}], \mathbb{Q}\in\mathcal{Q}^{(d)}(\nu,c(\nu),E), C\in\mathbb{R}^2} \mathbb{E}_{C,R,f,\mathbb{Q}}|\widehat{R}-R|^2 \leq \frac{C}{n^{1-\delta}}.$$

3.2. *The estimator of the density and of the center*

In this section, we consider the case of circular signals, that is $d = 2$. In this case, we can rewrite model (3) using one dimensional angles $U_i \in [0,1]$, as

$$X_i = C^\star + R^\star \begin{pmatrix} \cos(2\pi U_i) \\ \sin(2\pi U_i) \end{pmatrix}. \tag{10}$$

We shall focus on the following regularity classes. For any $L > 0$, $\beta > \frac{1}{2}$ and $\gamma > 0$, set

$$W_\beta(L) = \{f \in \mathbb{L}^2([0,1]) : \sum_{k=-\infty}^{\infty} |f_k|^2 |k|^{2\beta} \leq L^2\},$$

$$A_\gamma(L) = \{f \in \mathbb{L}^2([0,1]) : \sum_{k=-\infty}^{\infty} |f_k|^2 e^{2\gamma k} \leq L^2\},$$

where for any function $f \in \mathbb{L}^2([0,1])$, $(f_k)_{k\in\mathbb{Z}}$ is the sequence of Fourier coefficients of f:

$$f_k = \int_0^1 f(\theta) e^{2i\pi k\theta} d\theta, \quad k \in \mathbb{Z}.$$

We fix $\mathcal{F}$ as a compact subset of $\mathbb{L}^2[(0,1)]$ such that for all $f \in \mathcal{F}$, $\int_0^1 f(u)du = 1$, and containing as subsets all $W_\beta(L)$ and $A_\gamma(L)$ for all $\beta > \frac{1}{2}$, $\gamma > 0$, and $L \leq L_{\max}$ chosen. If a lower bound $\beta_0 > 1/2$ on β is known, we can choose $\mathcal{F} = W_{\beta_0}(L_{\max}) \cap \{f \in \mathbb{L}^2([0,1]) : \int_0^1 f(u)du = 1\}$. We shall now define an estimator of $f^\star$ using truncated Fourier expansions of $\widehat{f}$ defined in Section 2.3. For $N > 0$ an integer to be chosen, we define $T_N\widehat{f}$ the trigonometric polynomial estimator of $f^\star$:

$$\forall x \in (0,1), \quad T_N\widehat{f}(x) = \sum_{|k|\leq N} \widehat{f}_k e^{-2i\pi kx}. \tag{11}$$

Define now the maximum risk of the estimator for any class of densities $\mathcal{C}$ and any class of noise distribution $\mathcal{Q}$ as follows.

$$R\left[T_N\widehat{f}; \mathcal{C}; R_{\min}; R_{\max}; \mathcal{Q}\right] =$$

$$\sup_{f\in\mathcal{C}, R\in[R_{\min};R_{\max}], \mathbb{Q}\in\mathcal{Q}, C\in\mathbb{R}^2} \mathbb{E}_{C,R,f,\mathbb{Q}}\left(\int_0^1 (T_N\widehat{f}(x) - f(x))^2 dx\right).$$

The following theorem shows that a good choice of N leads to minimax adaptive estimation rate over the regularity classes $W_\beta(L)$ and controlled maximum risk over the regularity classes $A_\gamma(L)$.

Theorem 4. *For $\alpha \in (0, 1/2)$, set*

$$N = \left\lfloor \alpha \frac{\log n}{\log \log n} \right\rfloor .$$

Then for all $L > 0$, $\beta > 1/2$, (resp. $\gamma > 0$), for all $\nu \in (0, \nu_{est}]$, $c(\nu) > 0$, $E > 0$,

$$R\left[T_N \widehat{f}; W_\beta(L); R_{\min}; R_{\max}; \mathcal{Q}^{(2)}(\nu, c(\nu), E)\right] \leq L^2 \alpha^{-2\beta} \left(\frac{\log\log(n)}{\log(n)}\right)^{2\beta} (1+o(1)) \tag{12}$$

as n tends to infinity, and

$$R\left[T_N \widehat{f}; A_\gamma(L); R_{\min}; R_{\max}; \mathcal{Q}^{(2)}(\nu, c(\nu), E)\right] \leq \exp\left(-2\gamma \left(\alpha \frac{\log(n)}{\log\log(n)}\right)\right)(1+o(1)) \tag{13}$$

as n tends to infinity.

In [10], the author studies the estimation of the exploration density for noisy circular data on the unit circle (known radius) and with known noise distribution. Theorem 1 in [10] proves that for centered 2-dimensional Gaussian noise having variance $\sigma^2 I_2$, the minimax rate for estimating f over $W_\beta(L)$ is $L^2 \left(\frac{\log\log(n)}{\log(n)}\right)^{2\beta} (1 + o(1))$. Thus, our estimator is rate minimax adaptive to unknown radius, unknown noise distribution and unknown regularity over classes $W_\beta(L)$ for the signal, with a constant deteriorated by a factor at most $2^{2\beta}$. Theorem 2 in [10] proves that with the same noise, the minimax rate for estimating f over $A_\gamma(L)$ is $L^2 \exp\left(-2\gamma\left(\frac{\log(n)}{\log\log(n)}\right)\right)(1 + o(1))$. In our result, there is a loss in the upper bound for the rate of convergence of the maximum risk of our estimator in case of unknown radius and unknown noise distribution on classes $A_\gamma(L)$ for the signal.

Note that, as in [10] and [7], the convergence rates do not depend on the error distribution. This is in contrast to the standard density deconvolution problem where, when we know the error distribution, the optimal convergence rate depends on the rate of decrease of the Fourier transform of the noise distribution (over classes of noise distributions satisfying such decay), see [14] and references therein. Here, analyticity of the Fourier transform of the signal distribution allows us to dispense with the knowledge of the noise, in particular of the decay rate of its Fourier transform. It would be interesting, in the context of unknown noise, to recover noise dependent minimax risk by restricting the set of possible unknown noises. One way could be to make in our methodology $\nu = \nu_{\text{est}}$ go to infinity and to study the square integrated risk with $c_{\nu_{\text{est}}}$ having a precise decreasing behavior. But this would require new ideas as explained in the concluding section of [7].

Proof. For any $f \in \mathcal{F}$,

$$\int_0^1 \left(T_N\widehat{f}(x) - f(x)\right)^2 dx = \sum_{|k|\leq N} |f_k - \widehat{f}_k|^2 + \sum_{|k|>N} |f_k|^2,$$

so that for any $R \in (R_{\min}; R_{\max})$, $\mathbb{Q} \in \mathcal{Q}(\nu, c(\nu), E)$, $C \in \mathbb{R}^2$,

$$\mathbb{E}_{C,R,f,\mathbb{Q}}\left(\int_0^1 (T_N\widehat{f}(x) - f(x))^2 dx\right) = \mathbb{E}_{C,R,f,\mathbb{Q}}\left[\sum_{|k|\leq N} |f_k - \widehat{f}_k|^2\right] + \sum_{|k|>N} |f_k|^2.$$

The first term on the right hand side will be shown to be negligible with respect to the second term thanks to the following proposition, for which a detailed proof can be found in Section 7.4

Proposition 3. *Assume $f^\star \in \mathcal{F}$ and $R^\star \in (R_{\min}, R_{\max})$. For all $\nu \in (0, \nu_{est}]$ such that $c_\nu^\star > 0$, for all $\delta \in (0,1)$, there exists a constant $c > 0$ depending on δ, ν, $c_\nu^\star$, d, $R^\star$, $R_{\min}$, $R_{\max}$, and $E(\|Y\|^2)$ such that for all $x \geq 1$, $n \geq (1 \vee xc_1)/c_2$, with probability at least $1 - e^{-x}$,*

$$\sum_{|k|\leq N} |f_k^\star - \widehat{f}_k|^2 \leq c\left(\frac{2}{\nu R_{\min}}\right)^{2N} (N+1)(N!)^2 \left(\frac{x}{n^{1-\delta}} \vee \frac{x^2}{n^{2-2\delta}}\right). \tag{14}$$

Choose δ small enough so that $2\alpha < 1 - \delta$. Then for large enough n, for a constant $c > 0$,

$$\mathbb{E}_{C,R,f,\mathbb{Q}}\left[\sum_{|k|\leq N} |f_k - \hat{f}_k|^2\right] \leq c\ (\nu R_{\min})^{-2N}(N+1)2^{2N}(N!)^2 n^{-1+\delta},$$

and using the fact that, $\forall N \geq 1$, $N! \leq eN^{N+\frac{1}{2}}e^{-N}$, we finally have,

$$\mathbb{E}_{C,R,f,\mathbb{Q}}\left[\sum_{|k|\leq N} |f_k - \hat{f}_k|^2\right] \leq e^2 c\ (\nu R_{\min})^{-2N}(N+1)2^{2N}N^{2N+1}e^{-2N}n^{-1+\delta}. \tag{15}$$

The term at the right hand side of (15) is at most of order

$$\exp\left\{(2\alpha + \delta - 1)\ \log(n)\left[1 + o(1)\right]\right\}.$$

Now,

$$\sup_{f \in W_\beta(L)} \sum_{|k|>N} |f_k|^2 \leq L^2 N^{-2\beta}, \tag{16}$$

and

$$\sup_{f \in A_\gamma(L)} \sum_{|k|>N} |f_k|^2 \leq L^2 e^{-2\gamma N}. \tag{17}$$

Equation (12) follows from (15) and (16), and equation (13) follows from (15) and (17). □

Theorem 5. *Assume $f^\star \in \mathcal{F}$ and $R^\star \in (R_{\min}, R_{\max})$. Then for any $\delta \in (0,1)$, for all $\nu \in (0, \nu_{est}]$, $c(\nu) > 0$, $E > 0$,*

$$\sup_{f\in\mathcal{F}, R\in[R_{\min};R_{\max}], \mathbb{Q}\in\mathcal{Q}^{(2)}(\nu,c(\nu),E), C\in\mathbb{R}^2} \mathbb{E}_{C,R,f,\mathbb{Q}} \left\| \widehat{C} - C^\star \right\|^2 = O\left(n^{-1+\delta}\right).$$

Notice that we can not get exponential deviations for the empirical mean of the observations when nothing more is assumed about the noise apart having finite variance.

Proof. Notice that

$$f_1^\star = \int_0^1 \cos(2\pi u) f^\star(u) du + i \int_0^1 \sin(2\pi u) f^\star(u) du$$

so that

$$C^\star = \mathbb{E}[Y] - R^\star \begin{pmatrix} \mathrm{Re}(f_1^\star) \\ \mathrm{Im}(f_1^\star) \end{pmatrix},$$

and in the same way

$$\widehat{C} = \frac{1}{n}\sum_{l=1}^n Y_l - \widehat{R} \begin{pmatrix} \mathrm{Re}(\widehat{f}_1) \\ \mathrm{Im}(\widehat{f}_1) \end{pmatrix}.$$

Thus, using the triangle inequality, and the fact that $\widehat{R} \le R_{\max}$ and $|f_1^\star| \le 1$, we get

$$\|\widehat{C} - C^\star\|_2^2 \le 3 \left\| \frac{1}{n}\sum_{l=1}^n Y_l - \mathbb{E}[Y] \right\|_2^2 + 3|\widehat{R} - R^\star|^2 + 3R_{\max} \left\| \begin{pmatrix} \mathrm{Re}(f_1^\star) - \mathrm{Re}(\widehat{f}_1) \\ \mathrm{Im}(f_1^\star) - \mathrm{Im}(\widehat{f}_1) \end{pmatrix} \right\|_2^2.$$

The theorem follows from Theorem 3, Proposition 3 and the upper bound on the variance of the observations.

□

4. When the exploration density is known

In this section, we assume that $f^\star$ is known. By exchanging the role of the signal and of the noise, we can look at model (1) as a semi-parametric deconvolution problem in which the noise has known distribution (up to centering and radius) on a sphere. But we are able to estimate the radius and the center without solving the semi-parametric deconvolution problem, that is without estimating $\mathbb{Q}$. We estimate the radius using the contrast function $M_n(f^\star, R)$. Since this function is continuous, we can define

$$\widetilde{R} = \ \mathrm{Argmin}\ \{M_n(f^\star, R),\ R \in [R_{\min}; R_{\max}]\}.$$

If moreover $\mathbb{Q}^\star$ has finite first moment and is a centered distribution, then the estimator of the center is defined as

$$\widetilde{C} = \frac{1}{n}\sum_{i=1}^{n} Y_i - \widetilde{R}\int_{(0,1)^{d-1}} S(u) f^\star(u)du.$$

Theorem 6 states that $\sqrt{n}(\widetilde{R} - R^\star, \widetilde{C} - C^\star)$ converges in distribution as n tends to infinity to some centered Gaussian distribution. It will be a consequence of the lemma stated below. In the following, for $R \in [R_{\min}, R_{\max}]$, we omit $f^\star$ as an argument of M and M_n. We write M', M_n' their derivatives with respect to R and M'', M_n'' their second derivatives with respect to R.

Lemma 2. *The following results hold true under the assumptions of Theorem 6.*

(1) $\widetilde{R}$ is a consistent estimator of $R^\star$.
(2) There exists a matrix V such that $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i - E(Y_1), M_n'(R^\star)\right)$ converges in distribution to a centered Gaussian distribution with variance V.
(3) $M''(R^\star) \neq 0$ and for any random variable $R_n \in [R_{\min}, R_{\max}]$ converging in probability to $R^\star$, one has

$$M_n''(R_n) = M''(R^\star) + o_{\mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}}(1).$$

The proof of Lemma 2 is given in Section 7.5.
Define the $(d+1)\times(d+1)$ matrix

$$\Sigma = \begin{pmatrix} 0 & -\frac{1}{M''(R^\star)} \\ 1 & \frac{E(S(U))}{M''(R^\star)} \end{pmatrix} V \begin{pmatrix} 0 & 1 \\ -\frac{1}{M''(R^\star)} & \frac{E(S(U))^T}{M''(R^\star)} \end{pmatrix}$$

Theorem 6. *Assume that $\mathbb{Q}^\star$ has finite second moment and is a centered distribution. Then $\sqrt{n}(\widetilde{R} - R^\star, \widetilde{C} - C^\star)$ converges in distribution to a centered Gaussian distribution with variance Σ.*

The proof of Theorem 6 is detailed in Section 7.6.

5. Simulations

The aim of this section is to illustrate our method with examples for which the noise is not bounded. We choose $d = 2$ and we consider the model (10) with $R^\star = 3$, $C^\star = 0$ and $R^\star = 0.6$, $C^\star = 0$, with U, ε generated as follows.

(1) $U \sim \text{Unif}(0,1)$ and $\varepsilon \sim \mathcal{N}(0,(0.12)^2 I)$, figure 1 ($R^\star = 3$) and figure 2 ($R^\star = 0.6$).
(2) $U \sim \text{Unif}(0,1)$ and for $i \in \{1,2\}$ $\varepsilon^{(i)} \sim \frac{1}{2}\delta_{(-1)} + \frac{1}{2}\mathcal{E}\text{xp}\left\{\frac{1}{0.12}\right\}$, figure 3 ($R^\star = 3$) and figure 4 ($R^\star = 0.6$).
(3) $U \sim \text{Unif}(0,1)$ and $\varepsilon \sim \mathcal{N}(0, I)$, figure 5 ($R^\star = 3$) and figure 6 ($R^\star = 0.6$).

(4) $U \sim f^\star : x \in (0,1) \mapsto \frac{\exp\{\cos(2\pi x)\}}{\int_0^1 \exp\{\cos(2\pi u)\}du}$ and $\varepsilon \sim \mathcal{N}(\begin{pmatrix} -1.6 \\ 2.5 \end{pmatrix}, \begin{pmatrix} (0.2)^2 & 0 \\ 0 & (0.57)^2 \end{pmatrix})$, figure 7 ($R^\star = 3$).

(5) $U \sim f^\star : x \in (0,1) \mapsto \frac{\exp\{\cos(2\pi x)\}}{\int_0^1 \exp\{\cos(2\pi u)\}du}$ and $\varepsilon \sim \mathcal{N}(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} (0.2)^2 & 0 \\ 0 & (0.57)^2 \end{pmatrix})$, figure 8 ($R^\star = 3$).

For each case, we generate n observed points for $n \in V$ with

$$V = \{10^2, 2\cdot 10^2, 3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, \\ 7.5\cdot 10^4, 10^5, 3\cdot 10^5, 5\cdot 10^5, 8\cdot 10^5, 10^6\}.$$

In the case the exploration density is known, we estimate only the radius of the circle. When the exploration density is unknown, we estimate the radius and the exploration density. To numerically visualize that the center of the noise has no impact on the estimation of the radius and the exploratory density, we can look at figure 7 and figure 8 .

In practice, when we want to estimate the radius and the exploration density, we fix N and ν_{est} and we minimize $M_n(T_N f, R)$ for $f \in \mathcal{F}$ and $R \in [R_{\min}; R_{\max}]$. Since N is fixed, $f_0 = 1$ and f is a real function, we have for all $k \in \mathbb{N}$, $\overline{f_k} = f_{-k}$, this amounts to minimize a function of $(2N+1)$ variables. In our simulations, we noticed that the choice of N and ν_{est} does not significantly change the results thus the simulations are done with $N = \left\lfloor \frac{\log(n)}{\log(\log(n))} \right\rfloor$ and $\nu_{est} = 0.5$. Nevertheless, this point remains to be studied further, especially to apply the method to real-life data. For each figure, there are 6 plots,

Top left Scatter plot of the 10^6 observed points.

Top right Scatter plot of the 10^6 observed points.

Middle left Plot of $(\log|\widehat{R} - R^\star|, \log(n))_{n\in V}$ + the linear regression, when the density $f^\star$ is known and unknown.

Middle right We choose $W \subset V$ to better visualize the graph, and we plot $(\widehat{R}, n)_{n\in W}$, when the density $f^\star$ is known and unknown.

Bottom left Plot of $f^\star$ and $T_N\widehat{f}$ for $n \in \{3\cdot 10^3, 10^4, 5\cdot 10^4, 10^5, 10^6\}$.

Bottom right Plot of $(\log||f^\star - T_N\widehat{f}||_2^2, \log(n))_{n\in V}$.

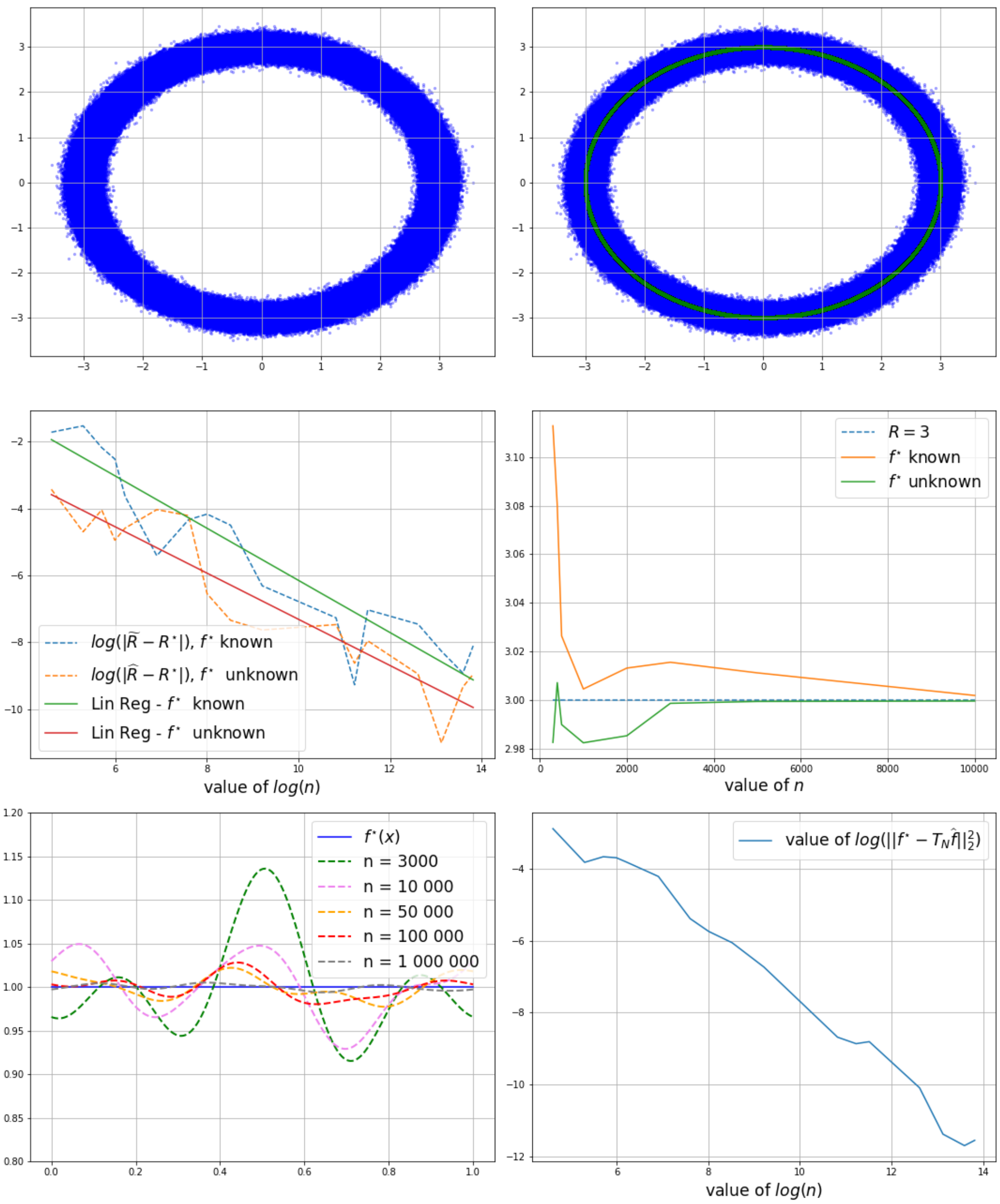

FIGURE 1. $R^\star = 3$, $U \sim Unif(0,1)$, $\varepsilon \sim \mathcal{N}(0,(0.12)^2 I)$ *and* $W = \{3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, 7.5\cdot 10^4, 10^5\}$

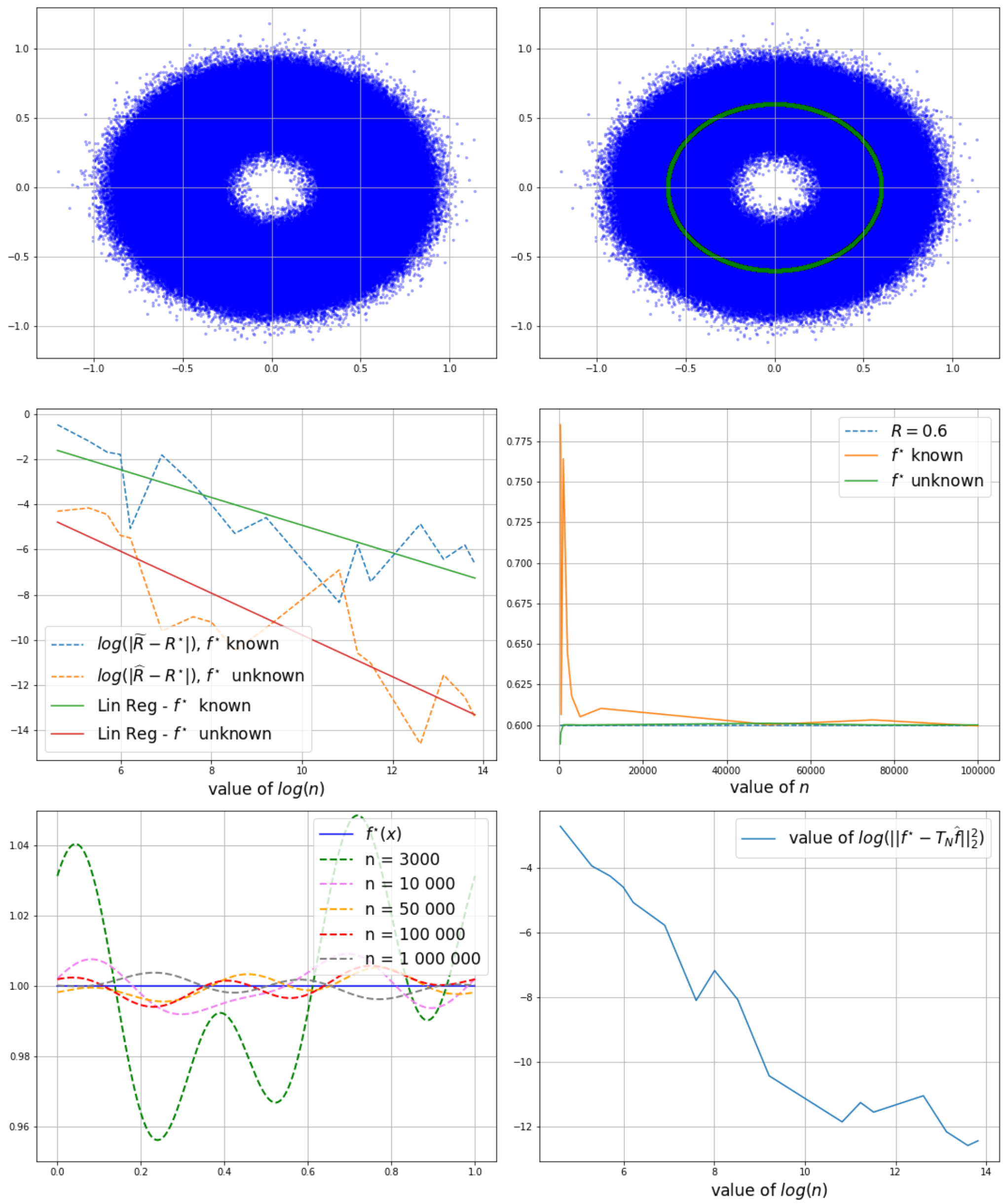

FIGURE 2. $R^\star = 0.6$, $U \sim Unif(0,1)$, $\varepsilon \sim \mathcal{N}(0,(0.12)^2 I)$ *and* $W = \{3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, 7.5\cdot 10^4, 10^5\}$

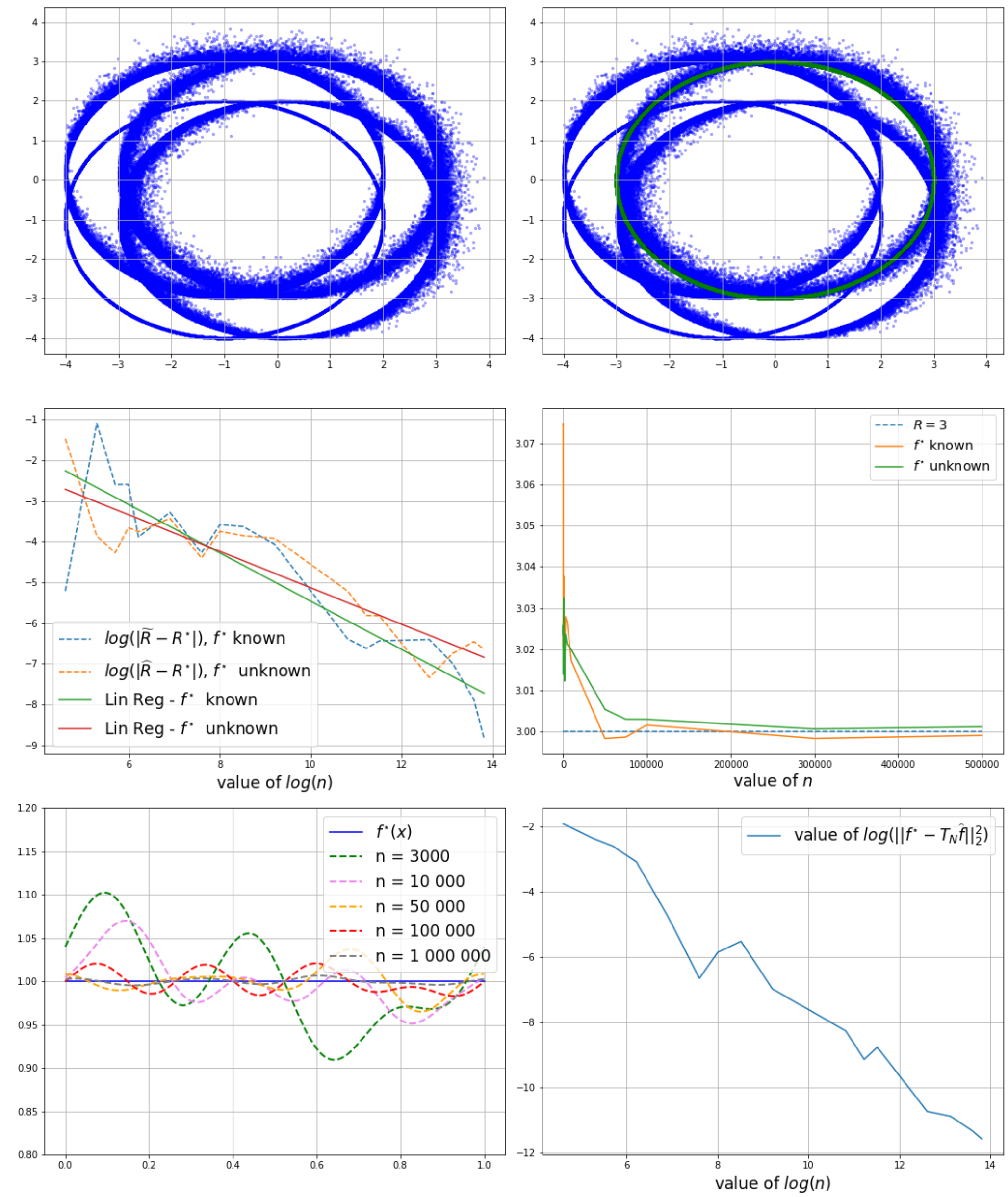

FIGURE 3. $R^\star = 3$, $U \sim Unif(0,1)$, for $i \in \{1,2\}$ $\varepsilon^{(i)} \sim \frac{1}{2}\delta_{(-1)} + \frac{1}{2}\mathcal{E}xp\left\{\frac{1}{0.12}\right\}$ and $W = \{3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, 7.5\cdot 10^4, 10^5, 3\cdot 10^5, 5\cdot 10^5\}$

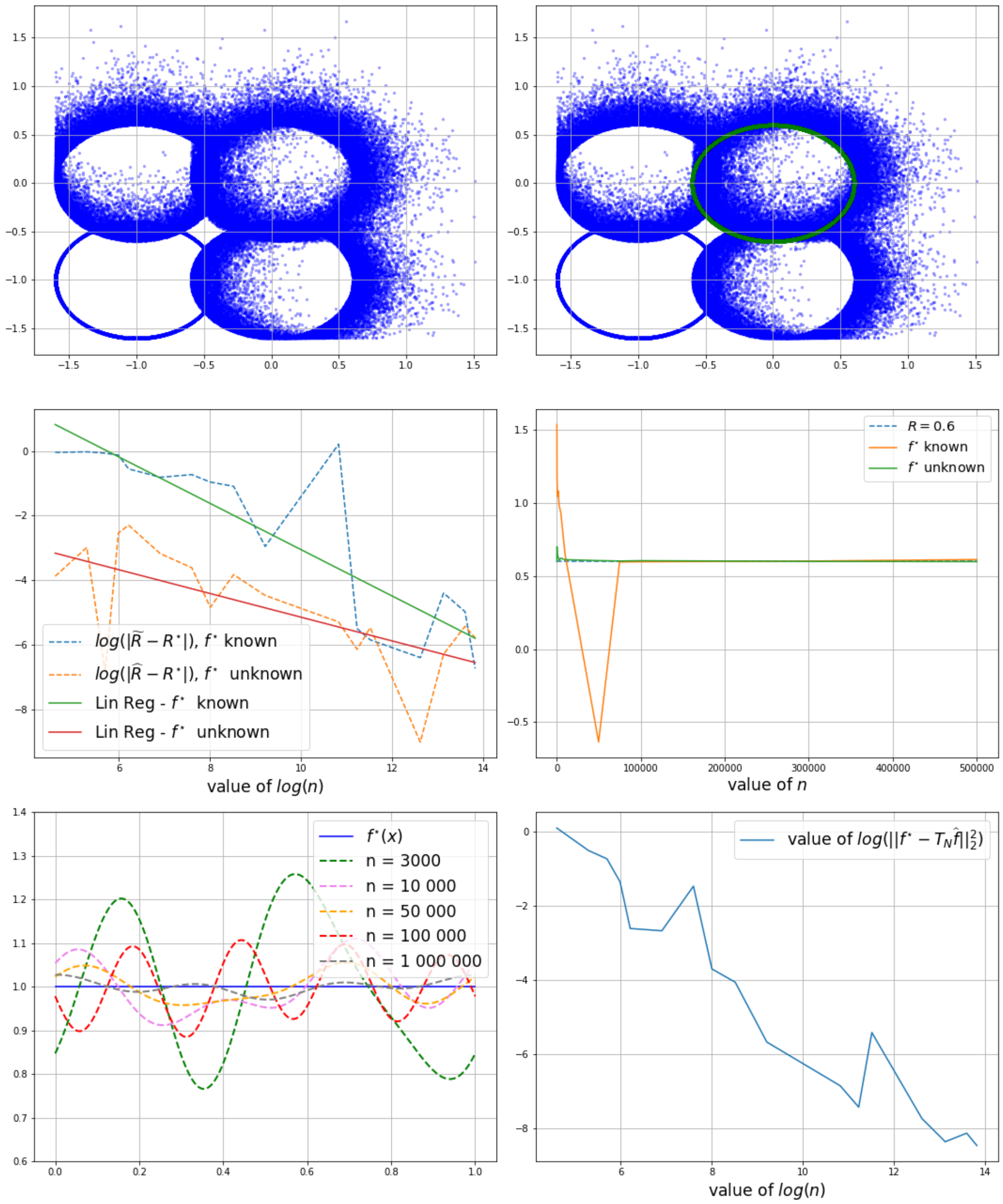

FIGURE 4. $R^{\star} = 0.6$, $U \sim Unif(0,1)$, for $i \in \{1,2\}$ $\varepsilon^{(i)} \sim \frac{1}{2}\delta_{(-1)} + \frac{1}{2}\mathcal{E}xp\left\{\frac{1}{0.12}\right\}$ and $W = \{3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, 7.5\cdot 10^4, 10^5, 3\cdot 10^5, 5\cdot 10^5\}$

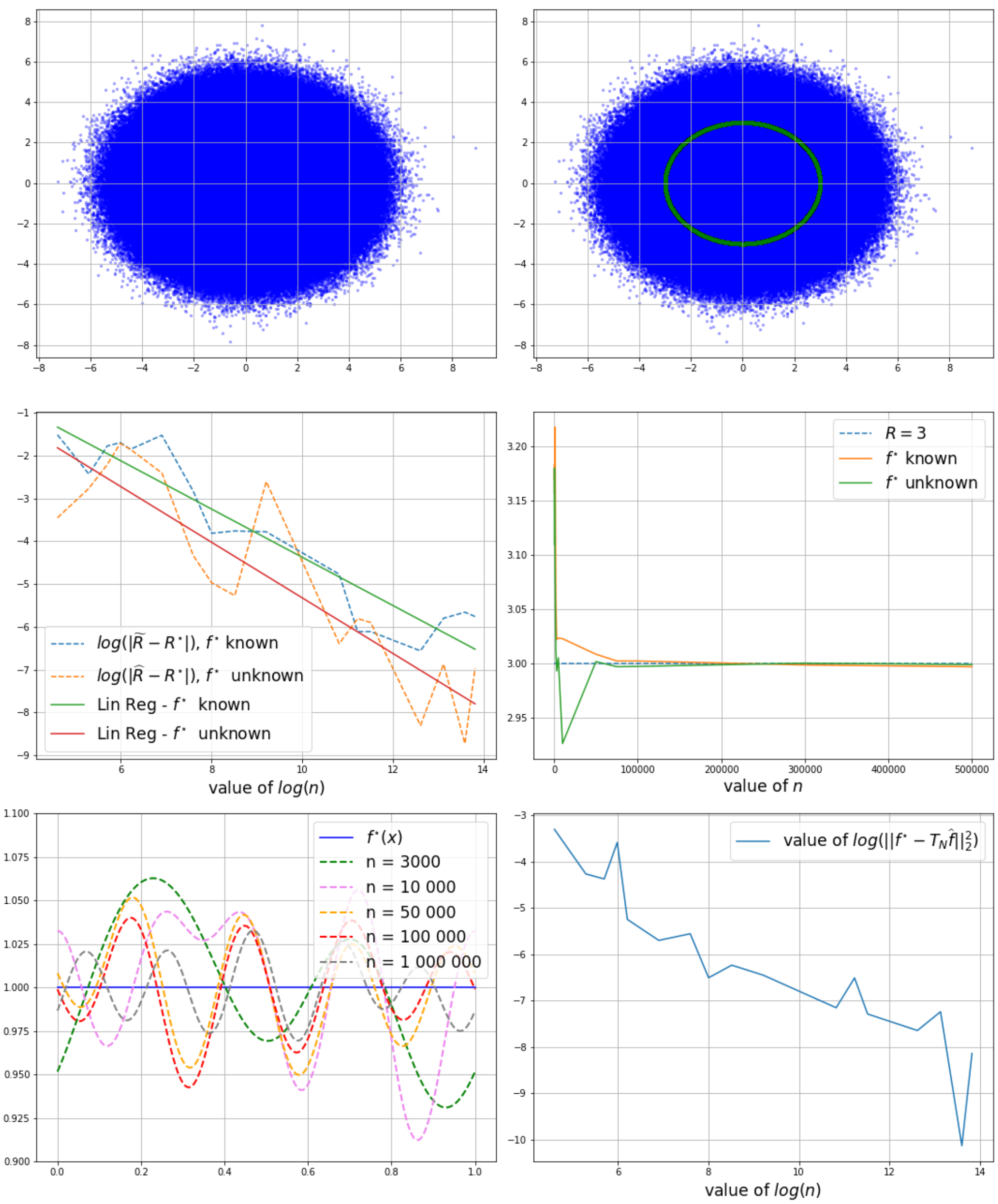

FIGURE 5. $R^\star = 3$, $U \sim Unif(0,1)$, $\varepsilon \sim \mathcal{N}(0, I)$ *and* $W = \{3 \cdot 10^2, 4 \cdot 10^2, 5 \cdot 10^2, 10^3, 2 \cdot 10^3, 3 \cdot 10^3, 5 \cdot 10^3, 10^4, 5 \cdot 10^4, 7.5 \cdot 10^4, 10^5, 3 \cdot 10^5, 5 \cdot 10^5\}$

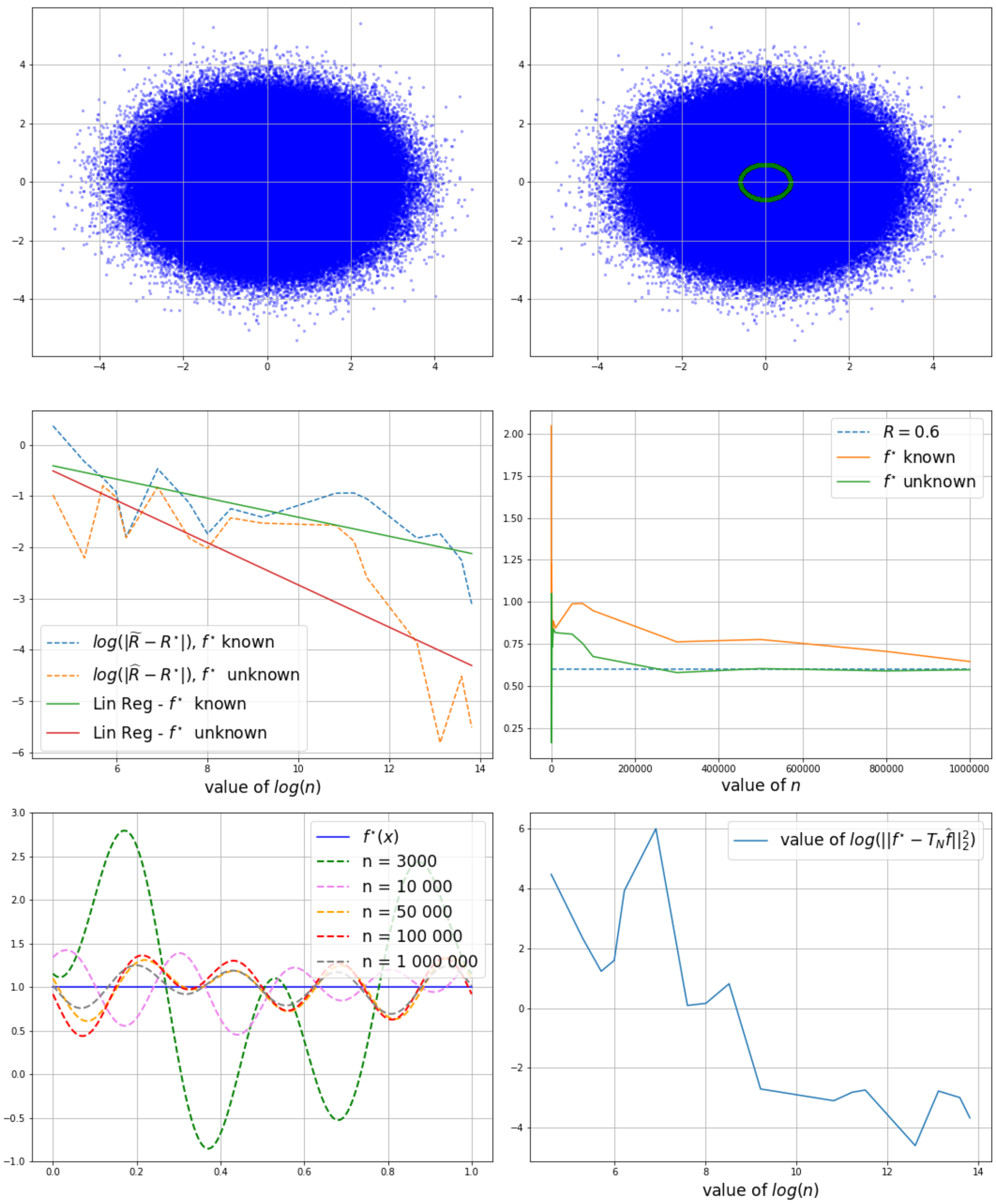

FIGURE 6. $R^{\star} = 0.6$, $U \sim Unif(0,1)$, $\varepsilon \sim \mathcal{N}(0, I)$ *and* $W = \{3 \cdot 10^2, 4 \cdot 10^2, 5 \cdot 10^2, 10^3, 2 \cdot 10^3, 3 \cdot 10^3, 5 \cdot 10^3, 10^4, 5 \cdot 10^4, 7.5 \cdot 10^4, 10^5, 3 \cdot 10^5, 5 \cdot 10^5, 8 \cdot 10^5, 10^6\}$

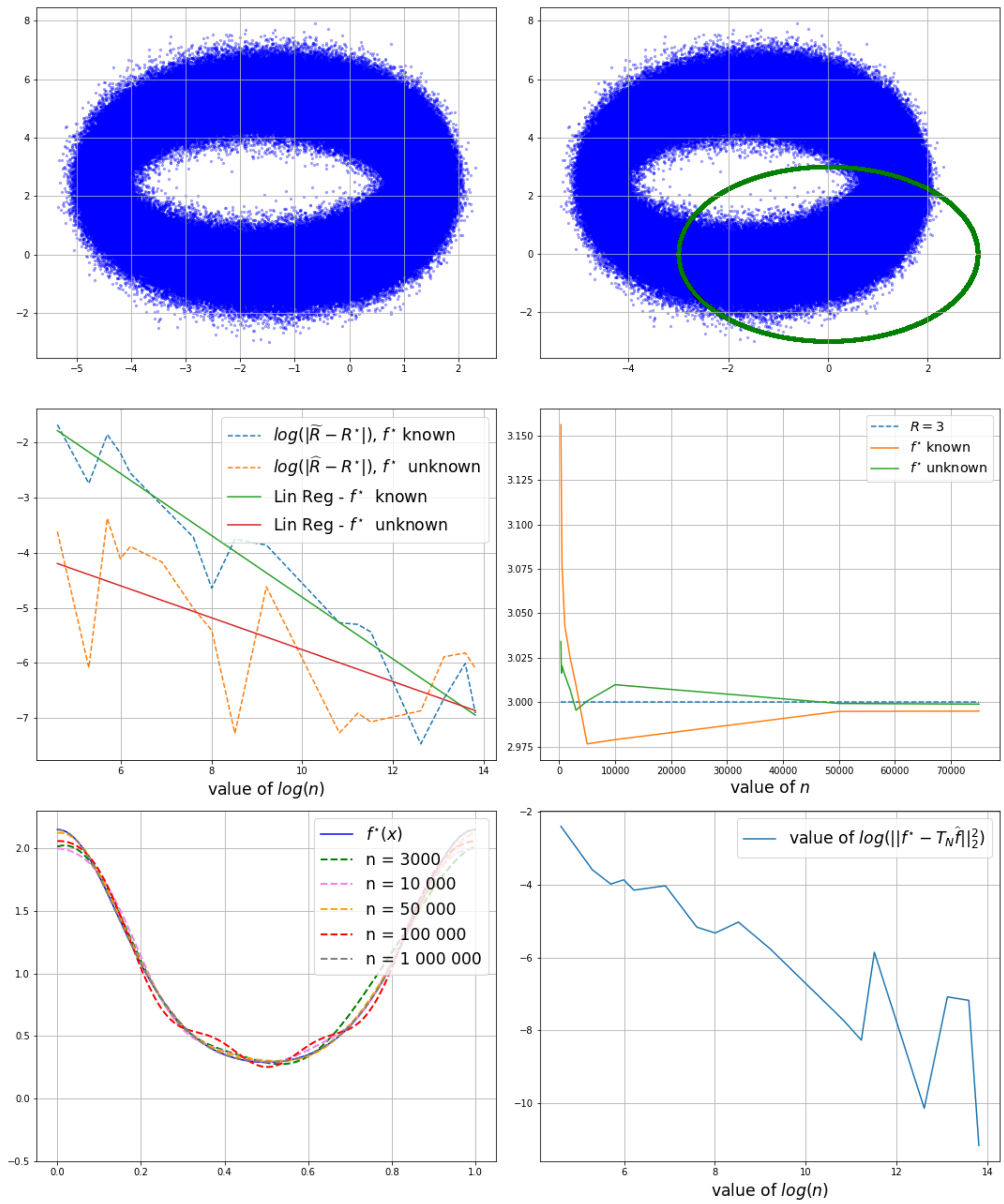

FIGURE 7. $R^{\star} = 3$, $U \sim f^{\star} : x \in (0,1) \mapsto \frac{\exp\{\cos(2\pi x)\}}{\int_0^1 \exp\{\cos(2\pi u)\}du}$, $\varepsilon \sim \mathcal{N}(\begin{pmatrix} -1.6 \\ 2.5 \end{pmatrix}, \begin{pmatrix} (0.2)^2 & 0 \\ 0 & (0.57)^2 \end{pmatrix})$ *and* $W = \{3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, 7.5\cdot 10^4\}$

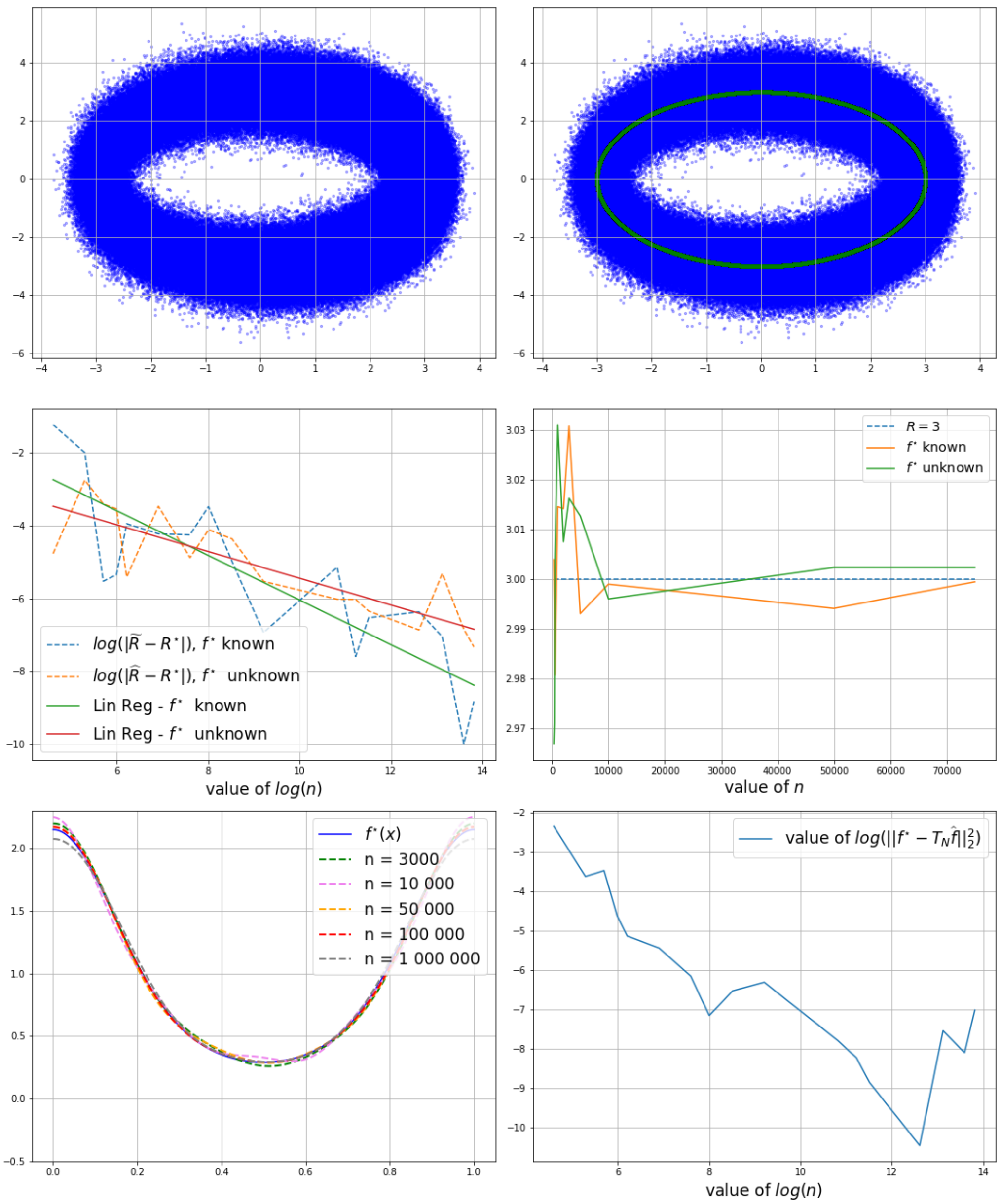

FIGURE 8. $R^\star = 0.6$, $U \sim f^\star : x \in (0,1) \mapsto \frac{\exp\{\cos(2\pi x)\}}{\int_0^1 \exp\{\cos(2\pi u)\}du}$, $\varepsilon \sim \mathcal{N}(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}(0.2)^2 & 0\\ 0 & (0.57)^2\end{pmatrix})$ *and* $W = \{3\cdot 10^2, 4\cdot 10^2, 5\cdot 10^2, 10^3, 2\cdot 10^3, 3\cdot 10^3, 5\cdot 10^3, 10^4, 5\cdot 10^4, 7.5\cdot 10^4\}$

The graph of $\log |R^\star - \widehat{R}|$ from figure 1 to figure 8 drive us to reasonably conjecture that the rate of convergence of $|R^\star - \widehat{R}|$ is the same when the density $f^\star$ is known and unknown.
We use Monte-Carlo to estimate M_n and the package *optimize.minimize* in *Python* to minimize M_n, that is, there is possibly a numerical bias that can explain the fluctuations on the values of $\widehat{R}$ as we can see in the figure 9 (left histogram) and figure 10. The histograms are computed with Monte-Carlo replications of 50 values of $\widehat{R}$ for $n = 10\ 000$ in the case of figure 1 and figure 3.

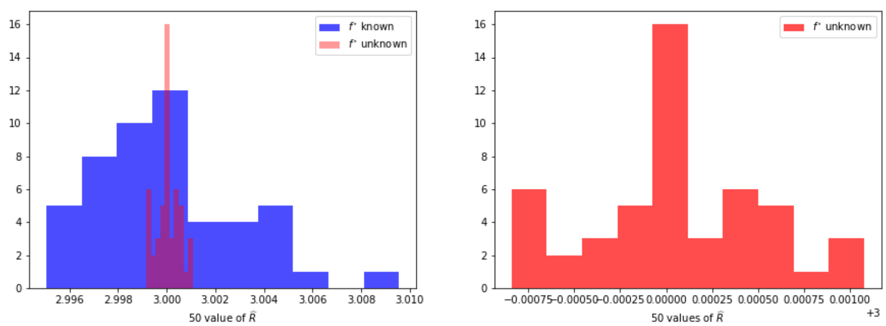

Figure 9. $U \sim$ *Unif*$(0,1)$ *and* $\varepsilon \sim \mathcal{N}(0,(0.12)^2 I)$

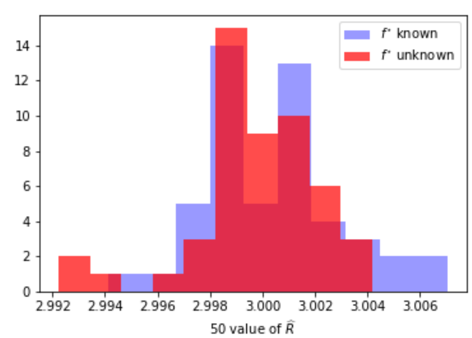

Figure 10. $U \sim$ *Unif*$(0,1)$ *and for* $i \in \{1,2\}$ $\varepsilon^{(i)} \sim \frac{1}{2}\delta_{(-1)} + \frac{1}{2}\mathcal{E}xp\left\{\frac{1}{0.12}\right\}$

Finally, for each $n \in V$ in the case of figure 1, we computed 30 values of $\widehat{R}$ denoted by $(\widehat{R}_k^n)_{1\le k\le 30}$ when the density $f^\star$ is known and unknown, we give the following table which gives the empirical mean squared error. The computation time to minimize M_n becomes long when the number of data increases, in particular when $f^\star$ is unknown. This is why we choose to make only 30 replications.

n	$f^\star$ known $\frac{1}{30}\sum_{k=1}^{30}\lvert R^\star - \widetilde{R}_k^n\rvert^2$	$f^\star$ unknown $\frac{1}{30}\sum_{k=1}^{30}\lvert R^\star - \widehat{R}_k^n\rvert^2$
10^2	$7.09 \cdot 10^{-2}$	$2.86 \cdot 10^{-3}$
$2 \cdot 10^2$	$3.87 \cdot 10^{-2}$	$1.02 \cdot 10^{-3}$
$3 \cdot 10^2$	$1.87 \cdot 10^{-2}$	$5.72 \cdot 10^{-4}$
$4 \cdot 10^2$	$1.5 \cdot 10^{-2}$	$6.15 \cdot 10^{-4}$
$5 \cdot 10^2$	$7.28 \cdot 10^{-3}$	$4.24 \cdot 10^{-4}$
10^3	$1.27 \cdot 10^{-3}$	$2.07 \cdot 10^{-4}$
$2 \cdot 10^3$	$3.64 \cdot 10^{-4}$	$9.97 \cdot 10^{-5}$
$3 \cdot 10^3$	$1.39 \cdot 10^{-4}$	$8.73 \cdot 10^{-5}$
$5 \cdot 10^3$	$1.31 \cdot 10^{-4}$	$5.29 \cdot 10^{-5}$
10^4	$6.99 \cdot 10^{-5}$	$6.63 \cdot 10^{-6}$
$5 \cdot 10^4$	$1.01 \cdot 10^{-5}$	$4.96 \cdot 10^{-7}$
$7.5 \cdot 10^4$	$7.75 \cdot 10^{-6}$	$2.72 \cdot 10^{-7}$
10^5	$7.63 \cdot 10^{-6}$	$1.96 \cdot 10^{-7}$
$3 \cdot 10^5$	$2.47 \cdot 10^{-6}$	$6.03 \cdot 10^{-8}$
$5 \cdot 10^5$	$1.88 \cdot 10^{-6}$	$3.27 \cdot 10p^{-8}$
$8 \cdot 10^5$	$1.37 \cdot 10^{-6}$	$2.32 \cdot 10^{-8}$
10^6	$1.02 \cdot 10^{-6}$	$1.78 \cdot 10^{-8}$

6. Discussion

In this paper, we proved that deconvolution of spherical data is possible without any knowledge of the distribution of the noise, and that the radius of the sphere can be recovered at nearly parametric rate. The question whether the rate $1/\sqrt{n}$ can be attained is still open. To get the almost parametric rate following the proposed analysis would require first to be able to strengthen the lower bound of M in (18). But in [7], getting a lower bound for M requires delicate arguments involving a technical truncation from which it is not possible to derive a quadratic lower bound. If ever such a lower bound can be proved, new ideas have to be developed. Also, we were able to prove the identifiability theorem for all possible densities on a circle, but in higher dimensions the proof holds only for densities that are positive near the origin. Extending the result to hold for any density for any d would be nice.

We also proved, for noisy data on a circle that the exploration density can be

recovered at the same minimax convergence rate on Sobolev regularity classes as when the noise distribution is known. The analysis we propose here does not extend to $d > 2$, and the question of the convergence rate for $d > 2$ remains unsolved.
Finally, we were able to run numerical simulations to estimate the radius and the exploratory density on simulated data, the results illustrate our theoretical findings. To apply the method to real-life data requires further work, both on the methodology side to find a data-driven strategy to choose N and ν_{est} and on the algorithmic side to improve on the computation time.
More generally, deconvolution of data coming from observations supported on a lower dimensional manifold and corrupted by additive noise has been investigated earlier for known noise in [9], see also [4]. The extension of the methodology proposed here to analyze those settings and to deal with unknown noise distribution will be developed in a further work. Understanding how to deal with noisy observations in topological data analysis is a challenging question, see for instance [1] and [2], and our solution for additive noise having independent components can be understood as a contribution in this perspective.

7. Proofs

7.1. Proof of Proposition 1

We shall denote $\|\cdot\|_{2,\nu}$ the $\mathbb{L}^2(B_\nu^m)$-norm and $\|\cdot\|_{\infty,\nu}$ the $\mathbb{L}^\infty(B_\nu^m)$-norm, where the dimension m may be d, d_1 or d_2 and is clear from the context.
Following the proof of Proposition A.2 in [7] and the proof of Proposition 24 in Appendix B.5 in [8], we easily get that there exist positive constants b, $\eta_1 < 1$ and $\eta_2 < 1$ depending only on ν, S, d, ρ such that for all $\Phi \in \Upsilon_{\rho,S}$,

$$\|h\|_{2,\nu} \le \eta_1 \Longrightarrow M(\Phi + h; \nu|\Phi) \ge c_\nu^4 \|h\|_{2,\nu}^{2+2\epsilon(\|h\|_{2,\nu})}, \tag{18}$$

with, for any $u \in (0, 1/e)$,

$$\epsilon(u) = \frac{b}{\log\log\frac{1}{u}},$$

and such that for any $t_1 \in \mathbb{R}^{d_1}$ and any $t_2 \in \mathbb{R}^{d_2}$,

$$\begin{aligned} \|h\|_{2,\nu} \le \eta_2 \quad \Longrightarrow \quad & \|h(\cdot,0)\|_{2,\nu}^2 \le \|h\|_{2,\nu}^{2-2\epsilon(\|h\|_{2,\nu})} \\ & \text{and } \|h(0,\cdot)\|_{2,\nu}^2 \le \|h\|_{2,\nu}^{2-2\epsilon(\|h\|_{2,\nu})}. \end{aligned} \tag{19}$$

We now fix $\eta = \eta_1 \wedge \eta_2$. Let Z_n be the random process defined, for all $t = (t_1, t_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, by

$$Z_n(t) = \sqrt{n}\left(\tilde{\phi}_n(t) - \Phi(t)\Phi_{\varepsilon^{(1)}}(t_1)\Phi_{\varepsilon^{(2)}}(t_2)\right)$$

Using explicit computation, straightforward upper bounds and (19) we easily get that there exists a constant C that depends only on ν_{est}, ρ and S such that if $h \in \Upsilon_{\rho,S}$ is such that $\|h\|_{2,\nu_{\text{est}}} \leq \eta$, then

$$\begin{aligned}
|(M_n(\Phi+h) - M(\Phi+h,\nu_{\text{est}}|\Phi)) - (M_n(\Phi) - M(\Phi,\nu_{\text{est}}|\Phi))| \\
\leq C\left[\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}} + \frac{\|Z_n\|^2_{\infty,\nu_{\text{est}}}}{n}\right]\cdot\left[\|h\|_{2,\nu_{\text{est}}}^{1-\epsilon(\|h\|_{2,\nu_{\text{est}}})} + \|h\|_{2,\nu_{\text{est}}}^{2-2\epsilon(\|h\|_{2,\nu_{\text{est}}})}\right] \\
\leq 6C\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}}\cdot\|h\|_{2,\nu_{\text{est}}}^{1-\epsilon(\|h\|_{2,\nu_{\text{est}}})}
\end{aligned} \tag{20}$$

since $\|Z_n\|_{\infty,\nu_{\text{est}}} \leq 2\sqrt{n}$ and $\|h\|_{2,\nu_{\text{est}}} \leq 1$. Let now $\widehat{h} = \widehat{\phi}_n - \Phi$. By using that $M_n(\Phi+\widehat{h}) \leq M_n(\Phi) + \frac{1}{n}$, $M(\Phi,\nu_{\text{est}}|\Phi) = 0$ and (20), we easily get

$$M\left(\Phi+\widehat{h},\nu_{\text{est}}|\Phi\right) \leq \frac{1}{n} + 3C\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}}\cdot\|h\|_{2,\nu_{\text{est}}}^{1-\epsilon(\|h\|_{2,\nu_{\text{est}}})}.$$

But for any $\nu \leq \nu_{\text{est}}$, $M(\cdot,\nu|\Phi) \leq M(\cdot,\nu_{\text{est}}|\Phi)$, so that using now (18) we get that for some constant C that depends only on ν, ν_{est}, ρ and S, as soon as $\|\widehat{h}\|_{2,\nu_{\text{est}}} \leq \eta$,

$$\begin{aligned}
\|\widehat{h}\|_{2,\nu}^{2+2\epsilon(\|\widehat{h}\|_{2,\nu})} &\leq \frac{C}{n} + C\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}}\cdot\|\widehat{h}\|_{2,\nu_{\text{est}}}^{1-\epsilon(\|\widehat{h}\|_{2,\nu_{\text{est}}})} \\
&\leq 2C\left(\frac{1}{n}\vee\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}}\cdot\|\widehat{h}\|_{2,\nu_{\text{est}}}^{1-\epsilon(\|\widehat{h}\|_{2,\nu_{\text{est}}})}\right).
\end{aligned} \tag{21}$$

In the case $\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}}\cdot\|\widehat{h}\|_{2,\nu_{\text{est}}}^{1-\epsilon(\|\widehat{h}\|_{2,\nu_{\text{est}}})} \leq \frac{1}{n}$, Proposition 1 is proven.
Otherwise, we relate $\|\widehat{h}\|_{2,\nu}$ to $\|\widehat{h}\|_{2,\nu_{\text{est}}}$. Using Lemma H.3 of [8] and following Section A.3 of [7] we get that there exists a constant D that depends only on ν, ν_{est} and S, for all $h \in \Upsilon_{\rho,S}$, for all integer $m \geq \rho d$,

$$\|h\|_{2,\nu_{\text{est}}} \leq D^m m^{-m/\rho+3d/2} + D^m m^{d/2}\|h\|_{2,\nu}.$$

By choosing $m = \frac{\rho\log(1/\|h\|_{2,\nu})}{\log\log(1/\|h\|_{2,\nu})}$ we get that for some constant C that depends only on ν, ν_{est}, ρ and S, for all $h \in \Upsilon_{\rho,S}$ small enough,

$$\|h\|_{2,\nu_{\text{est}}} \leq C\|h\|_{2,\nu}^{1-\tilde{\epsilon}(\|h\|_{2,\nu})}$$

where $\tilde{\epsilon}(u) = \frac{\log\log\log\frac{1}{u}}{\log\log\frac{1}{u}}$, which implies that $\epsilon(\|h\|_{2,\nu_{\text{est}}}) \leq 2\epsilon(\|h\|_{2,\nu})$ for small enough $\|h\|_{2,\nu}$. Then using (21) we finally get, for some constant C that depends only on ν, ν_{est}, ρ and S, that as soon as $\widehat{h}$ is small enough,

$$\|\widehat{h}\|_{2,\nu}^{1+4\epsilon(\|\widehat{h}\|_{2,\nu})+\tilde{\epsilon}(\|\widehat{h}\|_{2,\nu})(1-2\epsilon(\|\widehat{h}\|_{2,\nu}))} \leq C\frac{\|Z_n\|_{\infty,\nu_{\text{est}}}}{\sqrt{n}}. \tag{22}$$

The end of the proof follows from the fact that $\widehat{\phi}_n$ is uniformly consistent in $L^2([-\nu,\nu]^d)$, see [7] Appendix A.1, and the following deviation inequality which is proved in Appendix G of [8]. There exist a numerical constant c and a constant C that depends only on d, ν_{est} and $E(\|Y_1\|^2)$ such that for all $n \geq 1$ and $x > 0$, with probability at least $1 - 4e^{-x}$,

$$\|Z_n\|_{\infty,\nu_{\mathrm{est}}} \leq C + c\sqrt{x} + c\frac{x}{\sqrt{n}}. \tag{23}$$

Proposition 1 easily follows from the uniform consistency of $\widehat{\phi}_n$, (22) and (23).

7.2. Proof of Lemma 1

To begin with, for any $z_0 \in \mathbb{C}$ and $z \in \mathbb{C}^{d-1}$,

$$E\left[\exp\left(iz_0\widetilde{X}^{(1)} + iz^T\widetilde{X}^{(2)}\right)\right] = E\left[E\left[\exp\left(iz_0\widetilde{X}^{(1)}\right)|\widetilde{X}^{(2)}\right]\exp\left(iz^T\widetilde{X}^{(2)}\right)\right],$$

and usual arguments for multivariate analytic functions on $\mathbb{C}^{d-1}$ prove that $z \mapsto E\left[\exp\left(iz_0\widetilde{X}^{(1)} + iz^T\widetilde{X}^{(2)}\right)\right]$ is the null function if and only if $E\left[\exp\left(iz_0\widetilde{X}^{(1)}\right)|\widetilde{X}^{(2)}\right]$ is zero $P_{\widetilde{X}^{(2)}}$-a.s. In the same way, for any $z_0 \in \mathbb{C}^{d-1}$, $z \mapsto E\left[\exp\left(iz\widetilde{X}^{(1)} + iz_0^T\widetilde{X}^{(2)}\right)\right]$ is the null function if and only if $E\left[\exp\left(iz_0^T\widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]$ is zero $P_{\widetilde{X}^{(1)}}$-a.s. Also, the value of the center $C^\star$ can only change the function $E\left[\exp\left(z^T(\widetilde{X}^{(1)}, \widetilde{X}^{(2)})\right)\right]$, $z \in \mathbb{C}^d$, by a factor $\exp(z^TC^\star)$ which is non zero, so that we may assume $C^\star = 0$ to prove Lemma 1.
In the following, we write for all $u \in (0,1), \widetilde{S}^{(1)}(u) = \cos(2\pi u)$ and for all $u \in (0,1)^{d-1}$,

$$\widetilde{S}^{(2)}(u) = \begin{pmatrix} \sin(2\pi u^{(1)})\cos(\pi u^{(2)}) \\ \sin(2\pi u^{(1)})\sin(\pi u^{(2)})\cos(\pi u^{(3)}) \\ \vdots \\ \sin(2\pi u^{(1)})\sin(\pi u^{(2)})\cdots\sin(\pi u^{(d-2)})\cos(\pi u^{(d-1)}) \\ \sin(2\pi u^{(1)})\sin(\pi u^{(2)})\cdots\sin(\pi u^{(d-2)})\sin(\pi u^{(d-1)}) \end{pmatrix}.$$

We first prove that for any $z_0 \in \mathbb{C}$, $E\left[\exp\left(iz_0\widetilde{X}^{(1)}\right)|\widetilde{X}^{(2)}\right]$ is not $P_{\widetilde{X}^{(2)}}$-a.s. zero.
Since $f^\star$ is not identically zero, there exists a closed interval $[\alpha, \beta]$ subset of one of the four following intervals : $(0, \frac{1}{4})$, $(\frac{1}{4}, \frac{1}{2})$, $(\frac{1}{2}, \frac{3}{4})$, $(\frac{3}{4}, 1)$, a vector $a = (a^{(i)})_{1 \leq i \leq d-1} \in (\alpha, \beta) \times (0,1)^{d-2}$ (if $d = 2$, a is a real number in (α, β)) and $\zeta > 0$ such that if we define

$$I_1 = \{u = (u^{(i)})_{i \in \{1,\dots,d-1\}} \in (\alpha, \beta) \times (0,1)^{d-2} \ : \ \left\|\widetilde{S}^{(2)}(u) - \widetilde{S}^{(2)}(a)\right\|_2^2 < \zeta^2\},$$

then the restriction of $f^\star$ to I_1, $f^\star|_{I_1}$, is not the null function.
We choose ζ small enough such that, if we define $A \subset (-1,1)^{d-1}$ as $A = \widetilde{S}^{(2)}(I_1)$, we have that there exists $I_2 \subset (0,1)^{d-1}$ such that $I_1 \cap I_2 = \emptyset$ and $(\widetilde{S}^{(2)})^{-1}(A) =$

$I_1 \cup I_2$.
We define, for $i,j \in \{1,2\}$, the $\mathcal{C}^1$ diffeomorphisms $\eta_{i,j} : I_i \longrightarrow I_j$, $u \mapsto (S^{(2)})^{-1}(S^{(2)}(u))$, such that $\eta_{i,i} = Id|_{I_i}$ and $\eta_{i,j} \circ \eta_{j,i} = Id|_{I_i}$.
Note that we can explicitly calculate $\eta_{i,j}(x)$ for the different possible inclusions of $[\alpha,\beta]$:

1. for $[\alpha,\beta] \subset (0,\frac{1}{4})$ or $[\alpha,\beta] \subset (\frac{1}{4},\frac{1}{2})$, we have $\eta_{1,2}(u^{(1)},u^{(2)},\dots,u^{(d-1)}) = (\frac{1}{2} - u^{(1)},u^{(2)},\dots,u^{(d-1)})$,
2. and for $[\alpha,\beta] \subset (\frac{1}{2},\frac{3}{4})$ or $[\alpha,\beta] \subset (\frac{3}{4},1)$, we have $\eta_{1,2}(u^{(1)},u^{(2)},\dots,u^{(d-1)}) = (\frac{3}{2} - u^{(1)},u^{(2)},\dots,u^{(d-1)})$.

We now compute $E[\exp(iz_0\widetilde{X}^{(1)})|\widetilde{X}^{(2)}]1_{\widetilde{X}^{(2)}\in A}$. For any measurable bounded function ω on $(-R^\star,R^\star)^{d-1}$, we have

$$\begin{aligned}
&E[\exp(iz_0\widetilde{X}^{(1)})\omega(\widetilde{X}^{(2)})1_{\widetilde{X}^{(2)}\in A}] \\
&\quad= \int_{(\widetilde{S}^{(2)})^{-1}(A)} \omega(R^\star\widetilde{S}^{(2)}(u))\exp(iz_0R^\star\cos(2\pi u^{(1)}))f^\star(u)du \\
&\quad= \int_{I_1} \omega(R^\star\widetilde{S}^{(2)}(u))\exp(iz_0R^\star\cos(2\pi u^{(1)}))f^\star(u)du \\
&\qquad+ \int_{I_2} \omega(R^\star\widetilde{S}^{(2)}(u))\exp(iz_0R^\star\cos(2\pi u^{(1)}))f^\star(u)du.
\end{aligned}$$

Define the change of variables $u = \eta_{1,2}(v)$ in the second integral. Using the explicit definition of $\eta_{1,2}$ which is differentiable with Jacobian equal to 1 we get

$$\begin{aligned}
&E[\exp(iz_0\widetilde{X}^{(1)})\omega(\widetilde{X}^{(2)})1_{\widetilde{X}^{(2)}\in A}] \\
&\quad= \int_{I_1} \omega(R^\star\widetilde{S}^{(2)}(u))\bigg(\exp(iR^\star z_0\cos(2\pi u^{(1)}))f^\star(u) \\
&\qquad+ \exp(iR^\star z_0\cos(2\pi\eta_{1,2}(u)^{(1)}))f^\star(\eta_{1,2}(u))\bigg)du \\
&= \int_{I_1} \omega(R^\star\widetilde{S}^{(2)}(u))\frac{f^\star(u)}{f^\star(u)+f^\star(\eta_{1,2}(u))}\bigg(\exp(iR^\star z_0\cos(2\pi u^{(1)}))f^\star(u) \\
&\qquad+ \exp(iR^\star z_0\cos(2\pi\eta_{1,2}(u)^{(1)}))f^\star(\eta_{1,2}(u))\bigg)du \\
&+ \int_{I_1} \omega(R^\star\widetilde{S}^{(2)}(u))\frac{f^\star(\eta_{1,2}(u))}{f^\star(u)+f^\star(\eta_{1,2}(u))}\bigg(\exp(iR^\star z_0\cos(2\pi u^{(1)}))f^\star(u) \\
&\qquad+ \exp(iR^\star z_0\cos(2\pi\eta_{1,2}(u)^{(1)}))f^\star(\eta_{1,2}(u))\bigg)du.
\end{aligned}$$

Thus if we define $\nu_1 : A \longrightarrow I_1$ such that for all $u \in I_1$, $\nu_1(\widetilde{S}^{(2)}(u)) = u$, and $\nu_2 : A \longrightarrow I_2$ such that for all $u \in I_2$, $\nu_2(\widetilde{S}^{(2)}(u)) = u$, we get

$$E\left[\exp(iz_0\widetilde{X}^{(1)})|\widetilde{X}^{(2)}\right] 1_{\widetilde{X}^{(2)}\in A}$$
$$= \frac{1}{f^\star(\nu_1(\widetilde{X}^{(2)})) + f^\star(\nu_2(\widetilde{X}^{(2)}))}\Bigg(\exp(iR^\star z_0\cos(2\pi\nu_1(\widetilde{X}^{(2)}))^{(1)})f^\star(\nu_1(\widetilde{X}^{(2)}))$$
$$+ \exp(-iR^\star z_0\cos(2\pi\nu_2(\widetilde{X}^{(2)}))^{(1)})f^\star(\nu_2(\widetilde{X}^{(2)}))\Bigg)1_{\widetilde{X}^{(2)}\in A}.$$

Finally, $E\left[\exp(iz_0\widetilde{X}^{(1)})|\widetilde{X}^{(2)}\right] 1_{\widetilde{X}^{(2)}\in A}$ is null $\mathbb{P}_{\widetilde{X}^{(2)}}$-a.s if and only if for $f^\star du$ almost all $u \in I_1$,

$$\exp(iR^\star z_0\cos(2\pi u^{(1)}))f^\star(u) + \exp(iR^\star z_0\cos(2\pi\eta_{1,2}(u)^{(1)}))f^\star(\eta_{1,2}(u)) = 0,$$

that is for $f^\star du$ almost all $u \in I_1$,

$$\exp\left(iR^\star z_0\left(\cos(2\pi u^{(1)}) - \cos(2\pi\eta_{1,2}(u)^{(1)})\right)\right) = -\frac{f^\star(\eta_{1,2}(u))}{f^\star(u)}.$$

Since for almost all $u \in I_1$, $f^\star(u) \neq 0$, this would imply in particular that for almost all $u \in I_1$

$$R^\star\mathrm{Re}(z_0)\left(\cos(2\pi u^{(1)}) - \cos(2\pi\eta_{1,2}(u)^{(1)})\right) = \pi \ \ [\mathrm{mod} 2\pi], \tag{24}$$

which gives a contradiction.

Now, let us prove that for any $z_0 \in \mathbb{C}^{d-1}$, $E\left[\exp\left(iz_0^T\widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]$ is not $P_{\widetilde{X}^{(1)}}$-a.s. zero.
Let us first assume $d > 2$, since $f^\star$ is not identically zero, there exists a closed interval $[\alpha, \beta]$ in one of the four following intervals : $(0, \frac{1}{4})$, $(\frac{1}{4}, \frac{1}{2})$, $(\frac{1}{2}, \frac{3}{4})$, $(\frac{3}{4}, 1)$, such that $f|_{(\alpha,\beta)\times(0,1)^{d-2}}$ is not the null function.
Define $J_1 = (\alpha, \beta)$, $B = \widetilde{S}^{(1)}(J_1)$, and J_2 such that $J_2 \cap J_1 = \emptyset$ and $(\widetilde{S}^{(1)})^{-1}(B) = J_1 \cup J_2$. We define the $\mathcal{C}^1$ diffeomorphism $\sigma_{1,2} : J_1 \longrightarrow J_2$, $u \mapsto (\widetilde{S}^{(1)})^{-1}(\widetilde{S}^{(1)}(u))$, such that $\sigma_{1,1} = Id|_{I_1}$ and $\sigma_{1,2} \circ \sigma_{2,1} = Id|_{I_1}$. Note that we can explicitly calculate $\sigma_{1,2}(u)$, indeed, for $u \in J_1$, we have $\sigma_{1,2}(u) = 1 - u$. The reason of choosing J_1 in one of these four intervals is to have the decomposition of $(\widetilde{S}^{(1)})^{-1}(B)$ in exactly 2 disjoint open sets on which $\widetilde{S}^{(1)}$ is one to one.
For any bounded and measurable function ω on $(-R^\star, R^\star)$, we have

$$E\left[\exp\left(iz_0^T\widetilde{X}^{(2)}\right)\omega(\widetilde{X}^{(1)})1_{\widetilde{X}^{(1)}\in B}\right]$$
$$= \int_{J_1\times(0,1)^{d-2}} \omega(R^\star\cos(2\pi u^{(1)}))\exp(iR^\star z_0^\intercal\widetilde{S}^{(2)}(u))f^\star(u)du$$
$$+ \int_{J_2\times(0,1)^{d-2}} \omega(R^\star\cos(2\pi u^{(1)}))\exp(iR^\star z_0^\intercal\widetilde{S}^{(2)}(u))f^\star(u)du.$$

Define the following change of variables for the second integral, $u = \overline{\sigma}(v) = (\sigma_{1,2} \otimes Id_{(0,1)^{d-2}})(v)$, which has Jacobien equal to 1. Then

$$
\begin{aligned}
& E\left[\exp\left(iz_0^T \widetilde{X}^{(2)}\right)\omega(\widetilde{X}^{(1)})1_{\widetilde{X}^{(1)}\in B}\right] \\
= \int_{J_1\times(0,1)^{d-2}} & \omega(R^\star\cos(2\pi u^{(1)})) \\
& \left(\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u))f^\star(u) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\overline{\sigma}(u)))f^\star(\overline{\sigma}(u))\right)du \\
= \int_{J_1\times(0,1)^{d-2}} & \omega(R^\star\cos(2\pi u^{(1)}))\frac{f^\star(u)}{f^\star(u)+f^\star(\overline{\sigma}(u))} \\
& \left(\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u))f^\star(u) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\overline{\sigma}(u)))f^\star(\overline{\sigma}(u))\right)du \\
+ \int_{J_1\times(0,1)^{d-2}} & \omega(R^\star\cos(2\pi u^{(1)}))\frac{f^\star(\overline{\sigma}(u))}{f^\star(u)+f^\star(\overline{\sigma}(u))} \\
& \left(\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u))f^\star(u) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\overline{\sigma}(u)))f^\star(\overline{\sigma}(u))\right)du.
\end{aligned}
$$

We now define $\tau_1 : B \longrightarrow J_1$ such that for all $u \in J_1$, $\tau_1(\widetilde{S}^{(1)}(u)) = u$, and $\tau_2 : B \longrightarrow J_2$ such that for all $u \in J_2$, $\tau_2(\widetilde{S}^{(1)}(u)) = u$.

Since we assume (H2), when $d > 2$, we can choose $J_1 = (0, \zeta_0)$ with $\zeta_0 \leq \zeta^\star$ and $\zeta_0 < \frac{1}{4}$.

$$
\begin{aligned}
& E\left[\exp\left(iz_0^T \widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]1_{\widetilde{X}^{(1)}\in B} \\
&= \left(\int_{(0,1)^{d-2}} \frac{1}{f^\star(\tau_1(\frac{\widetilde{X}^{(1)}}{R^\star}),u) + f^\star(\tau_2(\frac{\widetilde{X}^{(1)}}{R^\star}),u)}\left[\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\tau_1(\frac{\widetilde{X}^{(1)}}{R^\star}),u))\right.\right. \\
& \left.\left. f^\star(\tau_1(\frac{\widetilde{X}^{(1)}}{R^\star}),u) + \exp(-iR^\star z_0^\intercal \widetilde{S}^{(2)}(\tau_2(\frac{\widetilde{X}^{(1)}}{R^\star}),u))f^\star(\tau_2(\frac{\widetilde{X}^{(1)}}{R^\star}),u)\right]du\right)1_{\widetilde{X}^{(1)}\in B}
\end{aligned}
$$

and $E\left[\exp\left(iz_0^T \widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]1_{\widetilde{X}^{(1)}\in B}$ is null $\mathbb{P}_{\widetilde{X}^{(1)}}$-a.s. if and only if for all $u \in J_1$,

$$
\int_{(0,1)^{d-2}} \frac{\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u,v))f^\star(u,v) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(1-u,v))f^\star(1-u,v)}{f^\star(u,v)+f^\star(1-u,v)}dv = 0.
$$

In particular, this implies that for all $u \in J_1$,

$$\int_{(0,1)^{d-2}} \frac{\cos(\mathrm{Re}(z_0)^\intercal \widetilde{S}^{(2)}(u,v))}{f^\star(u,v)+f^\star(1-u,v)} \Bigg(f^\star(u,v)\exp(-R^\star \mathrm{Im}(z_0)^\intercal \widetilde{S}^{(2)}(u,v)) + f^\star(1-u,v)\exp(R^\star \mathrm{Im}(z_0)^\intercal \widetilde{S}^{(2)}(u,v)) \Bigg) dv = 0.$$

But for small enough $u \in J_1$, $\cos(\mathrm{Re}(z_0)^\intercal \widetilde{S}^{(2)}(u,v))$ stays positive for all $v \in (0,1)^{d-2}$ which gives a contradiction.
When $d = 2$, applying an analogous reasoning, we get for any bounded and measurable function ω on $(-R^\star, R^\star)$, we have

$$\begin{aligned} &E\left[\exp\left(iz_0^T \widetilde{X}^{(2)}\right) \omega(\widetilde{X}^{(1)}) 1_{\widetilde{X}^{(1)}\in B}\right] \\ &\qquad = \int_{J_1} \omega(R^\star \cos(2\pi u)) \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u)) f^\star(u) du \\ &\qquad\qquad + \int_{J_2} \omega(R^\star \cos(2\pi u)) \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u)) f^\star(u) du. \end{aligned}$$

Define the following change of variables for the second integral, $u = \overline{\sigma}(v) = \sigma_{1,2}(v)$, which has Jacobien equal to 1. Then

$$\begin{aligned} & E\left[\exp\left(iz_0^T \widetilde{X}^{(2)}\right) \omega(\widetilde{X}^{(1)}) 1_{\widetilde{X}^{(1)}\in B}\right] \\ = \int_{J_1} & \omega(R^\star \cos(2\pi u)) \\ & \Bigg(\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u)) f^\star(u) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\overline{\sigma}(u))) f^\star(\overline{\sigma}(u)) \Bigg) du \\ = \int_{J_1} & \omega(R^\star \cos(2\pi u)) \frac{f^\star(u)}{f^\star(u) + f^\star(\overline{\sigma}(u))} \\ & \Bigg(\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u)) f^\star(u) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\overline{\sigma}(u))) f^\star(\overline{\sigma}(u)) \Bigg) du \\ + \int_{J_1} & \omega(R^\star \cos(2\pi u)) \frac{f^\star(\overline{\sigma}(u))}{f^\star(u) + f^\star(\overline{\sigma}(u))} \\ & \Bigg(\exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(u)) f^\star(u) + \exp(iR^\star z_0^\intercal \widetilde{S}^{(2)}(\overline{\sigma}(u))) f^\star(\overline{\sigma}(u)) \Bigg) du. \end{aligned}$$

We now define $\tau_1 : B \longrightarrow J_1$ such that for all $u \in J_1$, $\tau_1(\widetilde{S}^{(1)}(u)) = u$, and $\tau_2 : B \longrightarrow J_2$ such that for all $u \in J_2$, $\tau_2(\widetilde{S}^{(1)}(u)) = u$.

$$
\begin{aligned}
&E\left[\exp\left(iz_0^T\widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]1_{\widetilde{X}^{(1)}\in B}\\
&=\frac{1}{f^\star(\tau_1(\frac{\widetilde{X}^{(1)}}{R^\star}))+f^\star(\tau_2(\frac{\widetilde{X}^{(1)}}{R^\star}))}\Bigg(\exp(iR^\star z_0^\intercal\widetilde{S}^{(2)}(\tau_1(\frac{\widetilde{X}^{(1)}}{R^\star})))f^\star(\tau_1(\frac{\widetilde{X}^{(1)}}{R^\star}))\\
&\qquad\qquad+\exp(-iR^\star z_0^\intercal\widetilde{S}^{(2)}(\tau_2(\frac{\widetilde{X}^{(1)}}{R^\star})))f^\star(\tau_2(\frac{\widetilde{X}^{(1)}}{R^\star}))\Bigg)1_{\widetilde{X}^{(1)}\in B}
\end{aligned}
$$

and $E\left[\exp\left(iz_0^T\widetilde{X}^{(2)}\right)|\widetilde{X}^{(1)}\right]$ can not be null $\mathbb{P}_{\widetilde{X}^{(1)}}$-a.s. since it would require that for almost all $u\in J_1$,

$$\exp(iR^\star z_0^\intercal\widetilde{S}^{(2)}(u))f^\star(u)+\exp(iR^\star z_0^\intercal\widetilde{S}^{(2)}(1-u))f^\star(1-u)=0.$$

7.3. Proof of Proposition 2

For $f,f^\star\in\mathcal{F}$ and $R,R^\star\in[R_{\min},R_{\max}]$, denote $\theta=(f,R)$, $\theta^\star=(f^\star,R^\star)$, and define the distance d by $d(\theta,\theta^\star)=\left(\int_{(0,1)^{d-1}}(f(u)-f^\star(u))^2du\right)^{1/2}+|R-R^\star|$. First, using Lemma A.1 in [7] we get

$$\sup_{\theta\in\mathcal{F}\times[R_{\min},R_{\max}]}|M_n(\theta)-M(\theta)|=o_{\mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}}(1). \tag{25}$$

Then, using the continuity of M with respect to the distance d and the compacity of $\mathcal{F}\times[R_{\min},R_{\max}]$, using Theorem 2 we get that for any $\delta>0$,

$$\inf_{\theta\in\mathcal{F}\times[R_{\min},R_{\max}],\ d(\theta,\theta^\star)>\delta}M(\theta)>M(\theta^\star)=0. \tag{26}$$

Consistency of $\widehat{f}$ and $\widehat{R}$ follows from (25), (26) and Theorem 5.7 in [18]. Consistency of $\widehat{C}$ is then a consequence of the continuity theorem and the law of large numbers.

7.4. Proof of proposition 3

The functions $\Psi_{f,R}$ on $\mathbb{R}^2$ can be written as functions $\Phi_{f,R}$ on $[0,+\infty[\times[0,1)$ using polar representation. For any $r\geq 0$ and $\theta\in[0,1)$, define

$$\Phi_{f,R}(r,\theta)=\Psi_{f,R}(r\cos(2\pi\theta),r\sin(2\pi\theta)).$$

For all $r\geq 0$, let $(\lambda_p(r))_{p\in\mathbb{Z}}$ be the sequence of Fourier coefficients of $\Phi_{f,R}(r,\cdot)$,

$$\lambda_p(r)=\int_0^1\Phi_{f,R}(r,u)e^{2i\pi pu}du,\quad p\in\mathbb{Z}.$$

Using (III) in Section 8, we get that for all $r \geq 0$,

$$\begin{aligned}
\Phi_{f,R}(r,\theta) &= \int_0^1 f(u)\exp(irR\cos(2\pi u - 2\pi\theta))du \\
&= \sum_{p\in\mathbb{Z}} i^p J_p(rR)(\int_0^1 f(u)e^{2i\pi up}du)e^{-2i\pi p\theta} \\
&= \sum_{p\in\mathbb{Z}} i^p f_p J_p(rR)e^{-2i\pi p\theta},
\end{aligned}$$

so that for all $p \in \mathbb{Z}$,

$$\lambda_p(r) = i^p J_p(rR) f_p,$$

and also

$$\lambda_p^\star(r) = i^p J_p(rR^\star) f_p^\star, \quad \widehat{\lambda}_p(r) = i^p J_p(r\widehat{R})\widehat{f}_p,$$

where $(f_p^\star)_{p\in\mathbb{Z}}$ (resp. $(\widehat{f}_p)_{p\in\mathbb{Z}}$) are the Fourier coefficients of $f^\star$(resp. $\hat{f}$) and $(\widehat{\lambda}_p(r))_{p\in\mathbb{Z}}$ are the Fourier coefficients of $\Phi_{\widehat{f},\widehat{R}}(r,.)$. We have, using Parseval's identity,

$$\begin{aligned}
\int_0^1 |\Phi_{f^\star,R^\star}(r,\theta) - \Phi_{\widehat{f},\widehat{R}}(r,\theta)|^2 d\theta &= \sum_{k\in\mathbb{Z}} |\lambda_k^\star(r) - \widehat{\lambda}_k(r)|^2 \\
&= \sum_{k\in\mathbb{Z}} |f_k^\star J_k(rR^\star) - \widehat{f}_k J_k(r\widehat{R})|^2 \\
&\geq \sum_{|k|\leq N} |f_k^\star J_k(rR^\star) - \widehat{f}_k J_k(r\widehat{R})|^2.
\end{aligned}$$

We use the fact that $|a-b|^2 \geq \frac{|a|^2}{2} - |b|^2$ for all $a, b \in \mathbb{C}$, to get

$$\begin{aligned}
\int_0^1 |\Phi_{f^\star,R^\star}(r,\theta) - \Phi_{\widehat{f},\widehat{R}}(r,\theta)|^2 d\theta \geq \sum_{|k|\leq N} \frac{|f_k^\star - \widehat{f}_k|^2}{2} J_k(r\widehat{R})^2 \\
- \sum_{|k|\leq N} |f_k^\star|^2 |J_k(rR^\star) - J_k(r\widehat{R})|^2,
\end{aligned}$$

so that

$$\begin{aligned}
\sum_{|k|\leq N} |f_k^\star - \widehat{f}_k|^2 J_k(r\widehat{R})^2 \leq 2\int_0^1 |\Phi_{f^\star,R^\star}(r,\theta) - \Phi_{\widehat{f},\widehat{R}}(r,\theta)|^2 d\theta \\
+ 2\sum_{|k|\leq N} |f_k^\star|^2 |J_k(rR^\star) - J_k(r\widehat{R})|^2.
\end{aligned}$$

Then, for all $\nu \in (0, \nu_{\mathrm{est}}]$ such that $c_\nu^\star > 0$, we integrate from 0 to ν and we use (IV) in Section 8 to obtain

$$\sum_{|k|\leq N} |f_k^\star - \widehat{f}_k|^2 \int_0^\nu rJ_k(r\widehat{R})^2 dr \leq 2||\Psi_{f^\star,R^\star} - \Psi_{\widehat{f},\widehat{R}}||^2_{\mathbb{L}_2(\mathbb{D}_2(0,\nu))} + 2\sum_{|k|\leq N} |f_k^\star|^2 |R^\star - \widehat{R}|^2 \int_0^\nu r^3 dr.$$

Using corollary 1, Theorem 3 and the fact that $\sum_{|k|\leq N} |f_k^\star|^2 \leq \int_0^1 (f^\star(u))^2 du$ is uniformly upper bounded in the compact set $\mathcal{F}$, we have that there exists a constant $c > 0$ depending on δ, ν, $c_\nu^\star$, d, $R^\star$, $R_{\min}$, $R_{\max}$, and $E(\|Y\|^2)$ such that for all $x \geq 1$ and for c_1 and c_2 coming from Proposition 1, for all $n \geq (1 \vee xc_1)/c_2$, with probability at least $1 - e^{-x}$,

$$\sum_{|k|\leq N} |f_k^\star - \widehat{f}_k|^2 \int_0^\nu rJ_k(r\widehat{R})^2 dr \leq c \left(\frac{x}{n^{1-\delta}} \vee \frac{x^2}{n^{2-2\delta}} \right).$$

Using lemma 4, we finally have that with probability at least $1 - e^{-x}$,

$$\sum_{|k|\leq N} |f_k^\star - \widehat{f}_k|^2 \leq c\frac{32}{9\nu^2}(\nu\widehat{R})^{-2N}(N+1)2^{2N}(N!)^2 \left(\frac{x}{n^{1-\delta}} \vee \frac{x^2}{n^{2-2\delta}} \right).$$

We finally use $\widehat{R} \in [R_{\min}, R_{\max}]$ to end the proof.

7.5. Proof of lemma 2

The proof of (1) follows from the same arguments as in the proof of Proposition 2.
For all $(t_1, t_2) \in \mathbb{R} \times \mathbb{R}^{d-1}$, define

$$m_{n,R}(t_1,t_2) = \Psi_{f^\star,R}(t_1,t_2)\tilde{\psi}_n(t_1,0)\tilde{\psi}_n(0,t_2) - \tilde{\psi}_n(t_1,t_2)\Psi_{f^\star,R}(t_1,0)\Psi_{f^\star,R}(0,t_2)$$

and

$$m_R(t_1,t_2) = \Psi_{f^\star,R}(t_1,t_2)\Psi_{f^\star,R^\star}(t_1,0)\Psi_{f^\star,R^\star}(0,t_2) - \Psi_{f^\star,R^\star}(t_1,t_2)\Psi_{f^\star,R}(t_1,0)\Psi_{f^\star,R}(0,t_2),$$

such that $M_n(R) = \int_{B_{\nu_{\text{est}}} \times B_{\nu_{\text{est}}}^{d-1}} |m_{n,R}(t_1,t_2)|^2 dt_1 dt_2$ and
$M(R) = \int_{B_\nu \times B_\nu^{d-1}} |m_R(t_1,t_2)|^2 |\Phi_\epsilon(t_1,t_2)|^2 dt_1 dt_2$.

Let us prove (2). Differentiation of M_n gives

$$M_n'(R) = \int_{B_{\nu_{\text{est}}} \times B_{\nu_{\text{est}}}^{d-1}} \left(\frac{d}{dR}\{m_{n,R}(t_1,t_2)\}\overline{m_{n,R}(t_1,t_2)} + \frac{d}{dR}\{\overline{m_{n,R}(t_1,t_2)}\}m_{n,R}(t_1,t_2) \right) dt_1 dt_2,$$

where $\overline{z}$ denotes the complex conjugate of z. Since $\overline{m_{n,R}(t_1,t_2)} = m_{n,R}(-t_1,-t_2)$ we get

$$M_n'(R) = \int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}} \Bigg(\frac{d}{dR}\{m_{n,R}(t_1,t_2)\} m_{n,R}(-t_1,-t_2) + \frac{d}{dR}\{m_{n,R}(-t_1,-t_2)\} m_{n,R}(t_1,t_2) \Bigg) dt_1 dt_2.$$

Let $\mathbf{Z}_n$ be the random process defined for $(t_1,t_2)\in\mathbb{R}\times\mathbb{R}^{d-1}$ by

$$\mathbf{Z}_n(t_1,t_2) = \sqrt{n}\left(\tilde{\psi}_n(t_1,t_2) - \Psi_{f^\star,R^\star,}(t_1,t_2)\Phi_\epsilon(t_1,t_2)\right). \tag{27}$$

The random process $\mathbf{Z}_n$ converges weakly to a Gaussian process $(\mathbf{Z}(t_1,t_2))_{(t_1,t_2)}$ in the set of complex continuous functions endowed with the uniform norm. Using (27), $\tilde{\psi}_n(t_1,t_2) = \frac{1}{\sqrt{n}}\mathbf{Z}_n(t_1,t_2) + \Psi_{f^\star,R^\star}(t_1,t_2)\Phi_\epsilon(t_1,t_2)$, so that

$$\begin{aligned}
\sqrt{n}M_n'(R^\star) \quad = \quad & \int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}} C(t_1,t_2) \\
& \Bigg\{ \Psi_{f^\star,R^\star}(-t_1,-t_2)\Big[\mathbf{Z}_n(-t_1,0)\Psi_{f^\star,R^\star}(0,-t_2) + \mathbf{Z}_n(0,-t_2)\Psi_{f^\star,R^\star}(-t_1,0)\Big] \\
& - \mathbf{Z}_n(-t_1,-t_2)\Psi_{f^\star,R^\star}(-t_1,0)\Psi_{f^\star,R^\star}(0,-t_2)\Bigg\} dt_1 dt_2 \\
& + \int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}} C(-t_1,-t_2) \\
& \Bigg\{ \Psi_{f^\star,R^\star}(t_1,t_2)\Big[\mathbf{Z}_n(t_1,0)\Psi_{f^\star,R^\star}(0,t_2) + \mathbf{Z}_n(0,t_2)\Psi_{f^\star,R^\star}(t_1,0)\Big] \\
& - \mathbf{Z}_n(t_1,t_2)\Psi_{f^\star,R^\star}(t_1,0)\Psi_{f^\star,R^\star}(0,t_2)\Bigg\} dt_1 dt_2 + O_{\mathbb{P}}(\frac{1}{\sqrt{n}})
\end{aligned}$$

where all $O_{\mathbb{P}}$ (and later $o_{\mathbb{P}}$) are in $\mathbb{P}_{C^\star,R^\star,f^\star,\mathbb{Q}^\star}$ probability and $C(t_1,t_2)$ is defined by

$$C(t_1,t_2) = \Phi_\varepsilon(t_1,t_2)\frac{d}{dR}m_{R^\star}(t_1,t_2).$$

Now, the empirical process $\mathbf{Z}_n$ converges uniformly in distribution to a Gaussian process over the set of functions $\{Id, \exp(it^T\cdot), |t|\le\nu_{\text{est}}\}$, so that $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n Y_i - E(Y_1), M_n'(R^\star)\right)$ converges in distribution to $\mathcal{N}(0,V)$ as n goes to infinity for V the covariance matrix of the random vector.

Let us now prove (3). Twice differentiation of M gives

$$\begin{aligned} M''(R) &= \int_{B_\nu \times B_\nu^{d-1}} |\Phi_\epsilon(t_1,t_2)|^2 \left(\frac{d^2}{dR^2} m_R(t_1,t_2) m_R(-t_1,-t_2) + 2\frac{d}{dR} m_R(t_1,t_2) \right. \\ & \left. \frac{d}{dR} m_R(-t_1,-t_2) + \frac{d^2}{dR^2} m_R(-t_1,-t_2) m_R(t_1,t_2) \right) dt_1 dt_2. \end{aligned}$$

But $m_{R^\star}(t_1,t_2) = 0$ for all (t_1,t_2), so that

$$M''(R^\star) = 2 \int_{B_\nu \times B_\nu^{d-1}} \left| \frac{d}{dR} m_{R^\star}(t_1,t_2) \right|^2 |\Phi_\epsilon(t_1,t_2)|^2 dt_1 dt_2.$$

We shall prove $M''(R^\star) \neq 0$ by contradiction.
If it is not the case, we have, for almost all $(t_1,t_2) \in B_\nu \times B_\nu^{d-1}$, $\frac{d}{dR} m_{R^\star}(t_1,t_2)\Phi_\epsilon(t_1,t_2) = 0$. Now, there exists $r_\epsilon \in (0,\nu)$ such that for all $(t_1,t_2) \in B_{r_\epsilon} \times B_{r_\epsilon}^{d-1}$, $\Phi_\epsilon(t_1,t_2) \neq 0$. Since $\frac{d}{dR} m_{R^\star}$ is a continuous function on $\mathbb{C}^d$ we get $\frac{d}{dR} m_{R^\star}(t_1,t_2) = 0$ for all $(t_1,t_2) \in B_{r_\epsilon} \times B_{r_\epsilon}^{d-1}$, that is

$$\begin{aligned} &\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,t_2)\Psi_{f^\star,R^\star}(t_1,0)\Psi_{f^\star,R^\star}(0,t_2) \\ &\qquad = \Psi_{f^\star,R^\star}(t_1,t_2)\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,0)\Psi_{f^\star,R^\star}(0,t_2) \\ &\qquad\quad + \Psi_{f^\star,R^\star}(t_1,t_2)\Psi_{f^\star,R^\star}(t_1,0)\frac{d}{dR}\Psi_{f^\star,R^\star}(0,t_2), \end{aligned} \tag{28}$$

with

$$\Psi_{f^\star,R^\star}(t_1,t_2) = \int_{(0,1)^{d-1}} f^\star(u) \exp(iR^\star t^\intercal S(u)) du$$

and

$$\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,t_2) = i \int_{(0,1)^{d-1}} [t^\intercal S(u)] f^\star(u) e^{iR^\star t^\intercal S(u)} du.$$

But $\Psi_{f^\star,R^\star}$ and $\frac{d}{dR}\Psi_{f^\star,R^\star}$ are multivariate analytic functions, so that, using Lemma C.1 in [8], we have that (28) holds for all $(t_1,t_2) \in \mathbb{C}\times\mathbb{C}^{d-1}$. We shall now investigate the set of zeros of the functions $\Psi_{f^\star,R^\star}(\cdot,0)$ and $\frac{d}{dR}\Psi_{f^\star,R^\star}(\cdot,0)$. Let $t_1 \in \mathbb{C}$ be such that $\Psi_{f^\star,R^\star}(t_1,0) = 0$. Then by Lemma 1 it is possible to choose $t_2 \in \mathbb{C}^{d-1}$ such that $\Psi_{f^\star,R^\star}(t_1,t_2) \neq 0$, and also such that $\Psi_{f^\star,R^\star}(0,t_2) \neq 0$ since $\Psi_{f^\star,R^\star}(0,\cdot)$ is a multivariate analytic function having only isolated zeros. Equation (28) then leads to $\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,0) = 0$ so that the set of zeros of the function $\Psi_{f^\star,R^\star}(\cdot,0)$ is a subset of that of the function $\frac{d}{dR}\Psi_{f^\star,R^\star}(\cdot,0)$. Then, using Hadamard's factorization theorem (see [17] Chapter 4, Theorem 4.1), and the fact that $\Psi_{f^\star,R^\star}(\cdot,0)$ and $\frac{d}{dR}\Psi_{f^\star,R^\star}(\cdot,0)$ have exponential growth order 1, we get that there exists an entire function G of exponential growth order 1 such that for any $t_1 \in \mathbb{C}$,

$$\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,0) = \Psi_{f^\star,R^\star}(t_1,0)G(t_1).$$

Plugging into (28) we get that for all $(t_1,t_2)\in\mathbb{C}\times\mathbb{C}^{d-1}$,

$$\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,t_2)\Psi_{f^\star,R^\star}(0,t_2)=\Psi_{f^\star,R^\star}(t_1,t_2)G(t_1)\Psi_{f^\star,R^\star}(0,t_2) + \Psi_{f^\star,R^\star}(t_1,t_2)\frac{d}{dR}\Psi_{f^\star,R^\star}(0,t_2).$$

The same arguments applied for each coordinate of t_2 gives that there exists a multivariate anlytic function H such that for any $t_2\in\mathbb{C}^{d-1}$,

$$\frac{d}{dR}\Psi_{f^\star,R^\star}(0,t_2)=\Psi_{f^\star,R^\star}(0,t_2)H(t_2),$$

so that for all $(t_1,t_2)\in\mathbb{C}\times\mathbb{C}^{d-1}$,

$$\frac{d}{dR}\Psi_{f^\star,R^\star}(t_1,t_2)=\Psi_{f^\star,R^\star}(t_1,t_2)\left(G(t_1)+H(t_2)\right). \tag{29}$$

But for any $t\in\mathbb{C}^d$,

$$\frac{d}{dR}\Psi_{f^\star,R^\star}(t)=\frac{1}{R}\frac{d}{du}\Psi_{f^\star,R^\star}(ut),\ u\in\mathbb{R}$$

so that solving the derivative equation (29) we find that $\Psi_{f^\star,R^\star}(t_1,t_2)$ is a product of a function of t_1 only by a function of t_2 only, meaning that $S^{(1)}(U)$ and $S^{(2)}(U)$ are independent variables, which is not true and we get a contradiction. We conclude that $M''(R^\star)\neq 0$.

To end the proof of (3), for all $R\in[0,+\infty[$,

$$\begin{aligned}M_n''(R)-M_n''(R^\star)&=\int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}}|\tilde\psi_n(t_1,t_2)|^2[a_1(t_1,t_2,R)-a_1(t_1,t_2,R^\star)]dt_1dt_2\\&+\int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}}|\tilde\psi_n(t_1,0)|^2|\tilde\psi_n(0,t_2)|^2[a_2(t_1,t_2,R)-a_2(t_1,t_2,R^\star)]dt_1dt_2\\&+\mathrm{Re}\left\{\int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}}\tilde\psi_n(-t_1,t_2)\tilde\psi_n(t_1,0)\tilde\psi_n(0,t_2)[a_3(t_1,t_2,R)-a_3(t_1,t_2,R^\star)]dt_1,dt_2\right\},\end{aligned}$$

for functions a_1, a_2 and a_3 functions that are, for all (t_1,t_2), continuous in the variable R and uniformly upper bounded for bounded R. Since for all (t_1,t_2), $|\tilde\psi_n(t_1,t_2)|\leq 1$, we get that $|M_n''(R)-M_n''(R^\star)|$ is upper bounded by

$$\int_{B_{\nu_{\text{est}}}\times B_{\nu_{\text{est}}}^{d-1}}(|a_1(t_1,t_2,R)-a_1(t_1,t_2,R^\star)|+|a_2(t_1,t_2,R)-a_2(t_1,t_2,R^\star)| + |a_3(t_1,t_2,R)-a_3(t_1,t_2,R^\star)|)dt_1dt_2$$

from which, applying the continuity theorem, we deduce that $M_n''(R_n)-M_n''(R^\star)$ converges in probability to 0 whenever R_n is a random variable converging in probability to $R^\star$. Then, for any random variable $R_n\in[R_{\min},R_{\max}]$ converging in probability to $R^\star$, $M_n''(R_n)$ converges in probability to $M''(R^\star)$.

7.6. Proof of Theorem 6

Using Taylor expansion of M_n' near $R^\star$, there exists $R_n \in (\widetilde{R}, R^\star)$ such that

$$M_n'(\widetilde{R}) = M_n'(R^\star) + (R_n - R^\star) M_n''(R_n).$$

Using $M_n'(\widetilde{R}) = 0$ and Lemma 2 we get $M_n''(R_n) = M''(R^\star) + o_{\mathbb{P}}(1)$, so that

$$\sqrt{n}(\widetilde{R} - R^\star) = -\sqrt{n}\frac{M_n'(R^\star)}{M''(R^\star)}(1 + o_{\mathbb{P}}(1)).$$

We deduce that

$$\sqrt{n}\begin{pmatrix}\widetilde{R} - R^\star \\ \widetilde{C} - C^\star\end{pmatrix} = \begin{pmatrix}0 \\ \sqrt{n}(\frac{1}{n}\sum_{l=1}^n Y_l - \mathbb{E}[Y_1])\end{pmatrix} - \begin{pmatrix}1 \\ -\mathbb{E}[S(U)]\end{pmatrix}\frac{\sqrt{n}M_n'(R^\star)}{M''(R^\star)}(1 + o_{\mathbb{P}}(1))$$

and the conclusion follows.

8. Results on Bessel functions

Denote J_α the Bessel function of order $\alpha \in [0, +\infty[$.
We shall use the following results that can be found in [20].

(I) The Bessel function of order $\alpha \in [0, +\infty[$ can be represented as

$$J_\alpha(z) = \sum_{m=0}^{\infty}(-1)^m \frac{(z/2)^{\alpha+2m}}{m!\Gamma(\alpha+m+1)}, \quad z \in \mathbb{C},$$

where for all $z \in]0, +\infty[$, $\Gamma(z) = \int_{-\infty}^{+\infty} t^{z-1}e^{-t}dt$.

(II) For $k \geq 0$ and $z \in \mathbb{C}$

$$J_{-k}(z) = (-1)^k J_k(z).$$

(III) For $z \in \mathbb{C}$ and $\theta \in \mathbb{R}$

$$\exp(iz\cos(\theta)) = \sum_{k\in\mathbb{Z}} i^k J_k(z) e^{ik\theta}.$$

(IV) For $k \geq 0$ and $x, y > 0$

$$|J_k(x) - J_k(y)| \leq |x - y|.$$

Indeed, since, $J_k \in \mathcal{C}^1(0, +\infty)$, for $k > 0$, $J_k'(z) = \frac{1}{2}(J_{k-1}(z) - J_{k+1}(z))$, $J_0'(z) = -J_1(z)$ and $|J_k(x)| \leq 1$.

(V) For $\alpha \geq 1$ and $x > 0$

$$J_{\alpha+1}(x) = \frac{2\alpha}{x}J_\alpha(x) - J_{\alpha-1}(x).$$

We prove lemmas giving useful lower bounds.

Lemma 3. *For all $0 \leq x < 1$, for all $\alpha \in [0, +\infty[$, we have*

$$J_\alpha(x) \geq \frac{x^\alpha}{2^\alpha \Gamma(\alpha+1)}(1 - \frac{x^2}{4(\alpha+1)}).$$

Proof. Let $0 \leq x < 1$, for $\alpha \geq 0$, we have $J_\alpha(x) = \sum_{m=0}^{\infty}(-1)^m \frac{(x/2)^{\alpha+2m}}{m!\Gamma(\alpha+m+1)}$, so, if we expand the sum using that $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$:

$$J_\alpha(x) = \frac{x^\alpha}{2^\alpha \Gamma(\alpha+1)}(1 - \frac{x^2}{4(\alpha+1)} + \sum_{m \geq 2}^{\infty}(-1)^m \frac{(x/2)^{2m}}{m!(\alpha+1)\cdots(\alpha+m)}).$$

Since $0 \leq x < 1$, we have $\sum_{m=2}^{\infty}(-1)^m \frac{(x/2)^{2m}}{m!(\alpha+1)\cdots(\alpha+m)} \geq 0$ thus ,

$$J_k(x) \geq \frac{x^\alpha}{2^\alpha \Gamma(\alpha+1)}(1 - \frac{x^2}{4(\alpha+1)}).$$

□

Lemma 4. *For all $R > 0$ and $0 < \nu < 1/R$, for all $0 \leq k \leq N$, we have :*

$$\int_0^\nu r J_k(rR)^2 dr \geq \frac{9\nu^2}{32} \frac{(\nu R)^{2N}}{(N+1)2^{2N}(N!)^2}.$$

Proof. Let $R > 0$ and $0 \leq r \leq \nu < 1/R$. For all $k \geq 0$, since $0 \leq rR < 1$, we have from lemma 3,

$$J_k(rR) \geq \frac{(rR)^k}{2^k k!}(1 - \frac{(rR)^2}{4(k+1)}),$$

and

$$rJ_k(rR)^2 \geq \frac{r(rR)^{2k}}{2^{2k}(k!)^2}(1 - \frac{(rR)^2}{4(k+1)})^2.$$

Then,

$$\int_0^\nu rJ_k(rR)^2 dr \geq \int_0^\nu \frac{r(rR)^{2k}}{2^{2k}(k!)^2}(1 - \frac{(rR)^2}{4(k+1)})^2 dr \geq \frac{R^{2N}}{(2N+2)2^{2N}(N!)^2}(1 - \frac{(\nu R)^2}{4})^2 \nu^{2N+2}.$$

To conclude the proof, we use that $\nu < 1/R$, so that $(1 - \frac{(\nu R)^2}{4})^2 \geq \frac{9}{16}$, which gives the result.

□

Acknowledgments

This work is supported by a public grant overseen by the French National Research Agency (ANR) through the program UDOPIA, project funded by the ANR-20-THIA-0013-01 and DATAIA convergence institute (ANR-17-CONV-0003). Jérémie Capitao-Miniconi would also like to thank the IA Chair BisCottE (ANR-19-CHIA-0021-01), Elisabeth Gassiat would like to thank Institut Universitaire de France for supporting this project.

References

[1] Aamari, E. and Levrard, C. (2019). Nonasymptotic rates for manifold, tangent space and curvature estimation. *Ann. Statist.* **47** 177–204. MR3909931

[2] Aaron, C., Cholaquidis, A. and Cuevas, A. (2017). Detection of low dimensionality and data denoising via set estimation techniques. *Electron. J. Stat.* **11** 4596–4628. MR3724969

[3] Berman, M. and Culpin, D. (1986). The statistical behaviour of some least squares estimators of the centre and radius of a circle. *J. Roy. Statist. Soc. Ser. B* **48** 183–196. MR867996

[4] Brunel, V.-E., Klusowski, J. M. and Yang, D. (2021). Estimation of convex supports from noisy measurements. *Bernoulli* **27** 772–793. MR4255215

[5] Epstein, D. and Feldman, D. (2018). Quadcopter tracks quadcopter via real time shape fitting,. *IEEE Robot. Autom. Lett.* **3** 544–550.

[6] Epstein, D. and Feldman, D. (2020). Sphere fitting with applications to machine tracking,. *Algorithms* **13**.

[7] Gassiat, E., Le Corff, S. and Lehéricy, L. (2022). Deconvolution with unknown noise distribution is possible for multivariate signals. *Ann. Statist.* **50** 303–323.

[8] Gassiat, E., Le Corff, S. and Lehéricy, L. (2022). Supplementary to the paper "Deconvolution with unknown noise distribution is possible for multivariate signals". *Ann. Statist.*

[9] Genovese, C. R., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2012). Manifold estimation and singular deconvolution under Hausdorff loss. *Ann. Statist.* **40** 941–963. MR2985939

[10] Goldenshluger, A. (2002). Density deconvolution in the circular structural model. *J. Multivariate Anal.* **81** 360–375. MR1906385

[11] Johannes, J. and Schwarz, M. (2013). Adaptive circular deconvolution by model selection under unknown error distribution. *Bernoulli* **19** 1576–1611. MR3129026

[12] Kerkyacharian, G., Pham Ngoc, T. M. and Picard, D. (2011). Localized spherical deconvolution. *Ann. Statist.* **39** 1042–1068. MR2816347

[13] Lesouple, J., Pilastre, B., Altmann, Y. and Tourneret, J.-Y. (2021). Hypersphere Fitting From Noisy Data Using an EM Algorithm. *IEEE Signal Proc. Lett.* **28** 314-318.

[14] Meister, A. (2009). *Deconvolution problems in nonparametric statistics. Lecture Notes in Statistics* **193**. Springer-Verlag, Berlin. MR2768576

[15] Ovall, J. S. (2016). The Laplacian and mean and extreme values. *Amer. Math. Monthly* **123** 287–291.

[16] Pan, L., Chu, W.-S., Saragih, J. M., De la Torre, F. and Xie, M. (2011). Fast and robust circular object detection with probabilistic pairwise voting,. *IEEE Signal Process. Lett.* **18** 639–642.

[17] Stein, E. M. and Shakarchi, R. (2003). *Complex Analysis.* Princeton University Press, Princeton.

[18] van der Vaart, A. W. (1998). *Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge University Press, Cambridge. MR1652247

[19] Wahlström, N. and Özkan, E. (2015). Extended target tracking using Gaussian processes. *IEEE Trans. Signal Process.* **63** 4165–4178. MR3367006

[20] Watson, G. N. (1995). *A treatise on the theory of Bessel functions. Cambridge Mathematical Library*. Cambridge University Press, Cambridge Reprint of the second (1944) edition. MR1349110