



A test of Gaussianity based on the Euler characteristic of excursion sets

Elena Di Bernardino, Anne Estrade, José Rafael León

► To cite this version:

Elena Di Bernardino, Anne Estrade, José Rafael León. A test of Gaussianity based on the Euler characteristic of excursion sets. *Electronic Journal of Statistics*, 2017, 11 (1), pp.843-890. 10.1214/17-EJS1248 . hal-01345225v2

HAL Id: hal-01345225

<https://hal.science/hal-01345225v2>

Submitted on 21 Dec 2016

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

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

A test of Gaussianity based on the Euler characteristic of excursion sets

Elena Di Bernardino

*CNAM, Paris, EA4629, Département IMATH
292 rue Saint-Martin, Paris Cedex 03, France
e-mail: elena.di.bernardino@cnam.fr*

Anne Estrade

*MAP5 UMR CNRS 8145, Université Paris Descartes
45 rue des Saints-Pères, 75270 Paris Cedex 06, France
e-mail: anne.estrade@parisdescartes.fr*

José R. León*

*Escuela de Matemática, Facultad de Ciencias
Universidad Central de Venezuela, Av. Paseo de los Ilustres
Los Chaguaramos, Caracas 1040, Venezuela
e-mail: jose.leon@ciens.ucv.ve*

Abstract: In the present paper, we deal with a stationary isotropic random field $X : \mathbb{R}^d \rightarrow \mathbb{R}$ and we assume it is partially observed through some level functionals. We aim at providing a methodology for a test of Gaussianity based on this information. More precisely, the level functionals are given by the Euler characteristic of the excursion sets above a finite number of levels. On the one hand, we study the properties of these level functionals under the hypothesis that the random field X is Gaussian. In particular, we focus on the mapping that associates to any level u the expected Euler characteristic of the excursion set above level u . On the other hand, we study the same level functionals under alternative distributions of X , such as chi-square, harmonic oscillator and shot noise. In order to validate our methodology, a part of the work consists in numerical experimentations. We generate Monte-Carlo samples of Gaussian and non-Gaussian random fields and compare, from a statistical point of view, their level functionals. Goodness-of-fit p -values are displayed for both cases. Simulations are performed in one dimensional case ($d = 1$) and in two dimensional case ($d = 2$), using R.

MSC 2010 subject classifications: Primary 62G10 ; secondary 60G10, 60G15, 60G60.

Keywords and phrases: Test of Gaussianity, Gaussian fields, Excursion sets, Level sets, Euler characteristic, Crossings.

Introduction

The question of the Gaussianity of a phenomenon is a historical and fundamental problem in statistical literature. This type of information can be crucial in many application problems: oceanography and waves behavior, hydrology and climatology, agronomy, neurology and spike behavior, insurance and finance, astrophysics. For instance, in this latter application, during the last decade a large number of researchers joined efforts to decide whether the Cosmic Microwave Background (CMB) temperature is Gaussian or not.

The problem of determining whether an i.i.d random sample comes from a Gaussian distribution has been studied extensively. In the case where the mean and the variance of the random variable are known, one can use a classical goodness-of-fit test. However, if these parameters need to be estimated

*Partially supported by the INRIA International Chairs program

the test of Lillifors and a variant of the Cramer-Von Mises test, with estimated parameters, are well adapted. These tests are no more distribution-free and depend on the true value of the parameters. The p -values must be obtained by simulations.

The situation becomes more complicated when the sample comes from a stationary stochastic process satisfying some mixing conditions. For this type of problem, some tests have been designed to determine whether the one dimensional marginals are Gaussian. We can cite, by way of illustration: the Eps test [12] based on the empirical characteristic function; the test built by Lobato and Velasco [22] that uses a test of symmetry and kurtosis; the test of Subba and Gabr [30] where the bi-spectral density is the basis for the test. A remarkable exception is the test built in Cuesta-Albertos et al. [11]. There, the method for constructing the test uses a one dimensional random projection, after which the projected sample is subjected to a test that infers whether the one dimensional marginal is Gaussian. Actually, the random projection procedure allows one to test whether the whole distribution of the process is Gaussian or not, and hence is not limited to marginal distributions.

In the present article, we deal with a real valued stationary isotropic random field and we use the information given by level functionals of a single realisation of the process to build a test of Gaussianity. The level functional is the Euler characteristic (EC) of the excursion sets above some levels. Our first motivation comes from the article [1] Section 7 (Model identification). In the aforementioned paper, Adler suggests to use the EC of the excursion sets as a way to determine what is the actual distribution of the observed process. His words can describe better than ours the main goal *“Suppose that we are given a real valued random field f on \mathbb{R}^3 , which we observe on the unit cube $[0, 1]^3$. Our problem is that we are not certain what the distribution of the field is. For simplicity, we shall assume that the choices are Gaussian and χ^2 , where both are assumed to be stationary and isotropic. Then one way to choose between the two is to calculate, from the data, an empirical Euler characteristic curve,”* c.f. [1]. If the data is Gaussian, then the theoretical curve is a very precise one (see Figures 1 and 9, left panels), depending on the second spectral moment of the process and some other invariant quantities. Otherwise, if it is a χ^2 -process, Kramer oscillator process or a shot noise process, a completely different curve appears (see Figures 6, 8 left panel, 10 left panel, and 11). In what follows, we offer a methodology that tries to implement these ideas.

The idea of observing level functionals of a random field in order to infer some information on the distribution of the field is not new. In [20], Lindgren provides estimators for the second spectral moment of an univariate Gaussian process that are based on the number of upcrossings at various levels. In [8], Cabaña builds a test of isotropy, for a two dimensional random field, that is based on the perimeter and the area of excursion sets. In [31], the covariance function of a bivariate Gaussian field is inferred from the excursion sets Euler characteristic. The same idea has inspired many precursors working in materials science, see for instance [28], [24]. In those papers, the modelling of images or slices of a two-phases materials (even more complicated) is achieved by using a two dimensional stationary Gaussian field that has been thresholded at a certain level. The observed data are the lengths of the respective two phases along any line extracted from the image. The aim of these studies is to identify the Gaussian covariance function. Let us also mention [2] where the authors start from the observation of a neurological space-time signal at some moderate levels and deduce some parameters that help in estimating the probability of exceeding very high values. Not far from this thema, one can find the question of exceedances or the study of extreme values, when considering high levels (see, e.g., the seminal work of Rice [27] and [35]). We will not go further in that direction and, at the opposite, stay with the observation of moderate levels. In all the mentioned papers, the field that is under study is assumed to be Gaussian. On the contrary, in the present paper, Gaussianity is not assumed but has to be inferred without knowing the spatial correlation.

To be more precise, we aim at proving that the function that associates u to the mean excursions EC at level u provides a kind of signature of the distribution of the random field under study. Although a so complex information as the knowledge of the whole distribution cannot be summarize in a single function, our guess is that the shape of its graph could be enough to discriminate between Gaussianity and non Gaussianity. Our main tool will be a Central Limit Theorem for the EC of an excursion set of a stationary isotropic Gaussian random field. This asymptotic normality takes place when the domain grows to \mathbb{R}^d . The result is proved in [13] with the help of a Breuer-Major Theorem [25]. We will also need some generalisations, extensions and explicit computations of this result.

In particular, in the present work, we extend the results in [13] by showing that the random variables

$$Z_k^i = |T_i|^{-1/2} (\varphi(X, T_i, u_k) - \mathbb{E}[\varphi(X, T_i, u_k)]) \quad \text{for } i = 1, \dots, m; k = 1, \dots, p;$$

with $\varphi(X, T_i, u_k)$ the EC of a standard Gaussian field X associated to the level u_k in the domain T_i and $\mathbb{E}[\varphi(X, T_i, u_k)]$ the theoretical Gaussian mean of excursions EC at level u_k , are asymptotically jointly distributed as Gaussian when the disjoint domains T_i grow up and satisfy asymptotic independence (see Proposition 5 for further details). Furthermore, we provide a tractable expression of the associated asymptotic variances (see Propositions 2 and 8).

The aforementioned results are used to build a Gaussianity test for a standard random field X . Indeed,

I. if the **H0** hypothesis: “the random field X is Gaussian” holds true,

then the test statistic based on the sum of the scaled Z_k^i ’s follows an appropriate chi-squared distribution. This case is for instance, illustrated by the pretty QQ-plots alignments and high p -values of goodness-of-fit tests in Sections 4.1.3 and 5.1.

Conversely,

II. if the underlying standard random field X is not Gaussian with a given second spectral moment (in particular if X is a χ^2 or a Kramer oscillator process),

we deliberately center again, as in the *I.* case, the Z_k^i ’s variables by using the (wrong) theoretical Gaussian mean of excursions EC at level u_k (see Equation (6)) with the same second spectral moment. In this case, we obtain a very small goodness-of-fit p -values for the chi-squared distribution associated to the considered test statistics based on Z_k^i ’s (see Sections 4.2.1 and 4.2.2). Then we are able to reject the H_0 hypothesis under considered alternatives. The crucial point is that the theoretical mean of excursions EC has a very different shape in the Gaussian case with respect to the considered alternatives: χ^2 -process or a Kramer oscillator process. This evident difference allows us to easily discriminate between Gaussian and non Gaussian setting.

As mentioned above, we consider two alternative hypothesis versus the Gaussianity: χ^2 -process and Kramer oscillator process. These processes are chosen for two types of reasons. First, the mean excursions EC curve for the considered processes can be analytically known. Secondly, they are typical examples of models with specific properties: positive asymmetry or oscillation. This variety of behaviors can cover real-life situations in the applications. For the same types of reasons, we also consider a discontinuous shot noise process as an alternative. However, due to the non-smoothness property of this process, we separate its study in a specific section (see Section 6). A further discussion in this sense is postponed to the conclusion section.

Main contributions of the paper. Under the assumption of stationarity and isotropy for the Gaussian field X , we give a new explicit formula for the second moment of the excursions EC (see Proposition 1 and Proposition 2). We extend the results in [13] by showing a Central Limit Theorem for

the joint excursions EC concerned with different levels and disjoint domains (see Proposition 5). We also show finiteness of the asymptotic variance of the excursions EC of a chi-square field (see Proposition 7). A numerically tractable formula for the asymptotic variance in the univariate Gaussian case is given in Proposition 8. We propose a statistical methodology to test the Gaussianity distribution and we implement that on simulated data-sets to evaluate the finite-sample performance of our strategy.

Outline of the paper. Section 1 contains the general setting and the definition of the observation tool, namely the Euler characteristic of excursion sets. In Section 2, we focus on the Gaussian hypothesis. Assuming that the field under study X is a stationary isotropic Gaussian field, we give explicit formulas for the first two moments of the excursions EC. We also recall the Central Limit Theorem satisfied by the excursions EC when the domain tends to \mathbb{R}^d (see Theorem 4) and give an extension for the joint excursions EC concerned with different levels and disjoint domains. It allows us to describe the situation in term of a statistical model with observations whose distribution is known under the null hypothesis **H0**. Section 3 is concerned with the study of two alternative distributions of the observed random field: χ^2 and Kramer oscillator. In both cases, we give an explicit formula for the mean EC of excursion sets. Section 4 is devoted to numerical illustrations for univariate processes. We generate trajectory samples of stationary processes, Gaussian and non Gaussian, and compare the theoretical mean function of the excursions EC and the empirical one. We also build some chi-square statistics and associated goodness-of-fit tests in order to quantify the deviation between Gaussian and non Gaussian case. At last, in Section 5, we go further in the numerical study by considering two dimensional random fields. We generate Gaussian and χ^2 samples and compare with the theoretical situation. All the random generations and numerical computations are performed with **R**. In the last Section 6, we consider a shot noise process as a potential alternative, giving a preliminar study that could lead to a more general test of Gaussianity for non smooth processes. In order to keep the methodological spirit of the present paper, we have reported the technical proofs to the appendix section.

1. Setting

Let us start by introducing some definitions and terminology. Let X be a real valued random field defined on \mathbb{R}^d . We will sometimes call it a *multivariate process*. We say that it is *stationary* if its distribution is invariant under translations in \mathbb{R}^d and *isotropic* if its distribution is invariant under rotations in \mathbb{R}^d . Hence, the distribution of a stationary isotropic random field is invariant under the action of the isometry group of \mathbb{R}^d .

Assuming X is smooth enough, we write X' for its derivative, which is a random vector field from \mathbb{R}^d to \mathbb{R}^d . We also write X'' for its second derivative. For any $t \in \mathbb{R}^d$, we denote by $\nabla^2 X(t)$ the $\frac{1}{2}d(d+1)$ random vector that contains the upper coefficients of the symmetric Hessian matrix $X''(t)$ and by $\mathbf{X}(t)$ the $d + \frac{1}{2}d(d+1) + 1$ random vector $(X'(t), \nabla^2 X(t), X(t))$.

Let us assume now that X is stationary, isotropic and centered. If almost every realisation of X is of class C^1 , then the covariance matrix of $X'(0)$ has the following form

$$\mathbb{E}(X'(0)X'(0)^t) = \lambda I_d, \quad (1)$$

for some $\lambda > 0$ usually named as *second spectral moment*.

All over the paper, we consider a real valued random field X defined on \mathbb{R}^d that satisfies the following assumption.

Assumption (A): The random field X is stationary, isotropic, $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and

almost all realisations belong to $C^3(\mathbb{R}^d)$. For any fixed t in \mathbb{R}^d ,

the covariance matrix of the random vector $\mathbf{X}(t)$ has full rank.

At last, the covariance function r is such that,

$$\psi(t) \rightarrow 0 \text{ when } \|t\| \rightarrow +\infty \text{ and } \psi \in L^1(\mathbb{R}^d),$$

where $\psi(t) = \max \left(\left| \frac{\partial^{\mathbf{k}} r}{\partial t^{\mathbf{k}}}(t) \right| ; \mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}^d, k_1 + \dots + k_d \leq 4 \right)$.

Remark: Assumption **(A)** could appear too strong for the task described in this paper. Actually, we are not interesting in working under the weakest assumption, our aim is rather to provide a methodology that applies for any dimension d . Since the theoretical tools that we use are mainly derived from [13], we chose to work with the same assumption setting as in this paper.

Notations.

- for any $u \in \mathbb{R}$ and any compact $T \subset \mathbb{R}^d$, we call “excursion set of X above the level u within the domain T ” the following set $\{t \in T : X(t) \geq u\}$,
- $p_Z(\cdot)$ denotes the probability density function of any random vector Z (assuming it exists),
- $|\cdot|$ denotes without any ambiguity, either the absolute value, or the d -dimensional Lebesgue measure.

Euler characteristic.

The Euler characteristic of a compact domain K in \mathbb{R}^d can be heuristically defined in the case $d = 1$ as the number of connected components of K , or in the case $d = 2$ as the number of connected components minus the number of holes in K . In the case where K is an excursion set $\{t \in T : X(t) \geq u\}$, with T a rectangle in \mathbb{R}^d and u a real number, there exists a rather tractable formula that uses the theory of Morse functions (see [3] Chapter 9, for instance). It states that the Euler characteristic of $\{t \in T : X(t) \geq u\}$ is equal to a sum of two terms. The first one only depends on the restriction of X to the interior of T , it is given by the quantity $\varphi(X, T, u)$ defined in Equation (2) below. The second one exclusively depends on the behaviour of X on the l -dimensional faces of T , with $0 \leq l < d$. From now on, we focus on the term $\varphi(X, T, u)$, named as “modified Euler characteristic” in [13], and we still call it Euler characteristic (EC). It is defined by the following,

$$\varphi(X, T, u) = \sum_{k=0}^d (-1)^k \mu_k(T, u), \text{ where} \quad (2)$$

$$\mu_k(T, u) = \#\{t \in \overset{\circ}{T} : X(t) \geq u, X'(t) = 0, \text{index}(X''(t)) = d - k\},$$

with $\overset{\circ}{T}$ the interior of T and the “index” stands for the number of negative eigenvalues.

Special case 1: dimension one. When $d = 1$, writing $[0, T]$ instead of T for a while, Equation (2) becomes

$$\begin{aligned} \varphi(X, [0, T], u) &= \#\{\text{local maxima of } X \text{ above } u \text{ in } (0, T)\} \\ &\quad - \#\{\text{local minima of } X \text{ above } u \text{ in } (0, T)\}. \end{aligned} \quad (3)$$

Morse’s theorem says that this quantity is linked with the number of up-crossings,

$$U(X, [0, T], u) = \#\{t \in [0, T] : X(t) = u, X'(t) \geq 0\},$$

by the relation

$$\varphi(X, [0, T], u) + \mathbf{1}_{\{X(0) > u, X'(0) < 0\}} + \mathbf{1}_{\{X(T) > u, X'(T) > 0\}} = U(X, [0, T], u) + \mathbf{1}_{\{X(0) > u\}}.$$

Taking expectation in both expressions and using stationarity yield the next formula that we will use in the forthcoming sections, namely Sections 3 and 4,

$$\mathbb{E}[\varphi(X, [0, T], u)] = \mathbb{E}[U(X, [0, T], u)]. \quad (4)$$

Special case 2: dimension two. When $d = 2$, Equation (2) can be rewritten in the following way. With the notations introduced within this equation, $\mu_0(T, u)$ denotes the number of local maxima above u , $\mu_2(T, u)$ denotes the number of local minima above u and $\mu_1(T, u)$ the number of local saddle points above u . Hence,

$$\begin{aligned} \varphi(X, T, u) &= \#\{\text{local extrema of } X \text{ above } u \text{ in } \mathring{T}\} \\ &\quad - \#\{\text{local saddle points of } X \text{ above } u \text{ in } \mathring{T}\}. \end{aligned} \quad (5)$$

2. Under Gaussian hypothesis

In this section, we assume that X is Gaussian and satisfies Assumption **(A)** described in Section 1.

2.1. First two moments of the Euler characteristic of an excursion set

Let T be a cube in \mathbb{R}^d . This section is devoted to explicit formulas for the first two moments of $\varphi(X, T, u)$. They are based on the decomposition in (2) and on Rice formulas for the factorial moments of $\mu_k(T, u)$ (see for instance [3] Chapter 11 or [5] Chapter 6).

In particular, using the stationarity of X , the expectation can be computed as follows

$$\begin{aligned} \mathbb{E}[\varphi(X, T, u)] &= (-1)^d \int_T \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(t)) \det(X''(t)) | X'(t) = 0] p_{X'(t)}(0) dt \\ &= (-1)^d |T| (2\pi\lambda)^{-d/2} \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) \det(X''(0))], \end{aligned}$$

where we recall that λ is the second spectral moment of X , see (1). Moreover, it is proved in [3] Lemma 11.7.1, through a regression and due to Wick's formula, that

$$\mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) \det(X''(0))] = (-1)^d (2\pi)^{-1/2} \lambda^d H_{d-1}(u) e^{-u^2/2},$$

where H_k stands for the Hermite polynomial of order k . Hence, the next formula holds

$$\mathbb{E}[\varphi(X, T, u)] = |T| (2\pi)^{-(d+1)/2} \lambda^{d/2} H_{d-1}(u) e^{-u^2/2}. \quad (6)$$

In what follows, we will be particularly interested in the next function

$$C(u) = (2\pi)^{-(d+1)/2} \lambda^{d/2} H_{d-1}(u) e^{-u^2/2}, \quad (7)$$

that yields $\mathbb{E}[\varphi(X, T, u)] = |T| C(u)$. Equation (7) shows that $C(u)$ implicitly depends on X through its dimension parameter d and its second spectral moment λ . Whenever necessary in the next sections,

we will emphasize this dependence by writing $C(u) = C(u, \lambda)$.

For the second moment, a so nice formula as (6) seems to be out of reach. Nevertheless, in the next proposition, we express the second moment as an integral that can be numerically evaluated. Let us mention the paper [32] where a similar formula is established but concerning another modified Euler characteristic, namely the differential topology (DT) characteristic. In this reference, the authors also give the second moment of the excursions DT characteristic as an integral of conditional distribution functionals and they propose a Monte-Carlo simulation method to evaluate these functions.

We will use the following functions, defined for $u \in \mathbb{R}$ and $t \in \mathbb{R}^d$,

$$\left. \begin{aligned} g(u) &= \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) | \det(X''(0))|] \\ D(t) &= (2\pi)^{2d} \det(\lambda^2 I_d - r''(t)^2) \\ G(u, t) &= \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) \mathbf{1}_{[u, \infty)}(X(t)) \det(X''(0)) \det(X''(t)) | X'(0) = X'(t) = 0] \end{aligned} \right\} \quad (8)$$

In the one dimensional case, explicit formulas for the expectation functions $g(u)$ and $G(u, t)$ are given in the numerical Section 4 (see Proposition 8).

Proposition 1. *Assume that X is Gaussian and satisfies Assumption (A).*

Then, for any $u \in \mathbb{R}$, the map $t \mapsto G(u, t) D(t)^{-1/2}$ is integrable on any compact set in \mathbb{R}^d and

$$\mathbb{E}[\varphi(X, T, u)^2] = \int_{\mathbb{R}^d} |T \cap (T - t)| G(u, t) D(t)^{-1/2} dt + |T| (2\pi\lambda)^{-d/2} g(u).$$

Proof. Integrability comes from [13] Proposition 1.1 since $X \in C^3$. In order to compute the expectation of $\varphi(X, T, u)^2$, let us start with (2). It yields

$$\varphi(X, T, u)^2 = \sum_{0 \leq k \leq d} \mu_k(T, u) + \sum_{0 \leq k \leq d} \mu_k(T, u)(\mu_k(T, u) - 1) + \sum_{0 \leq k \neq l \leq d} (-1)^{k+l} \mu_k(T, u) \mu_l(T, u).$$

The expectation of the first term is equal to

$$\begin{aligned} \mathbb{E}\left[\sum_{0 \leq k \leq d} \mu_k(T, u)\right] &= \mathbb{E}[\#\{t \in T : X(t) \geq u, X'(t) = 0\}] \\ &= \int_T \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(t)) | \det X''(t)| | X'(t) = 0] p_{X'(t)}(0) dt \\ &= |T| g(u) p_{X'(0)}(0), \end{aligned}$$

where we have used Rice Formula to get the second line and stationarity as well as the independence between $X'(0)$ and $(X(0), X''(0))$ to get the third one.

For the second and third terms, we introduce, for $k, l = 1, \dots, d$ and $s, t \in \mathbb{R}^d$,

$$\begin{aligned} F_u^{k,l}(s, t) &= \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(s)) | \det X''(s)| \mathbf{1}_{D_k}(X''(s)) \\ &\quad \times \mathbf{1}_{[u, \infty)}(X(t)) | \det X''(t)| \mathbf{1}_{D_l}(X''(t)) | X'(s) = X'(t) = 0], \end{aligned}$$

where D_k denotes the set of symmetric matrices having index equal to $d - k$. Hence, Rice Formula for the second factorial moment allows us to obtain

$$\mathbb{E}\left[\sum_{0 \leq k \leq d} \mu_k(T, u)(\mu_k(T, u) - 1)\right] = \int_{T \times T} \left(\sum_{0 \leq k \leq d} F_u^{k,k}(s, t)\right) p_{X'(s), X'(t)}(0, 0) ds dt.$$

Given that $D_k \cap D_l = \emptyset$ for $k \neq l$, one can adapt the proof of the second moment Rice formula to get

$$\begin{aligned} \mathbb{E}[\sum_{0 \leq k \neq l \leq d} (-1)^{k+l} \mu_k(T, u) \mu_l(T, u)] \\ = \int_{T \times T} \left(\sum_{0 \leq k \neq l \leq d} (-1)^{k+l} F_u^{k,l}(s, t) \right) p_{X'(s), X'(t)}(0, 0) ds dt. \end{aligned}$$

Let us remark that

$$\sum_{0 \leq k \leq d} F_u^{k,k}(s, t) + \sum_{0 \leq k \neq l \leq d} (-1)^{k+l} F_u^{k,l}(s, t) = G(u, t - s).$$

It yields

$$\mathbb{E}[\varphi(X, T, u)^2] = \int_{\mathbb{R}^d} |T \cap (T - t)| G(u, t) p_{X'(0), X'(t)}(0, 0) dt + |T| g(u) p_{X'(0)}(0).$$

It remains to compute the probability density function of $(X'(0), X'(t))$. The covariance matrix of this vector is equal to $\begin{pmatrix} \lambda I_d & -r''(t) \\ -r''(t) & \lambda I_d \end{pmatrix}$ and so $p_{X'(0), X'(t)}(0, 0) = D(t)^{-1/2}$. Hence the result. \square

2.2. Asymptotic variance

In the next proposition, we let the cube T grow to \mathbb{R}^d and we give a formula for the asymptotic variance of $\varphi(X, T, u)$. Actually, we consider

$$T^{(N)} = \{Nt : t \in T\}$$

the image of a fixed cube T by the dilation $t \mapsto Nt$ and we let N grow to $+\infty$.

Proposition 2. *Assume that X is Gaussian and satisfies Assumption (A) and let T be a cube in \mathbb{R}^d . Then for any u in \mathbb{R} ,*

$$\lim_{N \rightarrow +\infty} \text{Var}[|T^{(N)}|^{-1/2} \varphi(X, T^{(N)}, u)] = V(u) < +\infty$$

with

$$V(u) = \int_{\mathbb{R}^d} (G(u, t) D(t)^{-1/2} - C(u)^2) dt + (2\pi\lambda)^{-d/2} g(u), \quad (9)$$

where $C(u)$, $g(u)$, $G(u, t)$, $D(t)$ have been defined in (7) and (8).

Proof. From Proposition 1, for a fixed cube T , we have

$$\begin{aligned} & \text{Var}[|T|^{-1/2} \varphi(X, T, u)] \\ &= \int_{\mathbb{R}^d} \frac{|T \cap (T - t)|}{|T|} G(u, t) D(t)^{-1/2} dt + (2\pi\lambda)^{-d/2} g(u) - |T|^{-1} (\mathbb{E} \varphi(X, T, u))^2 \\ &= \int_{\mathbb{R}^d} \frac{|T \cap (T - t)|}{|T|} (G(u, t) D(t)^{-1/2} - C(u)^2) dt + (2\pi\lambda)^{-d/2} g(u), \end{aligned}$$

where we have used the relation $|T|^2 = \int_{\mathbb{R}^d} |T \cap (T - t)| dt$ to get the last line. Hence, the asymptotic formula can easily be derived using Lebesgue dominated convergence theorem conditionally to the fact that $t \mapsto G(u, t) D(t)^{-1/2} - C(u)^2$ belongs to $L^1(\mathbb{R}^d)$. This point is the matter of the next lemma.

Lemma 3. *The map $t \mapsto G(u, t) D(t)^{-1/2} - C(u)^2$ belongs to $L^1(\mathbb{R}^d)$.*

The proof of Lemma 3 is postponed in the appendix section. \square

Beyond the existence of a finite asymptotic variance as stated in the previous proposition, $\varphi(X, T^{(N)}, u)$ satisfies a Central Limit Theorem.

Theorem 4 (Theorem 2.6 in [13]). *Assume that X is Gaussian and satisfies Assumption (A) and let T be a cube in \mathbb{R}^d . Then for any u in \mathbb{R} , the next convergence holds in distribution*

$$|T^{(N)}|^{-1/2} (\varphi(X, T^{(N)}, u) - \mathbb{E}[\varphi(X, T^{(N)}, u)]) \xrightarrow[N \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, V(u)), \quad (10)$$

where $\mathcal{N}(0, V(u))$ stands for the centered Gaussian distribution with variance $V(u)$ (see Equation (9)).

2.3. Disjoint domains and various levels

We now consider two domains T_1 and T_2 that are disjoint and two levels u_1 and u_2 that can be equal or not.

Proposition 5. *Assume that X is Gaussian and satisfies Assumption (A).*

(a) *Let T_1 and T_2 be two cubes in \mathbb{R}^d such that $|T_1| = |T_2|$ and $\text{dist}(T_1, T_2) > 0$ and let u_1 and u_2 belong to \mathbb{R} . For any integer $N > 0$, we introduce*

$$Z_i^{(N)} = |T_i^{(N)}|^{-1/2} (\varphi(X, T_i^{(N)}, u_i) - \mathbb{E}[\varphi(X, T_i^{(N)}, u_i)]) \text{ for } i = 1, 2.$$

As $N \rightarrow +\infty$, $(Z_1^{(N)}, Z_2^{(N)})$ converges in distribution to a centered Gaussian vector with diagonal covariance matrix $\begin{pmatrix} V(u_1) & 0 \\ 0 & V(u_2) \end{pmatrix}$ where $V(u_i)$ is prescribed by (9).

(b) *Let T be a cube in \mathbb{R}^d and let u_1 and u_2 belong to \mathbb{R} . For any integer $N > 0$, we introduce*

$$\zeta_i^{(N)} = |T^{(N)}|^{-1/2} (\varphi(X, T^{(N)}, u_i) - \mathbb{E}[\varphi(X, T^{(N)}, u_i)]) \text{ for } i = 1, 2.$$

As $N \rightarrow +\infty$, $(\zeta_1^{(N)}, \zeta_2^{(N)})$ converges in distribution to a centered Gaussian vector with covariance matrix $\begin{pmatrix} V(u_1) & V(u_1, u_2) \\ V(u_1, u_2) & V(u_2) \end{pmatrix}$, where $V(u_i)$ is prescribed by (9) and $V(u_1, u_2)$ is a finite real number.

Comment. In [13] Theorem 2.5, the covariance $V(u_1, u_2)$ is prescribed by a series expansion that, although convergent, is so awkward that it cannot be used to evaluate it in practise.

Proof of Proposition 5.

Concerning item (a), we first have to establish that the fields $Z_1^{(N)}$ and $Z_2^{(N)}$ are asymptotically independent. Intuitively, it comes from the fact that the distance between the cubes $T_1^{(N)}$ and $T_2^{(N)}$ goes to infinity and the covariance function of X has a sufficient rate of decay at infinity due to Assumption (A). Precisely, the asymptotic decorrelation is given in the next lemma.

Lemma 6. *Under assumptions of Proposition 5 and using the same notation, it holds that*

$$\text{Cov}(Z_1^{(N)}, Z_2^{(N)}) \xrightarrow[N \rightarrow +\infty]{} 0.$$

Lemma 6 is proved in the appendix section. By using this result and Theorem 4, we know that the covariance matrix of the random vector $(Z_1^{(N)}, Z_2^{(N)})$ tends to $\begin{pmatrix} V(u_1) & 0 \\ 0 & V(u_2) \end{pmatrix}$. One can use the same arguments as those in [13] (which were inspired by the Breuer-Major Theorem of [25]) to establish that any linear combination $xZ_1^{(N)} + yZ_2^{(N)}$ has a Gaussian limit in distribution. Item (b) is proved in [13], Theorem 2.5. \square

2.4. Statistical model

We are now able to build a test for the **H0** hypothesis: “the random field X is Gaussian”. Actually, we assume that X is observed through the family $(Y_k^i)_{1 \leq i \leq m; 1 \leq k \leq p}$ where

$$Y_k^i = \frac{\varphi(X, T_i, u_k)}{|T_i|}, \quad i = 1, \dots, m; \quad k = 1, \dots, p,$$

for disjoint domains T_1, \dots, T_m that have the same large volume and are at large distance one from each other, and various levels $u_1 \leq \dots \leq u_p$. Under these conditions, by using Proposition 5 and Equation (6), one can write the following approximation valid under **H0**,

$$Y_k^i \approx C(u_k, \lambda) + \epsilon_k^i, \quad i = 1, \dots, m; \quad k = 1, \dots, p, \quad (11)$$

where $(\epsilon_k^i)_{i,k}$ is a mp -dimensional centered Gaussian vector with covariance

$$\text{Cov}(\epsilon_k^i, \epsilon_l^j) = \text{Cov}(Y_k^i, Y_l^j) = \delta_{ij} V(u_k, u_l) / |T_i|. \quad (12)$$

The deterministic value $C(u_k, \lambda)$ is given by Equation (7). It depends on the level u_k and on the standard field X only through its second spectral moment λ . For $k = l$, $V(u_k, u_l) = V(u_k)$ is given by (9), and for $k \neq l$, since no explicit formula is available for $V(u_k, u_l)$ in practise, it has to be estimated. Using the statistical model in (11)-(12), we will focus on two particular cases.

- (a) **Diagonal case.** We take m disjoint domains T_1, \dots, T_m and m levels u_1, \dots, u_m , such that *each level u_k is associated to a single domain T_k* . In this setting we have m observations Y_1^1, \dots, Y_m^m and a diagonal covariance matrix equal to $\text{diag}(V(u_1)/|T_1|, \dots, V(u_m)/|T_m|)$.
- (b) **Crossed case.** We take m disjoint domains T_1, \dots, T_m and p levels u_1, \dots, u_p , such that *different levels u_k are associated to the same domain T_i* . In this setting we have mp observations $(Y_k^i)_{i,k}$ and their covariance matrix is given by (12).

3. Under alternative hypothesis

3.1. χ^2 hypothesis

In this section, we deal with χ^2 distributions instead of Gaussian ones. Let us fix s , a non negative integer, as the degrees of freedom.

We start with $\{X_i(\cdot)\}_{i=1}^s$, an independent sample of standard stationary Gaussian fields on \mathbb{R}^d that satisfy Assumption (A) of Section 1. We denote by r_X their covariance function and recall that $r_X(0) = 1$. Consider now the following stationary fields

$$\chi_s^2(\cdot) = \sum_{i=1}^s (X_i(\cdot))^2 \quad \text{and} \quad Z^{(s)}(\cdot) = \frac{1}{\sqrt{2s}} \left(\sum_{i=1}^s (X_i(\cdot))^2 - s \right).$$

Note that for any $t \in \mathbb{R}^d$, $\chi_s^2(t)$ is a chi-square random variable with s degrees of freedom. One get readily that

$$\mathbb{E}[Z^{(s)}(0)] = 0, \quad \text{Var}[Z^{(s)}(0)] = 1, \quad \mathbb{E}[Z^{(s)}(t)Z^{(s)}(0)] = r_X(t)^2.$$

Therefore, $Z^{(s)}$ also satisfies Assumption **(A)**. Moreover, its second spectral moment is equal to λ with $\lambda = -((r_X)^2)''(0) = -2r_X''(0)$.

We are now interested in the expectation of the Euler characteristic of

$$\{t \in T : Z^{(s)}(t) \geq u\} = \{t \in T : \chi_s^2(t) \geq s + u\sqrt{2s}\},$$

for a fixed cube $T \subset \mathbb{R}^d$ and a fixed level u in \mathbb{R} . A formula for the mean EC of excursion sets of χ^2 fields is given in [34], Theorem 3.5 (see also Theorem 15.10.1 in [3]). It applies in our context, although we have to handle carefully with the second spectral moment of the X_i 's, which is equal to $\lambda/2$. It yields

$$\begin{aligned} & \mathbb{E}[\varphi(Z^{(s)}, T, u)] \\ &= |T| \left(\frac{\lambda}{\pi} \right)^{d/2} \frac{e^{-(s+u\sqrt{2s})/2} (s + u\sqrt{2s})^{(s-d)/2}}{2^{(s-2+2d)/2} \Gamma(s/2)} P_{d,s}(s + u\sqrt{2s}) \mathbf{1}_{[0,\infty)}(s + u\sqrt{2s}), \end{aligned} \quad (13)$$

where $P_{d,s}(\cdot)$ is a polynomial of degree $d - 1$ with integer coefficients (see Section 3.3 in [34]). In particular, we stress that $P_{1,s}(u) = 1$ and $P_{2,s}(u) = u - s + 1$.

Let us recall that, in dimension one, the mean Euler characteristic of the excursion above the level u is equal to the mean number of upcrossings at level u , see Equation (4). With this point of view, Formula (13) can also be found in [29] for instance. Next proposition is concerned with the second moment in dimension one, in the same spirit as Proposition 2.

Proposition 7. *Let $d = 1$. Let us assume that the X'_i 's are one dimensional i.i.d. Gaussian processes that satisfy Assumption **(A)** for $i = 1, \dots, s$.*

Then, for any u in \mathbb{R} and T in $[0, +\infty)$, $\varphi(Z^{(s)}, [0, T], u)$ admits a finite second moment.

Moreover, there exists $v_s(u) \in [0, +\infty)$ such that

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \text{Var}[\varphi(Z^{(s)}, [0, T], u)] = v_s(u).$$

The proof of Proposition 7 can be found in the appendix section.

3.2. Kramer oscillator hypothesis

We work in dimension $d = 1$. Let us consider the following system of stochastic differential equations, well known as Kramer oscillator system,

$$\left. \begin{aligned} dQ(t) &= P(t)dt, \\ dP(t) &= \sigma dW(t) - (cP(t) + V'(Q(t)))dt, \end{aligned} \right\} \quad (14)$$

where $V(q) = a_0 q^4 - a_1 q^2$ for positive constants a_0 and a_1 , σ and c are also positive constants, and W is a Brownian motion. The asymptotic properties of such a system have been studied for instance in

[36]. Besides, it is well known that the Markov process (Q, P) has an invariant measure μ that can be written (up to a numerical constant factor) as

$$d\mu(p, q) = \exp\left(-\frac{2c}{\sigma^2}\left(\frac{p^2}{2} + V(q)\right)\right) dp dq.$$

From now on, we assume that $(Q(0), P(0))$'s distribution is proportional to μ , so that (Q, P) is stationary. Here, we are interested in the process Q . It is stationary, centered, but certainly not Gaussian since its distribution is proportional to $\exp\left(-\frac{2c}{\sigma^2}V(q)\right) dq$. Nevertheless, the distribution of its derivative process, $Q' = P$, is actually Gaussian with zero mean and variance equal to $\frac{\sigma^2}{2c}$. An application of Rice formula gives the mean number of upcrossings of Q in $[0, T]$ at any level u and hence, using (4), we get

$$\begin{aligned} \mathbb{E}[\varphi(Q, [0, T], u)] &= \mathbb{E}[U(Q, [0, T], u)] \\ &= T \left(\frac{1}{\sqrt{2\pi(\sigma^2/2c)}} \int_0^\infty p \exp\left(-\left(\frac{c}{\sigma^2}p^2\right)dp\right) p_Q(u) \right) \\ &= \frac{T\sigma}{2\sqrt{\pi c}} p_Q(u) \quad \text{with} \quad p_Q(u) = \frac{\exp\left(-\frac{2c}{\sigma^2}V(u)\right)}{\int_{\mathbb{R}} \exp\left(-\frac{2c}{\sigma^2}V(q)\right)dq}. \end{aligned} \quad (15)$$

Moreover, let us remark that a suitable choice of the parameters σ, c, a_0, a_1 allows us to get $\text{Var}(Q(0)) = 1$ and $\text{Var}(Q'(0)) = \lambda$, so that Q satisfies the same moments constraints as the generic process X in Section 1. Actually, it is sufficient to prescribe $\frac{\sigma^2}{2c} = \lambda$ and to solve the following non linear equation

$$\int_{\mathbb{R}} \exp\left(-\frac{2c}{\sigma^2}(a_0 q^4 - a_1 q^2)\right) dq = \int_{\mathbb{R}} q^2 \exp\left(-\left(\frac{2c}{\sigma^2}(a_0 q^4 - a_1 q^2)\right)\right) dq. \quad (16)$$

4. Univariate numerical illustrations

In this section, we focus on the one dimensional case and hence only deal with univariate processes. Both for Gaussian and alternative distributions, we first compare the theoretical formulas for the moments of excursions EC with the empirical ones, obtained by Monte Carlo simulations. Secondly, we perform the test statistics and calculate associated goodness-of-fit p -values. A graphical illustration of our test is also provided by drawing QQ-plots.

4.1. Gaussian process

4.1.1. First and second moments: $\mathbb{E}[\varphi(X, T, u)]$ and $V(u)$

We rewrite Equations (6) and (9) of Section 2 in the case $d = 1$ under **H0** hypothesis, i.e. when X is a stationary isotropic standard Gaussian process with covariance function r and second spectral moment $\lambda = -r''(0)$:

$$\mathbb{E}[\varphi(X, T, u)] = |T| C(u, \lambda) \quad \text{with} \quad C(u, \lambda) = (2\pi)^{-1} \lambda^{1/2} e^{-u^2/2}, \quad (17)$$

and $\text{Var}[\varphi(X, T, u)] \sim |T| V(u)$ as $|T| \rightarrow +\infty$, with

$$V(u) = \int_{\mathbb{R}} (G(u, t) D(t)^{-1/2} - C(u, \lambda)^2) dt + (2\pi\lambda)^{-1/2} g(u), \quad (18)$$

where $D(t) = (2\pi)^2 (\lambda^2 - r''(t)^2)$ and $g(u), G(u, t)$ are given by (8). In the one dimensional case, it yields

$$\begin{aligned} g(u) &= \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) |X''(0)|] \\ G(u, t) &= \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) \mathbf{1}_{[u, \infty)}(X(t)) X''(0) X''(t) | X'(0) = X'(t) = 0]. \end{aligned}$$

Our first task is now to provide formulas for $g(u)$ and $G(u, t)$ that can be numerically evaluated. They will obviously include Gaussian integrals and, therefore, we need to describe the Gaussian distributions that are involved.

On the one hand, introducing $\gamma = r^{(4)}(0)$, the covariance matrix of $(X(0), X''(0))$ is $\begin{pmatrix} 1 & -\lambda \\ -\lambda & \gamma \end{pmatrix}$.

On the other hand, the conditional distribution

$$\mathcal{L}(X(0), X(t), X''(0), X''(t) | X'(0) = X'(t) = 0)$$

is a 4-dimensional centered Gaussian distribution with covariance matrix given by

$$\Gamma(t) = C_{11}(t) - C_{12}(t)C_{22}(t)^{-1}C_{12}(t)^T,$$

where $C_{11}(t)$ is the covariance matrix of the vector $(X(0), X(t), X''(0), X''(t))$, $C_{22}(t)$ is the covariance matrix of the vector $(X'(0), X'(t))$ and $C_{12}(t)$ is the matrix of the covariances between those two vectors.

We still need some extra notations:

- ϕ denotes the standard Gaussian density, Φ the standard Gaussian distribution, and $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$;
- for any $(u, \alpha, m) \in \mathbb{R}^3$, we introduce

$$\begin{cases} Q^{(0)}(u, \alpha, m) = \int_u^\infty \bar{\Phi}(\alpha x + m) \phi(x) dx ; & Q^{(1)}(u, \alpha, m) = - \int_u^\infty \phi(\alpha x + m) \phi'(x) dx \\ Q^{(2)}(u, \alpha, m) = - \int_u^\infty \phi'(\alpha x + m) \phi(x) dx ; & Q^{(3)}(u, \alpha, m) = \int_u^\infty \bar{\Phi}(\alpha x + m) \phi''(x) dx. \end{cases}$$

In Proposition 8 below, we finally give explicit formulas for the functions g and G . They will be useful to give a numerical evaluation of $V(u)$ in (18) for various values of u .

Proposition 8. *Let X be a Gaussian process that satisfies Assumption (A).*

1. *For any u in \mathbb{R} ,*

$$g(u) = \lambda \phi(u) \left(2\Phi\left(\frac{\lambda u}{(\gamma - \lambda^2)^{1/2}}\right) - 1 \right) + \left(\frac{2\gamma}{\pi}\right)^{1/2} \bar{\Phi}\left(\left(\frac{\gamma}{\gamma - \lambda^2}\right)^{1/2} u\right).$$

2. *Let $u \in \mathbb{R}$ and $t \in \mathbb{R}$ be fixed. We denote by $(l_{ij})_{1 \leq i \leq j \leq 4}$ the coefficients of any lower triangular matrix $L(t)$ such that $L(t)L(t)^T = \Gamma(t)$. Then*

$$\begin{aligned} G(u, t) &= (l_{41}l_{31} + l_{42}l_{32} + l_{43}l_{33}) Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) \\ &\quad + (l_{42}l_{31} + l_{41}l_{32}) Q^{(1)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) \\ &\quad + l_{42}l_{32} Q^{(2)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) + l_{41}l_{31} Q^{(3)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right). \end{aligned}$$

The proof of this proposition is postponed in the appendix section.

In order to illustrate (17) and (18), we generate a 300 Monte-Carlo sample of a stationary standard Gaussian process with covariance function $r(t) = e^{-t^2}$. Note that it implies the second spectral moment $\lambda = 2$. By using the Cholesky function `chol` in R programm we are able to numerically evaluate $L(t)$ matrix, for all t . These quantities are necessary to build g , G and numerically approximate the integral that gives the asymptotic variance $V(u)$ (see Equation (18)). In order to evaluate $\varphi(X, T, u)$ on each realisation of X , we use Equation (3) for various values of u . Comparison between theoretical formulas and empirical counterparts are shown in Figure 1.

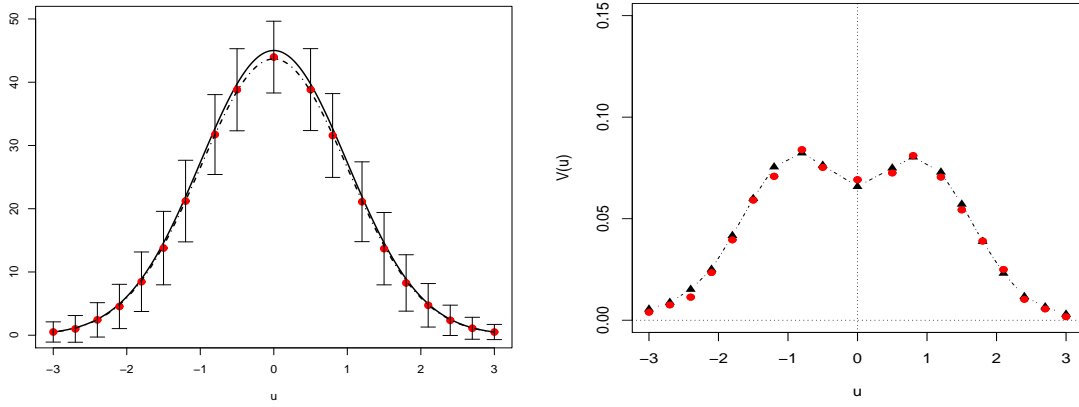


FIG 1. **Left:** Theoretical $u \mapsto \mathbb{E}[\varphi(X, T, u)]$ from Equation (17) for $|T| = 200$ (full line). We also display, for different levels u , the empirical counterpart $\widehat{\mathbb{E}}[\varphi(X, T, u)]$ (red dots) based on 300 Monte Carlo simulations. We also display the associated empirical intervals. The map $u \mapsto |T|C(u, \widehat{\lambda})$ from Equation (17), by using the estimated spectral moment $\widehat{\lambda}$ as in Section 4.1.2, is represented in dashed line. **Right:** Theoretical $u \mapsto V(u)$ from Equation (18) for various values of u (black triangles) and empirical variance of $|T|^{-1/2} \varphi(X, T, u)$ (red dots) based on 300 Monte Carlo simulations with $|T| = 200$ and the same values of u . $X(\cdot)$ is a **Gaussian univariate process** ($d = 1$) with $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and covariance function $r(t) = e^{-t^2}$. In this case $\lambda = 2$.

4.1.2. Estimation of the second spectral moment

It has been already quoted that the second spectral moment of X , denoted by λ , plays an important role in Equation (17). Actually, all the influence of X in $\mathbb{E}[\varphi(X, T, u)]$ is summarised through this parameter. From our statistical point of view, since we aim at inferring the distribution of X from observations of the excursions of X , parameter λ is unknown *a priori*. It has to be estimated from the observed level functionals X .

We estimate $\gamma = \lambda^{1/2}$ by using an unbiased estimator introduced by Lingren [20] for stationary zero-mean Gaussian processes. It is based on p different levels $u_1 < u_2 < \dots < u_p$ by the following prescription,

$$\widehat{\gamma} = \sum_{k=1}^p c_k \widehat{\gamma}_{u_k}, \quad \text{with } \widehat{\gamma}_{u_k} = 2\pi T^{-1} e^{u_k^2/2} \varphi(X, [0, T], u_k). \quad (19)$$

Actually, in the paper of Lingren, the number of up-crossings of level u_k in the interval $[0, T]$, namely $U(X, [0, T], u_k)$, is used instead of $\varphi(X, [0, T], u_k)$ (recall the analogy described in Section 1). As a

general rule, considering

$$p = 3, (u_1, u_2, u_3) = (-u, 0, u), \text{ with } u = \frac{2}{3} \sqrt{\text{Var}(X(t))}, c_1 = c_2 = c_3 = \frac{1}{3}$$

seems to be an acceptable choice (see discussion in [20]). In this case $\hat{\lambda}^{1/2} = \hat{\gamma} = \frac{1}{3} (\hat{\gamma}_{-u} + \hat{\gamma}_0 + \hat{\gamma}_u)$. An illustration of this estimation procedure is given in Figure 2. This estimation of $\lambda^{1/2}$ is also used in Figure 1 (left), dashed line.

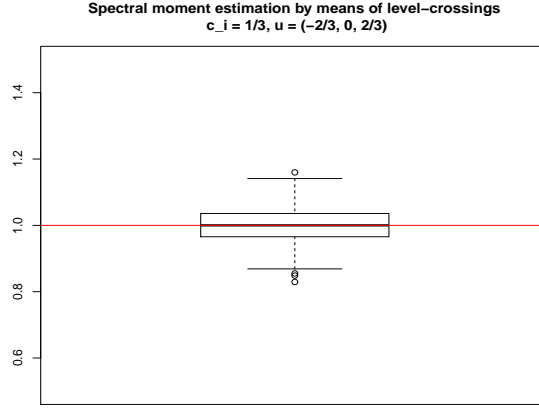


FIG 2. Boxplot, based on 300 Monte Carlo simulations, of the ratio between theoretical value $\lambda^{1/2}$ and estimated value $\hat{\gamma}$ given by (19). $X(\cdot)$ is a **Gaussian univariate process** ($d = 1$) with $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and covariance function $r(t) = e^{-t^2}$. In this case $\lambda = 2$. We choose $p = 3$, $c_1 = c_2 = c_3 = \frac{1}{3}$ and $u = (-2/3, 0, 2/3)$.

The same estimation strategy could be developed in the alternative hypothesis framework as soon as we have at our disposal a closed formula linking the mean EC with the second spectral moment (see for instance Equation (21) in the univariate chi-square case).

4.1.3. Chi-square statistics

In the following we will consider particular sub-cases of the two cases presented in Section 2.4, diagonal case (a) and crossed case (b). In particular, we focus on the four next models.

- (a) **Diagonal case.** Each level u_k is associated to a single domain T_k . We have m observations Y_1^1, \dots, Y_m^m in the statistical model (11).
 - (a.1) Case with same level u , i.e. $u_1 = \dots = u_m = u$, and m disjoint domains T_1, \dots, T_m .
 - (a.2) Case with different levels $u_1 < \dots < u_m$ and m disjoint domains T_1, \dots, T_m .
- (b) **Crossed case.** Different levels u_k can be associated to the same domain T_i . We have mp observations $(Y_k^i)_{1 \leq i \leq m, 1 \leq k \leq p}$ in (11).
 - (b.1) Case with one single domain T ($m = 1$) and p different levels $u_1 < \dots < u_p$.
 - (b.2) Case with m disjoint domains T_1, \dots, T_m and p different levels $u_1 < \dots < u_p$.

We will now associate chi-square statistics with each of the four aforementioned models. For the **diagonal case (a.1)**, we consider

$$F_{a1} := \sum_{i=1}^m \left(\frac{\varphi(X, T_i, u) - \mathbb{E}[\varphi(X, T_i, u)]}{\sqrt{|T_i| V(u)}} \right)^2 \quad \tilde{F}_{a1} := \sum_{i=1}^m \left(\frac{\varphi(X, T_i, u) - \widehat{\mathbb{E}}[\varphi(X, T_i, u)]}{\sqrt{\widehat{Var}(\varphi(X, T_i, u))}} \right)^2$$

where $\mathbb{E}[\varphi(X, T_i, u)] = |T_i| C(u, \lambda)$ is given by (17) and $V(u)$ is given by (18). Furthermore, $\widehat{\mathbb{E}}$ and \widehat{Var} in \tilde{F}_{a1} are respectively the empirical mean and the empirical variance on considered Monte-Carlo sample generations.

Under **H0** hypothesis, a consequence of (11)-(12) is that both random variables F_{a1} and \tilde{F}_{a1} are approximately χ_m^2 distributed, i.e. central chi-square with m degrees of freedom. We evaluate F_{a1} and \tilde{F}_{a1} on 300 Monte Carlo simulations. We choose $m = 3$ and $u = 1.2$. The QQ-plot comparison between the obtained empirical quantiles with the theoretical quantiles of a centered χ_m^2 distribution is gathered in Figure 3 (first and second panels).

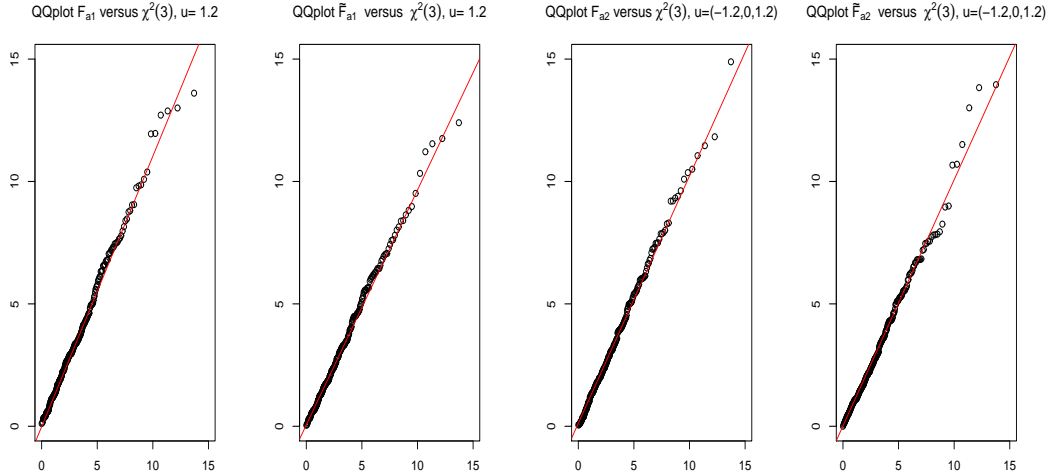


FIG 3. QQ-plot based on 300 Monte Carlo simulations. **First and second panels:** quantiles of F_{a1} , \tilde{F}_{a1} versus quantiles of the χ_m^2 distribution for $m = 3$ disjoint domains T_1, T_2, T_3 with $|T_i| = 200$ and a single level $u = 1.2$. **Third and fourth panels:** quantiles of F_{a2} , \tilde{F}_{a2} versus quantiles of the χ_m^2 distribution for $m = 3$ disjoint domains T_1, T_2, T_3 with $|T_i| = 200$ and different levels $u_1 = -1.2, u_2 = 0, u_3 = 1.2$. $X(\cdot)$ is a **Gaussian univariate process** ($d = 1$) with $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and covariance function $r(t) = e^{-t^2}$ with $\lambda = 2$. See associated goodness-of-fit test p -values in Table 1.

Consider now the **diagonal case (a.2)**. Let

$$F_{a2} := \sum_{i=1}^m \left(\frac{\varphi(X, T_i, u_i) - \mathbb{E}[\varphi(X, T_i, u_i)]}{\sqrt{|T_i| V(u_i)}} \right)^2 \quad \tilde{F}_{a2} := \sum_{i=1}^m \left(\frac{\varphi(X, T_i, u_i) - \widehat{\mathbb{E}}[\varphi(X, T_i, u_i)]}{\sqrt{\widehat{Var}(\varphi(X, T_i, u_i))}} \right)^2$$

where $\widehat{\mathbb{E}}$ and \widehat{Var} are respectively the empirical mean and the empirical variance on considered Monte Carlo sample generations. By using again the statistical model in (11)-(12), both F_{a2} and \tilde{F}_{a2} are approximately χ_m^2 distributed. An illustration is presented in Figure 3 (third and fourth panels), by choosing 300 Monte Carlo simulations, $m = 3$, $|T_i| = 200$ and different levels $u_1 = -1.2, u_2 = 0, u_3 =$

GOF test p -values for χ_m^2	Cramér-von Mises test	Anderson-Darling test
F_{a1}	0.7162	0.7519
\tilde{F}_{a1}	0.6379	0.7585
F_{a2}	0.8933	0.9576
\tilde{F}_{a2}	0.7518	0.7731

TABLE 1

Goodness-of-fit test p -values associated to test statistics F_{a1} , \tilde{F}_{a1} , F_{a2} and \tilde{F}_{a2} in Figure 3. The considered process $X(\cdot)$ is a **Gaussian univariate process** ($d = 1$) with $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and covariance function $r(t) = e^{-t^2}$.

1.2. Furthermore, in Table 1 we display the goodness-of-fit test p -values for F_{a1} , \tilde{F}_{a1} , F_{a2} and \tilde{F}_{a2} in Figure 3. Obtained p -values allow us to statistically accept the χ_m^2 distribution for these test statistics. Moreover, we highlight how p -values increase when one evaluates the EC curve for several levels u_i (see F_{a2} and \tilde{F}_{a2} cases).

For the **crossed case (b.1)**, let us now consider different levels $u_1 < \dots < u_p$ and one single domain T , i.e. $m = 1$. Let us define

$$F_{b1} := \|\Lambda^{-1}(\mathbf{Z} - \mathbb{E}[\mathbf{Z}])\|^2, \quad \tilde{F}_{b1} := \|\hat{\Lambda}^{-1}(\mathbf{Z} - \hat{\mathbb{E}}[\mathbf{Z}])\|^2,$$

where \mathbf{Z} is the p -dimensional Gaussian vector given by

$$\mathbf{Z} = \left(\frac{\varphi(X, T, u_k)}{|T|^{1/2}} \right)_{1 \leq k \leq p},$$

$\hat{\mathbb{E}}$ is the empirical mean on the Monte-Carlo simulations and the matrices Λ and $\hat{\Lambda}$ are defined below. Let $\Gamma = (V(u_k, u_l))_{1 \leq k, l \leq p}$ be the (theoretical) covariance matrix of \mathbf{Z} and let Λ stand for any square root of Γ . Similarly, let $\hat{\Gamma}$ be the empirical covariance matrix of \mathbf{Z} evaluated on the Monte-Carlo sample generations, and let $\hat{\Lambda}$ be any of its square root matrix. Hence, F_{b1} and \tilde{F}_{b1} are both approximately χ_{mp}^2 distributed, with $m = 1$. An illustration of the behaviour of \tilde{F}_{b1} is presented in Figure 4 (first panel), by choosing 300 Monte-Carlo simulations, $m = 1$, $|T| = 200$ and $p = 3$ different levels $u_1 = -1.5, u_2 = 0, u_3 = 1.5$.

Consider now the **crossed case (b.2)**. In this setting we have m disjoint domains T_1, \dots, T_m and p different levels $u_1 < \dots < u_p$. Let

$$F_{b2} := \sum_{i=1}^m \|\Lambda^{-1}(\mathbf{Z}^i - \mathbb{E}[\mathbf{Z}^i])\|^2, \quad \tilde{F}_{b2} := \sum_{i=1}^m \|\hat{\Lambda}_{(i)}^{-1}(\mathbf{Z}^i - \hat{\mathbb{E}}[\mathbf{Z}^i])\|^2,$$

where for any $i \in \{1, \dots, m\}$, \mathbf{Z}^i is the p -dimensional Gaussian vector given by

$$\mathbf{Z}^i = \left(\frac{\varphi(X, T_i, u_k)}{|T_i|^{1/2}} \right)_{1 \leq k \leq p}$$

and $\hat{\mathbb{E}}$ is the empirical mean on the Monte-Carlo simulations. Let $\Gamma = (V(u_k, u_l))_{1 \leq k, l \leq p}$ be the (theoretical) covariance matrix of \mathbf{Z}^i . Let Λ stand for any square root of Γ . Similarly, let $\hat{\Gamma}_{(i)}$ be the empirical covariance matrix of \mathbf{Z}^i evaluated on the Monte-Carlo sample generations, and let $\hat{\Lambda}_{(i)}$ be any of its

square root matrix. Hence, $\|\Lambda^{-1}(\mathbf{Z}^i - \mathbb{E}[\mathbf{Z}^i])\|^2$ and $\|\hat{\Lambda}_{(i)}^{-1}(\mathbf{Z}^i - \hat{\mathbb{E}}[\mathbf{Z}^i])\|^2$ are both approximately χ_p^2 distributed.

Moreover, since for $1 \leq i \neq j \leq m$, the Gaussian vectors \mathbf{Z}^i and \mathbf{Z}^j are independent, F_{b2} and \tilde{F}_{b2} are still centered χ^2 distributed with now mp degrees of freedom.

An illustration of the behaviour of \tilde{F}_{b2} is presented in Figure 4 (second panel), by choosing 300 Monte Carlo simulations, $m = 3$ disjoint domains T_1, T_2, T_3 with $|T_i| = 200$ and $p = 3$ different levels $u_1 = -1.5, u_2 = 0, u_3 = 1.5$. Furthermore, in Table 2 we display the goodness-of-fit test p -values for \tilde{F}_{b1} and \tilde{F}_{b2} in Figure 4. Obtained p -values allow us to statistically accept the chi-squared distribution for these test statistics.

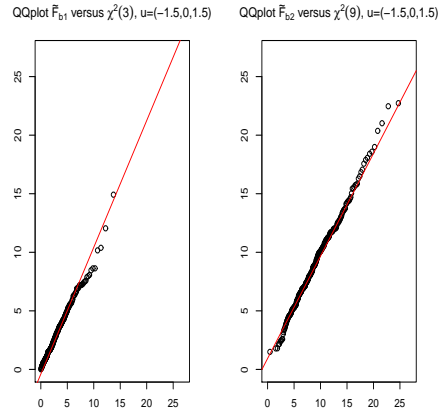


FIG 4. *QQ-plot based on 300 Monte Carlo simulations. **First panel:** quantiles of \tilde{F}_{b1} versus quantiles of the χ_p^2 distribution for $m = 1$, with $|T| = 200$ and $p = 3$ different levels $u_1 = -1.5, u_2 = 0, u_3 = 1.5$. **Second panel:** quantiles of \tilde{F}_{b2} versus quantiles of the χ_{mp}^2 distribution for $m = 3$ disjoint domains T_1, T_2, T_3 with $|T_i| = 200$ and $p = 3$ different levels $u_1 = -1.5, u_2 = 0, u_3 = 1.5$. $X(\cdot)$ is a **Gaussian univariate process** ($d = 1$) with $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and covariance function $r(t) = e^{-t^2}$ with $\lambda = 2$. See associated goodness-of-fit test p -values in Table 2.*

GOF test p -values for χ_{mp}^2	Cramér-von Mises test	Anderson-Darling test
\tilde{F}_{b1} ($m = 1, p = 3$)	0.7725	0.8548
\tilde{F}_{b2} ($m = 3, p = 3$)	0.9745	0.9886

TABLE 2

*Goodness-of-fit test p -values associated to test statistics \tilde{F}_{b1} and \tilde{F}_{b2} in Figure 4. The considered process $X(\cdot)$ is a **Gaussian univariate process** ($d = 1$) with $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$ and covariance function $r(t) = e^{-t^2}$.*

4.2. Alternative processes

In this section, we simulate trajectories of an alternative standard process Z . We compare it with a standard Gaussian process X having the same second spectral moment.

Similarly to Section 4.1.3, we now consider a test statistic. For sake of brevity, we only focus on the

diagonal case **(a.1)** with a unique level u and m disjoint domains T_1, \dots, T_m , and we introduce

$$D := \sum_{i=1}^m \left(\frac{\varphi(Z, T_i, u) - \mathbb{E}[\varphi(X, T_i, u)]}{\sqrt{\widehat{Var}(\varphi(Z, T_i, u_i))}} \right)^2, \quad (20)$$

where $\mathbb{E}[\varphi(X, T_i, u)]$ is the mean Euler characteristic of the excursions of X given by (6) with $d = 1$ and $\widehat{Var}(\varphi(Z, T_i, u_i))$ is the empirical variance on considered Monte-Carlo sample generations. In order to use this empirical variance, one has to be sure that the variance of the Euler characteristic of the excursions of Z is finite. We recall that this is the case when Z is chi-square distributed (see Proposition 7).

Under **H0** hypothesis, the D statistic is asymptotically distributed as a chi-square with m degrees of freedom.

4.2.1. First alternative: χ^2 process

In conformity with Section 3.1, we consider

$$Z^{(s)}(\cdot) = \frac{1}{\sqrt{2s}}(\chi_s^2(\cdot) - s) \quad \text{with} \quad \chi_s^2(\cdot) = \sum_{1 \leq i \leq s} X_i(\cdot)^2,$$

where the X_i 's are independent copies of a stationary standard Gaussian process with covariance function $r(t) = e^{-t^2/2}$. Hence, we get $\lambda = -2r''(0) = 2$ and then the obtained process $Z^{(s)}$ has the same variance and the same second spectral moment λ as the previous Gaussian one presented in Section 4.1.

In the first and third panels of Figure 5, we display the boxplot for the ratio between the empirical mean of 300 Monte Carlo values of $\varphi(Z^{(s)}, T, u)$ and the theoretical mean given by (13) in the case $d = 1$, i.e.,

$$\mathbb{E}[\varphi(Z^{(s)}, T, u)] = |T| \left(\frac{\lambda}{\pi} \right)^{1/2} \frac{(s + u\sqrt{2s})^{(s-1)/2}}{2^{s/2}\Gamma(s/2)} e^{-(s+u\sqrt{2s})/2} \mathbf{1}_{[0, \infty)}(s + u\sqrt{2s}). \quad (21)$$

This expectation has to be compared with the expectation given by (17) when X is a stationary standard Gaussian process with second spectral moment $\lambda = 2$. Therefore, in the second and fourth panels of Figure 5, we display the boxplot for the ratio between the empirical 300 Monte Carlo mean value of $\varphi(Z^{(s)}, T, u)$ and the Gaussian theoretical expectation given in (17).

Remark. When the degrees of freedom s tend to infinity, the CLT implies that $Z^{(s)}$ tends in distribution to a stationary centered Gaussian process with covariance function equal to $t \mapsto r(t)^2$, which implies a variance equal to 1 and a second spectral moment equal to 2 (exactly as the Gaussian process X considered in Section 4.1). On the other hand, using Stirling approximation for the Γ function, one can prove that the right-hand side of (21) tends to the right-hand side of (17), i.e.,

$$\mathbb{E}[\varphi(Z^{(s)}, T, u)] \rightarrow |T| \lambda^{1/2} (2\pi)^{-1} e^{-u^2/2}, \quad \text{when } s \rightarrow \infty.$$

See also Formula (3.4) in [29] for the same remark. The comparison between the second and fourth panels in Figure 5, as well as Figure 6, illustrate this convergence.

In Figure 7 below, we show the QQ-plot of the test statistic D (see (20)) versus the quantiles of χ_m^2 . Since the considered chi-square process $Z^{(s)}$ is not Gaussian, a deviation can be observed. In order

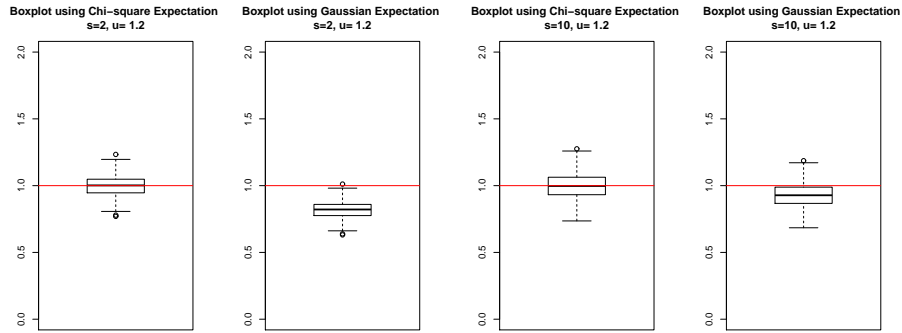


FIG 5. **First and third panels:** Boxplot for the ratio between the empirical 300 Monte Carlo values of $\varphi(Z^{(s)}, T, u)$ and the theoretical mean given by Equation (21). **Second and fourth panels:** Boxplot for the ratio between the same empirical values and the theoretical mean in the Gaussian case given by Equation (17). In both cases $\lambda = 2$. The degrees of freedom s is chosen equal to 2 and 10 respectively.

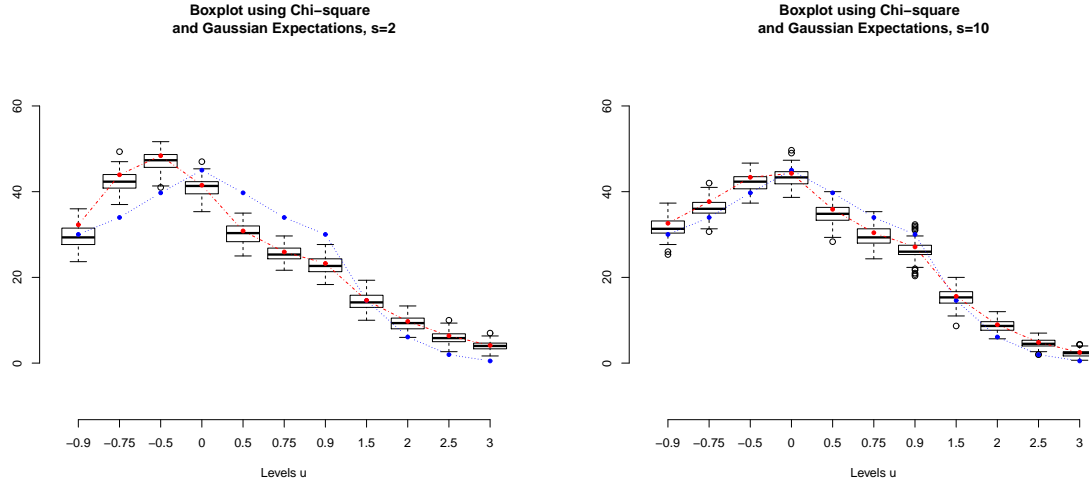


FIG 6. Boxplots of the empirical 300 Monte Carlo values of $\varphi(Z^{(s)}, T, u)$ for the considered **chi-square univariate process** $Z^{(s)}$ and different values of u . Red points represent the theoretical means given by Equation (21) for the same values of u ; blue ones are the Gaussian means given by Equation (17). In this case $\lambda = 2$, $|T| = 200$, $s = 2$ (left panel) and $s = 10$ (right panel).

to quantify how we are really able to reject the null hypothesis **H0** under the alternative, in Table 3 we display the generated p -values for the goodness-of-fit test of D . One can easily see that this test allows us to reject the fact that the considered underlying process $Z^{(s)}$ is Gaussian. This is particularly evident if one compares these p -values with those in Tables 1 and 2.

Figure 7 shows a huge deviation from the bisector of the first orthant in the case $s = 2$ whereas this deviation is less significant in the case $s = 10$. It emphasizes the convergence in distribution of $Z^{(s)}$ towards a Gaussian as s tends to infinity.

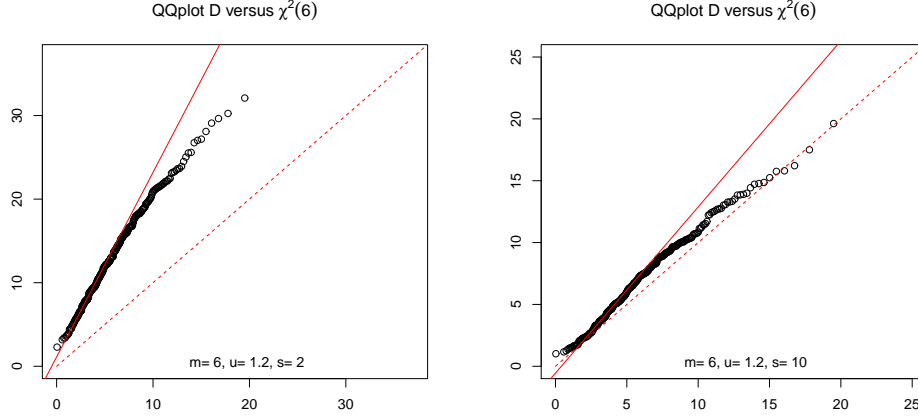


FIG 7. *QQ-plot: quantiles of D versus quantiles of the χ_m^2 distribution for fixed level $u = 1.2$ and $m = 6$ disjoint domains. Here $|T_i| = 200$ for $i \in \{1, \dots, m\}$. The dashed red line represents the bisector of the first orthant. We take 300 Monte Carlo simulations. The considered process $Z^{(s)}(\cdot)$ is a **chi-square univariate process** ($d = 1$) with $s = 2$ (first panel), $s = 10$ (second panel), $\mathbb{E}(Z^{(s)}(0)) = 0$, $\text{Var}(Z^{(s)}(0)) = 1$ and $\lambda = 2$. See associated goodness-of-fit test p -values in Table 3.*

GOF test p -values for χ_m^2	Cramér-von Mises test	Anderson-Darling test
D with $s = 2$	7e-08	2e-06
D with $s = 10$	0.0003	2e-06

TABLE 3

Goodness-of-fit test p -values associated to test statistics D in Figure 7 versus a χ_m^2 distribution with $m = 6$. The considered process $Z^{(s)}(\cdot)$ is a **chi-square univariate process** ($d = 1$) with $\mathbb{E}(Z^{(s)}(0)) = 0$, $\text{Var}(Z^{(s)}(0)) = 1$ and $\lambda = 2$.

4.2.2. Second alternative: Kramer oscillator process

In this section we generate a 300 Monte-Carlo sample of a Kramer oscillator process as defined in Section 3.2. In order to obtain a process Q with zero mean, unit variance and second spectral moment equal to 2, we solve Equation (16) and choose

$$\sigma = 2, \quad c = 1, \quad a_0 = 1, \quad a_1 = 2.3373, \quad C = 4.886.$$

The generation procedure is the following. We define a discretized schema to simulate the solution $(Q(t), P(t))$ of the stochastic differential system in (14). Actually, we use the Metropolization of the Euler-Verlet schema with a sufficient small discretization step. The interested reader is referred to Algorithm 2.11 (Generalized Hybrid Monte-Carlo) in [19].

The comparison between the (theoretical) expectation given by (15) and the empirical one is shown in left panel of Figure 8 below.

Furthermore, as we did in previous Section 4.2.1, we also consider the D test statistic. The QQ-plot of D versus χ_m^2 distribution is shown in right panel of Figure 8. Since the considered process Q is not Gaussian, a strong deviation is observed. It allows us to reject the **H0** hypothesis. To quantify this deviation we present the generated p -values for the goodness-of-fit test associated to D in Table 4.

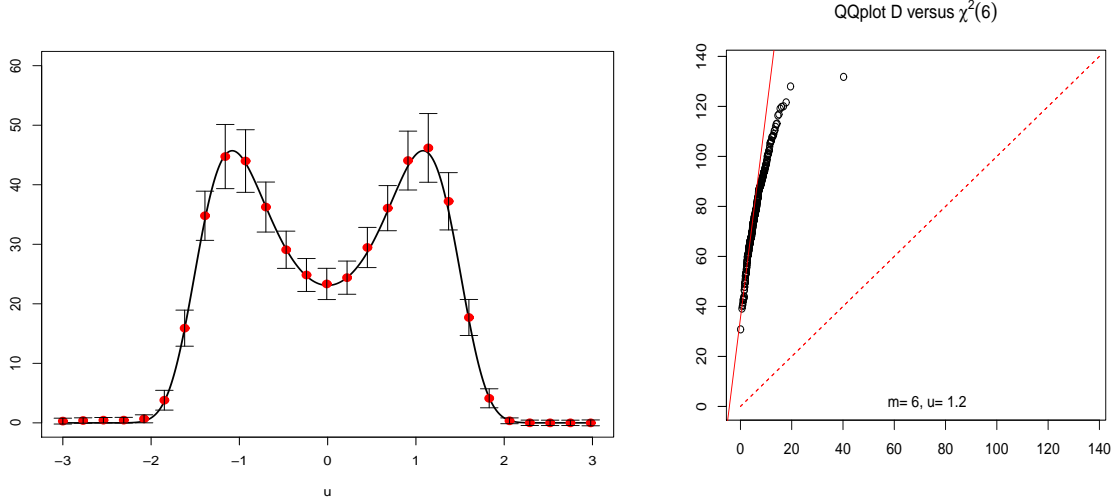


FIG 8. **Left:** Theoretical $u \mapsto \mathbb{E}[\varphi(Q, T, u)]$ from Equation (15) for $|T| = 200$ (full line). We also display, for different levels of u , the empirical counterpart $\widehat{\mathbb{E}}[\varphi(Q, T, u)]$ (red dots), with associated empirical intervals, based on 300 Monte Carlo simulations. **Right:** QQ-plot: quantiles of D versus quantiles of the χ_m^2 distribution for fixed level $u = 1.2$ and $m = 6$ disjoint domains. The dashed red line represents the bisector of the first orthant. Here $|T_i| = 200$ for $i \in \{1, \dots, m\}$. We take 300 Monte Carlo simulations. $Q(\cdot)$ is a **Kramer oscillator process** as defined in Section 3.2 with $\mathbb{E}(Q(0)) = 0$, $\text{Var}(Q(0)) = 1$ and second spectral moment $\lambda = 2$. See associated goodness-of-fit test p -value in Table 4.

GOF test p -values for χ_m^2	Cramér-von Mises test	Anderson-Darling test
D	0.000097	2e-06

TABLE 4

Goodness-of-fit test p -value associated to test statistic D in Figure 8 (right) versus a χ_m^2 distribution with $m = 6$. The considered process $Q(\cdot)$ is a **Kramer oscillator process** as defined in Section 3.2 with $\mathbb{E}(Q(0)) = 0$, $\text{Var}(Q(0)) = 1$ and second spectral moment $\lambda = 2$.

5. Bivariate numerical illustration

In this section, we focus on two dimensional random fields. For both Gaussian and chi-square distributions, we apply the same methodology as in the previous section. In particular, we use the test statistic D introduced in (20), its definition being not specific to dimension one.

5.1. Under H_0 hypothesis

We consider a stationary centered Gaussian random field $X = \{X(t) : t \in \mathbb{R}^2\}$. Its restriction to a finite regular grid included can be seen as a model for a grey level image. The modified Euler characteristic of an excursion set in the rectangle domain T above level u is given by Equation (5). On the other hand, Equation (6) in dimension $d = 2$ gives its expectation,

$$\mathbb{E}[\varphi(X, T, u)] = |T| C_2(u, \lambda) \quad \text{with} \quad C_2(u, \lambda) = (2\pi)^{-3/2} \lambda u e^{-u^2/2}. \quad (22)$$

In what follows, we generate a 300 Monte-Carlo sample of a bivariate stationary centered Gaussian random field X with covariance function $r(t) = e^{-\|t\|^2}$, $t \in \mathbb{R}^2$. In that case, $\mathbb{E}(X(0)) = 0$, $\text{Var}(X(0)) = 1$

and the second spectral moment λ is equal to 2. We use (5) in order to compute $\varphi(X, T, u)$ for a fixed cube T and various values of u . The local extremum points of X are given by the R function `extrema2dC` in the `EMD` package. To identify the saddle points, we find all the stationary points of X in the considered domain T (i.e., all the points with an associated null gradient function) and we exclude the points previously identified as local extremum points.

In the left panel of Figure 9, the comparison between the theoretical mean given by (22) and the empirical mean based on the simulations is illustrated.

In the right panel, we consider the chi-square statistic D introduced in (20) for a unique level u and m disjoint domains T_1, \dots, T_m . The considered process is now the Gaussian bivariate process X .

Furthermore, in Table 5, we display the goodness-of-fit test p -values for D . The obtained high p -values allow us to statistically accept the χ_m^2 distribution for the test statistic D .

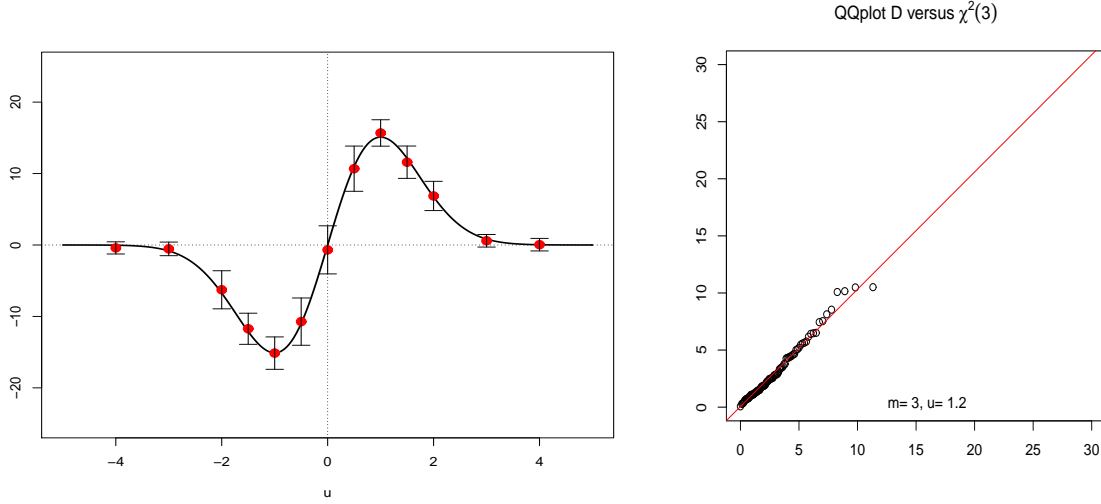


FIG 9. **Left:** Theoretical $u \mapsto \mathbb{E}[\varphi(X, T, u)]$ from Equation (22) for $|T| = 196$ (full line). We also display, for different levels u , the empirical counterpart $\hat{\mathbb{E}}[\varphi(X, T, u)]$ (red dots), with associated empirical intervals, based on 300 Monte Carlo simulations. **Right:** QQ-plot: quantiles of D versus quantiles of the χ_m^2 distribution for fixed level $u = 1.2$ and $m = 3$ disjoint domains. Here $|T_i| = 196$ for $i \in \{1, \dots, m\}$. We take 300 Monte Carlo simulations. $X(\cdot)$ is a **Gaussian bivariate process** ($d = 2$) with covariance function $r(t) = e^{-\|t\|^2}$, $t \in \mathbb{R}^2$. In this case $\mathbb{E}(X(0)) = 1$, $\text{Var}(X(0)) = 1$ and $\lambda = 2$. See associated goodness-of-fit test p -value in Table 5.

GOF test p -values for χ_m^2	Cramér-von Mises test	Anderson-Darling test
D	0.8879	0.8685

TABLE 5

Goodness-of-fit test p -value associated to test statistic D in Figure 9 (right) versus a χ_m^2 distribution with $m = 3$. The considered process $X(\cdot)$ is a **Gaussian bivariate process** ($d = 2$) with covariance function $r(t) = \exp(-\|t\|^2)$. In this case $\lambda = 2$.

5.2. Alternative: χ^2 bivariate process

As an alternative to hypothesis **H0**, let us consider

$$Z^{(s)}(\cdot) = \frac{1}{\sqrt{2s}}(\chi_s^2(\cdot) - s) \quad \text{with} \quad \chi_s^2(\cdot) = \sum_{1 \leq i \leq s} X_i(\cdot)^2,$$

where the X_i 's are independent copies of a centered stationary Gaussian two dimensional field with covariance function $r(t) = e^{-\|t\|^2/2}$. As described in Section 3.1, for any fixed $t \in \mathbb{R}^2$, $\chi_s^2(t)$ has a central χ^2 distribution with s degrees of freedom. Furthermore, the field $Z^{(s)}$ is centered, stationary and its covariance function is given by $r(t) = e^{-\|t\|^2}$. Hence its variance is equal to 1 and its second spectral moment is equal to $\lambda = -2r''(0) = 2$. The expectation of $\varphi(Z^{(s)}, T, u)$ is given by (13) with $d = 2$.

We generate a sample of 300 realisations of the bivariate process $Z^{(s)}$ for $s = 2$ on a fixed cube T . We use Equation (5) to empirically compute $\varphi(Z^{(2)}, T, u)$ for various values of u . On the other hand, Equation (13) yields the following in the case $d = s = 2$,

$$\mathbb{E}[\varphi(Z^{(2)}, T, u)] = |T| \left(\frac{\lambda}{4\pi} \right) (2u + 1) e^{-(u+1)} \mathbf{1}_{[0, \infty)}(u + 1). \quad (23)$$

In the left panel of Figure 10, we compare Equation (23) and its empirical counterpart based on the Monte Carlo simulations of $Z^{(2)}$. In the right panel, we show the QQ-plot of the test statistic D versus the χ_m^2 distribution, where D is computed with the realisations of $Z^{(2)}$. Since the considered chi-square process $Z^{(2)}$ is not Gaussian, a large deviation can be observed. Associated generated p -value for the goodness-of-fit test of D is given in Table 6. This quantifies how we are really able to reject the null hypothesis **H0** under the alternative.

GOF test p -values for χ_m^2	Cramér-von Mises test	Anderson-Darling test
D	9.4e-05	3e-06

TABLE 6

Goodness-of-fit test p -value associated to test statistic D in Figure 10 (right) versus a χ_m^2 distribution with $m = 3$. The considered process $Z^{(2)}(\cdot)$ is a **chi-square bivariate process** ($d = 2$) with covariance function $r(t) = e^{-\|t\|^2}$. In this case $\mathbb{E}(Z^{(2)})(0) = 1$, $\text{Var}(Z^{(2)})(0) = 1$ and $\lambda = 2$.

6. Non smooth processes: a case study

In this section, we sketch a preliminary study concerning non smooth processes. For this purpose, the shot noise process is a convenient toy model. We will only consider the one dimensional case pointing out two recent references that deal with dimension two, [7] and [18].

Let $d = 1$ and let us introduce the following shot noise process S ,

$$S(t) = \left(\sum_{\xi \in \Psi} \mathbf{1}_{[0, a]}(t - \xi) \right) - \nu a, \quad t \in \mathbb{R}, \quad (24)$$

where $a > 0$ and Ψ is a homogeneous Poisson point process on \mathbb{R} with intensity $\nu > 0$. The process S is clearly stationary with zero mean and variance equal to νa . Moreover, its values almost surely belong to the discrete set $\{k - \nu a, k \in \mathbb{N}\}$.

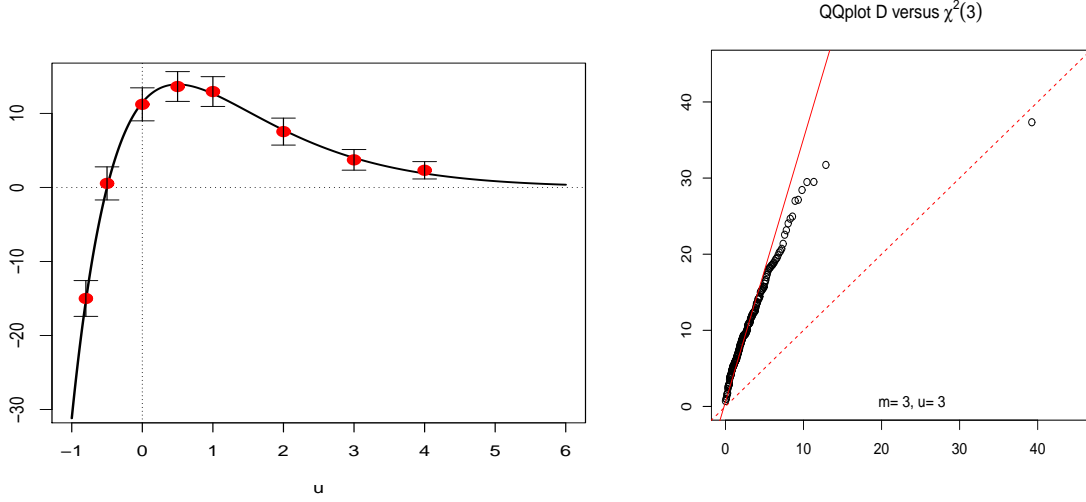


FIG 10. **Left:** Theoretical $u \mapsto \mathbb{E}[\varphi(Z^{(2)}, T, u)]$ from Equation (23) for $|T| = 196$ (full line). We also display, for different levels u , the empirical counterpart $\widehat{\mathbb{E}}[\varphi(Z^{(2)}, T, u)]$ (red dots), with associated empirical intervals, based on 300 Monte Carlo simulations. **Right:** QQ-plot: quantiles of D versus quantiles of the χ_m^2 distribution for fixed level $u = 3$ and $m = 3$ disjoint domains. Here $|T_i| = 196$ for $i \in \{1, \dots, m\}$. We take 300 Monte Carlo simulations. $Z^{(2)}(\cdot)$ is a **chi-square bivariate process** ($d = 2$) with covariance function $r(t) = e^{-\|t\|^2}$. In this case $\mathbb{E}(Z^{(2)})(0) = 1$, $\text{Var}(Z^{(2)}(0)) = 1$ and $\lambda = 2$. See associated goodness-of-fit test p -value in Table 6.

In what follows, we still aim at testing the Gaussianity hypothesis. If a whole trajectory of S could be observed, its discrete shape would clearly indicate the right choice between Gaussian and shot noise. We now assume that only excursion sets of S above a few levels are observed and we consider the number of “upcrossings”. Note that the notion of upcrossings has to be properly defined in this context, since S is not continuous. Actually, it is proved in [6] (see also [16] for a different approach) that for any level $u \in \mathbb{R} \setminus \{k - \nu a, k \in \mathbb{N}\}$ and any interval T ,

$$\mathbb{E}[U(S, T, u)] = |T| \nu e^{-\nu a} \sum_{k \geq 0} \frac{(\nu a)^k}{k!} \mathbf{1}_{\{k < u + \nu a < k+1\}}. \quad (25)$$

This formula, which obviously differs from (6), is an indicator that could enable to discriminate between a Gaussian or a shot noise model.

In the following, we generate 300 trajectories of such a process on a fixed interval T with $a = \nu = 1$. Since $a\nu = 1$, then $\text{Var}(S(0)) = 1$. A comparison between the theoretical expectation and the Monte-Carlo empirical mean is presented in Figure 11. One can observe the evident different shape of the mean EC curve compared to the Gaussian case (see Figure 1, left).

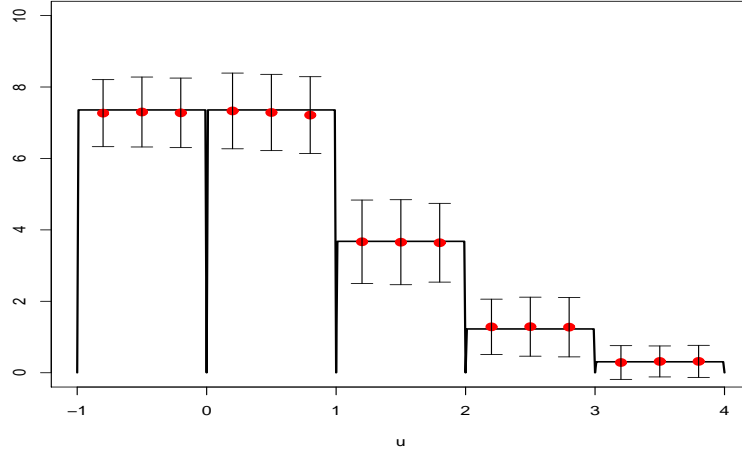


FIG 11. Theoretical $u \mapsto \mathbb{E}[U(S, T, u)]$ from Equation (25) for $|T| = 200$ (full line). We also display, for different levels u , the empirical counterpart $\hat{\mathbb{E}}[U(S, T, u)]$ (red dots) with associated empirical intervals based on 300 Monte Carlo simulations. $S(\cdot)$ is a **shot noise univariate process** ($d = 1$) defined by Equation (24) with $a = \nu = 1$.

7. Conclusion

In the present paper, we have proposed a methodology to test whether a random field defined on \mathbb{R}^d is Gaussian. The test statistics are computed from the observation of a single realisation, precisely the Euler characteristic of excursion sets at moderate levels is concerned. Our approach requires two theoretical ingredients that have to be valid under the Gaussian hypothesis: a Central Limit Theorem satisfied by the Euler characteristic of excursion sets when the domain tends to \mathbb{R}^d , and a closed formula that gives the mean Euler characteristic of the excursion sets. We have established both of them. We consider this methodological work as a first step, there are many open questions and possible future researches extending and going further this work.

There are two major domains where the Gaussianity hypothesis is really relevant and where our results could certainly be applied: neurology and cosmology. The neurological signals that are collected when studying brain activity are commonly observed through their level sets (see [35] and [21]). A natural question concerns the type of the signal distribution: Gaussian or Poisson? Gaussian or Oscillator? The test that we have built, together with the alternatives that we have considered, could be an appropriate tool for studying this problem. On the other hand, in the analysis of the Cosmic Microwave Background (CMB) radiation, the question of Gaussianity has been tackled in a huge amount of publications (see [26] for a recent overview). In that context, the random field under study is defined on the two dimensional celestial sphere. Hence, in order to be applied, our methodology should be extended beyond the Euclidean case. Let us mention that recent studies, like [10],[23],[9] for instance, contain theoretical results on high energy behavior of spherical random fields that could provide the required background for a test of Gaussianity in the same spirit as the one presented here.

Furthermore, it is not difficult to imagine that the same question of Gaussianity is actually asked in other real-life situations, and that the potential alternatives take various shapes, like a chi-square when the process under study is always positive, or an oscillator when the process under study presents an

almost periodic structure (sea waves) or a shot noise process when sudden peaks are observed. Our methodology could also serve as a goodness-of-fit test in the opposite way. For instance, in the area of geostatistics, the authors of [14] observe the Euler characteristic of excursion sets to detect abrupt changes on soil data sets. They use as an a priori model for the data a chi-square random field. A natural extension to this study could be a test of chi-square distribution versus not chi-square.

In order to go further in this direction, it would be necessary to establish Central Limit Theorems for the Euler characteristic of excursion sets for new classes of processes or random fields. Beside the present work, we implemented simulations that lead us to postulate that such a Central Limit Theorem could exist, in particular for the two alternative processes that we have considered.

At last, another appealing question concerns the hypothesis that the process under study is expected to satisfy, such as smoothness of the realisations and fast decay of the covariance function. As already mentioned at the beginning of the paper, Assumption **(A)** seems to strong at least in the one dimensional case. Indeed, let us remark that in that case, two important results of Section 2 remain valid under weaker assumptions. Actually, on the one hand, the Rice formula gives the mean number of upcrossings under the only assumption of a finite second spectral moment. On the other hand, it is well known that a sufficient condition for the number of upcrossings to have a finite second moment is the so-called Geman condition, which only requires that r'' , the second derivative of the covariance function of X , exists and satisfies some integrability condition near 0. Such considerations allows us to believe that the theoretical results that we have presented deserve to be studied under weaker assumptions, even in dimension greater than one. Another extension could concern non continuous random fields, like the shot noise processes.

Appendix: Proofs

A.1. Proof of Lemma 3.

The level u is fixed, so we do not refer to it in the following lines. For $t \in \mathbb{R}^d \setminus \{0\}$, we define $h(t) = G(u, t) D(t)^{-1/2} - C(u)^2$. We have to establish that h is integrable on \mathbb{R}^d . We already know that h is integrable on any compact set of \mathbb{R}^d , so it remains to study the behavior of $h(t)$ for large $\|t\|$. We write $K = \frac{1}{2}d(d+1) + 1$ and we introduce the function

$$(z, x) \in \mathbb{R}^K = \mathbb{R}^{\frac{1}{2}d(d+1)} \times \mathbb{R} \mapsto f(z, x) = \det(z) \mathbf{1}_{[u, \infty)}(x),$$

where $\det(z)$ stands for the determinant of the $d \times d$ symmetric matrix whose upper coefficient are given by the $\frac{1}{2}d(d+1)$ dimensional vector z .

Recall that we denote by $\mathbf{X}(\cdot)$ the stationary and isotropic random vector field $(X'(t), \nabla^2 X(t), X(t))$, which has dimension $D = d + \frac{1}{2}d(d+1) + 1 = d + K$. We have

$$\begin{aligned} h(t) &= \int_{\mathbb{R}^{2K}} f(\ddot{x}_1, x_1) f(\ddot{x}_2, x_2) \\ &\quad (p_{\mathbf{X}(0), \mathbf{X}(t)}(0, \ddot{x}_1, x_1, 0, \ddot{x}_2, x_2) - p_{\mathbf{X}(0)}(0, \ddot{x}_1, x_1) p_{\mathbf{X}(t)}(0, \ddot{x}_2, x_2)) dx_1 d\ddot{x}_1 dx_2 d\ddot{x}_2. \end{aligned}$$

Let Σ be the covariance matrix of $\mathbf{X}(0)$ and, for any $t \in \mathbb{R}^d$, let $\Gamma^{\mathbf{X}}(t)$ be the covariance matrix of $(\mathbf{X}(0), \mathbf{X}(t))$. The latter can be written by blocks, each of one having dimension $D \times D$, as follows

$$\Gamma^{\mathbf{X}}(t) = \begin{pmatrix} \Sigma & \Gamma_{12}(t) \\ \Gamma_{12}(t)^T & \Sigma \end{pmatrix}.$$

Moreover we know that all the terms of matrix $\mathbf{\Gamma}^{\mathbf{X}}(t)$ are uniformly bounded in absolute value by $\psi(t)$. Using Theorem 3.1 of [33] we get, for any $\mathbf{z}_1, \mathbf{z}_2$ in \mathbb{R}^D ,

$$\begin{aligned} & p_{\mathbf{X}(0), \mathbf{X}(t)}(\mathbf{z}_1, \mathbf{z}_2) - p_{\mathbf{X}(0)}(\mathbf{z}_1) p_{\mathbf{X}(t)}(\mathbf{z}_2) \\ &= p_{\mathbf{X}(0)}(\mathbf{z}_1) p_{\mathbf{X}(0)}(\mathbf{z}_2) \left(\sum_{\mathbf{J} \in \mathbb{N}^{D \times D}, |\mathbf{J}| > 0} H_{A_{\mathbf{J}}}(\mathbf{z}_1, \Sigma) H_{B_{\mathbf{J}}}(\mathbf{z}_2, \Sigma) \frac{\Gamma_{12}(t)^{\mathbf{J}}}{\mathbf{J}!} \right), \end{aligned}$$

where

- the sum runs over all $D \times D$ matrices $\mathbf{J} = (J_{jk})_{1 \leq j, k \leq D}$ with non negative integer entries and $|\mathbf{J}| = \sum_{1 \leq j, k \leq D} J_{jk} > 0$,
- $A_{\mathbf{J}}$ and $B_{\mathbf{J}}$ are D -dimensional indices defined by $(A_{\mathbf{J}})_j = \sum_{1 \leq k \leq D} J_{jk}$ and $(B_{\mathbf{J}})_k = \sum_{1 \leq j \leq D} J_{jk}$,
- $H_{\mathbf{k}}(\mathbf{z}, \Sigma) = \phi_{\Sigma}(\mathbf{z})^{-1} (-\frac{\partial}{\partial \mathbf{z}})^{\mathbf{k}} \phi_{\Sigma}(\mathbf{z})$ for any D -dimensional index \mathbf{k} and any $\mathbf{z} \in \mathbb{R}^D$, with ϕ_{Σ} the probability density function of any D -dimensional Gaussian vector $\mathcal{N}(0, \Sigma)$,
- $M^{\mathbf{J}} = \prod_{1 \leq j, k \leq D} (M_{jk})^{J_{jk}}$ for any matrix $M = (M_{jk})_{1 \leq j, k \leq D}$,
- and $\mathbf{J}! = \prod_{1 \leq j, k \leq D} (J_{jk})!$.

Thus

$$\begin{aligned} h(t) &= \int_{\mathbb{R}^{2K}} f(\mathbf{x}_1) f(\mathbf{x}_2) p_{\mathbf{X}(0)}(0, \mathbf{x}_1) p_{\mathbf{X}(0)}(0, \mathbf{x}_2) \\ &\quad \left(\sum_{\mathbf{J}, |\mathbf{J}| > 0} H_{A_{\mathbf{J}}}((0, \mathbf{x}_1), \Sigma) H_{B_{\mathbf{J}}}((0, \mathbf{x}_2), \Sigma) \frac{\Gamma_{12}(t)^{\mathbf{J}}}{\mathbf{J}!} \right) d\mathbf{x}_1 d\mathbf{x}_2. \end{aligned}$$

We can bound $|h(t)|$ in the following form

$$|h(t)| \leq \sum_{\mathbf{J}, |\mathbf{J}| > 0} \frac{\psi(t)^{|\mathbf{J}|}}{\mathbf{J}!} D(A_{\mathbf{J}}) D(B_{\mathbf{J}}), \quad (26)$$

where

$$\begin{aligned} D(A_{\mathbf{J}}) &= \int_{\mathbb{R}^K} |f(\mathbf{x}_1) H_{A_{\mathbf{J}}}((0, \mathbf{x}_1), \Sigma)| p_{\mathbf{X}(0)}(0, \mathbf{x}_1) d\mathbf{x}_1 \\ &\leq \mathbb{E} \left(|\tilde{\det}(\nabla^2 X(0)) H_{A_{\mathbf{J}}}((0, \nabla^2 X(0), X(0)), \Sigma)| \right) \\ &\leq \mathbb{E}(\tilde{\det}(\nabla^2 X(0))^2)^{1/2} \mathbb{E} \left(H_{A_{\mathbf{J}}}((0, \nabla^2 X(0), X(0)), \Sigma)^2 \right)^{1/2} \\ &\leq \psi(0)^d \mathbb{E} \left(H_{A_{\mathbf{J}}}((0, \nabla^2 X(0), X(0)), \Sigma)^2 \right)^{1/2}, \end{aligned} \quad (27)$$

where the last inequality is a consequence of expanding $\tilde{\det}(\nabla^2 X(0))$ as a multivariate polynomial function of degree d evaluated at the coordinates of the matrix $X''(0)$.

We now concentrate on $H_{A_{\mathbf{J}}}((0, \nabla^2 X(0), X(0)), \Sigma)$. Following [33], we have

$$H_{A_{\mathbf{J}}}((0, \mathbf{x}), \Sigma) = \mathbb{E} \left[\left(\Sigma^{-1}((0, \mathbf{x}) + iZ) \right)^{A_{\mathbf{J}}} \right],$$

where Z is any D -dimensional centered Gaussian random vector with covariance matrix Σ . Remember that Σ is the covariance matrix of $\mathbf{X}(0) = (X'(0), \nabla^2 X(0), X(0))$, so it can be factorized as $\Sigma = \begin{pmatrix} \lambda I_d & 0 \\ 0 & \Sigma_1 \end{pmatrix}$ where Σ_1 has size $K \times K$. Hence $\Sigma^{-1} = \begin{pmatrix} \lambda^{-1} I_d & 0 \\ 0 & \Sigma_1^{-1} \end{pmatrix}$ and Z can be expanded as

$Z = (Z_0, Z_1)$ where Z_0 is a d -dimensional centered Gaussian vector with covariance matrix λI_d and Z_1 is a K -dimensional centered Gaussian vector with covariance matrix Σ_1 , Z_0 and Z_1 being independent. Then

$$\Sigma^{-1}((0, \mathbf{x}) + iZ) = (i\lambda^{-1}Z_0, \Sigma_1^{-1}(\mathbf{x} + iZ_1)) ,$$

where the two blocks of coordinates are independent. So

$$H_{A_{\mathbf{J}}}((0, \mathbf{x}), \Sigma) = H_{A_{\mathbf{J}}^{(0)}}(0, \lambda I_d) H_{A_{\mathbf{J}}^{(1)}}(\mathbf{x}, \Sigma_1),$$

where the D -dimensional vector $A_{\mathbf{J}}$ is equal to $(A_{\mathbf{J}}^{(0)}, A_{\mathbf{J}}^{(1)})$ with blocks of respective size d and K .

For the first term, we note that $H_{A_{\mathbf{J}}^{(0)}}(0, \lambda I_d) = \lambda^{-\frac{1}{2}|A_{\mathbf{J}}^{(0)}|} H_{A_{\mathbf{J}}^{(0)}}(0)$ where $H_{\mathbf{k}}(\cdot)$ is the usual multidimensional Hermite polynomial of multi-order \mathbf{k} .

Concerning the second term, we compute $\mathbb{E} \left(H_{A_{\mathbf{J}}^{(1)}}((\nabla^2 X(0), X(0)), \Sigma_1)^2 \right)$ thanks to a formula for multivariate Gaussian integrals that can be found in [17] for instance (see equation (35) in this reference),

$$\begin{aligned} \mathbb{E} \left(H_{A_{\mathbf{J}}^{(1)}}((\nabla^2 X(0), X(0)), \Sigma_1)^2 \right) &= \int_{\mathbb{R}^K} H_{A_{\mathbf{J}}^{(1)}}(\mathbf{x}, \Sigma_1)^2 p_{\nabla^2 X(0), X(0)}(\mathbf{x}) d\mathbf{x} \\ &= 2^{|A_{\mathbf{J}}^{(1)}|-D} (A_{\mathbf{J}}^{(1)})! \operatorname{diag}(\Sigma_1^{-1})^{A_{\mathbf{J}}^{(1)}} \end{aligned}$$

where $\operatorname{diag}(B)$ denotes the vector containing the diagonal coefficients of any square matrix B .

Finally, since $H_{A_{\mathbf{J}}^{(0)}}(0)^2 \leq (A_{\mathbf{J}}^{(0)})!$,

$$\mathbb{E} (H_{A_{\mathbf{J}}}((0, \nabla^2 X(0), X(0)), \Sigma)^2) \leq (2\|\Sigma^{-1}\|)^{|\mathbf{J}|} (A_{\mathbf{J}})!$$

where we have used that $|A_{\mathbf{J}}^{(0)}| + |A_{\mathbf{J}}^{(1)}| = |\mathbf{J}|$ and $(A_{\mathbf{J}}^{(0)})! (A_{\mathbf{J}}^{(1)})! = (A_{\mathbf{J}})!$.

Note that for any D -dimensional index $\mathbf{k} = (k_1, \dots, k_D)$, we have $|\mathbf{k}|! \leq D^{|\mathbf{k}|} (\mathbf{k})!$ (actually, $\frac{|\mathbf{k}|!}{(\mathbf{k})!}$ is a multinomial coefficient). Hence $(A_{\mathbf{J}})_j! \leq D^{\sum_{1 \leq k \leq D} J_{jk}} \prod_{1 \leq k \leq D} J_{jk}!$ and so

$$(A_{\mathbf{J}})! = \prod_{1 \leq j \leq D} (A_{\mathbf{J}})_j! \leq D^{|\mathbf{J}|} \mathbf{J}!$$

Coming back to inequality (27) yields

$$D(A_{\mathbf{J}}) \leq \psi(0)^d (2D\|\Sigma^{-1}\|)^{|\mathbf{J}|/2} \sqrt{\mathbf{J}}!.$$

The same inequality can be established for $D(B_{\mathbf{J}})$. Hence, from (26) we get

$$|h(t)| \leq \psi(0)^{2d} \psi(t) \sum_{\mathbf{J}, |\mathbf{J}| > 0} (2D\|\Sigma^{-1}\| \psi(t))^{|\mathbf{J}|-1}.$$

The series in the right hand is convergent for $\|t\|$ large enough. Indeed we can choose A such that for $\|t\| > A$, we have $2D\|\Sigma^{-1}\| \psi(t) \leq \eta < 1$ and so the series is bounded by $(1 - \eta)^{-1}$.

The result follows since on the other hand, by assumption **(A)**, $\int_{\|t\| > A} \psi(t) dt < +\infty$. \square

A.2. Proof of Lemma 6.

Without loss of generality, we can assume that $|T_1| = |T_2| = 1$ and that there exists a and b in \mathbb{R}^d such that $T_1 = a + [0, 1)^d$ and $T_2 = b + [0, 1)^d$. Moreover, since the distance between T_1 and T_2 is

non negative, we have $\|b - a\|_\infty > 1$. In the following lines, we denote $\delta = \|b - a\|_\infty - 1$. We use the additivity of the Euler characteristic to expand $Z_1^{(N)}$ as

$$Z_1^{(N)} = N^{-d/2} \sum_{\mathbf{s} \in \mathcal{I}_N} \tilde{\varphi}(\theta_{Na+\mathbf{s}} X, [0, 1]^d, u_1),$$

where

$$\tilde{\varphi}(X, T, u) = \varphi(X, T, u) - \mathbb{E}[\varphi(X, T, u)],$$

$\theta_{\mathbf{s}}$ stands for the shift operator and \mathcal{I}_N for the set of integer valued indices $[[0, N - 1]]^d$. The same holds for $Z_2^{(N)}$. Hence, by stationarity

$$\begin{aligned} |\text{Cov}(Z_1^{(N)}, Z_2^{(N)})| &= N^{-d} \left| \sum_{\mathbf{s}, \mathbf{s}' \in \mathcal{I}_N} \mathbb{E}(\tilde{\varphi}(X, [0, 1]^d, u_1) \tilde{\varphi}(\theta_{N(b-a)+\mathbf{s}'-\mathbf{s}} X, [0, 1]^d, u_2)) \right| \\ &\leq \sum_{\mathbf{s} \in \mathcal{J}_N} |\mathbb{E}(\tilde{\varphi}(X, [0, 1]^d, u_1) \tilde{\varphi}(\theta_{N(b-a)+\mathbf{s}} X, [0, 1]^d, u_2))|, \end{aligned}$$

with $\mathcal{J}_N = [[-N + 1, N - 1]]^d$.

It is proved in [13] Proposition 1.3 that, for a cube T in \mathbb{R}^d and a level u in \mathbb{R} , the following expansion holds both *a.s.* and in $L^2(\Omega)$.

$$\tilde{\varphi}(X, T, u) = (-1)^d \sum_{q=1}^{\infty} \sum_{\mathbf{n} \in \mathbb{N}^D; |\mathbf{n}|=q} a(\mathbf{n}, u) \int_T H_{\mathbf{n}}(Y(t)) dt \quad (28)$$

where $D = d + \frac{1}{2}d(d+1) + 1$, $H_{\mathbf{n}}$ denotes the D dimensional Hermite coefficient of multi-order $\mathbf{n} \in \mathbb{N}^D$ and $Y(t) = \Lambda^{-1} \mathbf{X}(t)$ with Λ any square root of the covariance matrix of $\mathbf{X}(0)$. Moreover the Hermite coefficients $a(\mathbf{n}, u)$ are such that for any $q \geq 1$,

$$\sum_{\mathbf{n} \in \mathbb{N}^D; |\mathbf{n}|=q} a(\mathbf{n}, u)^2 \mathbf{n}! \leq C q^d \quad (29)$$

where C is some positive constant that only depends on d, Λ and u .

Hence, thanks to (28) and to the orthogonality of Hermite polynomials, we have for any $\tau \in \mathbb{R}^d$,

$$\begin{aligned} \mathbb{E}(\tilde{\varphi}(X, [0, 1]^d, u_1) \tilde{\varphi}(\theta_{\tau} X, [0, 1]^d, u_2)) \\ = \sum_{q=1}^{\infty} \int_{[0, 1]^d} \int_{[0, 1]^d} \mathbb{E}(F_q(Y(t_1), u_1) F_q(Y(\tau + t_2), u_2)) dt_1 dt_2, \end{aligned}$$

where $F_q(y, u) = \sum_{\mathbf{n} \in \mathbb{N}^D; |\mathbf{n}|=q} a(\mathbf{n}, u) H_{\mathbf{n}}(y)$, $y \in \mathbb{R}^D$. One can adapt Arcones inequality, for instance following the proof of [4] Lemma 1 step by step, to establish that there exists a constant $\kappa > 0$ depending on the covariance function of X such that

$$\begin{aligned} |\mathbb{E}(F_q(Y(t_1), u_1) F_q(Y(\tau + t_2), u_2))| \\ \leq \kappa^q \psi(\tau + t_2 - t_1)^q \sum_{\mathbf{n} \in \mathbb{N}^D; |\mathbf{n}|=q} \frac{1}{2} (a(\mathbf{n}, u_1)^2 + a(\mathbf{n}, u_2)^2) \mathbf{n}!. \end{aligned}$$

Let now $\eta > 0$ be fixed. Thanks to Assumption **(A)**, we can choose N large enough such that $\kappa\psi(\tau) \leq 1/2$ as soon as $\|\tau\| \geq N\delta$, $\int_{\|t\| \geq N} \psi(t) dt < \eta$ and $N\delta > 2$. Hence, using (29), we deduce that for $\|\tau\| \geq N\delta$,

$$\begin{aligned} |\mathbb{E}(\tilde{\varphi}(X, [0, 1]^d, u_1) \tilde{\varphi}(\theta_\tau X, [0, 1]^d, u_2))| &\leq C \sum_{q=1}^{\infty} q^d \kappa^q \int_{[0, 1]^d} \int_{[0, 1]^d} \psi(\tau + t_2 - t_1)^q dt_1 dt_2 \\ &\leq C \left(\sum_{q=1}^{\infty} q^d (1/2)^{q-1} \right) \int_{[-1, 1]^d} \psi(\tau + t) dt. \end{aligned}$$

Note that C stands here for a positive constant that may change from a line to another. Coming back to the covariance of $Z_1^{(N)}$ and $Z_2^{(N)}$ yields

$$\begin{aligned} |\text{Cov}(Z_1^{(N)}, Z_2^{(N)})| &\leq C \sum_{\mathbf{s} \in \mathcal{J}_N} \int_{[-1, 1]^d} \psi(N(b-a) + \mathbf{s} + t) dt \\ &\leq 2C \int_{\|t\| \geq N} \psi(t) dt \leq 2C\eta. \end{aligned}$$

Thus, Lemma 6 is proved. \square

A.3. Proof of Proposition 7.

Firstly, let us remark that, since $Z^{(s)}(\cdot) = \frac{1}{\sqrt{2s}}(\chi_s^2(\cdot) - s)$, it is sufficient to prove the statement of Proposition 7 for the process χ_s^2 instead of $Z^{(s)}$. Secondly, let us recall the relation (4), so that we can replace $\varphi(Z^{(s)}, [0, T], u)$ by $U(\chi_s^2, [0, T], u)$ in the proposition.

The finiteness of

$$V_s(T, u) := \mathbb{E}[U(\chi_s^2, [0, T], u)(U(\chi_s^2, [0, T], u) - 1)]$$

is proved in [29] under the following conditions,

$$r_X(t) = 1 - \lambda_2 \frac{t^2}{2} + \lambda_4 \frac{t^4}{4} + o(t^4) \quad \text{as } t \rightarrow 0 \quad \text{and} \quad r_X(t) = o(t^{-\alpha}) \quad \text{as } t \rightarrow \infty \quad (30)$$

for some positive finite constants $\lambda_2, \lambda_4, \alpha$ (see (4.1) and (4.2) in the cited paper). They are clearly satisfied in our case since Assumption **(A)** is in force. This establishes that $U(\chi_s^2, [0, T], u)$ admits a finite variance.

We now focus on the asymptotic variance as T goes to infinity. We will actually prove that

$$\lim_{T \rightarrow +\infty} \frac{1}{T} (V_s(T, u) - (\mathbb{E}[U(\chi_s^2, [0, T], u)])^2) < +\infty.$$

First, in a similar way as (13), let us remark that the expectation of $U(\chi_s^2, [0, T], u)$ is proportional to T ,

$$K_s(u) := \frac{1}{T} \mathbb{E}[U(\chi_s^2, [0, T], u)] = \left(\frac{\lambda}{\pi}\right)^{1/2} \frac{e^{-u/2} u^{(s-1)/2}}{2^{s/2} \Gamma(s/2)}. \quad (31)$$

On the other hand, introducing the function $\mathcal{G}(x) = x \mathbf{1}_{[0, \infty)}(x)$, Rice formula for the above factorial moment gives

$$V_s(T, u) = 2 \int_0^T (T-t) \mathcal{I}(t) dt,$$

where

$$\mathcal{I}(t) = \mathbb{E}[\mathcal{G}((\chi_s^2)'(0)) \mathcal{G}((\chi_s^2)'(t)) | \chi_s^2(0) = \chi_s^2(t) = u] p_{\chi_s^2(0), \chi_s^2(t)}(u, u).$$

For computing this conditional expectation, we begin with the following one that is easier to handle with. Let us define $\vec{X}(t) = (X_1(t), \dots, X_s(t))$ and, for $\mathbf{z} = (z_1, \dots, z_s)$ and $\mathbf{w} = (w_1, \dots, w_s)$ in \mathbb{R}^s ,

$$J(t, \mathbf{z}, \mathbf{w}) := \mathbb{E}[\mathcal{G}((\chi_s^2)'(0)) \mathcal{G}((\chi_s^2)'(t)) | \vec{X}(0) = \mathbf{z}, \vec{X}(t) = \mathbf{w}, \chi_s^2(0) = \chi_s^2(t) = u].$$

Let us fix \mathbf{z} and \mathbf{w} such that $\|\mathbf{z}\|^2 = \|\mathbf{w}\|^2 = u$. In that case, we have

$$J(t, \mathbf{z}, \mathbf{w}) = 4 \mathbb{E}[\mathcal{G}(\sum_{j=1}^s X_j(0) X_j'(0)) \mathcal{G}(\sum_{j=1}^s X_j(t) X_j'(t)) | \vec{X}(0) = \mathbf{z}, \vec{X}(t) = \mathbf{w}].$$

We will use, as is classical, the following regression model. For each $j \in \{1, \dots, s\}$, we write

$$\begin{aligned} X_j'(0) &= \alpha_1(t) X_j(0) + \beta_1(t) X_j(t) + \varepsilon_{1j}(t) \\ X_j'(t) &= \alpha_2(t) X_j(0) + \beta_2(t) X_j(t) + \varepsilon_{2j}(t). \end{aligned}$$

It is important to point out that, because of the independence, the coefficients above do not depend on j . We readily obtain that

$$\alpha_1(t) = \frac{r_X'(t) r_X(t)}{1 - r_X^2(t)}, \beta_1(t) = \frac{-r_X'(t)}{1 - r_X^2(t)}, \beta_2(t) = -\alpha_1(t), \alpha_2(t) = -\beta_1(t).$$

Thus

$$\begin{aligned} J(t, \mathbf{z}, \mathbf{w}) &= 4 \mathbb{E}[\mathcal{G}(\sum_{j=1}^s z_j(\alpha_1(t) z_j + \beta_1(t) w_j + \varepsilon_{1j}(t))) \mathcal{G}(\sum_{j=1}^s w_j(\beta_1(t) z_j + \alpha_1(t) w_j + \varepsilon_{2j}(t)))] \\ &= 4 \mathbb{E}[\mathcal{G}(\alpha_1(t) u + \beta_1(t) \langle \mathbf{w}, \mathbf{z} \rangle + \sqrt{u} \sigma_\varepsilon W(t)) \mathcal{G}(\alpha_1(t) u + \beta_1(t) \langle \mathbf{w}, \mathbf{z} \rangle + \sqrt{u} \sigma_\varepsilon Z(t))] \\ &= 4 \mathbb{E}[\mathcal{G}(d_u(t) + \sqrt{u} \sigma_\varepsilon(t) W(t)) \mathcal{G}(d_u(t) + \sqrt{u} \sigma_\varepsilon(t) Z(t))], \end{aligned}$$

where

$$\sigma_\varepsilon^2(t) = -r_X''(0) - (\alpha_1^2(t) + \beta_1^2(t) + 2\alpha_1(t)\beta_1(t)r_X(t))$$

is the variance of each $\varepsilon_{1,j}(t)$ or $\varepsilon_{2,j}(t)$,

$$W(t) = \frac{1}{\sqrt{u} \sigma_\varepsilon(t)} \sum_{i=1}^s w_i \varepsilon_{1,i}(t) \quad \text{and} \quad Z(t) = \frac{1}{\sqrt{u} \sigma_\varepsilon(t)} \sum_{i=1}^s z_i \varepsilon_{2,i}(t)$$

are independent $N(0, 1)$ random variables, and $d_u(t) = \alpha_1(t) u + \beta_1(t) \langle \mathbf{w}, \mathbf{z} \rangle$.

In particular, the latter shows that $J(t, \mathbf{z}, \mathbf{w})$ only depends on \mathbf{w} and \mathbf{z} through their scalar product. So, from now on, for ease of notation, we will write $J(t, \langle \mathbf{z}, \mathbf{w} \rangle)$.

We have $\beta_1(t) \rightarrow \infty$ when $t \rightarrow 0$ and $d_u(t) = (\alpha_1(t) + \beta_1(t)) u - \beta_1(t) u (1 - \frac{1}{u} \langle \mathbf{w}, \mathbf{z} \rangle) \xrightarrow[t \rightarrow 0]{} -\infty$, for all $\mathbf{w} \neq \mathbf{z}$. The latter convergence is equivalent to $-\frac{1}{t}$.

Moreover, $d_u(t) \rightarrow 0$ whenever $t \rightarrow \infty$, because of Assumption **(A)**.

Then we can write

$$\begin{aligned} J(t, \langle \mathbf{z}, \mathbf{w} \rangle) &= 4 u \sigma_\varepsilon^2(t) \mathbb{E}[\mathcal{G}(\frac{d_u(t)}{\sqrt{u} \sigma_\varepsilon(t)} + W(t)) \mathcal{G}(\frac{d_u(t)}{\sqrt{u} \sigma_\varepsilon(t)} + Z(t))] \\ &= 4 u \sigma_\varepsilon^2(t) \sum_{k=0}^{\infty} d_{uk}^2(t) k! (\text{Corr}(W(t), Z(t)))^k, \end{aligned} \tag{32}$$

where the last equality is a consequence of Mehler's formula. The coefficients $d_{uk}(t)$ are given by

$$\begin{aligned} d_{uk}(t) &= \frac{1}{k!} \int_{\mathbb{R}} \mathcal{G}\left(\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)} + y\right) H_k(y) \phi(y) dy \\ &= \frac{1}{k!} \int_{-\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}}^{\infty} \left(\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)} + y\right) H_k(y) \phi(y) dy, \end{aligned}$$

with ϕ the standard Gaussian density, and

$$\text{Corr}(W(t), Z(t)) = \frac{\langle \mathbf{z}, \mathbf{w} \rangle}{u\sigma_\varepsilon^2(t)} (-r_X''(t) - 2\alpha_1(t)\beta_1(t) - (\alpha_1^2 + \beta_1^2(t))r_X(t)).$$

Let us continue our study with the behaviour of $J(t, \langle \mathbf{z}, \mathbf{w} \rangle)$ for t close to 0. Let us start with the series. It holds

$$\begin{aligned} \sum_{k=0}^{\infty} d_{uk}^2(t) k! &= \int_{-\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}}^{\infty} \left(\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)} + y\right)^2 \phi(y) dy, \\ &\leq 2\left(\left(\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}\right)^2 \int_{-\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}}^{\infty} \phi(y) dy + \int_{-\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}}^{\infty} y^2 \phi(y) dy\right). \end{aligned}$$

For the last integral, we use the upper bound

$$\int_a^{\infty} y^2 \phi(y) dy \leq a\phi(a) + \frac{1}{a}\phi(a),$$

that yields, as $t \rightarrow 0$,

$$\sum_{k=0}^{\infty} d_{uk}^2(t) k! \leq 2\left(2\left(-\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}\right)\phi\left(\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}\right) + \left(-\frac{\sqrt{u}\sigma_\varepsilon(t)}{d_u(t)}\right)\phi\left(\frac{d_u(t)}{\sqrt{u}\sigma_\varepsilon(t)}\right)\right) = O(1). \quad (33)$$

Thus, for $0 < t < \delta$, we have $J(t, \langle \mathbf{z}, \mathbf{w} \rangle) \leq \mathbf{C} u \sigma_\varepsilon^2(t)$. Moreover, as this bound only depends on u , it yields immediately that,

$$\mathcal{I}(t) \leq \mathbf{C} u \sigma_\varepsilon^2(t), \quad \text{for } 0 < t < \delta.$$

We now focus on the behaviour of $\mathcal{I}(t)$ for t close to $+\infty$. We have

$$\begin{aligned} \mathcal{I}(t) &= \int_{\mathbb{R}^s \times \mathbb{R}^s} J(t, \langle \mathbf{z}, \mathbf{w} \rangle) p_{\vec{X}(0), \vec{X}(t), \chi_s^2(0), \chi_s^2(t)}(\mathbf{z}, \mathbf{w}, u, u) d\mathbf{z} d\mathbf{w} \\ &= \frac{u^{s-2}}{(2\pi)^s (1 - r_X^2(t))^{s/2}} \int_{\mathbb{S}^{s-1} \times \mathbb{S}^{s-1}} J(t, u\langle \tau_1, \tau_2 \rangle) e^{-u\left(\frac{1 - r_X(t)\langle \tau_1, \tau_2 \rangle}{1 - r_X^2(t)}\right)} d\sigma_{s-1}(\tau_1) d\sigma_{s-1}(\tau_2), \end{aligned}$$

where σ_{s-1} stands for the surface Lebesgue measure on \mathbb{S}^{s-1} . A consequence of Funk-Hecke formula says that, for any reasonable function F ,

$$\int_{\mathbb{S}^{s-1} \times \mathbb{S}^{s-1}} F(\langle \tau_1, \tau_2 \rangle) d\sigma_{s-1}(\tau_1) d\sigma_{s-1}(\tau_2) = \frac{\sigma_{s-1}(\mathbb{S}^{s-1})^2}{2W_{s-2}} \int_0^\pi F(\cos \psi) \sin^{s-2} \psi d\psi,$$

with W_s the Wallis integral, *i.e.* $W_s = \int_0^{\pi/2} \sin^s(\psi) d\psi$. It implies

$$\mathcal{I}(t) = \frac{4u^{s-2}}{(2\pi)^s (1 - r_X^2(t))^{s/2}} \frac{\sigma_{s-1}(\mathbb{S}^{s-1})^2}{2W_{s-2}} \int_0^\pi J(t, u \cos \psi) e^{-u\left(\frac{1 - r_X(t) \cos \psi}{1 - r_X^2(t)}\right)} \sin^{s-2} \psi d\psi.$$

Following (32), we now expand $J(t, x)$ as $J(t, x) := 4u\sigma_\varepsilon^2(t)\sum_{k=0}^\infty Q_k(t, x)$, so that

$$\begin{aligned} \frac{1}{T} V_s(u, T) &= 2 \int_0^T \left(1 - \frac{t}{T}\right) \mathcal{I}(t) dt \\ &= \int_0^T \left(1 - \frac{t}{T}\right) \frac{4u^{s-1}\sigma_\varepsilon^2(t)}{(2\pi)^s(1-r_X^2(t))^{s/2}} \frac{\sigma_{s-1}(\mathbb{S}^{s-1})^2}{W_{s-2}} \\ &\quad \times \int_0^\pi \sum_{k=0}^\infty Q_k(t, u \cos \psi_1) e^{-u(\frac{1-r_X(t)\cos\psi}{1-r_X^2(t)})} \sin^{s-2} \psi d\psi dt. \end{aligned}$$

In first place, let us study the limit when $t \rightarrow \infty$ of the $(k=0)$ -term in the integrand. We introduce the following notation,

$$\mathcal{I}_0(t) = \frac{4u^{s-1}\sigma_\varepsilon^2(t)}{(2\pi)^s(1-r_X^2(t))^{s/2}} \frac{\sigma_{s-1}(\mathbb{S}^{s-1})^2}{W_{s-2}} \int_0^\pi Q_0(t, u \cos \psi_1) e^{-u(\frac{1-r_X(t)\cos\psi}{1-r_X^2(t)})} \sin^{s-2} \psi d\psi.$$

We have

$$\mathcal{I}_0(t) \xrightarrow[t \rightarrow \infty]{} \frac{4u^{s-1}(-r_X''(0))}{(2\pi)^{s+1}} e^{-u} \sigma_{s-1}(\mathbb{S}^{s-1})^2 = K_s(u)^2,$$

given that $d_{u0}^2(t) \xrightarrow[t \rightarrow \infty]{} \frac{1}{2\pi}$, $\sigma_{s-1}(\mathbb{S}^{s-1}) = \frac{2\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2})}$ and using (31). Therefore, recalling the identity

$2 \int_0^T \left(1 - \frac{t}{T}\right) dt = T$, the first term of the asymptotic variance is

$$\begin{aligned} &\lim_{T \rightarrow \infty} 2 \int_0^T \left(1 - \frac{t}{T}\right) \mathcal{I}_0(t) dt - \frac{1}{T} (\mathbb{E}[U(\chi_s^2, T, u)])^2 \\ &= \lim_{T \rightarrow \infty} 2 \int_0^T \left(1 - \frac{t}{T}\right) \mathcal{I}_0(t) dt - T K_s(u)^2 \\ &= 2 \int_0^\infty (\mathcal{I}_0(t) - K_s(u)^2) dt. \end{aligned} \tag{34}$$

This last equality holds true if we show that the difference appearing into the integrand is bounded by an $L^1(\mathbb{R})$ function, outside of a compact interval.

Actually, under **(A)**, it is easy to prove that for t large enough, $|\mathcal{I}_0(t) - K_s(u)^2| \leq \mathbf{C} \left| \frac{d_u(t)}{\sqrt{u\sigma_\varepsilon(t)}} \right| \leq \mathbf{C}\psi(t)$. Hence, the limit (34) is established.

In the sequel we are going to study the asymptotic behaviour of the remaining terms. Let us introduce

$$\begin{aligned} \mathcal{R}(t) &:= \frac{4u^{s-1}\sigma_\varepsilon^2(t)}{(2\pi)^s(1-r^2(t))^{s/2}} \frac{\sigma_{s-1}(\mathbb{S}^{s-1})^2}{2W_{s-2}} \\ &\quad \times \int_0^\pi \sum_{k=1}^\infty Q_k(t, u \cos \psi_1) e^{-u(\frac{1-r_X(t)\cos\psi}{1-r_X^2(t)})} \sin^{s-2} \psi d\psi. \end{aligned}$$

So, by using (33), we obtain that for t large enough,

$$\left| \sum_{k=1}^\infty Q_k(t, u \cos \psi_1) \right| \leq |(\text{Corr}(W(t), Z(t)))| \sum_{k=1}^\infty d_{uk}^2(t) k! \leq \mathbf{C}\psi(t).$$

This bound and Assumption **(A)** entail

$$\lim_{T \rightarrow \infty} 2 \int_0^T (1 - \frac{t}{T}) \mathcal{R}(t) dt = 2 \int_0^\infty \mathcal{R}(t) dt. \quad (35)$$

Summing up (13), (34) and (35) yield

$$\lim_{T \rightarrow \infty} \frac{1}{T} (V_s(T, u) - (\mathbb{E}[U(\chi_s^2, [0, T], u)])^2) = 2 \int_0^\infty (\mathcal{I}_0(t) - K_s(u)^2) dt + 2 \int_0^\infty \mathcal{R}(t) dt < \infty.$$

□

A.4. Proof of Proposition 8.

1. *Computation of $g(u)$.* By Equation (8), we have

$$\begin{aligned} g(u) &= \mathbb{E}[\mathbf{1}_{[u, \infty)}(X(0)) |X''(0)|] \\ &= \frac{1}{2\pi(\gamma - \lambda^2)^{1/2}} \int_u^\infty \int_{\mathbb{R}} |\ddot{x}| e^{-\frac{1}{2(\gamma - \lambda^2)}(\gamma x^2 + 2\lambda x \ddot{x} + \ddot{x}^2)} d\ddot{x} dx \\ &= \frac{1}{2\pi(\gamma - \lambda^2)^{1/2}} \int_u^\infty e^{-\frac{1}{2}x^2} dx \int_{\mathbb{R}} |\ddot{x}| e^{-\frac{1}{2(\gamma - \lambda^2)}(\ddot{x} + \lambda x)^2} d\ddot{x} \end{aligned}$$

Let us split the inner integral into $\ddot{x} > 0$ and $\ddot{x} < 0$, i.e.

$$\frac{1}{\sqrt{2\pi}(\gamma - \lambda^2)^{1/2}} \int_{\mathbb{R}} |\ddot{x}| e^{-\frac{1}{2(\gamma - \lambda^2)}(\ddot{x} + \lambda x)^2} d\ddot{x} := I_1(x, \lambda, \gamma) + I_2(x, \lambda, \gamma),$$

where

$$\begin{aligned} I_1(x, \lambda, \gamma) &= \frac{1}{\sqrt{2\pi}(\gamma - \lambda^2)^{1/2}} \int_0^\infty \ddot{x} e^{-\frac{1}{2(\gamma - \lambda^2)}(\ddot{x} + \lambda x)^2} d\ddot{x} \\ I_2(x, \lambda, \gamma) &= \frac{1}{\sqrt{2\pi}(\gamma - \lambda^2)^{1/2}} \int_0^\infty \ddot{x} e^{-\frac{1}{2(\gamma - \lambda^2)}(\ddot{x} - \lambda x)^2} d\ddot{x} = I_1(x, -\lambda, \gamma). \end{aligned}$$

We make the following change of variable, $\frac{\ddot{x} + \lambda x}{(\gamma - \lambda^2)^{1/2}} = y$, to get

$$\begin{aligned} I_1(x, \lambda, \gamma) &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}}^\infty ((\gamma - \lambda^2)^{1/2} y - \lambda x) e^{-\frac{1}{2}y^2} dy \\ &= (\gamma - \lambda^2)^{1/2} \phi\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) - \lambda x \bar{\Phi}\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) \\ I_2(x, \lambda, \gamma) &= (\gamma - \lambda^2)^{1/2} \phi\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) + \lambda x \bar{\Phi}\left(-\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right). \end{aligned}$$

But it holds that $\bar{\Phi}(x) = \Phi(-x)$. Thus

$$I_1(x, \lambda, \gamma) + I_2(x, \lambda, \gamma) = 2(\gamma - \lambda^2)^{1/2} \phi\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) + \lambda x (2\Phi\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) - 1).$$

Summing up, we have

$$g(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{1}{2}x^2} \left[2(\gamma - \lambda^2)^{1/2} \phi\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) + \lambda x (2\Phi\left(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}\right) - 1) \right] dx.$$

Integrating by parts the second integral, we get that it is equal to

$$\lambda\phi(u)(2\Phi(\frac{\lambda u}{(\gamma - \lambda^2)^{1/2}}) - 1) + \frac{2\lambda^2}{\sqrt{2\pi}(\gamma - \lambda^2)^{1/2}} \int_u^\infty e^{-\frac{1}{2}x^2} \phi(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}) dx,$$

yielding

$$g(u) = \lambda\phi(u)(2\Phi(\frac{\lambda u}{(\gamma - \lambda^2)^{1/2}}) - 1) + \frac{2\gamma}{\sqrt{2\pi}(\gamma - \lambda^2)^{1/2}} \int_u^\infty e^{-\frac{1}{2}x^2} \phi(\frac{\lambda x}{(\gamma - \lambda^2)^{1/2}}) dx.$$

The second term is equal to

$$\begin{aligned} \frac{2\gamma}{2\pi(\gamma - \lambda^2)^{1/2}} \int_u^\infty e^{-\frac{1}{2}x^2} e^{-\frac{\lambda^2 x^2}{2(\gamma - \lambda^2)}} dx &= \frac{\gamma}{\pi(\gamma - \lambda^2)^{1/2}} \int_u^\infty e^{-\frac{\gamma x^2}{2(\gamma - \lambda^2)}} dx \\ &= (\frac{2\gamma}{\pi})^{1/2} \bar{\Phi}((\frac{\gamma}{(\gamma - \lambda^2)})^{1/2} u). \end{aligned}$$

Finally,

$$g(u) = \lambda\phi(u)(2\Phi(\frac{\lambda u}{(\gamma - \lambda^2)^{1/2}}) - 1) + (\frac{2\gamma}{\pi})^{1/2} \bar{\Phi}((\frac{\gamma}{(\gamma - \lambda^2)})^{1/2} u).$$

2. *Computation of $G(u, t)$.* Since $\Gamma(t)$ is the covariance matrix of $(X(0), X(t), X''(0), X''(t))$ conditioned to $(X'(0) = X'(t) = 0)$, one can write

$$G(u, t) = \mathbb{E}(\mathbf{1}_{Z_1 > u} \mathbf{1}_{Z_2 > u} Z_3 Z_4) \text{ where } \mathbf{Z} = (Z_1, Z_2, Z_3, Z_4) \stackrel{\mathcal{L}}{=} N(0, \Gamma(t)). \quad (36)$$

From now on, we remove the dependence on t . We begin by writing $\mathbf{Z} = L\mathbf{Y}$, where L is a lower triangular matrix such that $LL^T = \Gamma(t)$. Denoting by l_{ij} for $i \geq j$ the elements of L , we have

$$Z_1 = l_{11}Y_1 \quad Z_2 = l_{21}Y_1 + l_{22}Y_2 \quad Z_3 = \sum_{i=1}^3 l_{3i}Y_i \quad Z_4 = \sum_{i=1}^4 l_{4i}Y_i,$$

and the expectation (36) can be written as

$$G(u, t) = \int_{\mathbb{R}^4} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} (\sum_{i=1}^3 l_{3i}y_i) (\sum_{i=1}^4 l_{4i}y_i) \phi(\mathbf{y}) d\mathbf{y}.$$

By expanding the second sum, the integral can be written as the sum of the following terms (starting with index $i = 4$ term)

$$\begin{aligned} I_4 &= l_{44} \int_{\mathbb{R}^4} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} (\sum_{i=1}^3 l_{3i}y_i) y_4 \phi(\mathbf{y}) d\mathbf{y} = 0, \\ I_3 &= l_{43} \int_{\mathbb{R}^4} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} (\sum_{i=1}^3 l_{3i}y_i) y_3 \phi(\mathbf{y}) d\mathbf{y} \\ &= l_{43} l_{33} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} \phi(y_1) \phi(y_2) dy_1 dy_2 \left(\int_{\mathbb{R}} y_3^2 \phi(y_3) dy_3 \right) \\ &= l_{43} l_{33} Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) \end{aligned}$$

$$\begin{aligned}
I_2 &= l_{42} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} (l_{31}y_1y_2 + l_{32}[(y_2^2 - 1) + 1]) \phi(y_1)\phi(y_2)dy_1dy_2 \\
&= l_{42}l_{32} Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) + l_{42}l_{32} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} \phi(y_1)\phi''(y_2)dy_1dy_2 \\
&\quad + l_{42}l_{31} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} \phi'(y_1)\phi'(y_2)dy_1dy_2 \\
&= l_{42}l_{32} Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) - l_{42}l_{32} \int_{\frac{u}{l_{11}}}^{\infty} \phi'\left(\frac{u}{l_{22}} - \frac{l_{21}}{l_{22}}y_1\right)\phi(y_1)dy_1 \\
&\quad - l_{42}l_{31} \int_{\frac{u}{l_{11}}}^{\infty} \phi\left(\frac{u}{l_{22}} - \frac{l_{21}}{l_{22}}y_1\right)\phi'(y_1)dy_1 \\
&= l_{42}l_{32}(Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) + Q^{(2)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right)) + l_{42}l_{31} Q^{(1)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) \\
\\
I_1 &= l_{41} \int_{\mathbb{R}^4} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} \left(\sum_{i=1}^3 l_{3i}y_i\right)y_1\phi(\mathbf{y})d\mathbf{y} \\
&= l_{41} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} (l_{31}y_1 + l_{32}y_2)y_1\phi(y_1)\phi(y_2)dy_1dy_2 \\
&= l_{41}l_{32} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} \phi'(y_1)\phi'(y_2)dy_1dy_2 + l_{41}l_{31} Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) \\
&\quad + l_{41}l_{31} \int_{\mathbb{R}^2} \mathbf{1}_{\{l_{11}y_1 > u\}} \mathbf{1}_{\{l_{21}y_1 + l_{22}y_2 > u\}} \phi''(y_1)\phi(y_2)dy_1dy_2 \\
&= l_{41}l_{31}(Q^{(0)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right) + Q^{(3)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right)) + l_{41}l_{32} Q^{(1)}\left(\frac{u}{l_{11}}, -\frac{l_{21}}{l_{22}}, \frac{u}{l_{22}}\right).
\end{aligned}$$

In this manner, we get the result. \square

Acknowledgment. The authors sincerely express their gratitude to the referees for their valuable comments that help to motivate the study and make the paper clearer.

References

- [1] Adler R. (2008). Some New Random Field Tools for Spatial Analysis. *Stochastic Environmental Research and Risk Assessment*, **22**, 809-822.
- [2] Adler R.J., Bartz K., Kou S., Monod A. (2014). Estimating thresholding levels for random fields via Euler characteristics. *Preprint*.
- [3] Adler R.J., Taylor J.E. (2007). *Random Fields and Geometry*. Springer Monographs in Mathematics. Springer.
- [4] Arcones M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.*, vol 22, 2242-2274.
- [5] Azaïs J-M., Wschebor M. *Level Sets and Extrema of Random Processes and Fields*. Wiley (2009).
- [6] Biermé H., Desolneux A. (2012). A Fourier approach for the level crossings of Shot Noise processes with jumps. *J. Appl. Probab.*, **49**(1), 100–113.
- [7] Biermé H., Desolneux A. (2016). Level total curvature integral: Euler characteristic and 2D random fields. *Preprint*. <https://hal.archives-ouvertes.fr/hal-01370902>.

- [8] Cabaña E.M. (1987). Affine processes: a test of isotropy based on level sets. *SIAM Journal on Applied Mathematics*, **47**(4), 886-891.
- [9] Cammarota V., Marinucci D. (2016) A Quantitative Central Limit Theorem for the Euler-Poincare Characteristic of Random Spherical Eigenfunctions. *preprint* arXiv:1603.09588
- [10] Cheng D., Xiao Y. (2016). Excursion probability of Gaussian random fields on sphere. *Bernoulli*, **22**(2), 1113-1130.
- [11] Cuesta-Albertos J.A., Gamboa, F. and Nieto-Reyes, A. (2014). A random-projection based procedure to test if a stationary process is Gaussian. *Computational Statistics and Data Analysis*, **75**, 124-141.
- [12] Epps, T. W. (1987). Testing that a stationary time series is Gaussian. *Ann. Stat.* **15**(4), 1683-1698.
- [13] Estrade A., León J.R. (2016). A Central Limit Theorem for the Euler characteristic of a Gaussian excursion set. *Ann. of Probab.*, **44**(6), 3849-3878.
- [14] Gabriel E., Allard D., Bacro J-N. (2011). Estimating and testing zones of abrupt change for spatial data. *Stat. Comput.*, **21**, 107-120.
- [15] Geman D. *On the variance of the number of zeros of a stationary Gaussian process*. *Ann. Math. Stat.*, **43**, 977-982 (1972).
- [16] Hsing T. (1987). On the intensity of crossings by a shot noise process. *Adv. Appl. Probab.*, **19**, 743-745.
- [17] Kauderer M. (1993). Modes in n-dimensional first-order systems. *Journal of Math. Phys.*, **34**, p.4221.
- [18] Lachièze-Rey R. (2016). Two-dimensional Kac-Rice formula. Application to shot noise processes excursions. *preprint*. <https://arxiv.org/abs/1607.05467>.
- [19] Lelievre T., Rousset M. and Stolz G. (2010). *Free Energy Computations*. Imperial College Press.
- [20] Lindgren G. (1974). Spectral Moment Estimation by Means of Level Crossings. *Biometrika*, **61**(3), 401-418.
- [21] Lindner B. and Schimansky-Geier L. (1999) *Analytical approach to the stochastic FitzHugh-Nagumo system and coherence resonance*. Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics **60**, 7270-7276.
- [22] Lobato, I.N. and Velasco, C. (2004). A simple test of normality for time series. *Econometric Theory*, **20**(4), 671-689.
- [23] Marinucci D., Vадlamani S. (2016). High-Frequency Asymptotics for Lipschitz-Killing Curvatures of Excursion Sets on the Sphere. *Annals of Applied Probability*, **26**(1), 462-506.
- [24] Nott, D.J., Wilson, R.J. (2000). Multi-phase image modelling with excursion sets. *Signal Processing*, **80**, 125-139.
- [25] Nourdin I., Peccati G. and Podolskij M. (2011). Quantitative Breuer-Major theorems. *Stochastic Processes and their Applications*, **121**, 793-812.
- [26] Planck 2015 results XVI (2015). Isotropy and statistics of the CMB. *Astronomy and Astrophysics*. <http://arxiv.org/abs/1506.07135>
- [27] Rice S.O. (1945). Mathematical Analysis of Random Noise. *Bell System Technical Journal*, **23** and **24**.
- [28] Roberts, A.P., Torquato, S. (1999). Chord-distribution functions of three-dimensional random media: approximate first-passage times of Gaussian processes. *Phys. Rev. E* **59**(5), 4953-4963.
- [29] Sharpe K. (1978). Some properties of the crossings process generated by a stationary χ^2 process. *Adv. App. Prob.* **10**, 373-391.
- [30] Subba Rao, T. and Gabr, M.M. (1980). A test for linearity of stationary time series. *J. Time Ser. Anal.* **1**, 145-158.
- [31] Taheriyoun A.R. (2012). Testing the covariance function of stationary Gaussian random fields. *Statist. Probab. Lett.* **82**, 606-613.

- [32] Taheriyoun A.R., Shafie K., Jafari Jozani M. (2009). A note on the higher moments of the Euler characteristic of the excursion sets of random fields. *Statist. Probab. Lett.* **79**, 1074-1082.
- [33] Withers C. S., Nadarajah S. (2010). Expansions for the multivariate normal. *Journal of Mult. Anal.* **101**, 1311-1316.
- [34] Worsley K.J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2 , F and t fields. *Adv. Appl. Probab.* **26**, 13-42.
- [35] Worsley K.J. (1995). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Ann. Stat.* **23**, 640-669.
- [36] L. Wu (2001). Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stoch. Proc. and Appl.* **91**, 205-238.