Redundancy in Gaussian random fields

A Preprint

Valentin De Bortoli
CMLA - École normale supérieure Paris-Saclay, CNRS, Université Paris-Saclay, 94235 Cachan, France.

Agnès Desolneux
CMLA - École normale supérieure Paris-Saclay, CNRS, Université Paris-Saclay, 94235 Cachan, France.

Bruno Galerne
Institut Denis Poisson, Université d'Orléans, Université de Tours, CNRS

Arthur leclaire
Univ. Bordeaux, IMB, Bordeaux INP, CNRS, UMR 5251, F-33400 Talence, France.

November 13, 2018

Abstract

We introduce and study a notion of spatial redundancy in Gaussian random fields. we define similarity functions with some properties and give insight about their statistical properties in the context of image processing. We compute these similarity functions on local windows in random fields defined over discrete or continuous domains. We give explicit asymptotic Gaussian expressions for the distribution of similarity function random variables when computed over Gaussian random fields and illustrate the weaknesses of such Gaussian approximations by showing that the approximated probability of rare events is not precise enough, even for large windows. In the special case of the squared L^2 norm, non-asymptotic expressions are derived in both discrete and continuous periodic settings. A fast and accurate approximation is introduced using eigenvalues projection and moment methods.

( )

***K*eywords** Random fields, spatial redundancy, central limit theorem, law of large numbers, eigenvalues approximation, moment methods.

1 Introduction

Stochastic geometry [1, 2, 3] aims at describing the geometry of random structures based on the knowledge of the distribution of geometrical elementary patterns (point processes, random closed sets, etc.). When the considered geometrical elements are functions over some topological space, we can study the geometry of the associated random field. For example, centering a kernel function at each point of a Poisson point process gives rise to the notion of shot-noise random field [4, 5, 6]. We can then study the perimeter or the Euler-Poincaré characteristic of the excursion sets [7] among other properties [8]. In this article we will focus on the geometrical notion of redundancy in random fields, defined on local windows. We say that a local window in a random field is redundant if it is "similar" to other local windows in the same random field. The similarity of two local windows is defined as the output of some similarity function computed over these local windows. The lower is the output, the more similar are the local windows. We give some examples of similarity functions and illustrate their statistical properties with examples from image processing.

Identifying spatial redundancy is indeed a fundamental task in the field of image processing. For instance, in the context of denoising, Buades et al. in [9], propose the famous NL-means algorithm in which a noisy patch is replaced by the mean over all similar patches. Other examples can be found in the domains of inpainting [10] and video coding [11]. Spatial redundancy is also of crucial importance in exemplar-based texture synthesis, where we aim at sampling images with the same perceptual properties as an input exemplar texture. If Gaussian random fields [12, 13, 14, 15] give good visual results for input textures with no, or few, spatial redundancy, they fail when it comes to sampling structured textures (brick walls, fabric with repeated patterns, etc.) and more elaborated models are needed [16, 17]. In this work, we propose a statistical framework to get a better understanding of the spatial redundancy information. To be more precise, we aim at giving the explicit probability distribution function of the random variables associated to the output of similarity functions between local windows of random fields, in order to conduct rigorous statistical testing on the spatial redundancy in natural images.

In order to compute these explicit distributions we will consider specific random fields over specific topological spaces. First, the random fields will be defined either over $\mathbb{R}^d$ (or $\mathbb{T}^d$, where $\mathbb{T}^d$ is the d-dimensional torus, when considering periodicity assumptions on the field), or over $\mathbb{Z}^d$ (or $(\mathbb{Z}/(M\mathbb{Z}))^d$, with $M \in \mathbb{N}$ when considering periodicity assumptions on the field). The first case is the *continuous setting*, whereas the second one is the *discrete setting*. Since all the applications we are primarily interested in are image processing tasks we set $d = 2$. In image processing, the most common framework is the finite discrete setting, considering a finite grid of pixels. The discrete setting ($\mathbb{Z}^2$) can be used to define asymptotic properties when the size of images grows or when their resolution increases [18], whereas continuous settings are needed in specific applications where, for instance, rotation invariant models are required [19]. All the considered random fields will be Gaussian. This assumption will allow us to explicitly derive moments of some similarity functions computed on local windows of the random field. Once again another reason for this restriction comes from image processing. Indeed, given an input image, we can compute its first and second-order statistics. Sampling from the associated Gaussian random field gives examples of images which conserve the covariance structure but lose the global arrangement of the input image. Investigating redundancy of such fields is a first step towards giving a mathematical description of this lack of structure. The seminal work of Adler et al. [20] provides us the tools to perform geometrical analysis of such models. Mathematical properties of Gaussian random field models in image processing were described and investigated in [13].

The need of explicit distributions also restricts the choice of the similarity functions we are going to consider. Finding measurements which correspond to the one of our visual system is a long-standing problem in image processing. It was considered in the early days of texture synthesis and analyzed by Julesz [21, 22, 23] who formulated the conjecture that textures with similar first-order statistics (first conjecture) or that textures with similar first and second-order statistics (second conjecture) could not be discriminated by the human eye. Even if both conjectures were disproved [24], in a discrete setting, the work of Gatys et al. [16] suggests that second-order statistics of image features are enough to characterize a broad range of textures. To compute features on images we embed the image in a higher dimensional space. This operation can be conducted using linear filtering [25] or convolutional neural networks [16] for instance. Recent work examines the response of convolutional neural network to elementary geometrical pattern [26], giving insight about the perceptual properties of such a lifting. Another embedding is given by considering a square neighborhood, a patch, around each pixel. This embedding, is exploited in many image processing tasks such as inpainting [27], denoising [9, 28], texture synthesis [29, 30, 31, 32], etc. In order to avoid the curse of dimensionality small patches are considered. This justifies our definition of similarity functions on local windows. Recalling that, in this study, the distribution of the similarity functions must be explicit when computed on Gaussian random fields, we only consider similarity functions that are easy to compute over these random fields, using the properties of the scalar product associated to the L^2 norm.

In the special case of the L^2 norm, explicit distributions can be inferred even in the non-asymptotic case, *i.e.* when considering the finite discrete setting. This distribution is not easy to compute exactly since it requires the knowledge of some covariance matrix eigenvalues as well as an efficient method to compute cumulative distribution functions of quadratic forms of Gaussian random variables. We address both of these problems.

The paper is organized as follows. We introduce Gaussian random fields in general settings in Section 2.1. Similarity functions to be evaluated on these random fields, as well as their statistical properties, are described in Section 2.2. We give the asymptotic properties of these similarity functions in Gaussian random fields in the discrete setting in Section 3.1 and in the continuous setting in Section 3.2. It is shown in Section 3.3 that the Gaussian asymptotic approximation is valid only for large patches. This is not a reasonable situation since we aim at deriving local measurements. In order to bypass this problem we consider an explicit formulation of the probability distribution function for a particular similarity function, the L^2 norm. The computations are done in the finite discrete case in Section 4.1. We also derive an efficient algorithm to compute these probability distribution functions. Similar non-asymptotic expressions are given in the continuous case in Section 4.2.

2 Similarity functions and random fields

2.1 Gaussian random fields

Let $(\mathcal{A}, \mathcal{F}, \mathbb{P})$ be a probability space. Following [20], a random field over a topological vector space Ω is defined as a measurable mapping $U : \mathcal{A} \to \mathbb{R}^\Omega$. Thus, for all a in $\mathcal{A}$, $U(a)$ is a function over Ω and, for any $a \in \mathcal{A}$, and any $\mathbf{x} \in \Omega$, $U(a)(\mathbf{x})$ is a real number. For the sake of clarity we will omit a in what follows.

Assuming that U is a second-order random field, *i.e.* for any finite sequence $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in \Omega^n$ with $n \in \mathbb{N}$, the vector $(U(\mathbf{x}_1), \dots, U(\mathbf{x}_n))$ is square-integrable, we define the mean function of U, $m : \Omega \to \mathbb{R}$ as well as its covariance function, $C : \Omega^2 \to \mathbb{R}$ by for any $\mathbf{x}, \mathbf{y} \in \Omega^2$

$$m(\mathbf{x}) = \mathbb{E}\left[U(\mathbf{x})\right] \qquad \text{and} \quad C(\mathbf{x}, \mathbf{y}) = \mathbb{E}\left[(U(\mathbf{x}) - m(\mathbf{x}))(U(\mathbf{y}) - m(\mathbf{y}))\right] .$$

A random field U is said to be stationary if for any finite sequence $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in \Omega^n$ with $n \in \mathbb{N}$ and $\mathbf{t} \in \Omega$, the vector $(U(\mathbf{x}_1), \dots, U(\mathbf{x}_n))$ and $(U(\mathbf{x}_1 + \mathbf{t}), \dots, U(\mathbf{x}_n + \mathbf{t}))$ have same distribution. A second-order random field U over a topological vector field is said to be stationary in the weak sense if its mean function is constant and if for all $\mathbf{x}, \mathbf{y} \in \Omega$, $C(\mathbf{x}, \mathbf{y}) = C(\mathbf{x} - \mathbf{y}, \mathbf{0})$. In this case the covariance of U is fully characterized by its auto-covariance function $\Gamma : \Omega \to \mathbb{R}$ given for any $\mathbf{x} \in \Omega$ by

$$\Gamma(\mathbf{x}) = C(\mathbf{x}, 0) .$$

A random field U is said to be a Gaussian random field if, for any finite sequence $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in \Omega^n$ with $n \in \mathbb{N}$, the vector $(U(\mathbf{x}_1), \dots, U(\mathbf{x}_n))$ is Gaussian. The distribution of a Gaussian random field is entirely characterized by its mean and covariance functions. As a consequence, the notions of stationarity and weak stationarity coincide for Gaussian random fields.

Since the applications we are interested in are image processing tasks, we restrict ourselves (in the continuous setting) to the case where $\Omega = \mathbb{R}^2$. In Section 2.2 we will consider the Lebesgue integrals of random fields and thus need integrability condition for U over compact sets. Let $K = [a, b] \times [c, d]$ be a compact rectangular domain in $\mathbb{R}^2$. Imposing continuity conditions on the function C implies that $\int_K g(\mathbf{x})U(\mathbf{x})\mathrm{d}\mathbf{x}$ is well-defined as the quadratic mean limit, see [33]. for real-valued functions g over Ω such that $\int_{K \times K} g(\mathbf{x})g(\mathbf{y})C(\mathbf{x}, \mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}$ is finite, see [33]. However, we are interested in almost sure quantities and thus we want the integral to be defined almost surely over rectangular windows. Imposing the existence of a continuous modification of a random field, ensures the almost sure existence of Riemann integrals over rectangular windows. The following assumptions will ensure continuity almost surely, see Lemma 1 below and [20, 34]. We define $d : \Omega \times \Omega \to \mathbb{R}$ such that for any $\mathbf{x}, \mathbf{y} \in \Omega$

$$d(\mathbf{x}, \mathbf{y}) = \mathbb{E}\left[(U(\mathbf{x}) - U(\mathbf{y}))^2\right] = C(\mathbf{x}, \mathbf{x}) + C(\mathbf{y}, \mathbf{y}) - 2C(\mathbf{x}, \mathbf{y}) .$$

Assumption 1 (A1). *U is a second-order random field and there exist $M, \eta, \alpha > 0$ such that for any $\mathbf{x} \in \Omega$ and $\mathbf{y} \in \mathrm{B}(\mathbf{x}, \eta) \cap \Omega$ with $\mathbf{y} \neq \mathbf{x}$ we have*

$$d(\mathbf{x}, \mathbf{y}) \leqslant \frac{M\|\mathbf{x} - \mathbf{y}\|_2^2}{|\log(\|\mathbf{x} - \mathbf{y}\|_2)|^{2+\alpha}} .$$

This assumption can be considerably weakened in the case of a stationary Gaussian random field.

Assumption 2 (A2). *U is a stationary Gaussian random field and there exist $M, \eta, \alpha > 0$ such that for any $\mathbf{x} \in \Omega$ and $\mathbf{y} \in \mathrm{B}(\mathbf{x}, \eta) \cap \Omega$ with $\mathbf{y} \neq \mathbf{x}$ we have*

$$d(\mathbf{x}, \mathbf{y}) \leqslant \frac{M}{|\log(\|\mathbf{x} - \mathbf{y}\|_2)|^{1+\alpha}} .$$

Assumptions (A1) and (A2) both imply quadratic mean continuity of the random field U, *i.e.* for any $\mathbf{x} \in \Omega$, $\lim_{\mathbf{y} \to \mathbf{x}} d(\mathbf{x}, \mathbf{y}) = 0$. Note that the quadratic mean continuity implies the continuity of the covariance function and the mean function of the random field U. However, since assumptions (A1) and (A2) are stronger than the quadratic mean continuity we can show that each realization of the random field is continuous, up to a modification, see Lemma 1. The demonstration of this lemma can be found in [20] for the Gaussian case and in [34] for the general case.

Lemma 1 (Sample path continuity). *Assume (A1) or (A2). In addition suppose that for any $\mathbf{x} \in \Omega$, $m(\mathbf{x}) = 0$. Then there exists a modification of U, i.e. a random field $\tilde{U}$ such that for any $\mathbf{x} \in \Omega$, $\mathbb{P}\left[U(\mathbf{x}) = \tilde{U}(\mathbf{x})\right] = 1$, and for any $a \in \mathcal{A}$, $\tilde{U}(a)$ is continuous over Ω.*

In what follows we will replace U by its continuous modification $\tilde{U}$. Note that in the discrete case all random fields are continuous with respect to the discrete topology.

In Sections 3 and 4, we will suppose that U is a stationary Gaussian random field with zero mean. Asymptotic theorems derived in the next section remain true in broader frameworks, however restricting ourselves to stationary Gaussian random fields allows for explicit computations of asymptotic quantities in order to numerically assess the speed of convergence.

2.2 Similarity functions

In order to evaluate redundancy in random fields, we first need to derive a criterion for comparing random fields. We introduce similarity functions which take rectangular restrictions of random fields as inputs.

When comparing local windows of random fields (patches), two cases can occur. We can compare a patch with a patch extracted from the same image. We call this situation *internal matching*. Applications can be found in denoising [9] or inpainting [10] where the information of the image itself is used to perform the image processing task. On the other hand, we can compare a patch with a patch extracted from another image. We call this situation *template matching*. An application of this case is presented in the non-parametric exemplar-based texture synthesis algorithm proposed by Efros and Leung. [29].

The Euclidean distance is the usual way to measure the similarity between patches [28] but many other measurements exist, corresponding to different structural properties, see Figure 1. When they are defined, we introduce p-norms and angle measurements similarity functions.

Definition 1. *Let $P, Q \in \mathbb{R}^\omega$ with $\omega \subset \mathbb{R}^2$ or $\omega \subset \mathbb{Z}^2$. When it is defined we introduce*

(a) the L^p-similarity, $s_p(P,Q) = \|P-Q\|_p = \left(\int_{\mathbf{x}\in\omega} |P(\mathbf{x}) - Q(\mathbf{x})|^p \mathrm{d}\mu(\mathbf{x})\right)^{1/p}$, with $p \in (0,+\infty)$;

(b) the L^∞-similarity, $s_\infty(P,Q) = \sup_\omega(|P-Q|)$;

(c) the p-th power of the L^p-similarity, $s_{p,p}(P,Q) = s_p(P,Q)^p$, with $p \in (0,+\infty)$;

(d) the scalar product similarity, $s_{sc}(P,Q) = -\langle P, Q\rangle = \frac{1}{2}\left(s_{2,2}(P,Q) - \|P\|_2^2 - \|Q\|_2^2\right)$;

(e) the cosine similarity, $s_{\cos}(P,Q) = \frac{s_{sc}(P,Q)}{\|P\|_2\|Q\|_2}$, if $\|P\|_2\|Q\|_2 \neq 0$.

Depending on the case μ is either the Lebesgue measure on ω or the discrete measure on ω.

The locality of the measurements is ensured by the fact that these functions are defined on patches. Following conditions (1) and (3) in [35] we check that similarity functions (a), (c) and (e) satisfy the following properties

- (Symmetry) $s(P,Q) = s(Q,P)$;
- (Maximal self-similarity) $s(P,P) \leqslant s(P,Q)$;
- (Equal self-similarities) $s(P,P) = s(Q,Q)$.

Note that since s_{sc}, the scalar product similarity, is homogeneous in P, maximal self-similarity and equal self-similarity properties are not satisfied. In [35], the authors present many other similarity functions all relying on statistical properties such as likelihood ratios, joint likelihood criteria and mutual information kernels. The latter measurement is defined by a cosine in some feature space. In this paper we focus only on similarity functions defined directly in the spatial domain.

Definition 2 (Auto-similarity and template similarity)**.** *Let u and v be two functions defined over a domain $\Omega = \mathbb{R}^2$ or $\mathbb{Z}^2$. Let $\omega \subset \Omega$ be a patch domain. We introduce $P_\omega(u) = u|_\omega$, the restriction of u to the patch domain ω. When it is defined we introduce the auto-similarity with patch domain ω and offset $\mathbf{t} \in \mathbb{R}^2$ or $\mathbb{Z}^2$ such that $\mathbf{t} + \omega \subset \Omega$ by*

$$\mathcal{AS}_i(u, \mathbf{t}, \omega) = s_i\left(P_{\mathbf{t}+\omega}(u), P_\omega(u)\right),$$

where s_i corresponds to s_p with $p \in (0,+\infty]$, $s_{p,p}$ with $p \in (0,+\infty)$, s_{sc} or s_{cos}. In the same way, when it is defined, we introduce the template similarity with patch ω and offset $\mathbf{t}$ by

$$\mathcal{TS}_i(u, v, \mathbf{t}, \omega) = s_i\left(P_{\mathbf{t}+\omega}(u), P_\omega(v)\right).$$

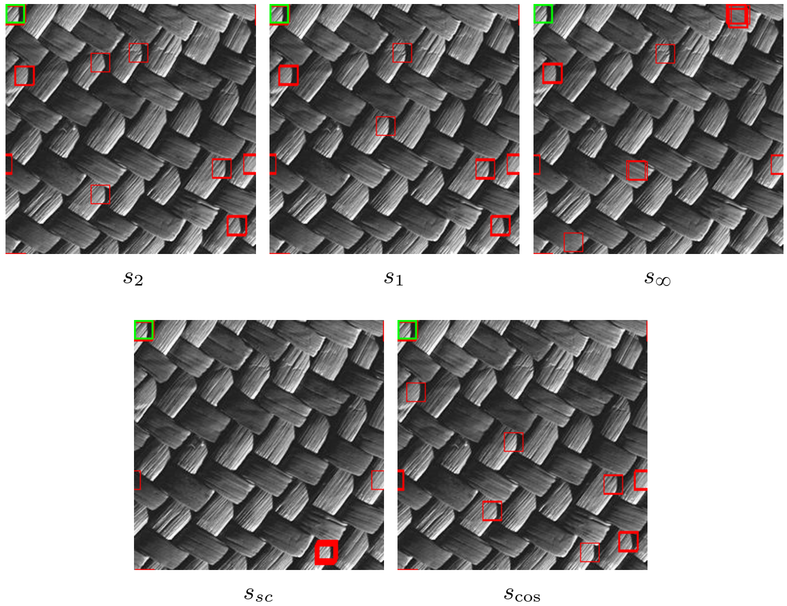


Figure 1: ***Structural properties of similarity functions.*** In this experiment we show in the 20×20 patch space the twenty closest patches to the green upper-left patch in the original image for different similarity functions. All introduced similarity functions, see Definition 1, correctly identify the structure of the patch, *i.e.* a large clear part with diagonal textures and a dark ray on the right side of the patch, except for s_∞ which is too sensitive to outliers. Indeed outliers have more importance for s_∞ than they have perceptually speaking. Similarities s_2 and s_1 have analogous behaviors and find correct regions. It might be noted that s_1 is more conservative as it identifies seven main different patches and s_2 identifies eight. Similarity s_{sc} is too sensitive to contrast and, as it finds a correct patch, it gives too much importance to illumination. The behavior of $s_{\cos}$ is interesting as it avoids some of the illumination problems encountered with the scalar product. The identified regions were also found with s_1 and s_2, but with the addition of a new one.

Note that in the finite discrete setting, *i.e.* $\Omega = (\mathbb{Z}/(M\mathbb{Z}))^2$ with $M \in \mathbb{N}$, the definition of $\mathcal{AS}$ and $\mathcal{TS}$ can be extended to any patch domain $\omega \subset \mathbb{Z}^2$ by replacing u by $\dot{u}$ its periodic extension to $\mathbb{Z}^2$.

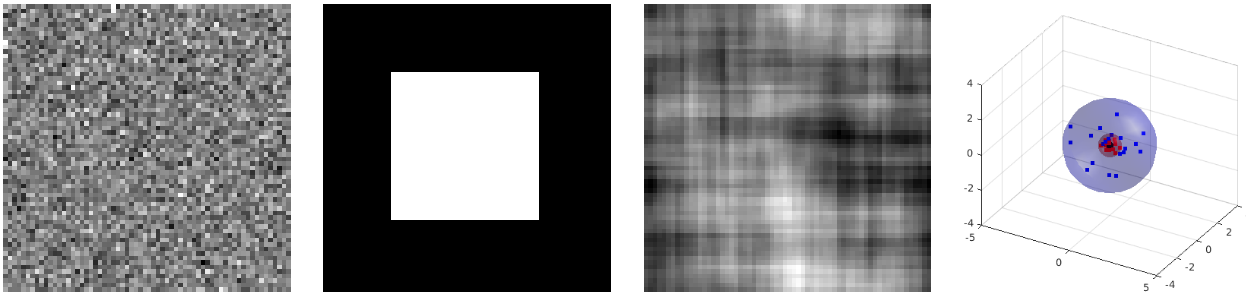
Suppose we evaluate the scalar product auto-similarity $\mathcal{AS}_{sc}(U, \mathbf{t}, \omega)$ with U a random field. Then the auto-similarity function is a random variable and its expectation depends on the second-order statistics of U. In the template case, the expectation of $\mathcal{TS}_{sc}(U, v, \mathbf{t}, \omega)$ depends on the first-order statistics of U. This shows that auto-similarity and template similarity can exhibit very different behaviors even for the same similarity functions.

For image processing tasks, the auto-similarity computes the local resemblance between a patch in u and a patch of the same size at another position in the same image corresponding to a shift by the offset $\mathbf{t}$, whereas the template similarity uses an image v as input and computes local resemblances between u and v.

In the discrete case, it is well-known that, due to the curse of dimensionality, the L^2 norm does not behave well in large-dimensional spaces and is a poor measure of structure. Thus, considering u and v two images, $s_2(u, v)$, the L^2 template similarity on full images, does not yield interesting information about the perceptual differences between u and v. The template similarity $\mathcal{TS}_2(u, v, \mathbf{0}, \omega)$ avoids this effect by considering patches which reduces the dimension of the data (if the cardinality of ω, denoted $|\omega|$, is small) and also allows for fast computation of similarity mappings, see Figure 1 for a comparison of the different similarity functions on a natural image.

In Figure 2, we investigate the behavior of the patch lifting operation on different Gaussian random fields. Roughly speaking, patches are said to be similar if they are clustered in the patch space. Using the Principal Component Analysis we illustrate that patches are more scattered in Gaussian white noise than in the Gaussian random field $U = f * W$ (with periodic convolution, *i.e.* $f * W(\mathbf{x}) = \sum_{\mathbf{y} \in \Omega} \dot{f}(\mathbf{y}) \dot{W}(\mathbf{x} - \mathbf{y})$ where $\dot{f}$ is the periodic extension of f to $\mathbb{Z}^2$), where W is a Gaussian white noise over Ω (a finite discrete grid) and f is the indicator function of a rectangle non reduced to a single point.

We continue this investigation in Figure 3 in which we present the closest patches (of size 10×10), for the L^2 norm, in two Gaussian random fields $U = f * W$ (where the convolution is periodic) for different functions f called spots. The



(a) (b) (c) (d)

Figure 2: ***Gaussian models and spatial redundancy.*** In this experiment we illustrate the notion of spatial redundancy in two models. In (A) we present a Gaussian white noise over a finite 64×64 discrete domain. (B) shows a function f with mean 0 and standard deviation $\sigma = 64^{-1}$. (C) presents a realization of the Gaussian random field defined by $f * W$ (with periodic convolution) where W is a Gaussian white noise over $\Omega = 64 \times 64$. Note that f was designed so that the two Gaussian random fields have the same gray-level distribution for each pixel. For each image we compute 64^2 9-dimensional vectors corresponding to the 64^2 vectors with initial point the top-left patch and final point any other patch (note that we handle the boundaries periodically), and patch domain $\omega = 3 \times 3$. These 9-dimensional vectors are projected in a 3 dimensional space using Principal Component Analysis. We only show the twenty shortest vectors for the Gaussian white noise model (in blue) and the Gaussian random field associated to function (B) (in red). The radius of the blue, respectively red, sphere represents the maximal L^2 norm of these 20 vectors in the Gaussian white noise model, respectively in model (C). Since the radius of the blue sphere is larger than the red one the points are more scattered in the patch space of (A) than in the patch space of (B). This implies that there is more spatial redundancy in (C) than in (A) which is expected.

more regular f is the more the patches are similar. Limit cases are $f = 0$ (all patches are constant) and $f = \delta_0$, *i.e.* $U = W$. Note that any stationary periodic Gaussian random field over $(\mathbb{Z}/(M\mathbb{Z}))^2$ with $M \in \mathbb{N}$ can be written as the convolution of some spot function f with a Gaussian white noise W [36]. We introduce the notion of autocorrelation. Let $f \in L^2(\mathbb{Z}^2)$. We denote by Γ_f the autocorrelation of f, *i.e.* $\Gamma_f = f * \check{f}$ where for any $\mathbf{x} \in \mathbb{Z}^2$, $\check{f}(x) = f(-x)$. In a more general setting we introduce the associated random field to a square-integrable function f as the stationary Gaussian random field U such that for any $\mathbf{x} \in \Omega$

$$\mathbb{E}\left[U(\mathbf{x})\right] = 0 , \qquad \text{and} \quad \Gamma(\mathbf{x}) = \Gamma_f(\mathbf{x}) .$$

In Figure 4, we compare the patch spaces of natural images and the one of their associated random fields. Since the associated Gaussian random fields lose all global structure, most of the spatial information is discarded. This situation can be observed in the patch space. In the natural images, patches containing the same highly spatial information (such as a white diagonal) are close for the L^2 norm. In Gaussian random field since this highly spatial information is lost, close patches for the L^2 norm are not necessarily perceptually close.

3 Asymptotic results

In this Section we aim at giving explicit asymptotic expressions for the probability distribution functions of the auto-similarity and the template similarity in both discrete and continuous settings. Using general versions of the law of large numbers and central limit theorem we will derive Gaussian asymptotic approximations.

We start by introducing two notions which will be crucial in order to derive a law of large numbers and a central limit theorem in broad settings. The R-independence, see Definition 3, ensures long-range independence whereas stochastic domination will replace integrability conditions in the standard law of large numbers or central limit theorem.

The notion of R-independence generalizes to $\mathbb{R}^2$ and $\mathbb{Z}^2$ the associated one-dimensional concept, see [37] and its extension to $\mathbb{N}^2$ [38], [39].

Definition 3 (R-independence)**.** *Let $\Omega = \mathbb{R}^2$ or $\Omega = \mathbb{Z}^2$ and V be a random field over Ω. Let $K_1, K_2 \subset \Omega$ be two compact sets, and $V|_{K_i}$ be the restriction of V to K_i, $i \in \{1, 2\}$. We say that V is R-independent, with $R \geqslant 0$, if $V|_{K_1}$ is independent from $V|_{K_2}$ as soon as $d_\infty(K_1, K_2) = \min_{\mathbf{x} \in K_1, \mathbf{y} \in K_2} \|\mathbf{x} - \mathbf{y}\|_\infty > R$.*

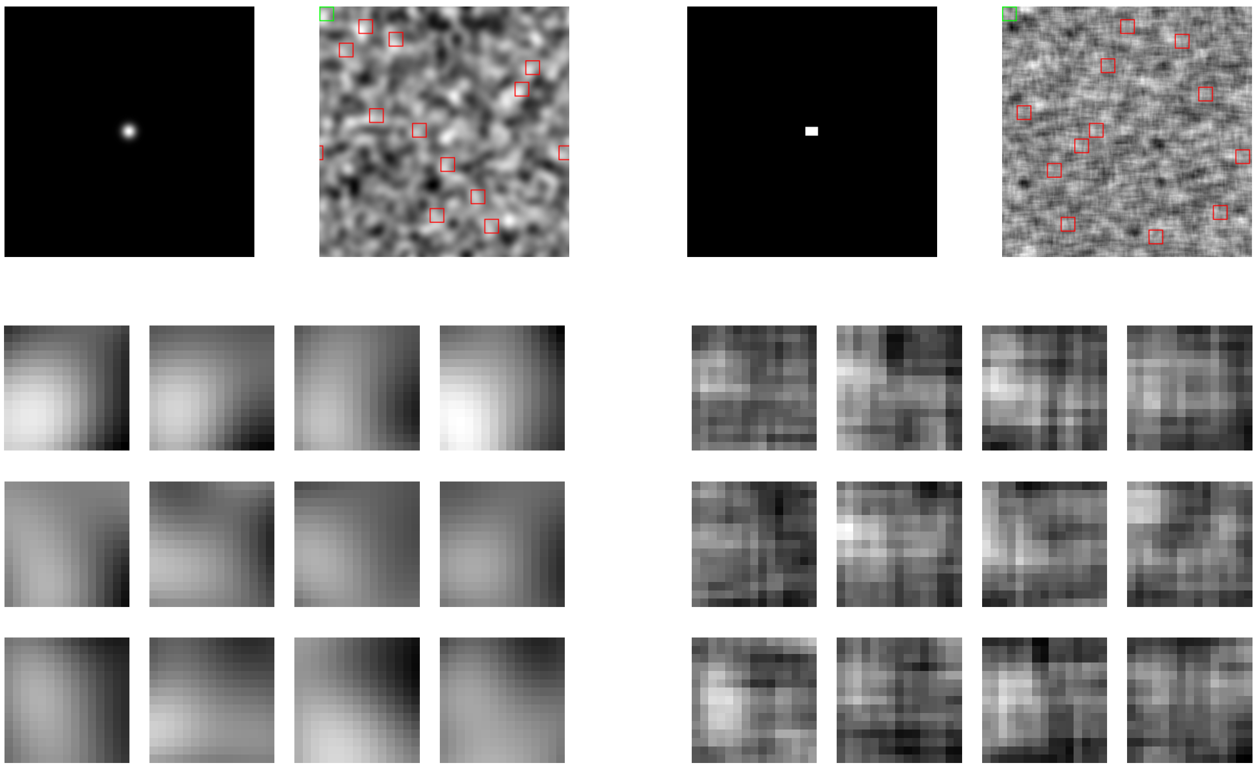


Figure 3: ***Patch similarity in Gaussian random fields.*** In this figure we show two examples of Gaussian random fields in the discrete periodic case. On the left of the first row we show a Gaussian spot f and a realization of the Gaussian random field $U = f * W$, where the convolution is periodic and W is a Gaussian white noise. Since the spot is isotropic so is the Gaussian random field. The smoothness of the realization of Gaussian random field is linked with the smoothness of the spot function f. The situation is different when considering a rectangular plateau spot (here the support of the spot is 7×5). The associated random field $U = f * W$ is no longer smooth nor isotropic. Images are displayed on the right of their respective of their respective spot. We illustrate the notion of clustering in the patch space. For each setting (Gaussian spot or rectangular spot) we present 12 patches of size 15×15. In each case the top-left patch is the top-left patch in the presented realization of the random field, shown in green. Following from the top to the bottom and the left to the right are the closest patches in the patch space for the L^2 norm. We discard patches which are spatially too close (we impose $\|\mathbf{x} - \mathbf{y}\|_\infty \geqslant 10$ for all indices of different patch domains). Note that since the random field with Gaussian spot is more regular than the one with rectangular spot, the 11 closest patches are more perceptually similar.

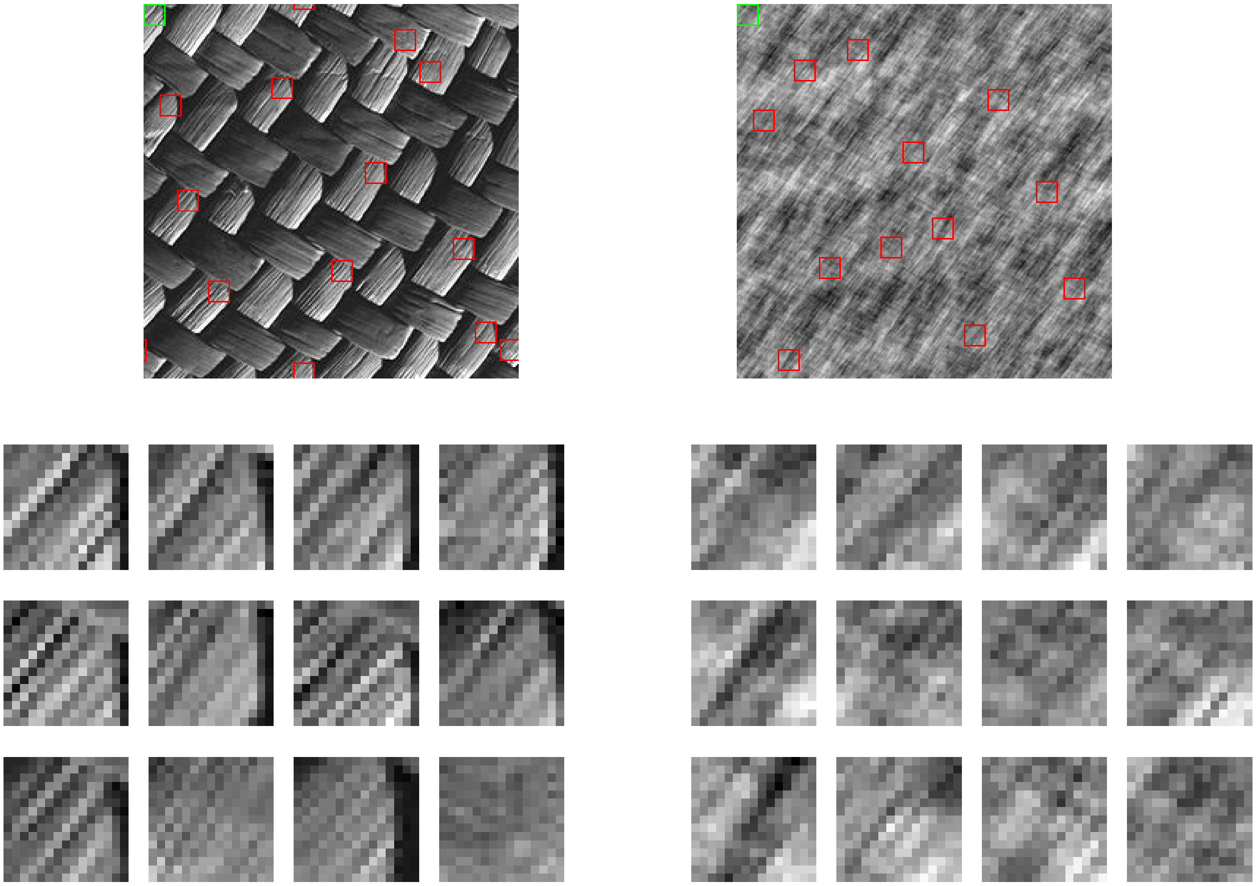


Figure 4: ***Natural images and Gaussian random fields.*** In this experiment we present the same image, f, which was used in Figure 1 and the associated Gaussian random field $U = f * W$, where the convolution is periodic and W is a Gaussian white noise. As in Figure 3 we present under each image the top-left patch (of size 15×15 and shown in green in the original images) and its 11 closest matches. We discard patches which are spatially too close (we impose $\|\mathbf{x} - \mathbf{y}\|_\infty \geqslant 10$ for all indices of different patch domains). Note that if a structure is clearly identified in the real image (black and white diagonals) and is retrieved in every patch, it is not as clear in the Gaussian random field. Contrast information seems to have more importance than structure information.

Note that in the case of $\Omega = \mathbb{Z}^2$, compacts sets K_1 and K_2 are finite sets of indices. This notion of R-independence will replace the traditional assumption of independence in asymptotic theorems.

Definition 4 (Uniform domination)**.** *Let $\Omega = \mathbb{R}^2$ or $\Omega = \mathbb{Z}^2$ and let $V, \tilde{V}$ be random fields over Ω. We say that:*

(a) $\tilde{V}$ uniformly stochastically dominates V if for any $\alpha \geqslant 0$ and $\mathbf{x} \in \Omega$, $\mathbb{P}\left[V(\mathbf{x}) \geqslant \alpha\right] \leqslant \mathbb{P}\left[\tilde{V}(\mathbf{x}) \geqslant \alpha\right]$;

(b) $\tilde{V}$ uniformly almost surely dominates V if for any $\mathbf{x} \in \Omega$, $V(\mathbf{x}) \leqslant \tilde{V}(\mathbf{x})$ almost surely.

Note that if $\tilde{V}$ uniformly almost surely dominates V then $\tilde{V}$ uniformly stochastically dominates $\tilde{V}$.

Supplementary hypotheses are required in the case of template matching since we use an exemplar input image v to compute $\mathcal{TS}_i(U, v, \mathbf{t}, \omega)$. Let $v \in \mathbb{R}^\Omega$, where Ω is equal to $\mathbb{R}^2$ or $\mathbb{Z}^2$. We denote by $(v_k)_{k \in \mathbb{N}}$ the sequence of the restriction of v to ω_k, with $\omega_k \subset \Omega$, extended to $\mathbb{Z}^2$ (or $\mathbb{R}^2$) by zero-padding, *i.e.* $v_k(\mathbf{x}) = 0$ for $x \notin \omega_k$. We suppose that $\lim_{k \to +\infty} |\omega_k| = +\infty$, where $|\omega_k|$ is the Lebesgue measure, respectively the cardinality, of ω_k if $\Omega = \mathbb{R}^2$, respectively $\Omega = \mathbb{Z}^2$. Note that the following assumptions are well-defined for both continuous and discrete settings.

Assumption 3 (A3)**.** *The function v is bounded on Ω.*

We also introduce the following assumption, ensuring the existence of spatial moments of any order for the function v.

Assumption 4 (A4). *For any $m, n \in \mathbb{N}$, there exist $\beta_m \in \mathbb{R}$ and $\gamma_{m,n} \in \mathbb{R}^\Omega$ such that*

(a) $\lim_{k\to+\infty} |\omega_k|^{1/2} \left(|\omega_k|^{-1} \int_{\omega_k} v_k^{2m}(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) - \beta_m\right) = 0$;

(b) for any $K \subset \Omega$ compact, $\lim_{k\to+\infty} \| |\omega_k|^{-1} v_k^{2m} * \check{v}_k^{2n} - \gamma_{m,n}\|_{\infty,K} = 0$,

with $\|\cdot\|_{\infty,K}$ such that for any $u \in \mathbb{R}^\Omega$, $\|u\|_{\infty,K} = \sup_{\mathbf{x}\in K} |u(\mathbf{x})|$. Note that in the case where Ω is discrete the uniform convergence on compact sets introduced in (b) is equivalent to the pointwise convergence.

Assumption 5 (A5). *There exists $\gamma \in \mathbb{R}^\Omega$ such that for any $K \subset \Omega$, compact,* $\lim_{k\to+\infty} \| |\omega_k|^{-1} v_k * \check{v}_k - \gamma\|_{\infty,K} = 0$.

3.1 Discrete case

In the discrete case, we consider a random field U over $\mathbb{Z}^2$ and compute local similarity measurements. The asymptotic approximation is obtained when the patch size grows to infinity. In Theorem 1 and Theorem 2 we obtain Gaussian asymptotic probability distribution in the auto-similarity case and in the template similarity case. In Proposition 1 and Proposition 2 we give explicit mean and variance for the Gaussian approximations.

Theorem 1 (Discrete case – asymptotic auto-similarity results). *Let $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}}$ be two positive increasing integer sequences and $(\omega_k)_{k\in\mathbb{N}}$ be the sequence of subsets such that for any $k \in \mathbb{N}$, $\omega_k = [\![0, m_k]\!] \times [\![0, n_k]\!]$. Let $f \in \mathbb{R}^{\mathbb{Z}^2}$, $f \neq 0$ with finite support, W a Gaussian white noise over $\mathbb{Z}^2$ and $U = f * W$. For $i = sc$, p or (p,p) with $p \in (0,+\infty)$ there exist $\mu_i, \sigma_i \in \mathbb{R}^{\mathbb{Z}^2}$ and $(\alpha_{i,k})_{k\in\mathbb{N}}$ a positive sequence such that for any $\mathbf{t} \in \mathbb{Z}^2\setminus\{\mathbf{0}\}$ we get*

(a) $\lim_{k\to+\infty} \frac{1}{\alpha_{i,k}} \mathcal{AS}_i(U,\mathbf{t},\omega_k) \underset{a.s}{=} \mu_i(\mathbf{t})$;

(b) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{\alpha_{i,k}} \mathcal{AS}_i(U,\mathbf{t},\omega_k) - \mu_i(\mathbf{t})\right) \underset{\mathcal{L}}{=} \mathcal{N}(0, \sigma_i(\mathbf{t}))$.

Proof. The proof is divided in two parts. First we show (a) and (b) for $i = p, p$ and extends the result to $i = p$. Then we show (a) and (b) for $i = sc$.

Let $p \in (0,+\infty)$, $\mathbf{t} \in \mathbb{Z}^2\setminus\{\mathbf{0}\}$ and define $V_{p,\mathbf{t}}$ the random field on $\mathbb{Z}^2$ by for any $\mathbf{x} \in \mathbb{Z}^2$, $V_{p,\mathbf{t}}(\mathbf{x}) = |U(\mathbf{x}) - U(\mathbf{x}+\mathbf{t})|^p$. We remark that for any $k \in \mathbb{N}$ we have

$$\mathcal{AS}_{p,p}(U,\mathbf{t},\omega_k) = \sum_{\mathbf{x}\in\omega_k} V_{p,\mathbf{t}}(\mathbf{x}) \ .$$

We first notice that U is R-independent with $R > 0$, see Lemma 2 in Appendix. Since for any $\mathbf{x} \in \mathbb{Z}^2$ we have that $V_{p,\mathbf{t}}(\mathbf{x})$ depends only on $U(\mathbf{x})$ and $U(\mathbf{x}+\mathbf{t})$ we have that $V_{p,\mathbf{t}}$ is $R_\mathbf{t} = R + \|\mathbf{t}\|_\infty$-independent. Since U is stationary, so is $V_{p,\mathbf{t}}$. The random field $V_{p,\mathbf{t}}$ admits moments of every order since it is the p-th power of the absolute value of a Gaussian random field. Thus $V_{p,\mathbf{t}}$ is a $R_\mathbf{t}$-independent second-order stationary random field. We can apply Lemma 3 and we get

(a) $\lim_{k\to+\infty} \frac{1}{|\omega_k|} \mathcal{AS}_{p,p}(U,\mathbf{t},\omega_k) \underset{a.s.}{=} \mu_{p,p}(\mathbf{t})$;

(b) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{|\omega_k|} \mathcal{AS}_{p,p}(U,\mathbf{t},\omega_k) - \mu_{p,p}(\mathbf{t})\right) \underset{\mathcal{L}}{=} \mathcal{N}(0, \sigma_{p,p}(\mathbf{t}))$.

with $\mu_{p,p}(\mathbf{t}) = \mathbb{E}[V_{p,\mathbf{t}}(\mathbf{0})]$ and $\sigma_{p,p}(\mathbf{t})^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}[V_{p,\mathbf{t}}(\mathbf{x}), V_{p,\mathbf{t}}(\mathbf{0})]$. By continuity of the p-th root over $[0,+\infty)$ we get (a) for $i = p$ with

$$\alpha_{p,k} = |\omega_k|^{1/p} \ , \qquad \mu_p(\mathbf{t}) = \mu_{p,p}(\mathbf{t})^{1/p} \ .$$

By Lemma 4 in Appendix we get that $\mathbb{E}\left[(U(\mathbf{0}) - U(\mathbf{t}))^2\right] = 2(\Gamma_f(\mathbf{0}) - \Gamma_f(\mathbf{t})) > 0$ thus $\mu_{p,p}(\mathbf{t}) = \mathbb{E}[V_{p,\mathbf{t}}(\mathbf{0})] > 0$. Since the p-th root is continuously differentiable on $(0,+\infty)$ we can apply the Delta method, see [40], and we get (b) for $i = p$ with

$$\alpha_{p,k} = |\omega_k|^{1/p} \ , \qquad \mu_p(\mathbf{t}) = \mu_{p,p}(\mathbf{t})^{1/p} \ , \qquad \sigma_p(\mathbf{t})^2 = \frac{1}{p^2}\sigma_{p,p}(\mathbf{t})^2 \mu_{p,p}(\mathbf{t})^{2/p-2} \ . \tag{1}$$

We now prove the theorem for $i = sc$. Let $\mathbf{t} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$ and define $V_{sc,\mathbf{t}}$ the random field on $\mathbb{Z}^2$ by for any $\mathbf{x} \in \mathbb{Z}^2$, $V_{sc,\mathbf{t}}(\mathbf{x}) = -U(\mathbf{x})U(\mathbf{x}+\mathbf{t})$. We remark that for any $k \in \mathbb{N}$ we have

$$\mathcal{AS}_{sc}(U, \mathbf{t}, \omega_k) = \sum_{\mathbf{x} \in \omega_k} V_{sc,\mathbf{t}}(\mathbf{x}) \ .$$

Since for any $\mathbf{x} \in \mathbb{Z}^2$, $V_{sc,\mathbf{t}}(\mathbf{x})$ depends only on $U(\mathbf{x})$ and $U(\mathbf{x}+\mathbf{t})$, we have that $V_{sc,\mathbf{t}}$ is $R_{\mathbf{t}} = R + \|\mathbf{t}\|_\infty$-independent. Since U is stationary, so is $V_{sc,\mathbf{t}}$. The random field $V_{sc,\mathbf{t}}$ admits moments of every order since it is a product of Gaussian random fields. Thus $V_{sc,\mathbf{t}}$ is a $R_{\mathbf{t}}$-independent second-order stationary random field. We can again apply Lemma 3 in Appendix and we get

(a) $\lim_{k \to +\infty} \frac{1}{|\omega_k|} \mathcal{AS}_{sc}(U, \mathbf{t}, \omega_k) \underset{a.s.}{=} \mu_{sc}(\mathbf{t})$;

(b) $\lim_{k \to +\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{|\omega_k|} \mathcal{AS}_{sc}(U, \mathbf{t}, \omega_k) - \mu_{sc}(\mathbf{t}) \right) \underset{\mathcal{L}}{=} \mathcal{N}\left(0, \sigma_{sc}(\mathbf{t})\right)$,

with $\mu_{sc}(\mathbf{t}) = \mathbb{E}\left[V_{sc,\mathbf{t}}(\mathbf{0})\right]$ and $\sigma_{sc}(\mathbf{t})^2 = \sum_{\mathbf{x} \in \mathbb{Z}^2} \mathrm{Cov}\left[V_{sc,\mathbf{t}}(\mathbf{x}), V_{sc,\mathbf{t}}(\mathbf{0})\right]$, which concludes the proof. □

In the following proposition we give explicit values for the constants involved in the law of large numbers and the central limit theorem derived in Theorem 1. We introduce the following quantities for $k, \ell \in \mathbb{N}$ and $j \in [\![0, k \wedge \ell]\!]$

$$q_\ell = \frac{(2\ell)!}{\ell!\, 2^\ell} \ , \qquad r_{j,k,\ell} = q_{k-j} q_{\ell-j} \binom{2k}{2j} \binom{2\ell}{2j} (2j)! \ . \tag{2}$$

We also denote $r_{j,\ell} = r_{j,\ell,\ell}$. Note that for all $\ell \in \mathbb{N}$, $r_{0,\ell} = q_L^2$ and $\sum_{j=0}^{\ell} r_{j,\ell} = q_{2\ell}$. We also introduce the following functions:

$$\Delta_f(\mathbf{t}, \mathbf{x}) = 2\Gamma_f(\mathbf{x}) - \Gamma_f(\mathbf{x}+\mathbf{t}) - \Gamma_f(\mathbf{x}-\mathbf{t}) \ , \quad \tilde{\Delta}_f(\mathbf{t}, \mathbf{x}) = \Gamma_f(\mathbf{x})^2 + \Gamma_f(\mathbf{x}+\mathbf{t})\Gamma_f(\mathbf{x}-\mathbf{t}) \ . \tag{3}$$

Note that Δ_f is a second-order statistic on the Gaussian field $U = f * W$ with W a Gaussian white noise over $\mathbb{Z}^2$, whereas $\tilde{\Delta}_f$ is a forth-order statistic on the same random field.

Proposition 1 (Explicit constants – Auto-similarity)**.** *In Theorem 1 we have the following constants for any* $\mathbf{t} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$.

(i) If $i = p$ *with* $p = 2\ell$ *and* $\ell \in \mathbb{N}$*, then for all* $k \in \mathbb{N}$*, we get that* $\alpha_{p,k} = |\omega_k|^{1/(2\ell)}$ *and*

$$\mu_p(\mathbf{t}) = q_\ell^{1/(2\ell)} \Delta_f(\mathbf{t}, \mathbf{0})^{1/2} \quad \text{and} \quad \sigma_p(\mathbf{t})^2 = \frac{q_\ell^{1/\ell - 2}}{(2\ell)^2} \sum_{j=1}^{\ell} r_{j,\ell} \left(\frac{\|\Delta_f(\mathbf{t}, \cdot)\|_{2j}}{\Delta_f(\mathbf{t}, \mathbf{0})} \right)^{2j} \Delta_f(\mathbf{t}, \mathbf{0}) \ .$$

(ii) If $i = sc$*, then for all* $k \in \mathbb{N}$*, we get that* $\alpha_{sc,k} = |\omega_k|$ *and*

$$\mu_{sc}(\mathbf{t}) = \Gamma_f(\mathbf{t}) \quad \text{and} \quad \sigma_{sc}(\mathbf{t})^2 = \sum_{\mathbf{x} \in \mathbb{Z}^2} \tilde{\Delta}_f(\mathbf{t}, \mathbf{x}) \ .$$

Proof. The proof is postponed to Appendix C. □

We now derive similar asymptotic properties in the template similarity case.

Theorem 2 (Discrete case – asymptotic template similarity results)**.** *Let* $(m_k)_{k \in \mathbb{N}}, (n_k)_{k \in \mathbb{N}}$ *be two positive increasing integer sequences and* $(\omega_k)_{k \in \mathbb{N}}$ *be the sequence of subsets such that for any* $k \in \mathbb{N}$*,* $\omega_k = [\![0, m_k]\!] \times [\![0, n_k]\!]$*. Let* $f \in \mathbb{R}^{\mathbb{Z}^2}$*,* $f \neq 0$ *with finite support,* W *a Gaussian white noise over* $\mathbb{Z}^2$*,* $U = f * W$ *and let* $v \in \mathbb{R}^{\mathbb{Z}^2}$*. For* $i = sc, p$ *or* (p, p) *with* $p = 2\ell$ *and* $\ell \in \mathbb{N}$*, if* $i = p$ *or* (p, p) *assume (A3) and (A4), if* $i = sc$ *assume (A3) and (A5). Then there exist* $\mu_i, \sigma_i \in \mathbb{R}$ *and* $(\alpha_{i,k})_{k \in \mathbb{N}}$ *a positive sequence such that for any* $\mathbf{t} \in \mathbb{Z}^2$ *we get*

(a) $\lim_{k \to +\infty} \frac{1}{\alpha_{i,k}} \mathcal{TS}_i(U, v, \mathbf{t}, \omega_k) \underset{a.s}{=} \mu_i$;

(b) $\lim_{k \to +\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{\alpha_{i,k}} \mathcal{TS}_i(U, v, \mathbf{t}, \omega_k) - \mu_i(\mathbf{t}) \right) \underset{\mathcal{L}}{=} \mathcal{N}\left(0, \sigma_i\right)$.

Note that contrarily to Theorem 1 we could not obtain such a result for all $p \in (0, +\infty)$. Indeed, in the general case the convergence of the sequence $\left(|\omega_k|^{-1}\mathbb{E}\left[\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k)\right]\right)_{k\in\mathbb{N}}$, which is needed in order to apply Theorem 5, is not trivial. Assuming that v is bounded it is easy to show that $\left(|\omega_k|^{-1}\mathbb{E}\left[\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k)\right]\right)_{k\in\mathbb{N}}$ is also bounded and we can deduce the existence of a convergent subsequence. In the general case, for Theorem 2 to hold with any $p \in (0, +\infty)$, we must verify that or any $\mathbf{t} \in \Omega$, there exist $\mu_{p,p}(\mathbf{t}) > 0$ and $\sigma_{p,p}(\mathbf{t}) \geqslant 0$ such that

(a) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{|\omega_k|}\mathbb{E}\left[\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k)\right] - \mu_{p,p}(\mathbf{t})\right) = 0$;

(b) $\lim_{k\to+\infty} \frac{1}{|\omega_k|} \operatorname{Var}\left[\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k)\right] = \sigma^2_{p,p}(\mathbf{t})$.

We now turn to the proof of Theorem 2.

Proof. The proof is divided in two parts. First we show (a) and (b) for $i = (p, p)$ and extends the result to $i = p$. Then we show (a) and (b) for $i = sc$.

Let $p \in (0, +\infty)$, $\mathbf{t} \in \mathbb{Z}^2$ and define $V_{p,\mathbf{t}}$ the random field on $\mathbb{Z}^2$ by for any $\mathbf{x} \in \mathbb{Z}^2$, $V_{p,\mathbf{t}}(\mathbf{x}) = |v(\mathbf{x}) - U(\mathbf{x} + \mathbf{t})|^p$. We remark that for any $k \in \mathbb{N}$ we have

$$\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k) = \sum_{\mathbf{x}\in\omega_k} V_{p,\mathbf{t}}(\mathbf{x}) \ .$$

By Lemma 2 in Appendix, U is R-independent with $R > 0$. Since for any $\mathbf{x} \in \mathbb{Z}^2$ we have that $V_{p,\mathbf{t}}(\mathbf{x})$ depends only on $U(\mathbf{x} + \mathbf{t})$ we also have that $V_{p,\mathbf{t}}$ is R-independent. We define the random field $V^\infty_{p,\mathbf{t}}$ by for any $\mathbf{x} \in \mathbb{Z}^2$, $V^\infty_{p,\mathbf{t}}(\mathbf{x}) = (\sup_{\mathbb{Z}^2} |v| + U(\mathbf{x}+\mathbf{t}))^p$. We have that $V^\infty_{p,\mathbf{t}}(\mathbf{x}) + \mathbb{E}\left[V^\infty_{p,\mathbf{t}}(\mathbf{0})\right]$ uniformly almost surely dominates $V_{p,\mathbf{t}}(\mathbf{x}) - \mathbb{E}\left[V_{p,\mathbf{t}}(\mathbf{x})\right]$. The random field $V^\infty_{p,\mathbf{t}}$ admits moments of every order since it is the p-th power of the absolute value of a Gaussian random field and is stationary because U is. Thus $V_{p,\mathbf{t}}$ is a $R_{\mathbf{t}}$-independent random field uniformly and $V_{p,\mathbf{t}}(\mathbf{x}) - \mathbb{E}\left[V_{p,\mathbf{t}}(\mathbf{x})\right]$ is stochastically dominated by $V^\infty_{p,\mathbf{t}}(\mathbf{x}) + \mathbb{E}\left[V^\infty_{p,\mathbf{t}}(\mathbf{0})\right]$, a second-order stationary random field. Using (A4) and Lemma 5 in Appendix, we can apply Theorem 5 and 6 and we get

(a) $\lim_{k\to+\infty} \frac{1}{|\omega_k|}\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k) \underset{a.s.}{=} \mu_{p,p}(\mathbf{t})$;

(b) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{|\omega_k|}\mathcal{TS}_{p,p}(U, v, \mathbf{t}, \omega_k) - \mu_{p,p}(\mathbf{t})\right) \underset{\mathcal{L}}{=} \mathcal{N}\left(0, \sigma_{p,p}(\mathbf{t})\right)$.

Note that since U is stationary we have for any $\mathbf{t} \in \mathbb{Z}^2$, $\mu_{p,p} = \mu_{p,p}(\mathbf{0}) = \mu_{p,p}(\mathbf{t})$ and $\sigma_{p,p} = \sigma_{p,p}(\mathbf{0}) = \sigma_{p,p}(\mathbf{t})$. By continuity of the p-th root over $[0, +\infty)$ we get (a) for $i = p$ with

$$\alpha_{p,k} = |\omega_k|^{1/p} \ , \qquad \mu_p = \mu^{1/p}_{p,p} \ .$$

By Lemma 5, we have that $\mu_{p,p} > 0$. Since the p-th root is continuously differentiable on $(0, +\infty)$ we can apply the Delta method and we get (b) for $i = p$ with

$$\alpha_{p,k} = |\omega_k|^{1/p} \ , \qquad \mu_p = \mu^{1/p}_{p,p} \ , \qquad \sigma^2_p = \sigma^2_{p,p}\mu^{2/p-2}_{p,p}/p^2 \ . \tag{4}$$

We now prove the theorem for $i = sc$. Let $\mathbf{t} \in \mathbb{Z}^2$ and define $V_{sc,\mathbf{t}}$ the random field on $\mathbb{Z}^2$ such that for any $\mathbf{x} \in \mathbb{Z}^2$, $V_{sc,\mathbf{t}}(\mathbf{x}) = -v(\mathbf{x})U(\mathbf{x} + \mathbf{t})$. We remark that for any $k \in \mathbb{N}$ we have

$$\mathcal{TS}_{sc}(U, v, \mathbf{t}, \omega_k) = \sum_{\mathbf{x}\in\omega_k} V_{sc,\mathbf{t}}(\mathbf{x}) \ .$$

It is clear that for any $k \in \mathbb{N}$, $\mathcal{TS}_{sc}(U, v, \mathbf{t}, \omega_k)$ is a R-independent Gaussian random variable with $\mathbb{E}\left[\mathcal{TS}_{sc}(U, v, \mathbf{t}, \omega_k)\right] = 0$ and

$$\operatorname{Var}\left[\mathcal{TS}_{sc}(U, v, \mathbf{t}, \omega_k)\right] = \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \mathbb{E}\left[V_{sc,\mathbf{t}}(\mathbf{x})V_{sc,\mathbf{t}}(\mathbf{y})\right] = \sum_{\mathbf{x},\mathbf{y}\in\omega_k} v(\mathbf{x})v(\mathbf{y})\Gamma_f(\mathbf{x} - \mathbf{y}) = \sum_{\mathbf{x}\in\mathbb{Z}^2} \Gamma_f(\mathbf{x})v_k * \check{v}_k(\mathbf{x}) \ ,$$

where we recall that v_k is the restriction of v to ω_k. The last sum is finite since $\operatorname{Supp}(f)$ finite implies that $\operatorname{Supp}(\Gamma_f)$ is finite. Using (A5) we obtain that for any $k \in \mathbb{N}$,

$$\sum_{\mathbf{x}\in\omega_k} \left(\mathbb{E}\left[V_{sc,\mathbf{t}}\right](\mathbf{x}) - \mu_{sc}\right) = 0 \ , \qquad \lim_{k\to+\infty} |\omega_k|^{-1} \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}\left[V_{sc,\mathbf{t}}(\mathbf{x}), V_{sc,\mathbf{t}}(\mathbf{y})\right] = \sigma^2_{sc} \ , \tag{5}$$

with $\mu_{sc} = 0$ and $\sigma^2_{sc} = \sum_{\mathbf{x}\in\mathbb{Z}^2} \Gamma_f(\mathbf{x})\tilde{v}(\mathbf{x})$ and. Since $V_{sc,\mathbf{t}}$ is a R-independent second-order random field using (5) we can apply Theorems 5 and 6 to conclude. $\square$

Proposition 2 (Explicit constants – template similarity)**.** *In Theorem 2 we have the following constants for any* $\mathbf{t} \in \mathbb{Z}^2$.

(i) If $i = p$ *with* $p = 2\ell$ *and* $\ell \in \mathbb{N}$*, then we get that* $\alpha_{p,k} = |\omega_k|^{\frac{1}{p}}$*, and*

$$\mu_p = \left(\sum_{j=0}^{\ell} \binom{2\ell}{2j} q_{\ell-j} \Gamma_f(\mathbf{0})^{-j} \beta_j \right)^{1/2\ell} \Gamma_f(\mathbf{0})^{1/2} \ ,$$

$$\sigma_p^2 = \left(\sum_{i,j=0}^{\ell} \binom{2\ell}{2i} \binom{2\ell}{2j} \sum_{m=1}^{\ell-i \wedge \ell-j} r_{m,\ell-i,\ell-j} \Gamma_f(\mathbf{0})^{-(i+j+2m)} \left\langle \Gamma_f^{2m}, \gamma_{i,j} \right\rangle \right) \left(\sum_{j=0}^{\ell} \binom{2\ell}{2j} q_{\ell-j} \Gamma_f(\mathbf{0})^{-j} \beta_j \right)^{1/\ell-2} \frac{\Gamma_f(\mathbf{0})}{(2\ell)^2} \ .$$

(ii) If $i = sc$ *then for all* $k \in \mathbb{N}$*, we get that* $\alpha_{sc,k} = |\omega_k|$ *and*

$$\mu_{sc} = 0 \ , \quad \sigma_{sc}^2 = \sum_{\mathbf{x} \in \mathbb{Z}^2} \Gamma_f(\mathbf{x}) \tilde{v}(\mathbf{x}) \ .$$

Proof. The proof is postponed to Appendix C. □

Note that limit mean and standard deviation do not depend on the offset anymore. Indeed, template similarity function are stationary in $\mathbf{t}$.

If v has finite support then (A4) holds with $\beta_i = 0$ and $\gamma_{i,j} = 0$ as soon as $i \neq 0$ or $j \neq 0$. Remarking that $\beta_0 = 1$ and $\gamma_{0,0} = 1$ we obtain that

$$\mu_p = q_\ell^{1/(2\ell)} \Gamma_f(\mathbf{0})^{1/2} \ , \quad \sigma_p^2 = \frac{q_\ell^{1/\ell-2}}{(2\ell)^2} \sum_{j=1}^{\ell} r_{j,\ell} \left(\frac{\|\Gamma_f\|_{2j}}{\Gamma_f(\mathbf{0})} \right)^{2j} \Gamma_f(\mathbf{0}) \ .$$

Limit mean and standard deviation in the p-norm template similarity do not depend on v. This result comes from the finite support property of v in Proposition 2 which implies that v is not considered in the similarity functions for large windows.

In both Theorem 1 and 2 we could have derived a law of large numbers for the cosine similarity function. Obtaining a central limit theorem with explicit constants, however, seems more technical, since it will require the use of a multidimensional version of the Delta method [40] in order to compute the asymptotic variance.

3.2 Continuous case

We now turn to the the continuous setting. Theorem 3, respectively Theorem 4, is the continuous counterpart of Theorem 1, respectively Theorem 2.

Theorem 3 (Continuous case – asymptotic auto-similarity results)**.** *Let* $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}}$ *be two positive increasing integer sequences and* $(\omega_k)_{k\in\mathbb{N}}$ *be the sequence of subsets such that for any* $k \in \mathbb{N}$*,* $\omega_k = [0, m_k] \times [0, n_k]$*. Let* U *be a zero-mean Gaussian random field over* $\mathbb{R}^2$ *with covariance function* Γ*. Assume (A2) and that* Γ *has finite support. For* $i \in \{sc, p, (p,p)\}$ *with* $p \in (0, +\infty)$ *there exist* $\mu_i, \sigma_i \in \mathbb{R}^{\mathbb{Z}^2}$ *and* $(\alpha_{i,k})_{k\in\mathbb{N}}$ *a positive sequence such that for any* $\mathbf{t} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ *we get*

(a) $\lim_{k\to+\infty} \frac{1}{\alpha_{i,k}} \mathcal{AS}_i(U, \mathbf{t}, \omega_k) \underset{a.s}{=} \mu_i(\mathbf{t})$;

(b) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{\alpha_{i,k}} \mathcal{AS}_i(U, \mathbf{t}, \omega_k) - \mu_i(\mathbf{t}) \right) \underset{\mathcal{L}}{=} \mathcal{N}(0, \sigma_i(\mathbf{t}))$.

Proof. The proof is the same as the one of Theorem 1 replacing Lemma 3 and Lemma 4 by Lemma 6 and Lemma 7.

□

Proposition 3 (Explicit constants – Continuous auto-similarity)**.** *Constants given in Proposition 1 apply to Theorem 3 provided that* Γ_f *is replaced by* Γ *in* (3).

Proof. The proof is the same as the one of Proposition 1. □

Theorem 4 (Continuous case – asymptotic template similarity results)**.** *Let $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}}$ be two positive increasing integer sequences and $(\omega_k)_{k\in\mathbb{N}}$ be the sequence of subsets such that for any $k \in \mathbb{N}$, $\omega_k = [0, m_k] \times [0, n_k]$. Let U be a zero-mean Gaussian random field over $\mathbb{R}^2$ with covariance function Γ. Assume (A2) and that Γ has finite support. For $i \in \{sc, p, (p, p)\}$ with $p \in (0, +\infty)$, if $i = p$ or (p, p) assume (A3) and (A4), if $i = sc$ assume (A3) and (A5). Then there exist $\mu_i, \sigma_i \in \mathbb{R}$ and $(\alpha_{i,k})_{k\in\mathbb{N}}$ a positive sequence such that for any $\mathbf{t} \in \mathbb{R}^2$ we get*

(a) $\lim_{k\to+\infty} \frac{1}{\alpha_{i,k}} \mathcal{TS}_i(U, v, \mathbf{t}, \omega_k) \underset{a.s.}{=} \mu_i$;

(b) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{\alpha_{i,k}} \mathcal{TS}_i(U, v, \mathbf{t}, \omega_k) - \mu_i(\mathbf{t}) \right) \underset{\mathcal{L}}{=} \mathcal{N}(0, \sigma_i)$.

Proof. The proof is the same as the one of Theorem 1. $\square$

Proposition 4 (Explicit constants – Continuous auto-similarity)**.** *Constants given in Proposition 2 apply to Theorem 4 provided that Γ_f is replaced by Γ in* (3)*.*

Proof. The proof is the same as the one of Proposition 2. $\square$

3.3 Speed of convergence

In the discrete setting, Theorem 1 justifies the use of a Gaussian approximation to compute $\mathcal{AS}_i(U, \mathbf{t}, \omega)$. However this asymptotic behavior strongly relies on the increasing size of the patch domains. We define the patch size to be $|\omega|$, the cardinality of ω, and the spot size $|\mathrm{Supp}\,(f)\,|$ to be the cardinality of the support of the spot f. The quantity of interest is the ratio $r = \frac{\text{patch size}}{\text{spot size}}$. If $r \gg 1$ then the Gaussian random field associated to f can be well approximated by a Gaussian white noise from the patch perspective. If $r \approx 1$ this approximation is not valid and the Gaussian approximation is no longer accurate, see Figure 5. We say that an offset $\mathbf{t}$ is *detected* in a Gaussian random field if $\mathcal{AS}_i(U, \mathbf{t}, \omega) \leqslant a(\mathbf{t})$ for some threshold $a(\mathbf{t})$. In the experiments presented in Figure 6 and Table 1 the threshold is given by the asymptotic Gaussian inverse cumulative distribution function evaluated at some quantile. The parameters of the Gaussian random variable are given by Proposition 1. We find that except for small spot sizes and large patches, *i.e.* $r \gg 1$, the approximation is not valid. More precisely, let $U = f * W$ with f a finitely supported function over $\mathbb{Z}^2$ and W a Gaussian white noise over $\mathbb{Z}^2$. Let $\omega \subset \mathbb{Z}^2$ and let Ω_0 be a finite subset of $\mathbb{Z}^2$. We compute $\sum_{\mathbf{t}\in\Omega_0} \mathbb{1}_{\mathcal{AS}_i(U,\mathbf{t},\omega)\leqslant a(\mathbf{t})}$, with $a(\mathbf{t})$ defined by the inverse cumulative distribution function of quantile $10/|\Omega_0|$ for the Gaussian $\mathcal{N}(\mu, \sigma)$ where μ, σ are given by Theorem 1 and Proposition 1. Note that $a(\mathbf{t})$ satisfies $\mathbb{P}\left[\mathcal{AS}_i(U, \mathbf{t}, \omega) \leqslant a(\mathbf{t})\right] \approx 10/|\Omega_0|$ if the approximation for the cumulative distribution function was correct. In other words, if the Gaussian asymptotic was always valid, we would have a number of detections equal to 10 independently of r. This is clearly not the case in Table 1. One way to interpret this is by looking at the left tail of the approximated distribution for $s_{2,2}$ and s_{sc} on Figure 5. For s_{sc} the histogram is above the estimated curve, see (a) in Figure 6 for example. Whereas for $s_{2,2}$ the histogram is under the estimated curve. Thus for s_{sc} we expect to obtain more detections than what is predicted whereas we will observe the opposite behavior for $s_{2,2}$. This situation is also illustrated for similarities s_2 and s_{sc} in Figure 6 in which we compare the asymptotic cumulative distribution function with the empirical one.

In the next section we address this problem by studying non-asymptotic cases for the $s_{2,2}$ auto-similarity function in both continuous and discrete settings.

4 A non-asymptotic case: internal Euclidean matching

4.1 Discrete periodic case

In this section Ω is a finite rectangular domain in $\mathbb{Z}^2$. We fix $\omega \subset \Omega$. We also define f a function over Ω. We consider the Gaussian random field $U = f * W$ (we consider the periodic convolution) with W a Gaussian white noise over Ω.

In the previous section, we derived asymptotic properties for similarity functions. However, a necessary condition for the asymptotic Gaussian approximation to be valid is for the spot size to be very small when compared to the patch size. This condition is not often met and non-asymptotic techniques must be developed. Some cases are easy. For instance it should be noted that the distribution of the s_{sc} template similarity, $\mathcal{TS}_i(U, v, \mathbf{t}, \omega)$, is Gaussian for every ω. We might also derive a non-asymptotic expression for the template similarity in the cosine case if the Gaussian model is a white noise model. In what follows we restrict ourselves to the auto-similarity framework and consider the square of the L^2 norm auto-similarity function, *i.e.* $\mathcal{AS}_{2,2}(u, \mathbf{t}, \omega)$. In this case we show that there exists an efficient method to compute the cumulative distribution function of the auto-similarity function in the non-asymptotic case.

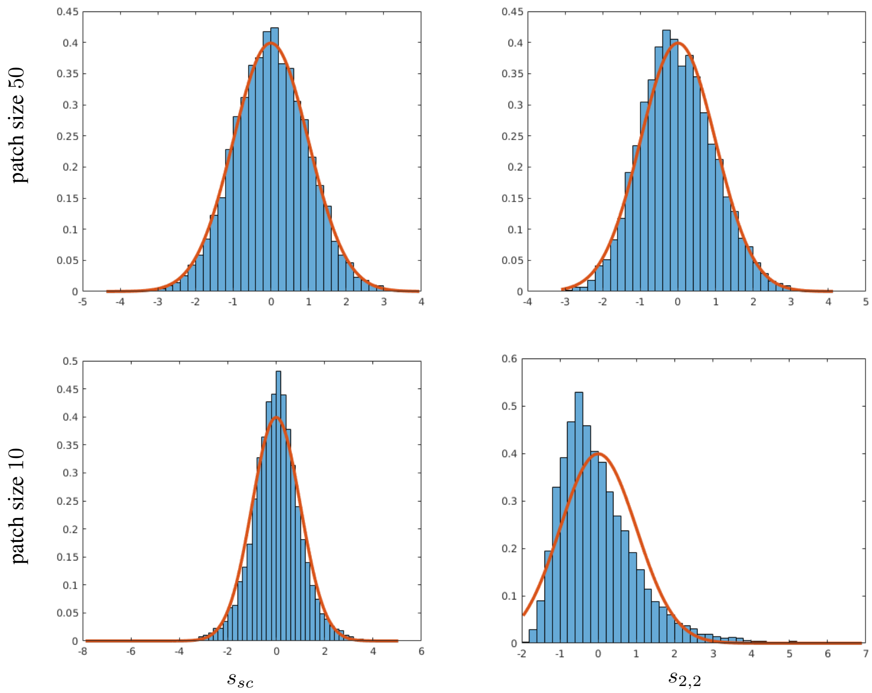
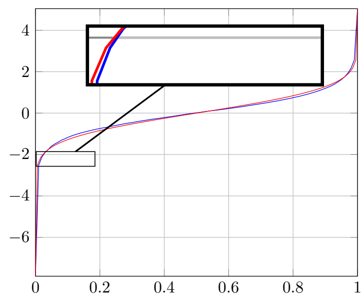


Figure 5: ***Gaussian moment matching.*** In this experiment, 10^4 realizations of 128×128 Gaussian images are computed with spot of size 5×5 (the spot is the indicator of this square). Scalar product auto-similarities and squared L^2 auto-similarities are computed for a fixed offset $(70, 100)$. Values are normalized to fit a standard Gaussian using the moment method. We plot the histograms of these normalized values. The red curve corresponds to the standard Gaussian $\mathcal{N}(0, 1)$. On the top row $r = 100 \gg 1$ and the Gaussian approximation is adapted. On the bottom row $r \approx 1$ and the Gaussian approximation is not valid. Note that the reasons for which the asymptotic approximations are not valid vary from one similarity function to another. In the case of s_{sc}, the modes coincide but the asymptotic distribution is not spiked enough around this mode. In the case of s_2 the asymptotic mode does not seem to coincide with the empirical one. Note also that using $s_{2,2}$ we impose that the similarity function takes its values in $[0, +\infty)$. This positivity constraint is omitted when considering a Gaussian approximation.

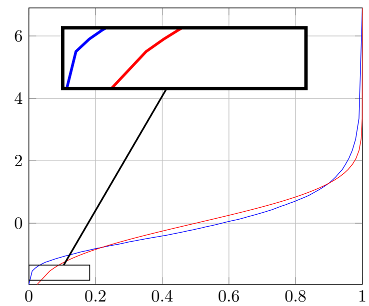
	5	10	15	20	40	70
1	0.3	1.4	3.2	4.6	7.4	9.0
2	0.3	0.4	1.2	2.2	5.8	8.5
5	0.3	0.4	0.4	0.5	1.3	4.1
10		0.4	0.5	0.5	0.4	1.4
15			0.5	0.5	0.5	0.5
20				0.5	0.5	0.5
25					0.5	0.5

	5	10	15	20	40	70
1	18.1	11.6	10.9	10.4	10.1	10.0
2	34.2	16.5	12.8	11.5	10.4	9.9
5	93.9	49.3	30.8	20.9	13.2	11.5
10		86.7	57.6	46.0	19.7	14.5
15			83.9	63.8	30.0	18.2
20				79.5	36.7	24.7
25					51.5	26.6

Table 1: ***Asymptotic properties.*** Number of detections with different patch domains from 5×5 to 70×70 and spot domains from 1×1 to 25×25 for the $s_{2,2}$ (left table) or s_{sc} (right table) auto-similarity function. We generate 5000 Gaussian random field images of size 256×256 for each setting (with spot the indicator of the spot domain). We set $\alpha = 10/256^2$. For each setting we compute $a(\mathbf{t})$ the inverse cumulative distribution function of $\mathcal{N}(\mu_i(\mathbf{t}), \sigma_i(\mathbf{t}))$ evaluated at quantile α, with μ_i and σ_i given by Proposition 1. For each pair of patch size and spot size we compute $\sum_{\mathbf{t} \in \Omega} \mathbb{1}_{\mathcal{AS}_i(u,\mathbf{t},\omega) \leqslant a(\mathbf{t})}$, namely the number of detections, for all the 5000 random fields realizations. The empirical averages are displayed in the table. If $\mathcal{AS}_i(u, \mathbf{t}, \omega)$ had Gaussian distribution with parameters given by Proposition 1 then the number in each cell would be $\sum_{\mathbf{t} \in \Omega} \mathbb{P}\left[\mathcal{AS}_i(U, \mathbf{t}, \omega) \leqslant a(\mathbf{t})\right] \approx 10$.



(a) s_{sc}



(b) s_2

Figure 6: ***Theoretical and empirical cumulative distribution function.*** This experiment illustrates the non-Gaussianity in Figure 5. Theorem 1 asserts that asymptotically the distribution of $\mathcal{AS}_i(U, \mathbf{t}, \omega)$ is Gaussian with parameters given by Proposition 1. In both cases, the red curve is the inverse cumulative distribution function of the standard Gaussian and the blue curve is the empirical inverse cumulative distribution function of normalized auto-similarity functions computed with 10^4 realizations of Gaussian models. Here we focus on the situation where the patch size is fixed to 10×10, *i.e.* $r \approx 1$. We present auto-similarity results obtained for $\mathbf{t} = (70, 100)$ and similarity function s_{sc} (on the left) and s_2 (on the right). We note that for rare events, see the magnified region, the theoretical inverse cumulative distribution function is above the empirical inverse cumulative distribution function. The opposite behavior is observed for similarity s_2. These observations are in accordance with the findings of Table 1.

Proposition 5 (Squared L^2 auto-similarity function exact probability distribution function)**.** *Let $\Omega = (\mathbb{Z}/M\mathbb{Z})^2$ with $M \in \mathbb{N}$, $\omega \subset \Omega$, $f \in \mathbb{R}^\Omega$ and $U = f * W$ where W is a Gaussian white noise over Ω. The following equality holds for any $\mathbf{t} \in \Omega$ up to a change of the underlying probability space*

$$\mathcal{AS}_{2,2}(U, \mathbf{t}, \omega) \underset{a.s}{=} \sum_{k=0}^{|\omega|-1} \lambda_k(\mathbf{t}, \omega) Z_k \ , \tag{6}$$

with Z_k independent chi-square random variables with parameter 1 and $\lambda_k(\mathbf{t}, \omega)$ the eigenvalues of the covariance matrix $C_\mathbf{t}$ associated with function $\Delta_f(\mathbf{t}, \cdot)$, see Equation (3)*, restricted to ω, i.e for any $\mathbf{x_1}, \mathbf{x_2} \in \omega$, $C_\mathbf{t}(\mathbf{x_1}, \mathbf{x_2}) = \Delta_f(\mathbf{t}, \mathbf{x_1} - \mathbf{x_2})$.*

Note that if we consider the Mahalanobis distance instead of the L^2 distance when computing the auto-similarity function then the equality in (6) is still valid with $\lambda_k(\mathbf{t}, \omega) = 1$ and the random variable $\mathcal{AS}_M(U, \mathbf{t}, \omega)$ (where we replace the square L^2 norm by the Mahalanobis distance) has the same distribution as a chi-square random variable with parameter $|\omega|$.

Proof. Let $\mathbf{t} \in \Omega$ and $V_\mathbf{t}$ defined for any $\mathbf{x} \in \Omega$ by $V_\mathbf{t}(\mathbf{x}) = U(\mathbf{x}) - U(\mathbf{x} + \mathbf{t})$. It is a Gaussian vector with mean 0 and covariance matrix C_V given for any $\mathbf{x_1}, \mathbf{x_2} \in \Omega$ by

$$C_V(\mathbf{x_1}, \mathbf{x_2}) = 2\Gamma_f(\mathbf{x_1} - \mathbf{x_2}) - \Gamma_f(\mathbf{x_1} - \mathbf{x_2} - \mathbf{t}) - \Gamma_f(\mathbf{x_1} - \mathbf{x_2} + \mathbf{t}) = \Delta_f(\mathbf{t}, \mathbf{x_1} - \mathbf{x_2}) \ .$$

The covariance of the random field $P_\omega(V_\mathbf{t})$, the restriction of $V_\mathbf{t}$ to ω, is defined by the restriction of $\Delta_f(\mathbf{t}, \cdot)$ to $\omega + (-\omega)$, in the sense of the Minkowski sum. This new covariance matrix, $C_\mathbf{t}$, is symmetric and the spectral theorem ensures that there exists an orthonormal basis $\mathcal{B}$ such that $C_\mathbf{t}$ is diagonal when expressed in $\mathcal{B}$. Thus we obtain that $P_\omega(V_\mathbf{t}) = \sum_{e_k \in \mathcal{B}} \langle P_\omega(V_\mathbf{t}), e_k \rangle e_k$. It is clear that, for any $k \in [\![0, |\omega| - 1]\!]$, $\langle P_\omega(V_\mathbf{t}), e_k \rangle$ is a Gaussian random variable with mean 0 and variance $e_k^T C_\mathbf{t} e_k = \lambda_k(\mathbf{t}, \omega) \geqslant 0$. We set $K = \{k \in [\![0, |\omega| - 1, \ \lambda_k(\mathbf{t}, \omega) \neq 0\}$ and define X a random vector in $\mathbb{R}^{|\omega|}$ such that

$$X_k = \lambda_k(\mathbf{t}, \omega)^{-1/2} \langle P_\omega(V_\mathbf{t}), e_k \rangle, \text{ if } k \in K \ , \qquad \text{and } X_{K_-} = Y \ ,$$

where X_{K_-} is the restriction of X to the indices of $K_- = [\![0, |\omega| - 1]\!] \backslash K$ and Y is a standard Gaussian random variable on $\mathbb{R}^{|K_-|}$ independent from the sigma field generated by $\{(X_k), k \in K\}$. By construction we have $\mathbb{E}[X_k X_\ell] = 0$ if $\ell \in K$ and $k \in K_-$, or $\ell \in K_-$ and $k \in K_-$. Suppose now that $k, \ell \in K$. We obtain that

$$\mathbb{E}[X_k X_\ell] = \lambda_k(\mathbf{t}, \omega)\lambda_\ell(\mathbf{t}, \omega)\mathbb{E}\left[e_k^T C_{\mathbf{t}} e_\ell\right] = 0 \ .$$

Thus X is a standard Gaussian random vector and we have $P_\omega(V_t) = \sum_{k=0}^{|\omega|-1} \lambda^{1/2}(\mathbf{t}, \omega) X_k e_k$, where the equality holds almost surely. We get that

$$\mathcal{AS}_{2,2}(U, \mathbf{t}, \omega) \underset{\text{a.s}}{=} \|P_\omega(V_{\mathbf{t}})\|_2^2 = \sum_{e_k \in \mathcal{B}} \langle P_\omega(V_{\mathbf{t}}), e_k \rangle^2 = \sum_{k=0}^{|\omega|-1} \lambda_k(\mathbf{t}, \omega) X_k^2 \ .$$

Setting $Z_k = X_k^2$ concludes the proof. □

Note that if $\omega = \Omega$ then we obtain that the covariance matrix $C_{\mathbf{t}}$ is block-circulant with circulant blocks and the eigenvalues are given by the discrete Fourier transform.

In order to compute the true cumulative distribution function of the auto-similarity square L^2 norm we need to: 1) compute the cumulative distribution function of a positive-weighted sum of independent chi-square random variable ; 2) compute the eigenvalues of a covariance matrix in $\mathcal{M}_{|\omega|}(\mathbb{R})$. Storing all covariance matrices for each offset $\mathbf{t}$ is not doable in practice. For instance considering a patch size of 10×10 and an image of size 512×512 we have approximately 2.6×10^9 coefficients to store, *i.e.* 10.5GB in float precision. In the rest of the section we suppose that $\mathbf{t}$ and ω are fixed and we denote by $C_{\mathbf{t}}$ the covariance matrix associated to the restriction of $\Delta_f(\mathbf{t}, \cdot)$ to $\omega + (-\omega)$, Proposition 5. In Proposition 6 we propose a method to efficiently approximate the eigenvalues of $C_{\mathbf{t}}$ by using its specific structure. Indeed, as a covariance matrix, $C_{\mathbf{t}}$ is symmetric and positive and, since its associated Gaussian random field is stationary, it is block-Toeplitz with Toeplitz blocks, *i.e.* is block-diagonally constant and each block has constant diagonals. In the one-dimensional case these properties translate into symmetry, positivity and Toeplitz properties of the covariance matrix. Proposition 6 is stated in the one-dimensional case for simplicity sake but two-dimensional analogous can be derived.

We recall that the Frobenius norm of a matrix of size $n \times n$ is the L^2 norm of the associated vector of size n^2.

Proposition 6 (Eigenvalues approximation). *Let b be a function defined over $[\![-(n-1), n-1]\!]$ with $n \in \mathbb{N} \backslash \{0\}$. We define $T_b(j, \ell) = b(j - \ell)$ for $j, \ell \in [\![0, n-1]\!]$. The matrix T_b is a circulant matrix if and only if b is n-periodic. T_b is symmetric if and only if b is symmetric. Let b be symmetric, defining $\Pi(T_b)$ the projection of T_b onto the set of symmetric circulant matrix for the Frobenius product, we obtain that*

1. *the projection satisfies $\Pi(T_b) = T_c$ with $c(j) = \left(1 - \frac{j}{n}\right) b(j) + \frac{j}{n} b(n-j)$ for all $j \in [\![0, n-1]\!]$ and c is extended by n-periodicity to $\mathbb{Z}$;*
2. *the eigenvalues of $\Pi(T_b)$ are given by $\left(2\,\mathrm{Re}(\hat{d}(j)) - b(0)\right)_{j \in [\![0, n-1]\!]}$ with $d(j) = \left(1 - \frac{j}{n}\right) b(j)$, and $\hat{d}$ is the discrete Fourier transform over $[\![0, n-1]\!]$;*
3. *let $(\lambda_j)_{j \in [\![1, n]\!]}$ be the sorted eigenvalues of T_b and $(\tilde{\lambda}_j)_{j \in [\![1, n]\!]}$ the sorted eigenvalues of $\Pi(T_b)$ (in the same order). For any $j \in [\![1, n]\!]$, we have $|\lambda_j - \tilde{\lambda}_j| \leqslant \|T_b - \Pi(T_b)\|_{\mathrm{Fr}}$;*
4. *if T_b is positive-definitive then $\Pi(T_b)$ is positive-definite.*

Proof. (1) Let T_c be an element of the symmetric circulant matrices set. Minimizing $\|T_b - T_c\|_{\mathrm{Fr}}^2$ in $c(j)_{j \in [\![0, n-1]\!]}$ we get that $c(j)$ satisfies for any $j \in [\![0, n-1]\!]$

$$c(j) = \underset{s \in \mathbb{R}}{\operatorname{argmin}} \left(2(n-j)(s - b(j))^2 + 2j(s - b(n-j))^2\right) \ ,$$

which gives the result.

(2) Since $T_c = \Pi(T_b)$ is circulant, its eigenvalues are given by the discrete Fourier transform of c. We have that if $i \neq 0$ then $c(i) = \dot{d}(j) + \dot{d}(-j)$ with $d(j) = \left(1 - \frac{j}{n}\right) b(j)$ and $\dot{d}$ its extension to $\mathbb{Z}$ by n-periodicity. We also have $c(0) = b(0)$. We conclude the proof by taking the discrete Fourier transform of c.

(3) The demonstration of the Lipschitz property on the sorted eigenvalues of symmetric matrices with respect to the L^2 matricial norm can be found in [41]. We conclude using the fact that the L^2 matricial norm is upper-bounded by the Frobenius norm.

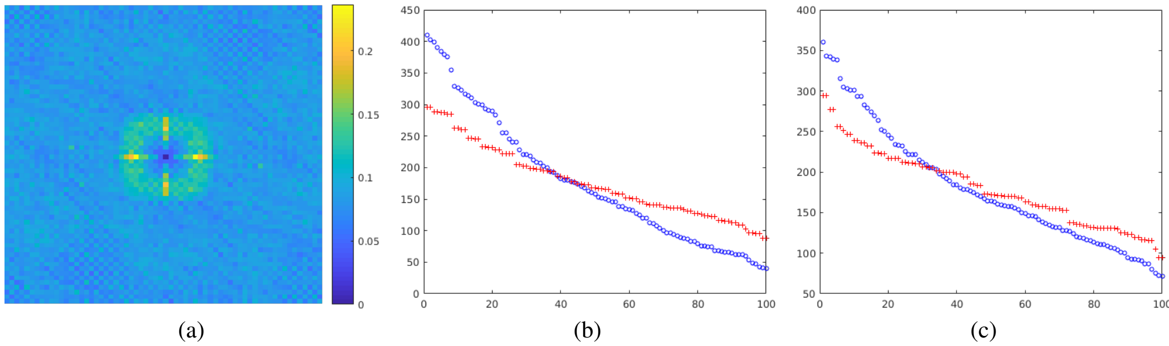


Figure 7: ***Eigenvalues approximation.*** We consider a Gaussian random field generated with $f * W$ with W a Gaussian white noise and f is a fixed realization of an independent Gaussian white noise over Ω. We consider patches of size 10×10 and study the approximation of the eigenvalues for the covariance matrix of the random field restricted to a domain of size 10×10, similarly to Proposition 5. (A) shows the Normalized Root-Mean Square Deviation between the eigenvalues computed with standard routines and the ones given by the approximation for each offset $\mathbf{t} \in \Omega$ with $\mathrm{NRMSD} = \frac{\left(\frac{1}{|\omega|}\sum_{k=0}^{|\omega|-1}|\tilde{\lambda}_k(\mathbf{t},\omega)-\lambda_k(\mathbf{t},\omega)|^2\right)^{1/2}}{\max(\lambda_k(\mathbf{t},\omega))_{k\in[\![0,|\omega|-1]\!]}-\min(\lambda_k(\mathbf{t},\omega))_{k\in[\![0,|\omega|-1]\!]}}$ with $\tilde{\lambda}_k(\mathbf{t},\omega)$ the two-dimensional approximation of the eigenvalues, for every possible offset in the image. Offset zero is at the center of the image. (B) and (C) illustrate the properties of Proposition 6. Blue circles correspond to the 100 eigenvalues computed with MATLAB routine for offset $(5,5)$ in (B), respectively $(10,20)$ in (C), and red crosses correspond to the 100 approximated eigenvalues for the same offsets. Note that a standard routine takes $273s$ for 10×10 patches on 256×256 images whereas it only takes $1.11s$ when approximating the eigenvalues using the discrete Fourier transform.

(4) This result is a special case of the spectrum contraction property of the projection proved in Theorem 2 of [42]. $\square$

In Figure 7 we display the behavior of the projection for the eigenvalues in the two-dimensional case. Computing the eigenvalues of the projection is done via Fast Fourier Transform (FFT) which is faster than standard routines, MATLABR2017a function `eig` for instance, for computing eigenvalues of Toeplitz matrices. The major cons of using such approximation is that it may not be valid for small offsets $\mathbf{t} \in \Omega$ as shown in Figure 7.

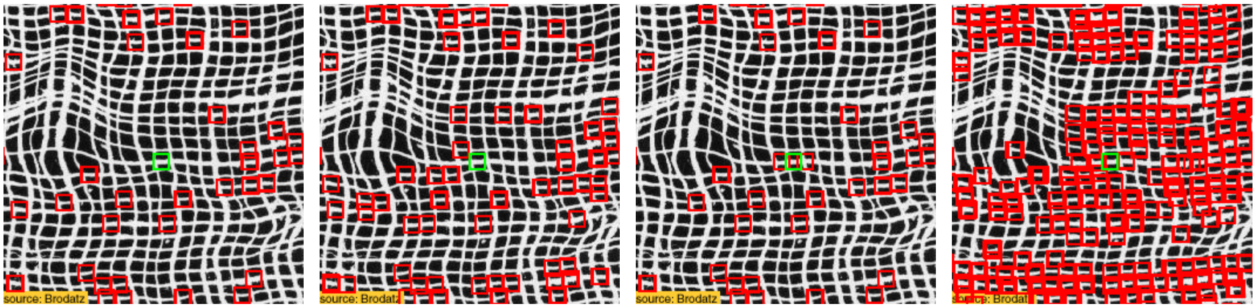
Suppose the approximation of the eigenvalues is valid, we need an efficient algorithm to compute the distribution of the associated positive-weighted sum of chi-square random variables in Equation (6). Exact computation is given in [43] but requires to compute heavy integrals. This exact method, named Imhof method in the following, will be used as a baseline for other algorithms. Numerous methods such as differential equations [44], series truncation [45], negative binomial mixtures [46] approaches were later introduced but all require stopping criteria such as truncation criteria which can be hard to set. We focus on cumulant methods which generalize and refine the Gaussian approximations used in Section 3. These methods rely on computing moments of the original distribution and then fitting a known probability distribution function to the objective distribution using these moments. Bodenham et al. in [47] show that the following methods can be efficiently computed:

- Gaussian approximation (discarded due to its poor results for small patches as illustrated in Section 3) ;
- Hall-Buckley-Eagleson [48, 49] (HBE), (three moments fitted Gamma distribution) ;
- Wood F [50] (three moments fitted Fischer-Snedecor distribution).

Other methods such as the Lindsay-Pilla-Basak-4, which relies on the computation of eight moments, are slower than HBE by a factor 350 at least, see [47], and thus are discarded. In Figure 8 we investigate the trade-off between computational speed and accuracy of these methods for the task of detection.

The experiments conducted in Figure 8 show that the HBE approximation does not give good results when evaluating the probability of rare events. This was already noticed by Bodenham et al. in [47] who stated that "Hall–Buckley–Eagleson

[0] $\mathrm{NRMSD} = \frac{\left(\frac{1}{|\omega|}\sum_{k=0}^{|\omega|-1}|\tilde{\lambda}_k(\mathbf{t},\omega)-\lambda_k(\mathbf{t},\omega)|^2\right)^{1/2}}{\max(\lambda_k(\mathbf{t},\omega))_{k\in[\![0,|\omega|-1]\!]}-\min(\lambda_k(\mathbf{t},\omega))_{k\in[\![0,|\omega|-1]\!]}}$ with $\tilde{\lambda}_k(\mathbf{t},\omega)$ the two-dimensional approximation of the eigenvalues, for every possible offset in the image.



(a) $1513s, n_d = 52$ (b) $200s, n_d = 116$ (c) $4.77s, n_d = 50$ (d) $4.64s, n_d = 671$

Figure 8: ***Similarity detection.*** In this figure we illustrate the accuracy of the different proposed approximations of the cumulative distribution function of $\mathcal{AS}_{2,2}(U, \mathbf{t}, \omega)$. We say that an offset $\mathbf{t}$ is *detected* in an image if $\mathcal{AS}_{2,2}(u, \mathbf{t}, \omega) \leqslant a(\mathbf{t})$ for some threshold $a(\mathbf{t}) \in \mathbb{R}$. In every image, in green we display the patch domain ω (in the center of the image) and in red we display the shifted patch domain for detected offsets with function $a(\mathbf{t})$ such that for any $\mathbf{t} \in \Omega$, $\mathbb{P}\left[\mathcal{AS}_{2,2}(U, \mathbf{t}, \omega) \leqslant a(\mathbf{t})\right] = 1/256^2$, where U is given by the Gaussian random field $f * W$ where f is the original image of fabric and W is a Gaussian white noise over $\Omega = 256 \times 256$. Approximations of the cumulative distribution function of $\mathcal{AS}_{2,2}(U, \mathbf{t}, \omega)$ lead to approximations of $a(\mathbf{t})$. The most precise approximation is given in (A) where the eigenvalues are computed using a MATLAB routine and the cumulative distribution function is given by the Imhof method. In (B) we approximate the eigenvalues using the projection described in Proposition 6. It yields twice as many detections. In (C) Wood F method is used instead of Imhof's yielding less detections but performing seven times faster. Interestingly errors seem to compensate and the obtained result with Wood F method is very close to the results obtained with the baseline algorithm in (A). In (D) HBE method is used instead of Imhof's, in this case we obtain too many detections, *i.e.* the approximation of the cumulative distribution function is not valid.

method is recommended for most practitioners [...]. However, [...], for very small probability values, either the Wood F or the Lindsay–Pilla–Basak method should be used".

4.2 Continuous periodic case

To conclude we show that a similar non-asymptotic study can be conducted in continuous settings.

Proposition 7 (Squared L^2 continuous auto-similarity function exact probability distribution function)**.** *Let $\Omega = \mathbb{T}^2$, $\omega \subset \Omega$ and let U be a zero-mean Gaussian random field on Ω with covariance function Γ. Assume (A2), then the following equality holds for any $\mathbf{t} \in \Omega$ up to a change of the underlying probability space*

$$\mathcal{AS}_{2,2}(U, \mathbf{t}, \omega) \underset{a.s}{=} \sum_{k \in \mathbb{N}} \lambda_k(\mathbf{t}, \omega) Z_k,$$

with Z_k independent chi-square random variables with parameter 1 and $\lambda_k(\mathbf{t}, \omega)$ the eigenvalues of the kernel $C_{\mathbf{t}}$ associated with function $\Delta(\mathbf{t}, \cdot) = 2\Gamma(\mathbf{t}) - \Gamma(\cdot + \mathbf{t}) - \Gamma(\cdot - \mathbf{t})$ restricted to ω, i.e. for any $\mathbf{x_1}, \mathbf{x_2} \in \omega$, $C_{\mathbf{t}}(\mathbf{x_1}, \mathbf{x_2}) = \Delta(\mathbf{t}, \mathbf{x_1} - \mathbf{x_2})$.

Proof. We consider the stationary Gaussian random field $P_\omega(V_{\mathbf{t}})$ over ω defined by the restriction to ω of $V_{\mathbf{t}}$ defined for any $\mathbf{x} \in \Omega$ by $V_{\mathbf{t}}(\mathbf{x}) = U(\mathbf{x}) - U(\mathbf{x} + \mathbf{t})$. The Karhunen-Loeve theorem [51] ensures the existence of $(\lambda_k(\mathbf{t}, \omega))_{k \in \mathbb{N}} \in \mathbb{R}_+^{\mathbb{N}}$, $(X_k)_{k \in \mathbb{N}}$ a sequence of independent normal Gaussian random variables and $(e_k)_{k \in \mathbb{N}}$ a sequence of orthonormal function over $L^2(\omega)$ such that

$$\lim_{n \to +\infty} \sup_{\mathbf{x} \in \omega} \mathbb{E}\left[|P_\omega(V_{\mathbf{t}})(\mathbf{x}) - \sum_{k=0}^{n} \sqrt{\lambda_k(\mathbf{t}, \omega)} e_k(\mathbf{x}) X_k|^2 \right] = 0 \ , \tag{7}$$

We define the sequence $(I_n)_{n\in\mathbb{N}} = \left(\int_\omega \left(\sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k\right)^2 \mathrm{d}\mathbf{x}\right)_{n\in\mathbb{N}}$. We have, using the Cauchy-Schwarz inequality on $L^2(\mathcal{A}\times\omega)$ and (7)

$$
\begin{aligned}
\mathbb{E}\left[|\mathcal{AS}_{2,2}(U,\mathbf{t},\omega) - I_n|\right] &\leqslant \mathbb{E}\left[\int_\omega |P_\omega(V_\mathbf{t})^2(\mathbf{x}) - \left(\sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k\right)^2 |\mathrm{d}\mathbf{x}\right] \\
&\leqslant \mathbb{E}\left[\int_\omega (P_\omega(V_\mathbf{t})(\mathbf{x}) - \sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k)^2 \mathrm{d}\mathbf{x}\right]^{1/2} \mathbb{E}\left[\int_\omega (P_\omega(V_\mathbf{t})(\mathbf{x}) + \sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k)^2 \mathrm{d}\mathbf{x}\right]^{1/2} \\
&\leqslant 2\mathbb{E}\left[\mathcal{AS}_{2,2}(U,\mathbf{t},\omega)\right]^{1/2} \int_\omega \mathbb{E}\left[(P_\omega(V_\mathbf{t})(\mathbf{x}) - \sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k)^2\right] \mathrm{d}\mathbf{x} \;, \qquad (8)
\end{aligned}
$$

where we used the Fubini theorem in the last inequality. Using the dominated convergence theorem in (8) with integral domination given by $\sup_{n\in\mathbb{N}} \sup_{\mathbf{x}\in\omega} \mathbb{E}\left[(P_\omega(V_\mathbf{t})(\mathbf{x}) - \sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k)^2\right]$ we conclude that $(I_n)_{n\in\mathbb{N}}$ converges to $\mathcal{AS}_{2,2}(U,\mathbf{t},\omega)$ in $L^1(\mathcal{A})$. Thus there exists a subsequence of $(I_n)_{n\in\mathbb{N}}$ which converges almost surely to $\mathcal{AS}_{2,2}(U,\mathbf{t},\omega)$. We also have $I_n = \int_\omega \left(\sum_{k=0}^n \sqrt{\lambda_k(\mathbf{t},\omega)} e_k(\mathbf{x}) X_k\right)^2 \mathrm{d}\mathbf{x} = \sum_{k=0}^n \lambda_k(\omega,k) X_k^2$ by orthonormality and thus the sequence $(I_n)_{n\in\mathbb{N}}$ is almost surely non-decreasing. We get that $(I_n)_{n\in\mathbb{N}}$ converges almost surely to $\mathcal{AS}_{2,2}(U,\mathbf{t},\omega)$ which can be rewritten as

$$
\mathcal{AS}_{2,2}(U,\mathbf{t},\omega) = \sum_{k\in\mathbb{Z}} \lambda_k(\mathbf{t},\omega) X_k^2 \quad \text{almost surely.}
$$

The characterization of $(\lambda_k(\mathbf{t},\omega), e_k(\mathbf{x}))$ is given by the Karhunen-Loeve theorem and $e_k(\mathbf{x})$ is solution of the following Fredholm equation for all $\mathbf{x}\in\omega$

$$
\int_\omega \Delta(\mathbf{t},\mathbf{x}-\mathbf{y}) e_k(\mathbf{y})\,\mathrm{d}\mathbf{y} = \lambda_k(\mathbf{t},\omega) e_k(\mathbf{x}) \;.
$$

Setting $Z_k = X_k^2$ concludes the proof. □

Note that if $\omega = \mathbb{T}^2$ then the solution of the Fredholm equation is given by the Fourier series of Γ.

Appendices

A Asymptotic theorems – discrete case

The following theorem is a two-dimensional law of large numbers with weak dependence assumptions. It is a slight modification of Corollary 4.1 (ii) in [38].

Theorem 5. *Let $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}}$ be two positive increasing integer sequences and $(\omega_k)_{k\in\mathbb{N}}$ be the sequence of subsets such that for any $k\in\mathbb{N}$, $\omega_k = [\![0, m_k]\!] \times [\![0, n_k]\!]$. Let V be a R-independent, with $R \geqslant 0$, random field over $\mathbb{Z}^2$ such that $|V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})]|$ is uniformly stochastically dominated by $\widehat{V}$, a second-order stationary random field over $\mathbb{Z}^2$. Then V is a second-order random field. In addition, assume that there exists $\mu\in\mathbb{R}$ such that $\lim_{k\to+\infty} |\omega_k|^{-1} \sum_{\mathbf{x}\in\omega_k} \mathbb{E}[V(\mathbf{x})] = \mu$. Then it holds that*

$$
\lim_{k\to+\infty} |\omega_k|^{-1} \sum_{\mathbf{x}\in\omega_k} V(\mathbf{x}) \underset{a.s}{=} \mu \;. \qquad (9)
$$

Proof. We suppose that for any $\mathbf{x}\in\mathbb{Z}^2$, $\mathbb{E}[V(\mathbf{x})] = 0$, otherwise we replace $V(\mathbf{x})$ by $V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})]$. In order to apply Corollary 4.1 (ii) in [38] we must check that:

(a) V is $R-$independent ;

(b) $|V|$ is uniformly stochastically dominated by a random field $\widetilde{V}$ and there exists $r \in [1, 2[$ such that for any $\mathbf{x} \in \mathbb{Z}^2$, $\mathbb{E}\left[\widetilde{V}^r(\mathbf{x}) \log^+(\widetilde{V}(\mathbf{x}))\right]$ is finite.

Item (a) is given in the statement of Theorem 5 and $|V|$ is uniformly stochastically dominated by the random field $\widetilde{V}_0$ defined for any $\mathbf{x} \in \mathbb{Z}^2$ by $\widetilde{V}_0(\mathbf{x}) = \widetilde{V}(\mathbf{0})$. Since $\mathbb{E}\left[\widetilde{V}(\mathbf{0})^2\right]$ is finite so is $\mathbb{E}\left[\widetilde{V}(\mathbf{0}) \log^+(\widetilde{V}(\mathbf{0}))\right]$ which implies (b). Then it holds that

$$\lim_{k \to +\infty} \sum_{\mathbf{x} \in \omega_k} (V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})]) \underset{a.s}{=} 0 \, .$$

Using that $\lim_{k \to +\infty} |\omega_k|^{-1} \sum\limits_{\mathbf{x} \in \omega_k} \mathbb{E}[U(\mathbf{x})] = \mu$ we conclude the proof. ☐

We now turn to an extension of the central limit theorem to two-dimensional random fields with weak dependence assumptions. This result is a consequence of [52, Theorem 2].

Theorem 6. *Under the hypotheses of Theorem 5 and assuming that there exist $\mu \in \mathbb{R}$ and $\sigma \geqslant 0$ such that*

(a) $\lim_{k \to +\infty} |\omega_k|^{-1/2} \sum_{\mathbf{x} \in \omega_k} (\mathbb{E}[V](\mathbf{x}) - \mu) = 0$;

(b) $\lim_{k \to +\infty} |\omega_k|^{-1} \sum_{\mathbf{x}, \mathbf{y} \in \omega_k} \mathrm{Cov}[V(\mathbf{x}), V(\mathbf{y})] = \sigma^2$.

Then it holds that

$$\lim_{k \to +\infty} |\omega_k|^{-1/2} \sum_{\mathbf{x} \in \omega_k} (V(\mathbf{x}) - \mu) \underset{\mathcal{L}}{=} \mathcal{N}(0, \sigma) \, , \tag{10}$$

with the convention that $\mathcal{N}(0, 0) = \delta_0$, the Dirac distribution at 0.

Proof. Using the same notations as in [52, Theorem 2] we first note that $\sigma_k^2 = |\omega_k|^{-1} \, \mathrm{Var}\left[\sum_{\mathbf{x} \in \omega_k} V(\mathbf{x})\right]$. Since V is R-independent each vertex of the dependency graph of V has its degree bounded by $(2R+1)^2$. Thus in order to apply [52, Theorem 2] we need to find a positive integer m such that

$$\lim_{k \to +\infty} |\omega_k|^{1/m} A_k / (|\omega_k|^{1/2} \sigma_k) = 0 \, ,$$

with A_k such that $\lim_{k \to +\infty} \sum_{\mathbf{x} \in \omega_k} \mathbb{E}\left[V(\mathbf{x})^2 \mathbb{1}_{|V(\mathbf{x})| > A_k}\right] / (|\omega_k| \sigma_k^2) = 0$. Since σ_k^2 is supposed to converge using hypothesis (b) the conditions reduce to find m and A_k such that

$$\lim_{k \to +\infty} |\omega_k|^{1/m - 1/2} A_k = 0 \, , \qquad \lim_{k \to +\infty} \sum_{\mathbf{x} \in \omega_k} \mathbb{E}\left[V(\mathbf{x})^2 \mathbb{1}_{|V(\mathbf{x})| > A_k}\right] / |\omega_k| = 0 \, .$$

Since $|V|$ is uniformly stochastically dominated almost surely by $\widetilde{V}$, a second-order stationary random field, we obtain the following stronger condition on A_k

$$\lim_{k \to +\infty} \frac{1}{|\omega_k|} \sum_{\mathbf{x} \in \omega_k} \mathbb{E}\left[V(\mathbf{x})^2 \mathbb{1}_{|V(\mathbf{x})| > A_k}\right] \leqslant \lim_{k \to +\infty} \mathbb{E}\left[\widetilde{V}(\mathbf{0})^2 \mathbb{1}_{\widetilde{V}(\mathbf{0}) > A_k}\right] = 0 \, .$$

Using the dominated convergence theorem this condition is satisfied for any A_k which tends to infinity. Thus setting, for example, $m = 4$ and $A_k = |\omega_k|^{1/8}$ we get using [52, Theorem 2]

$$\lim_{k \to +\infty} |\omega_k|^{-1/2} \left(\frac{\sum_{\mathbf{x} \in \omega_k} (V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})])}{\sigma_k} \right) \underset{\mathcal{L}}{=} \mathcal{N}(0, 1).$$

Since $\lim_{k \to +\infty} \sigma_k = \sigma$ by (b) we get that

$$\lim_{k \to +\infty} |\omega_k|^{-1/2} \sum_{\mathbf{x} \in \omega_k} (V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})]) \underset{\mathcal{L}}{=} \lim_{k \to +\infty} \mathcal{N}(0, \sigma_k) \, .$$

If $\lim_{k \to +\infty} \sigma_k = \sigma > 0$ then $\lim_{k \to +\infty} \mathcal{N}(0, \sigma_k) = \mathcal{N}(0, \sigma)$. If $\lim_{k \to +\infty} \sigma_k = 0$ then $\lim_{k \to +\infty} \mathcal{N}(0, \sigma_k) = \delta_0$. Using (a) we obtain that

$$\lim_{k \to +\infty} |\omega_k|^{-1/2} \sum_{\mathbf{x} \in \omega_k} (V(\mathbf{x}) - \mu) \underset{\mathcal{L}}{=} \mathcal{N}(0, \sigma) \, ,$$

with $\mathcal{N}(0, 0) = \delta_0$ if $\sigma = 0$. ☐

Lemma 2 explicits a class of Gaussian random fields over $\mathbb{Z}^2$ such that the R-independence property holds for some $R \geqslant 0$.

Lemma 2. *Let $f \in \mathbb{R}^{\mathbb{Z}^2}$ with finite support $\mathrm{Supp}\,(f) \subset [\![-r, r]\!]^2$, where $r \in \mathbb{N}$. Let W be a Gaussian white noise over $\mathbb{Z}^2$ and $V = f * W$ then V is a R-independent second-order random field with $R = 2r$.*

Proof. V is a Gaussian random field such that for any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^2$

$$\mathbb{E}\left[V(\mathbf{x})\right] = 0 \;, \qquad \mathrm{Cov}\left[V(\mathbf{x})V(\mathbf{y})\right] = \sum_{\mathbf{x}',\mathbf{y}' \in \mathbb{Z}^2} f(\mathbf{x}-\mathbf{x}')f(\mathbf{y}-\mathbf{y}')\,\mathrm{Cov}\left[W(\mathbf{x}'), W(\mathbf{y}')\right] = \Gamma_f(\mathbf{x}-\mathbf{y}) \;. \tag{11}$$

Note that since $\mathrm{Supp}\,(f) \subset [\![-r, r]\!]$ we have $\mathrm{Supp}\,(\Gamma_f) \subset [\![-R, R]\!]$ with $R = 2r$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^2$ such that $\|\mathbf{x}-\mathbf{y}\|_\infty > R$, using (11), we obtain

$$\mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{y})\right] = \Gamma_f(\mathbf{x}-\mathbf{y}) = 0 \;. \tag{12}$$

Let $K_1, K_2 \subset \mathbb{Z}^2$ two finite sets with $d(K_1, K_2)_\infty > R$ and consider $V|_{K_i}$ the restriction of V to K_i for $i = \{1, 2\}$. Using (12), we get that for any $\mathbf{x} \in K_1$, $\mathbf{y} \in K_2$ we have

$$\mathrm{Cov}\left[V|_{K_1}(\mathbf{x}), V|_{K_2}(\mathbf{y})\right] = 0 \;.$$

As a consequence, $\mathrm{Cov}\left[V|_{K_1}, V|_{K_2}\right] = 0$ and $V|_{K_1}$ and $V|_{K_2}$ are uncorrelated. Since $V|_{K_1}, V|_{K_2}$ are Gaussian random fields we get that $V|_{K_1}, V|_{K_2}$ are R-independent. □

Lemma 3 give specific conditions on random fields in order for Theorems 5 and 6 to hold.

Lemma 3. *Let $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}}$ be two positive increasing integer sequences and $(\omega_k)_{k\in\mathbb{N}}$ be the sequence of subsets given by for any $k \in \mathbb{N}$, $\omega_k = [\![0, m_k]\!] \times [\![0, n_k]\!]$. Let V be a R-independent, with $R \geqslant 0$, second-order stationary random field over $\mathbb{Z}^2$. Then for all $k \in \mathbb{N}$*

- *(a)* $|\omega_k|^{-1} \sum_{\mathbf{x}\in\omega_k} \mathbb{E}\left[V(\mathbf{x})\right] = \mathbb{E}\left[V(\mathbf{0})\right]$;
- *(b)* $\lim_{k\to+\infty} |\omega_k|^{-1} \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{y})\right] = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{0})\right]$.

In addition, Equations (9) *and* (10) *hold with $\mu = \mathbb{E}\left[V(\mathbf{0})\right]$ and $\sigma = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{0})\right]$ which is finite.*

Proof. Item (a) is immediate by stationarity. Concerning (b), for any $k \in \mathbb{N}$ we have by stationarity

$$|\omega_k|^{-1} \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{y})\right] = |\omega_k|^{-1} \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \mathrm{Cov}\left[V(\mathbf{x}-\mathbf{y}), V(\mathbf{0})\right] = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{0})\right] g_k(\mathbf{x}) \;,$$

where $g_k \in \mathbb{R}^{\mathbb{Z}^2}$ satisfies for any $\mathbf{x} \in \mathbb{Z}^2$, $g_k(\mathbf{x}) = |\omega_k|^{-1}\mathbb{1}_{\omega_k} * \check{\mathbb{1}}_{\omega_k}(\mathbf{x})$. For any $k \in \mathbb{N}$, $\mathbf{x} \in \mathbb{Z}^2$ we have $0 \leqslant g_k(\mathbf{x}) \leqslant 1$ and $\lim_{k\to+\infty} g_k(\mathbf{x}) = 1$. For any $\mathbf{x} \in \mathbb{Z}^2$ such that $\|\mathbf{x}\|_\infty > R$, $\mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{0})\right] = 0$ and then $\sum_{\mathbf{x}\in\mathbb{Z}^2} |\mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{0})\right]| < +\infty$. Using the dominated convergence theorem we get that

$$\lim_{k\to+\infty} |\omega_k|^{-1} \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{y})\right] = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}\left[V(\mathbf{x}), V(\mathbf{0})\right] \;,$$

We obtain Equations (9) and (10) by applying Theorems 5 and 6. □

Lemma 4. *Let $f \in \mathbb{R}^{\mathbb{Z}^2}$, $f \neq 0$, a function with finite support. Then it holds for any $\mathbf{t} \in \mathbb{Z}^2$, $\Gamma_f(\mathbf{t}) \leqslant \Gamma_f(\mathbf{0})$, with equality if and only if $\mathbf{t} = \mathbf{0}$.*

Proof. For any $\mathbf{t} \in \mathbb{Z}^2$, let $\tau_{\mathbf{t}} f = f(\cdot + \mathbf{t})$. By the definition of the autocorrelation Γ_f and using the Cauchy-Schwarz inequality we get that for any $\mathbf{t} \in \mathbb{Z}^2$

$$\Gamma_f(\mathbf{t}) = \langle \tau_{\mathbf{t}} f, f \rangle \leqslant \|f\|_2^2 \leqslant \Gamma_f(\mathbf{0}) \;,$$

with equality if and only if $f = \alpha\tau_{\mathbf{t}} f$, with $\alpha \neq 0$ since $f \neq 0$. This implies that $\mathrm{Supp}\,(\tau_{\mathbf{t}}(f)) = \mathrm{Supp}\,(f)$. As a consequence $\mathbf{t} = \mathbf{0}$, which concludes the proof. □

The following lemma ensures that items (a) and (b) in Theorem 6 are satisfied in the template similarity case when imposing summability conditions over v.

Lemma 5. *Under the hypotheses of Theorem 1, assuming (A4) with $\ell \in \mathbb{N}$ and $p = 2\ell$. There exist $\mu_{p,p} > 0$ and $\sigma_{p,p} \geqslant 0$ such that for any $\mathbf{t} \in \Omega$*

(a) $\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{|\omega_k|}\mathbb{E}\left[\mathcal{TS}_{p,p}(U,v,\mathbf{t},\omega_k)\right] - \mu_{p,p}(\mathbf{t})\right) = 0$;

(b) $\lim_{k\to+\infty} \frac{1}{|\omega_k|} \operatorname{Var}\left[\mathcal{TS}_{p,p}(U,v,\mathbf{t},\omega_k)\right] = \sigma_{p,p}^2(\mathbf{t})$.

Proof. (a) For any $k \in \mathbb{N}$ we have that

$$\begin{aligned}
\mathbb{E}\left[\mathcal{TS}_{p,p}(U,v,\mathbf{t},\omega_k)\right] &= \sum_{\mathbf{x}\in\omega_k} \mathbb{E}\left[(v(\mathbf{x}) - U(\mathbf{x}+\mathbf{t}))^{2\ell}\right] \\
&= \sum_{j=0}^{2\ell} \binom{2\ell}{j} \sum_{\mathbf{x}\in\omega_k} (-1)^j v(\mathbf{x})^j \mathbb{E}\left[U(\mathbf{x})^{2\ell-j}\right] \\
&= \sum_{j=0}^{\ell} \binom{2\ell}{2j} \sum_{\mathbf{x}\in\omega_k} v(\mathbf{x})^{2j} \mathbb{E}\left[U(\mathbf{x})^{2(\ell-j)}\right] = \sum_{j=0}^{\ell} \binom{2\ell}{2j} \mathbb{E}\left[U(\mathbf{0})\right]^{2(\ell-j)} \sum_{\mathbf{x}\in\omega_k} v(\mathbf{x})^{2j} .
\end{aligned}$$

Let $\mu_{p,p} = \sum_{j=0}^{\ell} \binom{2\ell}{2j} \mathbb{E}\left[U(\mathbf{0})\right]^{2(\ell-j)} \beta_j$ and using (a) of (A4) we get that

$$\lim_{k\to+\infty} |\omega_k|^{\frac{1}{2}} \left(\frac{1}{|\omega_k|}\mathbb{E}\left[\mathcal{TS}_{p,p}(U,v,\mathbf{t},\omega_k)\right] - \mu_{p,p}(\mathbf{t})\right) = 0 .$$

Now since $\mu_{p,p} \geqslant \mathbb{E}\left[U(\mathbf{0})^{2\ell}\right] \geqslant \mathbb{E}\left[U(\mathbf{0})^2\right]^{\ell} \geqslant \Gamma_f(\mathbf{0}) > 0$ we have that $\mu_{p,p} > 0$.

(b) For any $k \in \mathbb{N}$ we have that

$$\begin{aligned}
\operatorname{Var}\left[\mathcal{TS}_{p,p}(U,v,\mathbf{t},\omega_k)\right] &= \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}\left[(U(\mathbf{x}) - v(\mathbf{x}))^{2\ell}, (U(\mathbf{y}) - v(\mathbf{y}))^{2\ell}\right] \\
&= \sum_{\mathbf{x},\mathbf{y}\in\omega_k} \sum_{i,j=0}^{\ell} \binom{2\ell}{2i}\binom{2\ell}{2j} v(\mathbf{x})^{2i} v(\mathbf{y})^{2j} \operatorname{Cov}\left[U(\mathbf{x})^{2(\ell-i)}, U(\mathbf{y})^{2(\ell-j)}\right] \\
&= \sum_{\mathbf{x},\mathbf{y}\in\mathbb{Z}^2} \sum_{i,j=0}^{\ell} \binom{2\ell}{2i}\binom{2\ell}{2j} v_k(\mathbf{x})^{2i} v_k(\mathbf{x}+\mathbf{y})^{2j} \operatorname{Cov}\left[U(\mathbf{y})^{2(\ell-i)}, U(\mathbf{0})^{2(\ell-j)}\right] \\
&= \sum_{i,j=0}^{\ell} \binom{2\ell}{2i}\binom{2\ell}{2j} \left\langle v_k^{2i} * \check{v}_k^{2j}, \operatorname{Cov}\left[U(\cdot)^{2(\ell-i)}, U(\mathbf{0})^{2(\ell-j)}\right]\right\rangle .
\end{aligned}$$

Let $\sigma_{p,p} = \sum_{i,j=0}^{\ell} \binom{2\ell}{2i}\binom{2\ell}{2j} \left\langle \gamma_{i,j}, \operatorname{Cov}\left[U(\cdot)^{2(\ell-i)}, U(\mathbf{0})^{2(\ell-j)}\right]\right\rangle$. Using (b) in (A4) we can conclude.

□

Note that this lemma is also valid in the continuous case.

B Asymptotic theorems – continuous case

We now turn to the the continuous setting. We start by stating the continuous counterparts of Theorems 5 and 6. The following theorem, given here for completeness, can be found with different assumptions (in the one-dimensional case) in [33].

Theorem 7. *Let $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}}$ be two positive increasing integer sequences and $(\omega_k)_{k\in\mathbb{N}}$ be the sequence of subsets given by for any $k \in \mathbb{N}$, $\omega_k = [0, m_k] \times [0, n_k]$. Let V be a R-independent, with $R \geqslant 0$, random field over $\mathbb{R}^2$ such that $|V(\mathbf{x}) - \mathbb{E}\left[V(\mathbf{x})\right]|$ is uniformly stochastically dominated by $\widetilde{V}$, a second-order stationary random field over $\mathbb{R}^2$. Then V is a second-order random field. In addition, assume V is sample path continuous and that there exists $\mu \in \mathbb{R}$ given by $\lim_{k\to+\infty} |\omega_k|^{-1} \int_{\mathbf{x}\in\omega_k} \mathbb{E}\left[V(\mathbf{x})\right] \mathrm{d}\mathbf{x} = \mu$. Then it holds that*

$$\lim_{k\to+\infty} |\omega_k|^{-1} \int_{\mathbf{x}\in\omega_k} V(\mathbf{x}) \mathrm{d}\mathbf{x} \underset{a.s.}{=} \mu . \tag{13}$$

Proof. Without loss of generality we can suppose that for any $\mathbf{x} \in \Omega$, $\mathbb{E}[V(\mathbf{x})] = 0$. Let $(\sigma_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ such that for any $k \in \mathbb{N}$ we have

$$\sigma_k^2 = \mathbb{E}\left[\left(k^{-2}\int_{\Omega_k} V(\mathbf{x})\mathrm{d}\mathbf{x}\right)^2\right] , \tag{14}$$

with $\Omega_k = [0,k]^2$. Since V is R-independent, for any $\mathbf{x}, \mathbf{y} \in \Omega$ such that $\|\mathbf{x}-\mathbf{y}\|_\infty > R$, we have $C(\mathbf{x},\mathbf{y}) = 0$. Hence for k large enough we obtain

$$\int_{\Omega_k}\int_{\Omega_k} C(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \leqslant \int_{\mathbf{x}\in\Omega_k}\int_{\|\mathbf{y}\|_\infty \leqslant R} |C(\mathbf{x},\mathbf{x}+\mathbf{y})|\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{x} \leqslant k^2|\bar{\mathrm{B}}_\infty(0,R)| \sup_{\Omega_k\times\bar{\mathrm{B}}_\infty(0,R)} |C(\mathbf{x},\mathbf{x}+\mathbf{y})| . \tag{15}$$

Using that $\widetilde{V}$ uniformly stochastically dominates $|V|$, the stationarity of $\widetilde{V}$, and the Cauchy-Schwarz inequality, we obtain for any $\mathbf{x}, \mathbf{y} \in \Omega$,

$$|C(\mathbf{x},\mathbf{x}+\mathbf{y})| = |\mathbb{E}[V(\mathbf{x})V(\mathbf{x}+\mathbf{y})]| \leqslant \mathbb{E}\left[\widetilde{V}^2(\mathbf{x})\right]^{1/2} \mathbb{E}\left[\widetilde{V}^2(\mathbf{x}+\mathbf{y})\right]^{1/2} \leqslant \mathbb{E}\left[\widetilde{V}^2(\mathbf{0})\right] . \tag{16}$$

Combining (14), (15) and (16) we get that for any $k \in \mathbb{N}$

$$\sigma_k^2 \leqslant Mk^{-2} ,$$

with $M = |\bar{\mathrm{B}}_\infty(0,R)|\mathbb{E}\left[\widetilde{V}^2(\mathbf{0})\right]$. Thus the series $\sum_{k\in\mathbb{N}} \sigma_k^2$ converges and $\sum_{k\in\mathbb{N}}\left(k^{-2}\int_{\Omega_k} V(\mathbf{x})\mathrm{d}\mathbf{x}\right)^2$ is finite almost surely. This proves that $\lim_{k\to+\infty} k^{-2}\int_{\Omega_k} V(\mathbf{x})\mathrm{d}\mathbf{x} = 0$ almost surely. In order to conclude for the general case it remains to prove that

$$\lim_{k\to+\infty} \sup_{\Omega_k\subset\omega\subset\Omega_{k+1}} \left| |\omega|^{-1}\int_\omega V(\mathbf{x})\mathrm{d}\mathbf{x} - |\omega_k|^{-1}\int_{\Omega_k} V(\mathbf{x})\mathrm{d}\mathbf{x}\right| \underset{a.s.}{=} 0 .$$

This is proved in the one-dimensional case in [53] and extends to the two-dimensional situation. $\square$

The following theorem is an application of [54, Theorem 1.7.1].

Theorem 8. *Under the hypotheses of Theorem 7 and assuming that there exist $\mu \in \mathbb{R}$ and $\sigma \geqslant 0$ such that*

(a) $\lim_{k\to+\infty} |\omega_k|^{-1/2}\int_{\mathbf{x}\in\omega_k} (\mathbb{E}[V](\mathbf{x}) - \mu)\,\mathrm{d}\mathbf{x} = 0$;

(b) $\lim_{k\to+\infty} |\omega_k|^{-1}\int_{\mathbf{x},\mathbf{y}\in\omega_k} \mathrm{Cov}[V(\mathbf{x}), V(\mathbf{y})]\,\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} = \sigma^2$.

Then it holds that

$$\lim_{k\to+\infty} |\omega_k|^{-1/2}\int_{\mathbf{x}\in\omega_k} (V(\mathbf{x}) - \mu)\,\mathrm{d}\mathbf{x} \underset{\mathcal{L}}{=} \mathcal{N}(0,\sigma) . \tag{17}$$

Proof. Upon considering for any $\mathbf{x} \in \mathbb{R}^2$, $V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})]$ instead of $V(\mathbf{x})$ we can suppose that for any $\mathbf{x} \in \mathbb{R}^2$, $\mathbb{E}[V(\mathbf{x})] = 0$. In [54, Theorem 1.7.1] setting the weight function to be $\mathbb{1}_{\omega_k}/|\omega_k|^{1/2}$, condition V is immediately fulfilled. The following equality implies that Condition IV is satisfied

$$\lim_{k\to+\infty} \mathbb{E}\left[\left(\int_{\omega_k} V(\mathbf{x})\mathrm{d}\mathbf{x}\right)^2\right]/|\omega_k| = \lim_{k\to+\infty}\int_{\mathbf{x},\mathbf{y}\in\omega_k} C(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}/|\omega_k| = \sigma^2 ,$$

where the first equality is obtained using the Fubini-Lebesgue theorem. Indeed since $|V|$ is uniformly stochastically dominated by $\widetilde{V}$ we have by the Fubini-Tonelli theorem

$$\mathbb{E}\left[\int_{\omega_k}\int_{\omega_k} |V(\mathbf{x})V(\mathbf{y})|\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\right] = \int_{\omega_k}\int_{\omega_k} \mathbb{E}[|V(\mathbf{x})V(\mathbf{y})|]\,\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \leqslant |\omega_k|^2\mathbb{E}\left[\widetilde{V}(\mathbf{0})^2\right] < +\infty .$$

Condition VI is ensured because V is a second-order random field and is R-independent. We get that

$$\lim_{k\to+\infty} |\omega_k|^{1/2}\left(\int_{\omega_k} (V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})])\mathrm{d}\mathbf{x}/|\omega_k|\right) \underset{\mathcal{L}}{=} \mathcal{N}(0,\sigma^2) .$$

Using that $\lim_{k\to+\infty} |\omega_k|^{-1/2}\int_{\omega_k} (\mathbb{E}[V(\mathbf{x})] - \mu)\mathrm{d}\mathbf{x} = 0$ we conclude the proof. $\square$

The following lemmas are the continuous versions of Lemma 3 and 4.

Lemma 6. *Let $(m_k)_{k\in\mathbb{N}}$, $(n_k)_{k\in\mathbb{N}}$ be two positive increasing integer sequences and $(\omega_k)_{k\in\mathbb{N}}$ be the sequence of subsets such that for any $k \in \mathbb{N}$, $\omega_k = [0, m_k] \times [0, n_k]$. Let V be a R-independent, with $R \geqslant 0$, second-order stationary random field over $\mathbb{R}^2$. Assume that V is sample path continuous, then for all $k \in \mathbb{N}$*

(a) $|\omega_k|^{-1} \int_{\mathbf{x}\in\omega_k} \mathbb{E}[V(\mathbf{x})] \, d\mathbf{x} = \mathbb{E}[V(\mathbf{0})]$;

(b) $\lim_{k\to+\infty} |\omega_k|^{-1} \int_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{y})] \, d\mathbf{x}d\mathbf{y} = \int_{\mathbf{x}\in\mathbb{R}^2} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{0})] \, d\mathbf{x}$.

In addition, (13) *and* (17) *hold with $\mu = \mathbb{E}[V(\mathbf{0})]$ and $\sigma = \int_{\mathbf{x}\in\mathbb{R}^2} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{0})] \, d\mathbf{x}$.*

Proof. (a) Using the Fubini-Tonelli theorem we obtain that for any $k \in \mathbb{N}$

$$\mathbb{E}\left[\int_{\mathbf{x}\in\omega_k} |V(\mathbf{x})| d\mathbf{x}\right] = \int_{\mathbf{x}\in\omega_k} \mathbb{E}[|V(\mathbf{x})|] \, d\mathbf{x} = \int_{\mathbf{x}\in\omega_k} \mathbb{E}[|V(\mathbf{0})|] \, d\mathbf{x} = |\omega_k| \mathbb{E}[|V(\mathbf{0})|] < +\infty \ .$$

Thus, using the Fubini-Lebesgue theorem we get (a) for any $k \in \mathbb{N}$.

(b) For any $k \in \mathbb{N}$ we have by stationarity

$$|\omega_k|^{-1} \int_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{y})] \, d\mathbf{x}d\mathbf{y} = |\omega_k|^{-1} \int_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}[V(\mathbf{x}-\mathbf{y}), V(\mathbf{0})] \, d\mathbf{x}d\mathbf{y} \ .$$

By the Fubini-Lebesgue theorem we obtain that for any $k \in \mathbb{N}$,

$$|\omega_k|^{-1} \int_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{y})] \, d\mathbf{x} = \int_{\mathbf{x}\in\mathbb{R}^2} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{0})] \, g_k(\mathbf{x}) d\mathbf{x} \ ,$$

where $g_k \in L^\infty(\mathbb{R}^2)$ satisfies for any $\mathbf{x} \in \mathbb{R}^2$, $g_k(\mathbf{x}) = |\omega_k|^{-1} \mathbb{1}_{\omega_k} * \check{\mathbb{1}}_{\omega_k}(\mathbf{x})$. For any $k \in \mathbb{N}$, $\mathbf{x} \in \mathbb{R}^2$ we have $0 \leqslant g_k(\mathbf{x}) \leqslant 1$ and $\lim_{k\to+\infty} g_k(\mathbf{x}) = 1$. For any $\mathbf{x} \in \mathbb{R}^2$ such that $\|\mathbf{x}\|_\infty > R_{\mathbf{t}}$, $\operatorname{Cov}[V(\mathbf{x}), V(\mathbf{0})] = 0$ and then

$$\int_{\mathbf{x}\in\mathbb{R}^2} |\operatorname{Cov}[V(\mathbf{x}), V(\mathbf{0})]| d\mathbf{x} < +\infty \ .$$

Using the dominated convergence theorem we get that

$$|\omega_k|^{-1} \int_{\mathbf{x},\mathbf{y}\in\omega_k} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{y})] \, d\mathbf{x}d\mathbf{y} = \int_{\mathbf{x}\in\mathbb{R}^2} \operatorname{Cov}[V(\mathbf{x}), V(\mathbf{0})] \, d\mathbf{x} \ ,$$

If V is R-independent we conclude the proof by applying Theorem 7 and 8.

□

Lemma 7. *Let Γ be a function over $\mathbb{R}^2$, $\Gamma \neq 0$, such that for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, $C(\mathbf{x}, \mathbf{y}) = \Gamma(\mathbf{x}-\mathbf{y})$ with C the covariance function of V a second-order random field over $\mathbb{R}^2$. Assume that Γ has finite support. Then it holds for any $\mathbf{t} \in \mathbb{R}^2$, $\Gamma(\mathbf{t}) \leqslant \Gamma(\mathbf{0})$, with equality if and only if $\mathbf{t} = \mathbf{0}$.*

Proof. Upon replacing for any $\mathbf{x} \in \mathbb{R}^2$, $V(\mathbf{x})$ by $V(\mathbf{x}) - \mathbb{E}[V(\mathbf{x})]$ we suppose that $\mathbb{E}[V(\mathbf{x})] = 0$. Using the Cauchy-Schwarz inequality and the stationarity of V we get for any $\mathbf{t} \in \mathbb{R}^2$ and $\mathbf{x} \in \mathbb{R}^2$

$$\Gamma(\mathbf{t}) = \mathbb{E}[V(\mathbf{x}+\mathbf{t})V(\mathbf{x})] \leqslant \mathbb{E}\left[V(\mathbf{x}+\mathbf{t})^2\right]^{1/2} \mathbb{E}\left[V(\mathbf{x})^2\right]^{1/2} \leqslant \mathbb{E}\left[V(\mathbf{x})^2\right] \leqslant \Gamma(\mathbf{0}) \ .$$

with equality if and only if $V(\mathbf{x}+\mathbf{t}) = \alpha(\mathbf{x})V(\mathbf{x})$ with $\alpha(\mathbf{x}) \in \mathbb{R}$. Since V is stationary and $V \neq 0$ we get that for any $\mathbf{x}, \mathbf{t} \in \mathbb{R}^2$, $\mathbb{E}\left[V(\mathbf{x}+\mathbf{t})^2\right] = \mathbb{E}\left[V(\mathbf{x})^2\right] > 0$. Thus $\alpha(\mathbf{x})^2 = 1$ and for all $n \in \mathbb{N}$, $V(n\mathbf{t}) = \pm V(\mathbf{0})$. If $\mathbf{t} \neq 0$ then there exists $n \in \mathbb{N}$ such that $n\mathbf{t} \notin \operatorname{Supp}(\Gamma)$ and then we have

$$0 = \Gamma(n\mathbf{t}) = \mathbb{E}[V(n\mathbf{t})V(\mathbf{0})] = \pm\mathbb{E}\left[V(\mathbf{0}^2)\right] \neq 0 \ ,$$

which is absurd. Thus the equality in the inequality holds if and only if $\mathbf{t} = 0$. □

C Explicit constants

In order to derive precise constants in Theorem 1 and 2 we use the following lemma which is a consequence of the Isserlis formula [55].

Lemma 8. *Let U and V be two zero-mean, real-valued Gaussian random variable and $k, \ell \in \mathbb{N}$. We have*

$$\mathbb{E}\left[U^{2k}V^{2\ell}\right] = \sum_{j=0}^{k\wedge\ell} r_{j,k,\ell}\mathbb{E}\left[U^2\right]^{k-j}\mathbb{E}\left[V^2\right]^{\ell-j}\mathbb{E}\left[UV\right]^{2j} \quad \textit{and} \quad \mathrm{Cov}\left[U^{2k}V^{2\ell}\right] = \sum_{j=1}^{k\wedge\ell} r_{j,k,\ell}\mathbb{E}\left[U^2\right]^{k-j}\mathbb{E}\left[V^2\right]^{\ell-j}\mathbb{E}\left[UV\right]^{2j} ,$$

with $r_{j,k,\ell}$ defined by (2)*.*

Proof. Let $k, \ell \in \mathbb{N}$. Using Isserlis formula [55] we obtain that $\mathbb{E}\left[U^{2k}V^{2\ell}\right]$ is the sum over all the partitions in pairs of $\{\underbrace{U, \dots, U}_{2k \text{ times}}, \underbrace{V, \dots, V}_{2\ell \text{ times}}\}$ of the product of the expectations given by a pair partition. Given a pair partition we identify three different cases, $\{U, U\}$, $\{V, V\}$ and $\{U, V\}$. We only need to count the number of times each case appears in the sum. We denote the number of $\{U, U\}$ couples in a given pair partition p by $n_{U,U}(p)$. In the same fashion we define $n_{U,V}(p)$ and $n_{V,V}(p)$. We have $2k = 2n_{U,U}(p) + n_{U,V}(p)$ which proves that $n_{U,V}(p)$ is even. We denote by $\mathscr{P}_j$ the number of pair partitions p such that $n_{U,V}(p) = 2j$, with $j \in [\![0, k \wedge \ell]\!]$.

The cardinality of $\mathscr{P}_j$ is given by $r_{j,k,\ell}$. Indeed, in order to select $2j$ pair $\{U, V\}$ we select $2j$ elements among $2k$ (selection of replicates of U), same for V which gives $\binom{2k}{2j}\binom{2\ell}{2j}$ possibilities. Considering all the bijections between these elements we construct all the possible $2j$ pairs $\{U, V\}$. Given $2j$ pairs $\{U, V\}$ we must construct $k-j$ pairs $\{U, U\}$ and $\ell - j$ pairs $\{V, V\}$ in order to obtain a pair partition of $\mathscr{P}_j$. The number of pairs partition of a set with $\ell - j$ elements is given $q_{\ell-j}$. As a consequence we obtain for all $j \in [\![0, k \wedge \ell]\!]$

$$|\mathscr{P}_j| = q_{k-j}q_{\ell-j}\binom{2k}{2j}\binom{2\ell}{2j}(2j)! = r_{j,k,\ell} .$$

Summing over $j \in [\![0, k \wedge \ell]\!]$ we obtain all the possible pair partition and we get

$$\mathbb{E}\left[U^{2k}V^{2\ell}\right] = \sum_{j=0}^{k\wedge\ell} r_{j,k,\ell}\mathbb{E}\left[U^2\right]^{k-j}\mathbb{E}\left[V^2\right]^{\ell-j}\mathbb{E}\left[UV\right]^{2j} . \tag{18}$$

Using that $r_{0,k} = q_k^2$, respectively $r_{0,\ell} = q_\ell^2$ and $\mathbb{E}\left[U^{2k}\right] = q_k\mathbb{E}\left[U\right]^k$, respectively $\mathbb{E}\left[V^{2\ell}\right] = q_\ell\mathbb{E}\left[V\right]^\ell$, we obtain that the first term in the right-hand side sum of (18) is equal to $\mathbb{E}\left[U^{2k}\right]\mathbb{E}\left[V^{2\ell}\right]$. Hence by removing this term we obtain the covariance and conclude the proof. ☐

Proof of Proposition 1. The proof is divided in two parts. First we consider the case $i = p$ then the case $i = sc$.

Let $i = p$ with $p = 2\ell$ and $\ell \in \mathbb{N}$, $\mathbf{t} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$ and $V_\mathbf{t}$ the Gaussian random field given for any $\mathbf{x} \in \mathbb{Z}^2$ by $V_\mathbf{t}(\mathbf{x}) = U(\mathbf{x}) - U(\mathbf{x} + \mathbf{t})$. Note that for all $\mathbf{x} \in \mathbb{Z}^2$ we have $V_\mathbf{t}(\mathbf{x})^{2\ell} = V_{p,\mathbf{t}}(\mathbf{x})$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^2$ we have

$$\mathbb{E}\left[V_\mathbf{t}(\mathbf{x})\right] = 0 , \qquad \mathrm{Cov}\left[V_\mathbf{t}(\mathbf{x}), V_\mathbf{t}(\mathbf{y})\right] = 2\Gamma_f(\mathbf{x}-\mathbf{y}) - \Gamma_f(\mathbf{x}-\mathbf{y}-\mathbf{t}) - \Gamma_f(\mathbf{x}-\mathbf{y}+\mathbf{t}) = \Delta_f(\mathbf{t}, \mathbf{x}-\mathbf{y}) , \tag{19}$$

with Δ_f given by (3). We show in proof of Theorem 1, see Equation (1), that for any $\mathbf{t} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$

$$\mu_p(\mathbf{t}) = \mathbb{E}\left[V_\mathbf{t}^{2\ell}(\mathbf{0})\right]^{1/2\ell} , \qquad \sigma_p(\mathbf{t})^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}\left[V_\mathbf{t}^{2\ell}(\mathbf{x}), V_\mathbf{t}^{2\ell}(\mathbf{0})\right]\mathbb{E}\left[V_\mathbf{t}^{2\ell}(\mathbf{0})\right]^{1/\ell-2}/(2\ell)^2 . \tag{20}$$

Combining (19), (20) and Lemma 8 we get that

(a) $\mu_p(\mathbf{t}) = q_{2\ell}^{1/(2\ell)}\Delta_f(\mathbf{t}, \mathbf{0})^{1/2}$;

(b) $\sigma_p(\mathbf{t})^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2}\left(\sum_{j=1}^{\ell} r_{j,\ell}\Delta_f(\mathbf{t}, \mathbf{0})^{2(\ell-j)}\Delta_f(\mathbf{t}, \mathbf{x})^{2j}\right) q_\ell^{1/\ell-2}\Delta_f(\mathbf{t}, \mathbf{0})^{1-2\ell}/(2\ell)^2$.

Exchanging the sums in (b) we get $\sigma_p(\mathbf{t})^2 = \frac{q_\ell^{1/\ell-2}}{(2\ell)^2}\sum_{j=1}^{\ell} r_{j,\ell}\left(\frac{\|\Delta_f(\cdot,\mathbf{t})\|_{2j}}{\Delta_f(\mathbf{0},\mathbf{t})}\right)^{2j}\Delta_f(\mathbf{0}, \mathbf{t})$.

Let $i = sc$, $\mathbf{t} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$ and $V_{sc,\mathbf{t}}$ be a Gaussian random field given for any $\mathbf{x} \in \mathbb{Z}^2$, by $V_{\mathbf{t}}(\mathbf{x}) = U(\mathbf{x})U(\mathbf{x}+\mathbf{t})$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^2$ we have

$$\mathbb{E}\left[V_{sc,\mathbf{t}}(\mathbf{x})\right] = \Gamma_f(\mathbf{t}) \ , \qquad \mathrm{Cov}\left[V_{sc,\mathbf{t}}(\mathbf{x}), V_{sc,\mathbf{t}}(\mathbf{y})\right] = \Gamma_f(\mathbf{x}-\mathbf{y}) - \Gamma_f(\mathbf{x}-\mathbf{y}-\mathbf{t})\Gamma_f(\mathbf{x}-\mathbf{y}+\mathbf{t}) = \Delta_f^m(\mathbf{t}, \mathbf{x}-\mathbf{y}) \ , \tag{21}$$

with Δ_f^m given by (3). We show in the proof of Theorem 1, see (1), that for any $\mathbf{t} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}$

$$\mu_{sc}(\mathbf{t}) = \mathbb{E}\left[V_{sc,\mathbf{t}}(\mathbf{0})\right] \ , \qquad \sigma_{sc}(\mathbf{t})^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2} \mathrm{Cov}\left[V_{sc,\mathbf{t}}(\mathbf{x}), V_{sc,\mathbf{t}}(\mathbf{0})\right] \ . \tag{22}$$

Combining (21) and (22) we get that

(a) $\mu_{sc}(\mathbf{t}) = \Gamma_f(\mathbf{t})$;

(b) $\sigma_{sc}(\mathbf{t})^2 = \sum_{\mathbf{x}\in\mathbb{Z}^2} \Delta_f^m(\mathbf{t}, \mathbf{x})$.

□

Proof of Proposition 2. The proof is divided in two parts. First we treat the case $i = p$ then the case $i = sc$.

Let $p = 2\ell$ with $\ell \in \mathbb{N}$. Lemma 5 gives us that

$$\mu_{p,p} = \sum_{j=0}^{\ell} \binom{2\ell}{2j} \mathbb{E}\left[U(\mathbf{0})\right]^{2(\ell-j)} \beta_j \quad \text{and} \quad \sigma_{p,p} = \sum_{i,j=0}^{\ell} \binom{2\ell}{2i}\binom{2\ell}{2j} \left\langle \gamma_{i,j}, \mathrm{Cov}\left[U(\cdot)^{2(\ell-i)}, U(\mathbf{0})^{2(\ell-j)}\right]\right\rangle \ .$$

Using Lemma 8 we obtain that

$$\mu_{p,p} = \Gamma_f(\mathbf{0})^{\ell} \sum_{j=0}^{\ell} \binom{2\ell}{2j} q_{\ell-j} \Gamma_f(\mathbf{0})^{-j} \beta_j \quad \text{and} \quad \sigma_{p,p} = \sum_{i,j=0}^{\ell} \binom{2\ell}{2i}\binom{2\ell}{2j} \sum_{m=1}^{\ell-i\wedge\ell-j} r_{m,k,\ell} \left\langle \gamma_{i,j}, \Gamma_f^{2m}\right\rangle \Gamma_f(\mathbf{0})^{2\ell-i-j-2m} \ .$$

We conclude using (4).

For $i = sc$, the result is given in the proof of Theorem 2.

□

References

[1] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S. Kendall, and Joseph Mecke. *Stochastic geometry and its applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, third edition, 2013.

[2] Adrian Baddeley, Ege Rubak, and Rolf Turner. *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London, 2015.

[3] Rolf Schneider and Wolfgang Weil. *Stochastic and integral geometry*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2008.

[4] D. J. Daley. The definition of a multi-dimensional generalization of shot noise. *J. Appl. Probability*, 8:128–135, 1971.

[5] John Rice. On generalized shot noise. *Advances in Appl. Probability*, 9(3):553–565, 1977.

[6] Stephen O Rice. Mathematical analysis of random noise. *Bell System Technical Journal*, 23(3):282–332, 1944.

[7] Hermine Biermé and Agnès Desolneux. On the perimeter of excursion sets of shot noise random fields. *Ann. Probab.*, 44(1):521–543, 2016.

[8] Robert J. Adler, Gennady Samorodnitsky, and Jonathan E. Taylor. Excursion sets of three classes of stable random fields. *Adv. in Appl. Probab.*, 42(2):293–318, 2010.

[9] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20-26 June 2005, San Diego, CA, USA*, pages 60–65, 2005.

[10] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. Image Processing*, 13(9):1200–1212, 2004.

[11] J. Jain and A. Jain. Displacement measurement and its application in interframe image coding. *IEEE Transactions on Communications*, 29(12):1799–1808, December 1981.

[12] Jarke J. van Wijk. Spot noise texture synthesis for data visualization. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1991, Providence, RI, USA, April 27-30, 1991*, pages 309–318, 1991.

[13] Bruno Galerne, Yann Gousseau, and Jean-Michel Morel. Random phase textures: Theory and synthesis. *IEEE Trans. Image Processing*, 20(1):257–267, 2011.

[14] Arthur Leclaire. *Champs à phase aléatoire et champs gaussiens pour la mesure de netteté d'images et la synthèse rapide de textures*. PhD thesis, Université Paris Descartes, 2015. 2015USPCB041.

[15] Gui-Song Xia, Sira Ferradans, Gabriel Peyré, and Jean-François Aujol. Synthesizing and mixing stationary gaussian texture models. *SIAM J. Imaging Sciences*, 7(1):476–508, 2014.

[16] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. Texture synthesis using convolutional neural networks. In *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pages 262–270, 2015.

[17] Yang Lu, Song-Chun Zhu, and Ying Nian Wu. Learning FRAME models using CNN filters. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pages 1902–1910, 2016.

[18] J. Bruna and S. Mallat. Multiscale Sparse Microcanonical Models. *ArXiv e-prints*, January 2018.

[19] Michael Unser, Philippe Thévenaz, and Leonid P. Yaroslavsky. Convolution-based interpolation for fast, high-quality rotation of images. *IEEE Trans. Image Processing*, 4(10):1371–1381, 1995.

[20] Robert J. Adler. *The geometry of random fields*. John Wiley & Sons, Ltd., Chichester, 1981. Wiley Series in Probability and Mathematical Statistics.

[21] John I. Yellott. Implications of triple correlation uniqueness for texture statistics and the julesz conjecture. *J. Opt. Soc. Am. A*, 10(5):777–793, May 1993.

[22] Bela Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91, 1981.

[23] Béla Julesz. Visual pattern discrimination. *IRE Trans. Information Theory*, 8(2):84–92, 1962.

[24] Persi Diaconis and David Freedman. On the statistics of vision: The julesz conjecture. *Journal of Mathematical Psychology*, 24(2):112 – 138, 1981.

[25] Javier Portilla and Eero P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–70, 2000.

[26] Alasdair Newson, Andrés Almansa, Yann Gousseau, and Saïd Ladjal. Taking Apart Autoencoders: How do They Encode Geometric Shapes ? working paper or preprint, January 2018.

[27] Kaiming He and Jian Sun. Image completion approaches using the statistics of similar patches. *IEEE Trans. Pattern Anal. Mach. Intell.*, 36(12):2423–2435, 2014.

[28] Marc Lebrun, Antoni Buades, and Jean-Michel Morel. A Nonlocal Bayesian Image Denoising Algorithm. *SIAM J. Imaging Sciences*, 6(3):1665–1688, 2013.

[29] Alexei A. Efros and Thomas K. Leung. Texture synthesis by non-parametric sampling. In *ICCV*, pages 1033–1038, 1999.

[30] Alexei A. Efros and William T. Freeman. Image quilting for texture synthesis and transfer. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2001, Los Angeles, California, USA, August 12-17, 2001*, pages 341–346, 2001.

[31] Elizaveta Levina and Peter J. Bickel. Texture synthesis and nonparametric resampling of random fields. *Ann. Statist.*, 34(4):1751–1773, 2006.

[32] Lara Raad, Agnès Desolneux, and Jean-Michel Morel. A conditional multiscale locally gaussian texture synthesis algorithm. *Journal of Mathematical Imaging and Vision*, 56(2):260–279, 2016.

[33] Georg Lindgren. *Stationary stochastic processes*. Chapman & Hall/CRC Texts in Statistical Science Series. CRC Press, Boca Raton, FL, 2013. Theory and applications.

[34] Jürgen Potthoff. Sample properties of random fields. II. Continuity. *Commun. Stoch. Anal.*, 3(3):331–348, 2009.

[35] Charles-Alban Deledalle, Loïc Denis, and Florence Tupin. How to compare noisy patches? Patch similarity beyond Gaussian noise. *Int. J. Comput. Vis.*, 99(1):86–102, 2012.

[36] David Mumford and Agnès Desolneux. *Pattern theory*. Applying Mathematics. A K Peters, Ltd., Natick, MA, 2010. The stochastic analysis of real-world signals.

[37] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, third edition, 1995. A Wiley-Interscience Publication.

[38] Ulrich Stadtmüller and Le Van Thanh. On the strong limit theorems for double arrays of blockwise M-dependent random variables. *Acta Math. Sin. (Engl. Ser.)*, 27(10):1923–1934, 2011.

[39] Ferenc Móricz, Ulrich Stadtmüller, and Monika Thalmaier. Strong laws for blockwise m-dependent random fields. *J. Theoret. Probab.*, 21(3):660–671, 2008.

[40] Harald Cramér. *Mathematical methods of statistics*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1999. Reprint of the 1946 original.

[41] Philippe G. Ciarlet. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1982.

[42] Raymond H. Chan, Xiao-Qing Jin, and Man-Chung Yeung. The circulant operator in the Banach algebra of matrices. *Linear Algebra Appl.*, 149:41–53, 1991.

[43] J. P. Imhof. Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48:419–426, 1961.

[44] A. W. Davis. A differential equation approach to linear combinations of independent chi-squares. *J. Amer. Statist. Assoc.*, 72(357):212–214, 1977.

[45] Samuel Kotz, N. L. Johnson, and D. W. Boyd. Series representations of distributions of quadratic forms in normal variables. II. Non-central case. *Ann. Math. Statist.*, 38:838–848, 1967.

[46] Samuel D. Oman and S. Zacks. A mixture approximation to the distribution of a weighted sum of chi-squared variables. *Journal of Statistical Computation and Simulation*, 13(3-4):215–224, 1981.

[47] Dean A. Bodenham and Niall M. Adams. A comparison of efficient approximations for a weighted sum of chi-squared random variables. *Stat. Comput.*, 26(4):917–928, 2016.

[48] Peter Hall. Chi squared approximations to the distribution of a sum of independent random variables. *Ann. Probab.*, 11(4):1028–1036, 1983.

[49] Michael J. Buckley and G.K. Eagleson. An approximation to the distribution of quadratic forms in normal random variables. *Australian Journal of Statistics*, 30A(1):150–159, 1988.

[50] Andrew T. A. Wood. An f approximation to the distribution of a linear combination of chi-squared variables. *Communications in Statistics - Simulation and Computation*, 18(4):1439–1456, 1989.

[51] Roger G. Ghanem and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.

[52] Svante Janson. Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs. *Ann. Probab.*, 16(1):305–312, 1988.

[53] Harald Cramér and M. R. Leadbetter. *Stationary and related stochastic processes*. Dover Publications, Inc., Mineola, NY, 2004. Sample function properties and their applications, Reprint of the 1967 original.

[54] Alexander V. Ivanov and Nikolai N. Leonenko. *Statistical analysis of random fields*, volume 28 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1989. With a preface by A. V. Skorokhod, Translated from the Russian by A. I. Kochubinskiĭ.

[55] Leon Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1-2):134–139, 1918.