A Multiscale Wavelet-based Test for Isotropy of Random Fields on a Regular Lattice

Kevin Thon*, Marc Geilhufe, and Donald B. Percival

Abstract—A test for isotropy of images modeled as stationary or intrinsically stationary random fields on a lattice is developed. The test is based on wavelet theory, and can operate on the horizontal and vertical scale of choice, or on any combination of scales. Scale is introduced through the wavelet variances (sometimes referred to as the wavelet power spectrum), which decompose the variance over different horizontal and vertical spatial scales. The method is more general than existing tests for isotropy, since it handles intrinsically stationary random fields as well as second-order stationary fields. The performance of the method is demonstrated on samples from different (an)isotropic random fields, and compared to three existing methods. It is competitive with or outperforms existing methods since it consistently rejects close to the nominal level for isotropic fields while having a rejection rate for anisotropic fields comparable to the existing methods in the stationary case, and superior in the intrinsic case. As practical examples, paper density images of handsheets and mammogram images are analyzed.

Index Terms—Maximal overlap discrete wavelet transform, wavelet variance, isotropy, anisotropy, random fields.

EDICS Category: TEC-MRS

I. INTRODUCTION

The use of random fields is ubiquitous in the study of natural processes and in image analysis. A common assumption, usually made for simplicity and computational purposes, is that the field in question is isotropic [46], meaning that the variability in the random field is independent of direction; however, many processes have a clear directional component. For example, an image of ocean waves would likely have a clear directionality determined by the dominant wind direction [1], as would ripples in sediment surfaces in shallow water formed by waves [34]. As part of seabed texture characterization, it is of interest to determine the presence of anisotropy in images of the seafloor obtained by autonomous underwater vehicles with high-resolution, side-looking sonars [10, 15, 17]. Anisotropy is indicative of areas of sand ripples, and knowledge of their location will lead to improvements in automatic target recognition. The visibility of sand ripples in sonar images is dependent on the viewing angle and can be impaired if observed from a vehicle on a path perpendicular to the sand ripple field's major orientation. In radiographic images of bone the presence of directionalities has been shown to be a useful predictor of fracture risk [3], and in [12], the level of anisotropy of plastic food wrapping is connected to it having a higher tendency for micro-organisms to adhere. A discussant for this paper specifically pointed out the need for a test for isotropy. In modern cosmology a topic currently receiving much attention is whether the cosmic microwave background radiation is isotropic [7]. The presence of anisotropies challenges the standard cosmological model and points to the importance of a test for isotropy for understanding the very nature of our universe (the random field resides on a spherical grid in this case).

The examples above include both stationary and intrinsically stationary random field models, and in this paper we will present a new way to test for isotropy of both models. The proposed method is based on the variances of the coefficients from the two-dimensional maximal overlap discrete wavelet transform (MODWT), a non-subsampled and shift invariant version of the discrete wavelet transform. Originally wavelet variance was used in the study of time series, decomposing the variance of the time series over different temporal scales [41]. In the context of spatial processes or images, the wavelet variance can be used to decompose the variance over different spatial scales. Papers in this journal that have investigated this topic include [48], where wavelet variances are used for the classification and segmentation of textures, and [40], where many of the statistical properties of wavelet variances are worked out and alternative variance decompositions are given. Wavelet variances have also been used in, e.g., the analysis of soil properties and variations [32, 31], aerial photographs of lake patterns [38], sea ice synthetic aperture images [18], and texture analysis based on locally stationary wavelet fields [14].

Currently there exist several approaches for assessing the presence of anisotropy. Arguably the most common is a visual evaluation based on variograms (a more general construct than the covariance for expressing variability in a random field) for different lags plotted in a rose diagram to find the anisotropy axis [23]. The use of variograms is natural to identify and model isotropy or anisotropy of random fields [8, 6], and several formal statistical tests for isotropy exist that are based on comparing variogram values over different lags [20, 46, 35]. Tests have also been developed using the periodogram to assess certain symmetry conditions [44, 33]. In [39], a test based on bivariate circular statistics and subsampling is derived to detect anisotropy, and in [5] a test for isotropy for affine processes based on level sets is developed. All of the tests

K. Thon was with Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, 9038 Tromsø, Norway. He is now with Center for Clinical Documentation and Evaluation (SKDE), P.O. Box 6, 9038 Tromsø (e-mail: kevin.thon@skde.no)

M. Geilhufe was with Department of Mathematics and Statistics, Faculty of Science and Technology, University of Tromsø, 9037 Tromsø, Norway. He is now with Norwegian Defence Research Establishment (FFI), P.O. Box 25, 2027 Kjeller, Norway (e-mail: marc.geilhufe@ffi.no)

D. B. Percival is with Applied Physics Laboratory and the Department of Statistics, University of Washington, Seattle, WA 98195, USA (e-mail: dbp@apl.washington.edu)

above have been developed for stationary fields. For intrinsically stationary fields the only formal test we know of uses a fractional Brownian field as a null hypothesis and is based on evaluation of the Hurst index in different directions [43].

The main contribution of this paper is the development of a scale-based test for isotropy of random fields. Our test differs from existing ones in that it can be tuned to include exactly the scales of interest for the application at hand and that it can handle both intrinsically stationary and stationary random fields. The performance of our test will be demonstrated on realizations from (intrinsically) stationary random fields with different underlying covariance functions (or variograms in the intrinsic case), ranging from weakly to strongly correlated processes. We compare our proposed tests to the methods described in [20, 33, 43]. We focus on these three particular methods because they are statistical tests especially developed for random fields on a regular two-dimensional lattice, and they represent conceptually different ways of assessing isotropy, from variograms [20] and periodograms [33] in the stationary case to directional Hurst indices [43] in the intrinsically stationary case of fractional Brownian fields.

The remainder of the paper is organized as follows. Section II introduces necessary spatial concepts needed for the assessment of isotropy. Theoretical background on the two-dimensional MODWT wavelet coefficients and estimators of the wavelet variance is given in Section III. The test for isotropy including derivation of necessary estimators is developed in Section IV. Examples of (Gaussian) isotropic and anisotropic random fields with different correlation structures are presented in Sections V, which are used in Section VI to check the validity of the asymptotics used for deriving the test statistic, and in Section VII for a simulation study including a comparison of the presented approach with several existing methods. Finally we investigate the presence of isotropy in two practical examples using paper density images in Section VIII and mammogram images in Section IX (the former are modeled as stationary fields, and the latter, as intrinsically stationary fields). Our findings are summarized and discussed in Section X.

II. ISOTROPY

Let $\{X_{u,v} : (u,v) \in \mathbb{Z}^2\}$ denote a random field or random process on a lattice. A random field is *second-order stationary* (also known as *weak* or *wide-sense stationary*) when it has constant mean and the covariance between its elements only depends on the relative displacement of the locations; i.e., $C(\boldsymbol{\kappa}) = s_{X,\kappa_1,\kappa_2} = \mathrm{cov}(X_{u,v}, X_{u+\kappa_1, v+\kappa_2})$ for all lags $\boldsymbol{\kappa} = (\kappa_1, \kappa_2) \in \mathbb{Z}^2$. We will henceforth refer to second-order stationarity simply as *stationarity*. A more general class of random fields is that of *intrinsic* random fields, which are defined in terms of their increments being stationary. Following [29, 40], the random field $\{X_{u,v}\}$ is *intrinsically stationary* of order d if all increment processes of order d are stationary. In this paper we will solely deal with intrinsic stationarity of order *1* (henceforth referred to as *intrinsic stationarity*); i.e., the increments have constant mean, usually expressed as 0, and finite variance, which we express in terms

of the *semivariogram* $\gamma(\boldsymbol{\kappa}) = \gamma_{X,\kappa_1,\kappa_2} = \frac{1}{2}\mathrm{var}(X_{u,v} - X_{u+\kappa_1, v+\kappa_2})$ for all $\boldsymbol{\kappa} = (\kappa_1, \kappa_2) \in \mathbb{Z}^2$. The quantity $2\gamma(\boldsymbol{\kappa})$ is known as *variogram*.

If the semivariogram of a (intrinsically) stationary process $\{X_{u,v}\}$ is only a function of magnitude and not direction, i.e., $\gamma(\boldsymbol{\kappa}) = \gamma(||\boldsymbol{\kappa}||)$, then $\{X_{u,v}\}$ is called *isotropic*, otherwise *anisotropic*. Note that for stationary processes $\gamma(\boldsymbol{\kappa}) = C(0) - C(\boldsymbol{\kappa})$ [8] and for stationary isotropic processes $C(\boldsymbol{\kappa}) = C(||\boldsymbol{\kappa}||)$. For intrinsically stationary processes the variance is allowed to be infinite; i.e., the covariance $C(\boldsymbol{\kappa})$ need not exist in the intrinsic case. Anisotropy occurs when the underlying process evolves differentially in space. As [8] notes, sometimes a linear transformation can correct anisotropy; i.e., there exists a positive definite matrix $\mathbf{B}$ (rotation and scaling) for which the anisotropic semivariogram $\gamma(\boldsymbol{\kappa}) = \gamma_0\left(\sqrt{\boldsymbol{\kappa}^T \mathbf{B} \boldsymbol{\kappa}}\right)$ for all lags $\boldsymbol{\kappa} \in \mathbb{Z}^2$, where γ_0 is an isotropic semivariogram. This case is commonly known as *geometric anisotropy*. There also exist anisotropy types which are not geometric. These are often collectively referred to as *zonal anisotropy* [26], although there is some ambiguity as to the precise definition and classification of the non-geometric anisotropy types. For instance in [6] the anisotropy types are classified as geometric and zonal, but in addition a third type is introduced which is simply referred to as *other anisotropies*. An alternative and more detailed classification of types of anisotropy was proposed by [49]. In this paper we will for simplicity exclusively consider geometric anisotropy. The performance for other types of anisotropies is discussed in [16].

III. WAVELET VARIANCE BACKGROUND

Let $\{h_{1,l} : l = 0, \dots, L-1\}$ and $\{g_{1,l} : l = 0, \dots, L-1\}$ denote, respectively, the unit level one-dimensional MODWT wavelet and scaling filters. The filters are from the commonly used Daubechies class, and are of even length L and have $L/2$ vanishing moments [41, 11]. This class is divided into two families of filters, the Debauchies extremal phase filters $D(L)$ and the least asymmetric filters $LA(L)$. The simplest Daubechies wavelet is $D(2)$, more commonly known as the Haar filter. The Haar wavelet filter $\{1/2, -1/2\}$ yields a contrast between adjacent data values, while the scaling filter $\{1/2, 1/2\}$ produces an average. In this paper we will focus on Haar, $D(4)$ and $LA(8)$ filters. The MODWT filters for level j are of length $L_j = (2^j - 1)(L - 1) + 1$. They are obtained by operations on the unit-level filters that can be computed with (inverse) Fourier transforms and a normalization such that $\sum_l h_{j,l}^2 = \sum_l g_{j,l}^2 = 1/2^j$ (see, e.g., [41] for more details).

Building a tensor product of the filters on a lattice process $\{X_{u,v}\}$ yields the two-dimensional wavelet coefficients for levels j and j' [40]; i.e.,

$$W_{j,j'}(u,v) = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} h_{j,l} h_{j',l'} X_{u-l,v-l'}, \qquad (1)$$

$$U_{j,j'}(u,v) = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} g_{j,l} h_{j',l'} X_{u-l,v-l'}, \qquad (2)$$

$$V_{j,j'}(u,v) = \sum_{l=0}^{L_j-1} \sum_{l'=0}^{L_{j'}-1} h_{j,l} g_{j',l'} X_{u-l,v-l'}, \qquad (3)$$

Equations (1) – (3) describe wavelet-wavelet (ww), scaling-wavelet (sw) and wavelet-scaling (ws) coefficients. They differ by the application of either the wavelet or scaling filter to the rows and/or columns of $\{X_{u,v}\}$.

Let $\tau_j = 2^{j-1}$ denote the scale associated with level j. The notion of scale is fundamental in wavelet analysis and is related to how much data are effectively covered (or spanned) by a particular filter. A scaling filter $\{g_{j,l}\}$ produces weighted averages over a scale of $2\tau_j$, whereas a wavelet filter essentially produces differences of adjacent weighted averages, each of which has a scale of τ_j. Hence, the sw process $\{U_{j,j',u,v}\}$ is associated with changes at scale $\tau_{j'} = 2^{j'-1}$ of the row averages corresponding to scale $2\tau_j$. Similarly, the ws process $\{V_{j,j',u,v}\}$ is associated with changes at scale $\tau_j = 2^{j-1}$ of the column averages corresponding to scale $2\tau_{j'}$. The ww process $\{W_{j,j',u,v}\}$ is associated with changes along the rows at scale τ_j and along the columns at scale $\tau_{j'}$.

The coefficients can be calculated efficiently by a two-dimensional pyramid algorithm described in [18] for all possible combinations of levels j and j'. Note that it is common in the literature to focus only on the case where $j = j'$; i.e., the same scales are considered for the horizontal and vertical direction (see, e.g., [19]). Following the convention in [40, 18] we will refer to this as the diagonal case, which is not to be confused with the nomenclature in, e.g., [32], where the term refers to the part of the decomposition involving ww coefficients.

Let C stand for the coefficients listed in Equations (1) – (3), i.e., W, U or V. Then the variance of C for levels j, j' is defined as $v_{C,j,j'}^2 = \text{var}\{C_{j,j'}(u,v)\}$. For an (intrinsically) stationary field $\{X_{u,v}\}$ with known semivariogram $\gamma_{X,\kappa_1,\kappa_2}$ the wavelet variances can be written as

$$v_{W,j,j'}^2 = -\sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} h_{j,l'} h_{j',k} h_{j',k'} \gamma_{X,l-l',k-k'}, \qquad (4)$$

$$v_{U,j,j'}^2 = -\sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} g_{j,l} h_{j,l'} g_{j',k} h_{j',k'} \gamma_{X,l-l',k-k'}, \qquad (5)$$

$$v_{V,j,j'}^2 = -\sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} g_{j,l'} h_{j',k} g_{j',k'} \gamma_{X,l-l',k-k'}. \qquad (6)$$

When $\{X_{u,v}\}$ is stationary, we can express Equations (4) – (6) equivalently in terms of its covariance function s_{X,κ_1,κ_2},

$$v_{W,j,j'}^2 = \sum_{l,l'=0}^{L_j-1} \sum_{k,k'=0}^{L_{j'}-1} h_{j,l} h_{j,l'} h_{j',k} h_{j',k'} s_{X,l-l',k-k'}.$$

Similar forms exist for $v_{U,j,j'}^2$ and $v_{V,j,j'}^2$.

As shown in [40] an unbiased estimator for the wavelet variances of a realization of $\{X_{u,v}\}$ on a finite grid of size $N \times M$ is given by

$$\hat{v}_{C,j,j'}^2 = \frac{1}{N_j M_{j'}} \sum_{u=L_j-1}^{N-1} \sum_{v=L_{j'}-1}^{M-1} C_{j,j'}^2(u,v), \qquad (7)$$

where $N_j = N - L_j + 1$, $M_{j'} = M - L_{j'} + 1$ (this presumes the level j is small enough so that $N_j > 0$ and $M_{j'} > 0$).

IV. Test for isotropy

For an isotropic (intrinsically) stationary process $\gamma_{X,\kappa_1,\kappa_2} = \gamma\left(\sqrt{\kappa_1^2 + \kappa_2^2}\right) = \gamma_{X,\kappa_2,\kappa_1}, (\kappa_1, \kappa_2) \in \mathbb{Z}^2$. A necessary condition for isotropy is therefore that

$$\gamma_{X,\kappa_1,\kappa_2} = \gamma_{X,\kappa_2,\kappa_1}. \qquad (8)$$

For a realization from a random field $\{X_{u,v}\}$, we will test the validity of Eq. (8) by looking at the variances of the wavelet coefficients from Equations (1) – (3), which capture distinct features in $\{X_{u,v}\}$. The sw coefficients $U_{j,j'}$ are constructed from the original image by applying differencing filters of length $L_{j'}$ in the horizontal direction and averaging filters of length L_j vertically. The ws coefficients $V_{j,j'}$, on the other hand are constructed by applying averaging filters of length $L_{j'}$ in the horizontal direction and differencing filters of length L_j vertically. When the random field is isotropic, the variance of the sw coefficients indexed j, j' should therefore be equal to the variance of the ws coefficients indexed j', j; i.e. $v_{U,j,j'}^2 = v_{V,j',j}^2$. The ww coefficients $W_{j,j'}$ are built somewhat differently. Here averages over diagonal blocks vs. counter diagonal blocks are created leading to the symmetry condition $v_{W,j,j'}^2 = v_{W,j',j}^2$ if the field in question is isotropic. These relations are also readily deduced from Equations (4) – (6) when the condition in Eq. (8) is fulfilled, in which case the following relations are valid under the null hypothesis H_0 of isotropy:

$$v_{U,j,j'}^2 = v_{V,j',j}^2 \quad \Leftrightarrow \quad \log\left(\frac{v_{U,j,j'}^2}{v_{V,j',j}^2}\right) = 0, \qquad (9)$$

$$v_{W,j,j'}^2 = v_{W,j',j}^2 \quad \Leftrightarrow \quad \log\left(\frac{v_{W,j,j'}^2}{v_{W,j',j}^2}\right) = 0. \qquad (10)$$

A test for isotropy can therefore be constructed using any of the ratios above, or any combination of them, replacing the variances with their estimators from Eq. (7) and using large sample results to approximate their distributions. The estimators are shown in [40] to be unbiased and asymptotically normal. Hence the delta method [13] may be used to find the asymptotic distribution of the log-ratios in question, leading to a statistical test for isotropy. The log-ratios are chosen as a test statistic rather than the ratios themselves since a test that is simultaneous over several ratios can readily be developed through simple matrix operations for the log-ratio case (this is not the case for the ratios directly). In the subsections that follow, we first derive the test for single ratios and subsequently a simultaneous test over all ratios of interest (Matlab code for the tests is available from the authors).

Prior to developing the proposed test, we note one important potential weakness, namely that the log-ratios will be zero for the geometrically anisotropic case when the direction of anisotropy is $45°$ to the image axes, since the condition in Eq. (8) will be met for this anomalous case. To guard against this pitfall, we propose a two-stage test where both the image and a $45°$ rotated version of the image are tested.

This will, for images that are anisotropic, guarantee that the direction of anisotropy cannot be $45°$ to the image axes in both stages of the test. With regard to the significance level, α, of the resulting two-stage test, we apply a simple Bonferroni correction. Hence, we set the significance level of the individual tests to $\alpha_1 = \alpha/2$, thus making sure that P(Rejection in at least one test $|H_0) \leq \alpha$. Note that in a practical application of the test, the first stage will often be sufficient (see Section VII-D for details).

A. Delta method: Single ratio

From [40] we know that for a stationary or intrinsically stationary Gaussian field the large sample distribution of $\hat{\boldsymbol{\beta}} = \left[\hat{v}_{U,j,j'}^2, \hat{v}_{V,j',j}^2\right]^T$ is that of a normal with mean $\boldsymbol{\beta} = \left[v_{U,j,j'}^2, v_{V,j',j}^2\right]^T$ and whose covariance matrix can be deduced as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{U_{jj'}}^2 & \sigma_{U_{jj'}V_{j'j}} \\ \sigma_{U_{jj'}V_{j'j}} & \sigma_{V_{j'j}}^2 \end{bmatrix}.$$

The diagonal elements of $\boldsymbol{\Sigma}$ are the large sample variances of the wavelet variance estimators (Eq. (7)), while the off-diagonal elements are the large sample covariances between them. A general expression for the elements of $\boldsymbol{\Sigma}$ is given in Eq. (17) in Section IV-C. The delta method (see e.g., [13], pages 37–39) says that if the large sample distribution of $\hat{\boldsymbol{\beta}}$ is that of a normal with mean $\boldsymbol{\beta}$ and covariance matrix $\boldsymbol{\Sigma}$, then the large sample distribution of a nonlinear function $g(\hat{\boldsymbol{\beta}})$ is that of a normal with mean $g(\boldsymbol{\beta})$ and variance $\nabla g(\boldsymbol{\beta})^T \boldsymbol{\Sigma} \nabla g(\boldsymbol{\beta})$. Letting $g([x \ y]^T) = \log(x/y)$, we find that for large sample sizes $\log\left(\hat{v}_{U,j,j'}^2/\hat{v}_{V,j',j}^2\right)$ is approximately normal with mean $\log\left(v_{U,j,j'}^2/v_{V,j',j}^2\right)$ and variance

$$\frac{\sigma_{U_{jj'}}^2}{(v_{U,j,j'}^2)^2} - 2\frac{\sigma_{U_{jj'}V_{j'j}}}{v_{U,j,j'}^2 v_{V,j',j}^2} + \frac{\sigma_{V_{j'j}}^2}{(v_{V,j',j}^2)^2}. \tag{11}$$

A test for isotropy can now be designed for the single-ratio case:

$$H_0: \ \log\left(v_{U,j,j'}^2/v_{V,j',j}^2\right) = 0; \ H_1: \ \log\left(v_{U,j,j'}^2/v_{V,j',j}^2\right) \neq 0.$$

Under the null hypothesis the log-ratio is approximately normal with mean zero and variance given by Eq. (11), so using

$$\hat{\chi}^2 = \frac{v_{U,j,j'}^2 v_{V,j',j}^2 \log\left(\hat{v}_{U,j,j'}^2/\hat{v}_{V,j',j}^2\right)^2}{\sigma_{U_{jj'}}^2 - 2 \sigma_{U_{jj'}V_{j'j}} + \sigma_{V_{j'j}}^2} \xrightarrow{D} \chi_1^2 \tag{12}$$

as a test statistic, we find that, for significance level α, H_0 will be rejected when $\hat{\chi}^2 > q_{\chi_1^2}(\alpha)$, where $q_{\chi_1^2}(\alpha)$ denotes the upper (100α)th percentile of the chi-square distribution with one degree of freedom. In Eq. (12) we have used the fact that $v_{U,j,j'}^2 = v_{V,j',j}^2$ under the null hypothesis.

In a practical implementation of the test, the unknown quantities must be replaced by estimates. The development of these estimators is the subject of Section IV-C.

B. Delta method: Multiple ratios

We will first show how to develop the test based on two distinct ratios $\hat{v}_{U,j,j'}^2/\hat{v}_{V,j',j}^2$ and $\hat{v}_{U,k,k'}^2/\hat{v}_{V,k',k}^2$, and later describe how the method generalizes to include as many ratios as desired. First we define the four element vector $\hat{\boldsymbol{\beta}} = [\hat{v}_{U,j,j'}^2, \hat{v}_{V,j',j}^2, \hat{v}_{U,k,k'}^2, \hat{v}_{V,k',k}^2]^T$. Again we can utilize results from [40] to state that the large sample distribution of $\hat{\boldsymbol{\beta}}$ is that of a normal with mean $\boldsymbol{\beta}$ and covariance matrix

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} \sigma_{U_{jj'}}^2 & \sigma_{U_{jj'}V_{j'j}} & \sigma_{U_{jj'}U_{kk'}} & \sigma_{U_{jj'}V_{k'k}} \\ \sigma_{U_{jj'}V_{j'j}} & \sigma_{V_{j'j}}^2 & \sigma_{V_{j'j}U_{kk'}} & \sigma_{V_{j'j}V_{k'k}} \\ \sigma_{U_{jj'}U_{kk'}} & \sigma_{U_{kk'}V_{j'j}} & \sigma_{U_{kk'}}^2 & \sigma_{U_{kk'}V_{k'k}} \\ \sigma_{U_{jj'}V_{k'k}} & \sigma_{V_{k'k}V_{j'j}} & \sigma_{V_{k'k}U_{kk'}} & \sigma_{V_{k'k}}^2 \end{bmatrix}.$$

$\boldsymbol{\beta}$ is defined analogously to $\hat{\boldsymbol{\beta}}$ but with the actual variances in place of the estimates. Proceeding with the multivariate version of the delta method we find that, with $g(x) = \log(x)$, the large sample distribution of $\log(\hat{\boldsymbol{\beta}})$ is that of a multivariate normal with mean $\log(\boldsymbol{\beta})$ and covariance matrix $\boldsymbol{\Sigma}_2$ given by

$$\boldsymbol{\Sigma}_2 = \text{diag}(\boldsymbol{\beta} \oslash (-1))\boldsymbol{\Sigma}_1 \text{diag}(\boldsymbol{\beta} \oslash (-1)). \tag{13}$$

Here $\oslash$ denotes the operation of raising each element in the vector to the given power, and the diag operator creates a matrix with its vector input on the diagonal and zeros elsewhere. Next, introducing a differencing matrix $\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ gives us the vector of estimated log-ratios,

$$\hat{\boldsymbol{\theta}} = \mathbf{B}\log(\hat{\boldsymbol{\beta}}) = \left[\log(\hat{v}_{U,j,j'}^2/\hat{v}_{V,j',j}^2), \ \log(\hat{v}_{U,k,k'}^2/\hat{v}_{V,k',k}^2)\right]^T.$$

Using standard results from multivariate theory (see, e.g., [25]), we find that $\hat{\boldsymbol{\theta}}$ is approximately bivariate normal with mean $\boldsymbol{\theta} = \left[\log(v_{U,j,j'}^2/v_{V,j',j}^2), \ \log(v_{U,k,k'}^2/v_{V,k',k}^2)\right]^T$ and covariance $\boldsymbol{\Sigma} = \mathbf{B}\boldsymbol{\Sigma}_2\mathbf{B}^T = \begin{bmatrix} \Sigma_{j,j',j,j'} & \Sigma_{j,j',k,k'} \\ \Sigma_{j,j',k,k'} & \Sigma_{k,k',k,k'} \end{bmatrix}$ with elements

$$\Sigma_{j,j',k,k'} = \frac{\sigma_{U_{jj'}U_{kk'}}}{v_{U,j,j'}^2 v_{U,k,k'}^2} + \frac{\sigma_{V_{j'j}V_{k'k}}}{v_{V,j',j}^2 v_{V,k',k}^2} - \frac{\sigma_{U_{jj'}V_{k'k}}}{v_{U,j,j'}^2 v_{V,k',k}^2} - \frac{\sigma_{V_{j'j}U_{kk'}}}{v_{V,j',j}^2 v_{U,k,k'}^2}.$$

This allows us to formulate a test for isotropy which is simultaneous over two ratios,

$$H_0: \begin{bmatrix} \log(v_{U,j,j'}^2/v_{V,j',j}^2) \\ \log(v_{U,k,k'}^2/v_{V,k',k}^2) \end{bmatrix} = \mathbf{0}; \ H_1: \begin{bmatrix} \log(v_{U,j,j'}^2/v_{V,j',j}^2) \\ \log(v_{U,k,k'}^2/v_{V,k',k}^2) \end{bmatrix} \neq \mathbf{0},$$

where $\mathbf{0}$ is the null vector of length 2. Under the null hypothesis $E[\hat{\boldsymbol{\theta}}] = \mathbf{0}$, so again using standard results from multivariate theory,

$$\hat{\chi}^2 = \hat{\boldsymbol{\theta}}^T \boldsymbol{\Sigma}^{-1} \hat{\boldsymbol{\theta}} \tag{14}$$

is distributed as χ_2^2, the chi-square distribution with 2 degrees of freedom. Using $\hat{\chi}^2$ as the test statistic, we find that for significance level α, H_0 will be rejected when $\hat{\chi}^2 > q_{\chi_2^2}(\alpha)$, the upper (100α)th percentile of the χ_2^2 distribution.

To create a simultaneous test over any subset of ratios, the vector $\hat{\boldsymbol{\beta}}$ is constructed by placing pairwise the corresponding

estimates of the nominators and denominators from Eq. (9) and/or (10) for any choice of levels j, j'. With r ratios chosen, this defines $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1, \cdots, \hat{\beta}_{2r}]^T$ of length $2r$. The large sample distribution of $\hat{\boldsymbol{\beta}}$ is that of a normal with mean $\boldsymbol{\beta}$ and covariance matrix $\boldsymbol{\Sigma}_1$ of size $2r \times 2r$ defined in terms of the ratios used in $\hat{\boldsymbol{\beta}}$. $\boldsymbol{\beta}$ is defined analogously to $\hat{\boldsymbol{\beta}}$ but with the actual variances in place of the estimates. The delta method again results in the large sample distribution of $\log(\hat{\boldsymbol{\beta}})$ being that of a multivariate normal with mean $\log(\boldsymbol{\beta})$ and covariance matrix $\boldsymbol{\Sigma}_2$ given by Eq. (13). With the differencing matrix $\mathbf{B}$ from above expanded to size $r \times 2r$ in the obvious way, we find that the vector of estimated log-ratios, $\hat{\boldsymbol{\theta}} = \mathbf{B}\log(\hat{\boldsymbol{\beta}})$, is approximately multinormal with mean $\boldsymbol{\theta}$, the vector of r true log-ratios, and covariance $\boldsymbol{\Sigma} = \mathbf{B}\boldsymbol{\Sigma}_2\mathbf{B}^T$ of dimension $r \times r$ with elements

$$\Sigma_{s,t} = \frac{\sigma_{\beta_{2s}, \beta_{2t}}}{\beta_{2s}\beta_{2t}} + \frac{\sigma_{\beta_{2s-1}, \beta_{2t-1}}}{\beta_{2s-1}\beta_{2t-1}} - \frac{\sigma_{\beta_{2s}, \beta_{2t-1}}}{\beta_{2s-1}\beta_{2t-1}} - \frac{\sigma_{\beta_{2s-1}, \beta_{2t}}}{\beta_{2s-1}\beta_{2t}}.$$

Since the general term for $\Sigma_{s,t}$ should cover any conceivable combination of ww, ws, and sw-variances, we have to express it with dependence on the respective elements in $\boldsymbol{\beta}$. The hypothesis test is $H_0 : \boldsymbol{\theta} = \mathbf{0}$ versus $H_1 : \boldsymbol{\theta} \neq \mathbf{0}$, where $\mathbf{0}$ denotes the null vector of length r. The test statistic $\hat{\chi}^2$ from Eq. (14) now follows a chi-square distribution with r degrees of freedom. Thus, for significance level α, H_0 will be rejected when $\hat{\chi}^2 > q_{\chi_r^2}(\alpha)$.

Note that there is a large number of ratio subsets available to base a test on. Letting J denote the maximum horizontal and vertical scale index considered, there are $2^{(3J^2-J)/2} - 1$ distinct subsets that can be formed from $r = (3J^2 - J)/2$ distinct ratios (e.g., for J = 4 this gives $r = 22$ ratios and 4194303 possible subsets). In Section VI we will explore which subsets can reasonably be expected to be useful in a practical setting.

C. Estimating the covariances

With $C_{j,j'}$ and $D_{k,k'}$ denoting the coefficients $U_{j,j'}$, $V_{j,j'}, W_{j,j'}$, or $U_{k,k'}$, $V_{k,k'}$, $W_{k,k'}$, respectively, the isotropy test statistics involve both the unknown wavelet variances $v_{C_{j,j'}}^2$ or $v_{D_{k,k'}}^2$ and the unknown covariance between their estimators. In practice, these quantities must be estimated. In [40], estimators are developed for the variances, and for the variance of the variance estimators, but not for the covariance between the variance estimators (sw, ws and ww). Here we will derive the more general case of the large sample covariances between the variance estimators, from which the variance of the variance estimator is a special case.

Applying Eq. (7), $\text{cov}\left\{\hat{v}_{C,j,j'}^2, \hat{v}_{D,k,k'}^2\right\}$ equals

$$\text{cov}\left\{ \frac{1}{N_j M_{j'}} \sum_{a=L_j-1}^{N-1} \sum_{b=L_{j'}-1}^{M-1} C_{j,j',a,b}^2, \frac{1}{N_k M_{k'}} \sum_{c=L_k-1}^{N-1} \sum_{d=L_{k'}-1}^{M-1} D_{k,k',c,d}^2 \right\}$$
$$= \frac{1}{N_j M_{j'} N_k M_{k'}} \sum_{a=L_j-1}^{N-1} \sum_{b=L_{j'}-1}^{M-1} \sum_{c=L_k-1}^{N-1} \sum_{d=L_{k'}-1}^{M-1} \text{cov}\left\{ C_{j,j',a,b}^2, D_{k,k',c,d}^2 \right\}.$$

Since the coefficients $C_{j,j',a,b}$ and $D_{k,k',c,d}$ are zero mean normal random variables, we can use the Isserlis theorem [24],

from which it follows that

$$\text{cov}\left\{ C_{j,j',a,b}^2, D_{k,k',c,d}^2 \right\} = 2\text{cov}^2\left\{ C_{j,j',a,b}, D_{k,k',c,d} \right\}. \quad (15)$$

When the underlying image is an intrinsically stationary Gaussian field, the coefficients need not be zero mean for the Haar wavelet. In the case of a non-Gaussian field, a central limit argument says that Gaussianity becomes a more viable approximation as the filter lengths increase [36]. For this reason D(4) filters or longer might be preferred over the Haar wavelet when the image size permits it, and in general the results at higher scales can be trusted more than those of the lower scales. There is, however, a trade-off in practical applications between the validity of Eq. (15), favoring long filters and higher scales, and the fact that the number of coefficients available to estimate the quantities of interest decrease as a function of the filter length L and level j.

Denoting the cross-covariance function $s_{C_{j,j'}D_{k,k'}}(t,t') = \text{cov}\left\{ C_{j,j',u,v}, D_{k,k',u+t,v+t'} \right\}$, we may write $\text{cov}\{\hat{v}_{C,j,j'}^2, \hat{v}_{D,k,k'}^2\}$ as

$$\frac{2}{N^+ M^+} \sum_{t'=-(M^--1)}^{M^--1} \sum_{t=-(N^--1)}^{N^--1} \left(1 - \frac{|t'|}{M^-}\right)\left(1 - \frac{|t|}{N^-}\right) s_{C_{j,j'}D_{k,k'}}^2(t,t') + R, \quad (16)$$

where $N^+ = \max(N_j, N_k)$, $M^+ = \max(M_{j'}, M_{k'})$, $N^- = \min(N_j, N_k)$, and $M^- = \min(M_{j'}, M_{k'})$. The exact form of the remainder term R is intricate, but is similar for all cases. As $N \to \infty$ and $M \to \infty$ this term becomes negligible, and following the reasoning in [27], a Cesàro sum argument (see, e.g., [47]) allows the double summation in Eq. (16) to be approximated as

$$\sigma_{C_{j,j'}D_{k,k'}} = \frac{2}{N^+ M^+} \sum_{t=-\infty}^{\infty} \sum_{t'=-\infty}^{\infty} s_{C_{j,j'}D_{k,k'}}^2(t,t') = \frac{2A_{j,j',k,k'}}{N^+ M^+}, \quad (17)$$

where, by Parseval's theorem,

$$A_{j,j',k,k'} = \sum_{t=-\infty}^{\infty} \sum_{t'=-\infty}^{\infty} s_{C_{j,j'}D_{k,k'}}^2(t,t') = \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |S_{C_{j,j'}D_{k,k'}}(f,f')|^2 \, df df', \quad (18)$$

$S_{C_{j,j'}D_{k,k'}}(f,f')$ being the cross-spectral density between $C_{j,j'}$ and $D_{k,k'}$. Since the ws and sw coefficients are obtained by linear filtering operations on the same underlying field, their magnitude squared coherence is unity, so we can write

$$|S_{C_{j,j'},D_{k,k'}}(f,f')|^2 = S_{C_{j,j'},C_{j,j'}}(f,f')S_{D_{k,k'},D_{k,k'}}(f,f'),$$

with $S_{C_{j,j'},C_{j,j'}}(f,f')$ and $S_{D_{k,k'},D_{k,k'}}(f,f')$ denoting the spectral density functions (SDFs) of $C_{j,j'}$ and $D_{k,k'}$, respectively.

An estimator for the covariance in Eq. (17) can be found by replacing the cross-spectral density in Eq. (18) by a standard estimator. We use the cross-periodogram. With $\hat{S}_{\cdot,\cdot}^{(p)}(f,f')$ denoting the (cross-)periodogram, the identity

$$|\hat{S}_{C_{j,j'},D_{k,k'}}^{(p)}(f,f')|^2 = \hat{S}_{C_{j,j'},C_{j,j'}}^{(p)}(f,f')\hat{S}_{D_{k,k'},D_{k,k'}}^{(p)}(f,f')$$

follows immediately from basic definitions. Accordingly,

$$E\left\{ |\hat{S}_{C_{j,j'},D_{k,k'}}^{(p)}(f,f')|^2 \right\}$$
$$\approx S_{C_{j,j'},C_{j,j'}}(f,f')S_{D_{k,k'},D_{k,k'}}(f,f') + |S_{C_{j,j'},D_{k,k'}}(f,f')|^2$$
$$= 2|S_{C_{j,j'},D_{k,k'}}(f,f')|^2, \quad (19)$$

where the approximation follows from the asymptotic unbiasedness of the periodogram and from the fact that

$$\operatorname{cov}\left\{\hat{S}^{(p)}_{C_{j,j'},C_{j,j'}}(f,f'), \hat{S}^{(p)}_{D_{k,k'},D_{k,k'}}(f,f')\right\} \approx |S_{C_{j,j'},D_{k,k'}}(f,f')|^2$$

(see, e.g., [2, 42]). The (cross-)periodogram and the standard biased (cross-)covariance estimator

$$\hat{s}_{C_{j,j'}D_{k,k'}}(t,t') = \frac{1}{N^- M^-}\sum_u \sum_v C_{j,j',u,v} D_{k,k',t+u,t'+v}$$

are a Fourier pair. This fact, along with Equations (18) and (19) says that

$$\hat{A}_{j,j',k,k'} = \frac{1}{2}\sum_{t=-(N^--1)}^{N^--1}\sum_{t'=-(M^--1)}^{M^--1} \hat{s}^2_{C_{j,j'}D_{k,k'}}(t,t')$$

is an approximately unbiased estimator for $A_{j,j',k,k'}$. This makes $\hat{\sigma}_{C_{j,j'}D_{k,k'}} = \frac{2\hat{A}_{j,j',k,k'}}{N^+M^+}$ an asymptotically unbiased estimator for $\sigma_{C_{j,j'}D_{k,k'}}$.

V. RANDOM FIELDS USED IN SIMULATION STUDIES

For the simulation studies in Sections VI and VII we consider *exponential*, *spherical* and *fractional Brownian* fields,

$$\gamma_{X,\kappa_1,\kappa_2} = \lambda\left(1 - \phi^{||\boldsymbol{\kappa}||}\right), \tag{20}$$

$$\gamma_{X,\kappa_1,\kappa_2} = \begin{cases} \lambda\left(\frac{3}{2m}||\boldsymbol{\kappa}|| - \frac{1}{2m^3}||\boldsymbol{\kappa}||^3\right) & \text{if } 0 \le ||\boldsymbol{\kappa}|| \le m \\ \lambda & \text{otherwise,} \end{cases} \tag{21}$$

$$\gamma_{X,\kappa_1,\kappa_2} = \lambda||\boldsymbol{\kappa}||^{2H}, \tag{22}$$

with scale parameter λ and smoothness parameters ϕ, m or H, where $\phi, H \in (0,1)$ and $m > 0$. Higher values of the smoothness parameter result in smoother random fields. Eq. (22) characterizes fractional Brownian fields (FBFs), which are intrinsically stationary, whereas exponential and spherical fields are stationary and can be characterized equivalently via their covariance functions.

Throughout our study we generate samples using the R package RandomFields [45]. We consider random fields with scale $\lambda = 1$ that are weakly, moderately and strongly correlated; i.e., the respective smoothness parameters are $\phi = \{0.125, 0.5, 0.875\}$, $m = \{2,5,8\}$ and $H = \{0.125, 0.5, 0.875\}$. For random fields the *range* is defined as the minimum distance at which points become uncorrelated, and for fields where this happens only asymptotically, the *effective range* is defined as the minimum distance at which the covariance function reaches 0.05 times its value at lag zero. For the exponential fields the ϕ values above correspond to effective ranges $r = \{1.44, 4.32, 22.43\}$, while for the spherical fields m denotes the ranges. FBFs are not even asymptotically uncorrelated, so neither the range nor the effective range is well defined. Here the Hurst index H is linearly related to the fractal dimension D of FBFs via $D = 3 - H$ [8]. Fig. 1 shows example realizations of 128×128 random fields.

Geometric anisotropy is introduced by a rotation matrix $\mathbf{R}$ with rotation angle θ and an axis scaling matrix $\mathbf{S}$, i.e.,

$$\mathbf{R} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} a_u & 0 \\ 0 & a_v \end{bmatrix}.$$
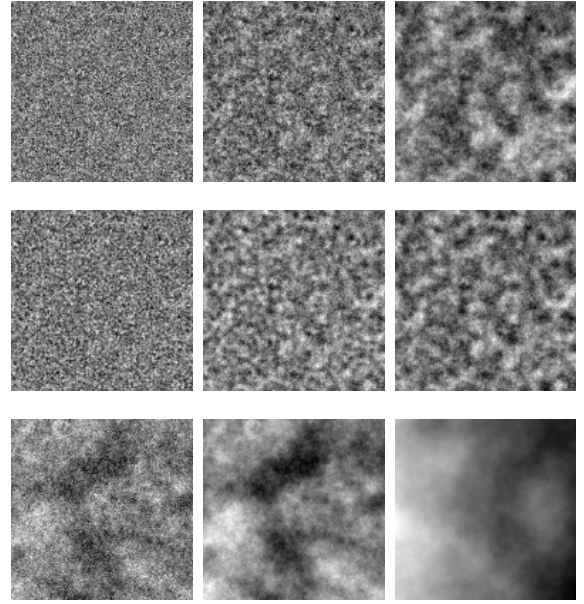

Figure 1. Realizations of random fields from Equations (20) – (22). Upper row: exponential fields with $\phi = \{0.125, 0.5, 0.875\}$, middle row: spherical fields with $m = \{2,5,8\}$, lower row: FBFs with $H = \{0.125, 0.5, 0.875\}$. For purposes of illustration, all realizations are generated starting with the same random seed, and all images have been normalized to the same range.
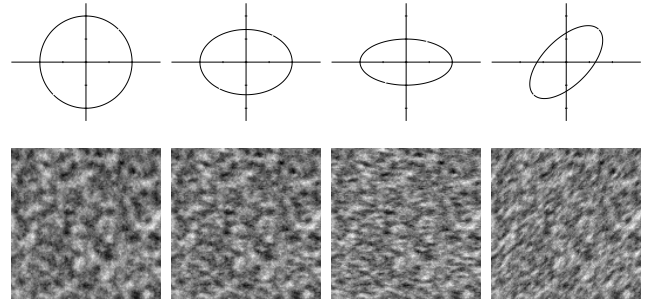

Figure 2. Realizations from spherical random fields with $m = 8$ corresponding to the constant semivariogram contours immediately above.

We replace $||\boldsymbol{\kappa}||$ in Equations (20) – (22) with $||\boldsymbol{\kappa}||_{\mathbf{B}} = \left(\sqrt{\boldsymbol{\kappa}^T \mathbf{B} \boldsymbol{\kappa}}\right)$, where $\mathbf{B} = \mathbf{R}^T \mathbf{S}^T \mathbf{S} \mathbf{R}$. Denote $\mathbf{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{B}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and $\mathbf{B}_3 = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Matrix $\mathbf{B}_1$ ensures isotropy, while $\mathbf{B}_2$ and $\mathbf{B}_3$ lead to anisotropy with ratio between major and minor anisotropy axis of $\sqrt{2}:1$ and $2:1$ respectively. For the anisotropic cases, we also consider four rotations of matrices $\mathbf{B}_2$ and $\mathbf{B}_3$, namely, $11.25°$, $22.5°$, $33.75°$ and $45°$. The lower row of Fig. 2 shows realizations of anisotropic fields with $m = 8$ on a 128×128 grid, with the corresponding elliptical constant semivariogram contours immediately above. For comparability, these realizations are generated with the same random seed as in Fig. 1.

VI. DISTRIBUTION UNDER THE NULL HYPOTHESIS

The tests developed in Section IV rely on asymptotic results both for finding the distribution of the test statistics under the null hypothesis and in finding unbiased estimators of the covariances needed to construct them. Here the validity of

these asymptotic results will be explored through simulations from the nine different isotropic random fields shown in Fig. 1, each involving 1000 realizations on a regular 512×512 lattice. We chose this image size to demonstrate the performance of the simultaneous test over multiple log-ratios and for longer filters. We only present the results for the D(4) wavelet with a maximal horizontal and vertical scale index of $J = 4$ (the Haar and LA(8) wavelets yield similar results, the only exception being the Haar-based test for FBFs, as discussed in Section IV-C).

We assess how well our test statistics agree with their theoretical distributions using Kolmogorov–Smirnov (K-S) tests with significance level $\alpha = 5\%$, and with visualization via histograms and Q-Q plots. Figures 3, 5 and 6 display these for the isotropic exponential random field with $\phi = 0.5$ (other random fields with different smoothness parameters have qualitatively similar results). In the Q-Q plots, there is a solid line with a slope of unity passing through the origin. A Q-Q plot fluctuating about this line indicates a good match between the samples and the theoretical distribution. If, however, the Q-Q plot follows a line with a different slope and/or does not pass through the origin, the samples are from a scaled and/or shifted version of the theoretical distribution. Scaling and shifting matter here since we wish to set thresholds for the test based on a theoretical asymptotic distribution, so deviations from the solid line are important. The dashed lines indicate the confidence bands for population distribution functions based upon the K-S goodness of fit test for a fully specified distribution. We do not reject the null hypothesis if the solid line lies entirely within the confidence bands.

As pointed out in Section IV-B, there is a prohibitive number of ratio subsets that can be formed from the full set of ratios from Equations (9) and (10), so it is not practical to comprehensively study all possibilities. Instead, we concentrate on the single-ratio and two-ratio tests; however, our study of two-ratio tests suggests multiple-ratio tests worthy of consideration. As a cautionary note, we also consider a test consisting of all possible ratios, for which it is unrealistic to expect asymptotic results to hold for sample sizes encountered in practice.

We begin by comparing the empirical distribution of the test statistic $\hat{\chi}^2$ in the single-ratio case (Eq. (12)) to its asymptotic χ_1^2 distribution. We do so by applying a level $\alpha = 5\%$ Kolmogorov–Smirnov goodness of fit test on all 22 single ratio test statistics calculated from 1000 samples from each random field above. There are 198 goodness of fit tests in all, out of which only 12 (6.1%) resulted in rejection, suggesting that the empirical distribution of $\hat{\chi}^2$ is well approximated by a χ_1^2 distribution. Even when the K-S test rejects, the visual fit is still good, as is evident from the worst case scenario shown in Fig. 3. Here we plot the histogram along with the probability density function (pdf) of the χ_1^2 distribution, as well as the Q-Q plot, for the test statistic with the lowest observed p-value ($p = 0.043$, which happened for $\hat{\chi}^2 = \log(\hat{v}_{W,1,3}^2/\hat{v}_{W,3,1}^2)$ for the exponential field with $\phi = 0.5$). While the Q-Q plot shows that the tail is not in keeping with the hypothesized pdf, the histogram fits nicely. Here the rejection level of the isotropy test is 4.0% for $\alpha = 5\%$, which indicates that the χ_1^2 distribution fits the empirical distribution even when the K-S
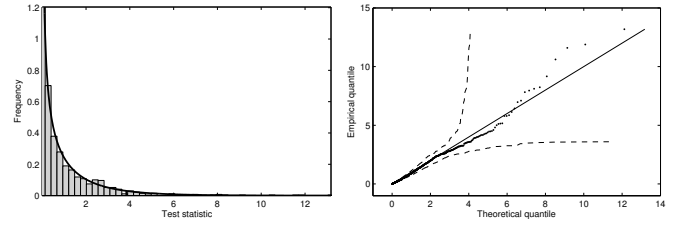

Figure 3. Left: Histogram of the test statistic based on the $\hat{v}_{W,1,3}^2/\hat{v}_{W,3,1}^2$ ratio, along with the theoretical χ_1^2 distribution for isotropic exponential fields of size 512×512 with $\phi = 0.5$. Right: Q-Q plot of the corresponding test statistic, where the theoretical distribution is represented by the solid line, while the sample quantiles are shown dotted and the two dashed lines display the upper and lower confidence band.

test rejects. The fit improves with increasing p-values.

Next, we consider the simultaneous test over two ratios. Using all possible ratio combinations for the nine random fields considered yields 2079 K-S tests, and out of these 212 (10.2%) are rejected. This is higher than to be expected if all test statistics were well described by a χ_2^2 distribution. In Fig. 4 the K-S test rejections are displayed in terms of the specific ratio combinations. A clear pattern emerges (particularly for the sw/ws ratios): the number of rejections is high when the wavelet filters in both ratios operate on the same scale $\tau_{j'}$ (with the exception of the lowest level $j' = 1$). The effect is compounded when one of the scaling filter indices is close or equal to j'. For a simultaneous test with two ratios, as many as ten covariances are estimated, and some of the estimates are highly correlated. Searching for the strongest correlations reveals that the correlation between $\hat{\sigma}_{U_{jj'}U_{jj'}}$ and $\hat{\sigma}_{U_{kj'}U_{kj'}}$ as well as between $\hat{\sigma}_{V_{j'j}V_{j'j}}$ and $\hat{\sigma}_{V_{j'k}V_{j'k}}$ stand out, i.e., the cases where the differencing is performed on the same scale $\tau_{j'}$. The correlation coefficient between $\hat{\sigma}_{V_{j'j}U_{jj'}}$ and $\hat{\sigma}_{V_{j'k}U_{kj'}}$ is also high in this case but the values are negligibly small compared to the above mentioned. We have also simulated from 4096×4096 exponential random fields. The correlations are still present, so for image sizes of practical interest this effect persists. Since the *inverse* of Σ is used for calculating the test statistic, it is a non-trivial task to determine how small errors in the covariance estimates (and also wavelet variance estimates) collectively propagate into the test statistic, even for the simplest simultaneous test using only two ratios; however, [30] illustrates that small deviations in dependent estimates can lead to large overall errors.

We can conclude that a simultaneous test should not include more than one sw/ws ratio operating on the same wavelet filter scale $\tau_{j'}$. Thus, with a maximum horizontal and vertical scale index of J, at most J sw/ws ratios should be combined in a simultaneous test. An obvious choice that fulfills this criterion uses only the diagonal $\hat{v}_{U,j,j}^2/\hat{v}_{V,j,j}^2$ ratios, here with $j = 1,...,4$. That this is a reasonable choice is supported by the K-S test not rejecting the test statistic being χ_4^2 distributed for any of the nine fields. Also, the visual fit in Fig. 5 is remarkably good, as evidenced by the Q-Q plot and the agreement between the histogram and the theoretical χ_4^2 distribution.

Another choice is to use all $\hat{v}_{W,j,j'}^2/\hat{v}_{W,j',j}^2$ ratios. In this case the K-S test rejects the hypothesis that the test statistic is χ_6^2 distributed for one out of the nine fields. While nine is
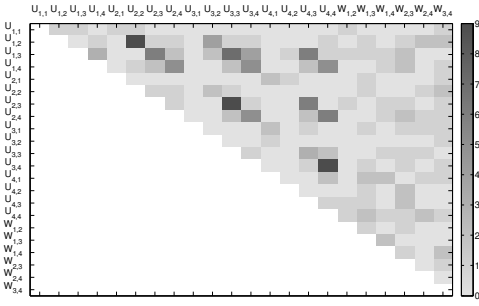
Figure 4. Number of K-S test rejections of two-ratio test statistics based on 1000 samples for each of the nine random fields. The point indexed $(U_{j,j'}, U_{k,k'})$, $(U_{j,j'}, W_{k,k'})$ or $(W_{j,j'}, W_{k,k'})$ refers to, respectively, the number of K-S rejections of the test statistic based on subset $\{\hat{v}^2_{U,j,j'}/\hat{v}^2_{V,j',j}, \hat{v}^2_{U,k,k'}/\hat{v}^2_{V,k',k}\}$, $\{\hat{v}^2_{U,j,j'}/\hat{v}^2_{V,j',j}, \hat{v}^2_{W,k,k'}/\hat{v}^2_{W,k',k}\}$ or $\{\hat{v}^2_{W,j,j'}/\hat{v}^2_{W,j',j}, \hat{v}^2_{W,k,k'}/\hat{v}^2_{W,k',k}\}$ being distributed as χ^2_2.
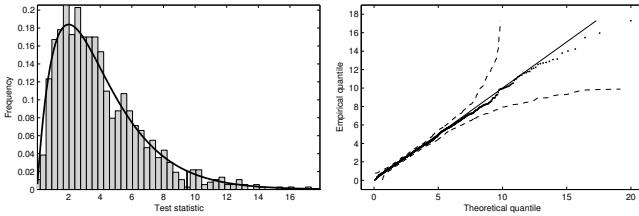

Figure 5. As Fig. 3, now based on all four diagonal sw/ws ratios, along with the theoretical χ^2_4 distribution.

a small number to base any conclusions on, a single rejection is not inconsistent with the claim that the distribution of the test statistic is well described by a χ^2_6 distribution. Choosing only four out of six ww/ww ratios gives 15 distinct ratio subsets. Applying the K-S test on the test statistics based on these for each of the nine fields, gives seven rejections out of 135 tests (5.2%). This indicates that the isotropy test based on ww/ww ratios performs well under the null hypothesis, particularly if the number of ratios included is limited. In general, caution should be exercised in including ratios in a simultaneous test. Consider, e.g., the test based on all diagonal $\hat{v}^2_{U,j,j}/\hat{v}^2_{V,j,j}$ ratios and all $\hat{v}^2_{W,j,j'}/\hat{v}^2_{W,j',j}$ ratios. Although these perform well separately, the K-S test rejects the resulting test statistic being χ^2_{10} distributed for seven out of the nine fields.

Finally, we consider a simultaneous test based on all 22 ratios available. The K-S test now rejects the null hypothesis for all nine fields. The upper row in Figure 6 shows the histogram of the test statistic along with the pdf of the theoretical χ^2_{22} distribution, as well as the Q-Q plot. The Q-Q plot and the mismatch between the histogram and the theoretical distribution both indicate that the distribution of the test statistic is not well approximated by a χ^2_{22} distribution. This is to be expected due to the large number of ratios, of which several have wavelet filters operating on the same scale. The poor performance is likely due to correlations amongst the covariance estimators, a view that is consistent with the bottom row in Fig. 6. Here we use the theoretical Σ in forming the test statistic $\hat{\chi}^2$ of Eq. (14) rather than estimating the components of Σ. There is a near perfect match between the samples and the theoretical χ^2_{22} distribution. The K-S test, in this artificial setting, does not reject any of the nine instances.
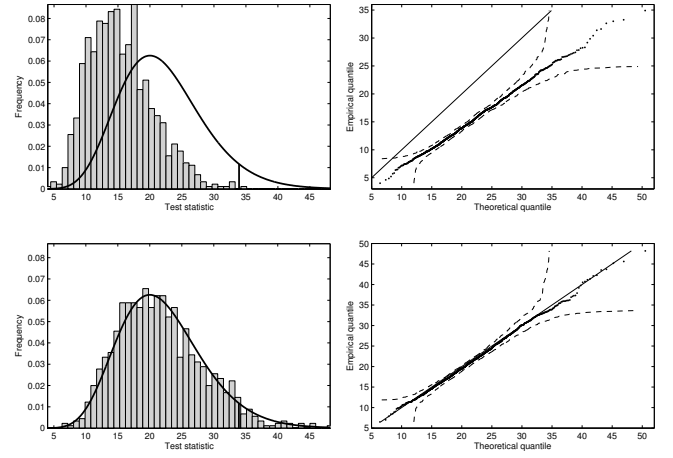

Figure 6. As Fig. 3, now based on all 22 ratios with the estimated (upper row) and theoretical covariances (lower row), along with the theoretical χ^2_{22} distribution.

In summary, the large-sample distribution of Eq. (14) is a reasonably accurate approximation for finite sample sizes of practical interest if care is taken in choosing the ratios involved. Replacing true covariance/variance values with estimates has a deleterious effect on the distribution of the test statistic; however, this effect is small if the number of ratios is limited and if use of subsets including more than one $\hat{v}^2_{U,j,j'}/\hat{v}^2_{V,j',j}$ with the same j' is avoided. In addition to single-ratio tests, we recommend using the simultaneous test based on either all diagonal sw/ws ratios or all ww/ww ratios (or any subset of these).

VII. RESULTS AND COMPARISON WITH OTHER METHODS

A. Study design

We set up the simulation study as follows. For each correlation structure defined in Equations (20) – (22) with parameters as given in Section V and (an)isotropy matrix $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$, we generate 1000 samples on $N \times N$ grids for $N = \{20, 40, 128\}$. For FBFs the largest grid will be $N = 512$. For $\mathbf{B}_2$ and $\mathbf{B}_3$ we also consider fields with $11.25°$, $22.5°$, $33.75°$ and $45°$ rotations of the major anisotropy axis. Due to space limitations we only present averages $\overline{\mathbf{B}}_2$ and $\overline{\mathbf{B}}_3$ over all five rotations in Tables I and II, but we do discuss the performance at different rotations in the text.

We compare the performance of the wavelet-based isotropy test with approaches described by Guan, Sherman and Calvin (GSC) [20], Lu and Zimmerman (LZ) [33], and Richard and Bierme (RB) [43]. We chose these three methods because, as opposed to simple qualitative visual aids, they are statistical tests especially developed for random fields on a regular two-dimensional lattice. The GSC and LZ tests are applicable to stationary random fields, while method RB is explicitly designed for FBFs. GSC use the property that under isotropy the values of the variogram at two different lags with the same Euclidean distance are equal. The LZ method [33] tests for a necessary condition for isotropy based on directional symmetries assessed by periodogram ratios using a Cramér-von Mises test statistic [9]. The RB method tests if the Hurst index H in FBFs is the same in different directions.

In our implementation of the other methods we have used a fixed parametrization based on the recommendations in the relevant papers. To not unfairly bias the comparison in our favor, we stick with the same wavelet-based test statistic for all fields and grid sizes. Specifically, we use the test statistic based on the first diagonal single level ratio $\hat{v}_{U,1,1}^2/\hat{v}_{V,1,1}^2$ and the D(4) wavelet. For images as small as 20×20 we cannot use more than a single ratio since the number of coefficients available to estimate the required variances and covariances is so low. We chose the D(4) wavelet because it is the smallest wavelet that works in the intrinsic case, (use of the Haar would likely be preferable for the smallest stationary fields because of the increase in the number of coefficients). For the image rotation necessary for the second stage of the test, the `Matlab` function `imrotate` is used. Here, bilinear and bicubic interpolation are both to be preferred over the nearest neighbor method [19], since the latter can distort the covariances in the original image. Bilinear and bicubic interpolation give, in our examples, practically identical results, so we have chosen bilinear interpolation for its speed and because it assures values inside the original range.

For a well tuned test, the *type I error*, i.e., rejection of the null hypothesis for isotropic fields, should correspond to the significance level α (set here to 5%). The *type II error*, i.e., not rejecting anisotropic fields, should ideally be zero. The *power* of the test (one minus the type II error) is ideally one.

B. Stationary case

Table I shows the rejection rates in the simulation study for the exponential and spherical fields presented in Section V. The first row in each block of rows corresponds to our wavelet-based isotropy test; the second row, to the GSC method [20]; and the third, to method LZ [33]. For the isotropic case $\mathbf{B}_1$, numbers in bold reflect that the anticipated $\alpha = 5\%$ significance level is within the 95% confidence region of the rejection rate. The confidence region is based on the Clopper–Pearson method for a binomial confidence interval with trial parameter 1000 and "success" probability equaling the number of rejections. Table I shows that our test has the 5% point within its confidence region in all cases except for the spherical field with moderate correlation $m = 5$ on the smallest grid size $N = 20$, where it is slightly above the bounds (the upper limit is 6.44%). For GSC the 5% point is outside the binomial confidence bounds for the type I error in 4 of 9 grid and parameter choices, both for the exponential and spherical fields. With 12.4% and 12.6% rejection of isotropic exponential and spherical fields, respectively, the type I error is particularly high for strongly correlated samples on the smallest grids. On larger grids ($N = 128$) the type I error shrinks, with a tendency to be smaller than the significance level. The LZ test is outside the confidence bounds in 3 of 9 cases for both types of fields. Interestingly, isotropic exponential fields with a high level of smoothness (i.e., $\phi = 0.875$) lead to a rejection rate $> 10\%$ no matter the grid size.

Our wavelet-based test correctly rejects anisotropic fields more often for cases with larger grids and/or with higher smoothness parameters. Already on a 40×40 grid high rejection rates are achieved for medium and strongly correlated and

– in the spherical case – even for the weakly correlated fields, while on a 128×128 grid all those realizations are correctly rejected. The test performs better on the anisotropic $\mathbf{B}_3$ fields of different rotations than on the corresponding $\mathbf{B}_2$ fields. In general, the results become progressively worse as the rotation approaches $45°$. None of these results are surprising. For larger grids, more data are available, so the performance of the test statistic should improve. Grid size is part of the reason why performance deteriorates with increasing rotation: the closer the rotation is to $45°$, the more we have to rely on the second stage of the test which operates on a rotated image of smaller size $N^* \times N^*$, with $N^* = \lfloor N/\sqrt{2} \rfloor$. While errors introduced by the interpolation might affect the results, high rejection rates achieved for the rotated versions on larger grids indicate that the interpolation is not adversely affecting the properties of the resulting fields. That the test performs better on $\mathbf{B}_3$ fields than on $\mathbf{B}_2$ fields is to be expected by considering Fig. 2. The larger the ratio between major and minor axis of the constant semivariogram contours, the larger the difference will be between $v_{U,j,j'}^2$ and $v_{V,j',j}^2$, and between $v_{W,j,j'}^2$ and $v_{W,j',j}^2$ as calculated by Equations (4) – (6) using the corresponding semivariograms (only the $v_{U,1,1}^2/v_{V,1,1}^2$ ratio is used in Table I, but the statement holds in general). This causes the log-ratios to deviate more from zero in the $\mathbf{B}_3$ case, representing a more pronounced violation of the null hypothesis. The test should perform worse for weakly correlated fields as there is little structure there to begin with (see the first column of Fig. 1). In particular, the exponential field with $\phi = 0.125$ is visually very similar to random noise. Hence, introducing stretching and rotation to the covariance structure of the weakly correlated fields does not affect the resulting fields nearly as much as for the moderately and strongly correlated fields, with predictable results for the power of the test.

For the anisotropic fields based on $\mathbf{B}_2$, the performance of the wavelet-based approach is comparable with the GSC test and has power consistently higher than or equal that of the LZ method. With rotation of the major anisotropy axis, the power of all tests decreases, particularly on smaller grids and for weakly correlated random fields. This effect persists for the more pronounced anisotropic fields based on $\mathbf{B}_3$.

C. Intrinsically stationary case

Table II shows the rejection rates in the simulation study for the FBFs presented in Section V. The first row in each block of rows corresponds to our wavelet-based isotropy test, the second row to the RB method [43]. Our method performs on par with the stationary case. The type II error vanishes quickly with increasing grid size and correlation. The significance level of the test is within the binomial confidence bounds of the rejection rates for all considered grid sizes and smoothness parameters, with the exception of one case ($N = 40$ and $H = 0.875$), where the rejection rate is slightly above the upper confidence bound. On the other hand, on small grids the RB method has rejection rates for anisotropic FBFs which are barely above the type I error of the isotropic case. Basically, for $N = 20$ and 40 no difference in the results for isotropic and anisotropic fields is visible. Since in [43] the examples are on

Table I
REJECTION RATES FOR ISOTROPIC ($\mathbf{B}_1$) AND ANISOTROPIC ($\overline{\mathbf{B}}_2\&\overline{\mathbf{B}}_3$) EXPONENTIAL AND SPHERICAL FIELDS OF SIZE N. FIRST ROW IN EACH BLOCK: METHOD FROM THIS PAPER, SECOND ROW: METHOD GSC [20], THIRD ROW: METHOD LZ [33]. BOLD NUMBERS FOR $\mathbf{B}_1$ INDICATE TYPE I ERROR WITHIN THE 95% BINOMIAL CONFIDENCE REGION, WHILE FOR $\overline{\mathbf{B}}_2\&\overline{\mathbf{B}}_3$ THEY INDICATE THE METHOD WITH THE HIGHEST POWER.

		$N=20$			$N=40$			$N=128$		
		$\mathbf{B}_1$	$\overline{\mathbf{B}}_2$	$\overline{\mathbf{B}}_3$	$\mathbf{B}_1$	$\overline{\mathbf{B}}_2$	$\overline{\mathbf{B}}_3$	$\mathbf{B}_1$	$\overline{\mathbf{B}}_2$	$\overline{\mathbf{B}}_3$
Exponential, $\phi=$	0.125	**4.4**	8.4	**12.7**	**4.8**	**23.9**	**39.3**	**4.1**	82.6	91.1
		6.1	**9.1**	11.6	4.0	16.3	28.5	3.2	**90.4**	**98.0**
		5.0	7.4	9.0	4.9	15.0	24.0	**4.6**	77.5	89.5
	0.500	**5.0**	**37.2**	72.2	**5.4**	87.3	99.6	**5.4**	100	100
		6.7	33.7	**73.6**	4.4	**93.0**	**100**	3.8	100	100
		5.4	25.4	61.5	4.6	79.3	99.7	**5.7**	100	100
	0.875	**5.2**	**56.1**	95.4	3.8	**98.1**	100	**6.2**	100	100
		12.4	52.6	**95.8**	**5.3**	**99.5**	100	3.0	100	100
		10.5	34.1	79.1	10.7	89.1	100	10.7	100	100
Spherical, $m=$	2	**5.6**	**36.3**	**58.6**	**5.2**	79.2	84.1	**5.0**	100	100
		8.8	33.6	49.0	4.9	**89.9**	**89.1**	3.7	100	100
		5.0	27.9	45.5	5.2	78.3	80.7	4.3	100	100
	5	**6.5**	**61.8**	**97.1**	**5.6**	98.5	100	**5.4**	100	100
		12.1	53.2	95.5	5.1	**99.8**	100	4.4	100	100
		6.7	38.6	85.4	6.1	94.8	100	5.3	100	100
	∞	**6.1**	**61.2**	**96.6**	**5.9**	98.7	100	**5.2**	100	100
		12.6	53.3	96.3	7.1	**99.4**	100	**4.9**	100	100
		9.0	34.6	84.2	10.2	92.1	100	**6.3**	100	100

Table II
AS TABLE I, NOW FOR FBFS, COMPARING THE WAVELET-BASED TEST (FIRST ROW IN EACH BLOCK) TO METHOD RB [43] (SECOND ROW).

		$N=20$			$N=40$			$N=512$		
		$\mathbf{B}_1$	$\overline{\mathbf{B}}_2$	$\overline{\mathbf{B}}_3$	$\mathbf{B}_1$	$\overline{\mathbf{B}}_2$	$\overline{\mathbf{B}}_3$	$\mathbf{B}_1$	$\overline{\mathbf{B}}_2$	$\overline{\mathbf{B}}_3$
FBFs, $H=$	0.125	**4.6**	**9.4**	**23.0**	**3.7**	**27.4**	**72.3**	**5.1**	**100**	**100**
		10.1	8.7	10.9	10.6	10.9	13.3	7.1	21.7	48.2
	0.500	**5.9**	**58.9**	**95.6**	**5.1**	**98.5**	**100**	**4.5**	**100**	**100**
		4.5	5.2	7.4	4.1	6.0	8.9	2.9	12.5	48.9
	0.875	**6.3**	**83.8**	**99.8**	**6.7**	**99.96**	**100**	**6.3**	**100**	**100**
		2.6	2.0	2.3	1.2	1.7	2.4	0.4	1.1	10.0

Table III
REJECTION RATES BASED ON 1000 REALIZATIONS OF ANISOTROPIC RANDOM FIELDS OF SIZE 256×256 WITH ROTATION OF THE MAJOR ANISOTROPY AXIS OF $42°$, $43°$ AND $44°$.

		$\mathbf{B}_2$			$\mathbf{B}_3$		
		$42°$	$43°$	$44°$	$42°$	$43°$	$44°$
Exp.	$\phi=0.125$	25.6	14.9	7.8	31.9	17.7	6.8
	$\phi=0.500$	96.3	70.1	23.6	100	98.4	57.0
	$\phi=0.875$	99.8	90.1	34.0	100	100	85.6
Sph.	m=2	96.9	76.1	24.3	100	96.0	46.5
	m=5	99.7	89.8	39.4	100	100	88.9
	m=8	99.8	88.0	37.0	100	100	87.1
FBFs	$H=0.125$	29.4	15.7	9.1	73.1	38.9	14.5
	$H=0.500$	99.4	90.7	34.6	100	100	86.8
	$H=0.875$	100.0	99.1	56.3	100	100	98.0

the 512×512 case, we used thresholds given in [43], which were calculated in an manner analogous to what we described above, but only for fields with Hurst index ranging from 0.1 to 0.7 (this might explain why the number of rejections for isotropic fields with $H=0.5$ is not within the confidence bounds for the given significance level).

D. Practical considerations

In this paper we propose a multiscale test for isotropy of random fields, but, to capture the case of geometric anisotropy oriented $45°$ to the axes, we must test both the original image and one with a $45°$ rotation. This two-stage procedure is not without problems, since the second stage uses a cropped and interpolated version of the original image. This can cause problems in interpreting results on different scales, as well as potentially distorting the structure of the underlying field (our simulation study in Sections VII-B and VII-C indicate that in practice the interpolation does not have a detrimental influence, but there is no theoretical support for this claim).

In Table III we investigate how close to $45°$ the anisotropy axis has to be for our test to fail. The table shows the rejection rates for 1000 realizations of anisotropic exponential, spherical and fractional Brownian fields of size 256×256 with $42°$, $43°$ and $44°$ anisotropy axes. We follow the setup of the simulation study and use the D(4) wavelet and the $\log(\hat{v}^2_{U,1,1}/\hat{v}^2_{V,1,1})$ test statistic. The test correctly rejects close to 100% of the anisotropic fields with a rotation angle of $42°$, and for more pronounced anisotropy $\mathbf{B}_3$ also $43°$, with the exception of weakly correlated exponential and fractional Brownian fields (i.e., $\phi=H=0.125$). Thus, for sufficiently large images, the power of the first stage is high enough to capture anisotropy oriented close to $45°$ for all but weakly correlated fields. In practical applications, it is unlikely for structures to be oriented exactly on $45°$ axes, so performing a single stage test should usually suffice. Testing only the original image has a computational advantage but, more importantly, the results on different scales should be easier to interpret.

VIII. PAPER DENSITY EXAMPLE

In this section we apply our wavelet-based isotropy test to a real-world example using relative mass density plots of paper handsheets with size 256×256 from the School of Environmental and Forest Sciences, University of Washington. Fig. 7 shows two examples of this type of image, which is used

a 512×512 grid, we also include this case in our comparison. Here the type II error further shrinks, but nevertheless the method is clearly worse than the wavelet-based approach.

It is not surprising that the RB method has problems in detecting fields with rotated major anisotropy axis. Of the several test statistics proposed in [43], there is one that in principle should be able to capture rotation. We looked into this test statistic, but the results were worse than those for the RB test statistic reported in Table II. The same two-stage approach we used in the wavelet-based test might improve the RB test, but this modification was not explored in [43].

The variance of the RB test statistic, and hence the rejection bounds for a chosen size α of the test, cannot be estimated on an image to image basis. For a particular image size simulations from a large number of isotropic fractional Brownian fields with different Hurst indices are required to determine the threshold for the test. For generating Table II we determined the thresholds by sampling 1000 isotropic fields for 9 values of the Hurst index H ranging from 0.1 to 0.9 in steps of 0.1. In particular, for grids with $N=\{20,40\}$, we calculated the standard deviations of the test statistic for each of the 9 indices and used the mean of the standard deviations to set the threshold. FBFs with $H=0.5$ (the mean Hurst index) are close to the significance level $\alpha=5\%$, whereas the type I error increases for smaller H and decreases for larger H, in both cases outside the binomial confidence bounds. For

Table IV
P-VALUES FOR SINGLE RATIO TESTS STATISTICS ON IMAGES A (UPPER ROW IN EACH BLOCK) AND B (LOWER ROW) FROM FIG. 7. BOLDFACE INDICATES REJECTION FOR $\alpha = 5\%$.

$\hat{v}^2_{U,1,1}/\hat{v}^2_{V,1,1}$	$\hat{v}^2_{U,1,2}/\hat{v}^2_{V,2,1}$	$\hat{v}^2_{U,1,3}/\hat{v}^2_{V,3,1}$	$\hat{v}^2_{U,1,4}/\hat{v}^2_{V,4,1}$	$\hat{v}^2_{U,1,5}/\hat{v}^2_{V,5,1}$
0.18	**< 0.0001**	**< 0.0001**	**0.0005**	0.38
0.018	**0.035**	0.26	0.80	0.12

$\hat{v}^2_{U,2,1}/\hat{v}^2_{V,1,2}$	$\hat{v}^2_{U,2,2}/\hat{v}^2_{V,2,2}$	$\hat{v}^2_{U,2,3}/\hat{v}^2_{V,3,2}$	$\hat{v}^2_{U,2,4}/\hat{v}^2_{V,4,2}$	$\hat{v}^2_{U,2,5}/\hat{v}^2_{V,5,2}$
0.28	**< 0.0001**	**< 0.0001**	**0.0002**	0.42
0.032	**0.032**	0.22	0.60	0.12

$\hat{v}^2_{U,3,1}/\hat{v}^2_{V,1,3}$	$\hat{v}^2_{U,3,2}/\hat{v}^2_{V,2,3}$	$\hat{v}^2_{U,3,3}/\hat{v}^2_{V,3,3}$	$\hat{v}^2_{U,3,4}/\hat{v}^2_{V,4,3}$	$\hat{v}^2_{U,3,5}/\hat{v}^2_{V,5,3}$
0.86	**< 0.0001**	**< 0.0001**	**< 0.0001**	0.19
0.039	**0.004**	0.20	0.40	0.13

$\hat{v}^2_{U,4,1}/\hat{v}^2_{V,1,4}$	$\hat{v}^2_{U,4,2}/\hat{v}^2_{V,2,4}$	$\hat{v}^2_{U,4,3}/\hat{v}^2_{V,3,4}$	$\hat{v}^2_{U,4,4}/\hat{v}^2_{V,4,4}$	$\hat{v}^2_{U,4,5}/\hat{v}^2_{V,5,4}$
0.84	**< 0.0001**	**< 0.0001**	**< 0.0001**	0.063
0.13	**0.006**	0.22	0.43	0.19

$\hat{v}^2_{U,5,1}/\hat{v}^2_{V,1,5}$	$\hat{v}^2_{U,5,2}/\hat{v}^2_{V,2,5}$	$\hat{v}^2_{U,5,3}/\hat{v}^2_{V,3,5}$	$\hat{v}^2_{U,5,4}/\hat{v}^2_{V,4,5}$	$\hat{v}^2_{U,5,5}/\hat{v}^2_{V,5,5}$
0.31	**0.0002**	**< 0.0001**	**< 0.0001**	**0.022**
0.064	**0.016**	0.12	0.30	0.13

$\hat{v}^2_{W,1,2}/\hat{v}^2_{W,2,1}$	$\hat{v}^2_{W,1,3}/\hat{v}^2_{W,3,1}$	$\hat{v}^2_{W,1,4}/\hat{v}^2_{W,4,1}$	$\hat{v}^2_{W,1,5}/\hat{v}^2_{W,5,1}$	$\hat{v}^2_{W,2,3}/\hat{v}^2_{W,3,2}$
< 0.0001	**< 0.0001**	0.091	0.22	**< 0.0001**
0.004	0.18	**0.013**	0.86	0.56

$\hat{v}^2_{W,2,4}/\hat{v}^2_{W,4,2}$	$\hat{v}^2_{W,2,5}/\hat{v}^2_{W,5,2}$	$\hat{v}^2_{W,3,4}/\hat{v}^2_{W,4,3}$	$\hat{v}^2_{W,3,5}/\hat{v}^2_{W,5,3}$	$\hat{v}^2_{W,4,5}/\hat{v}^2_{W,5,4}$
0.0002	**0.001**	**< 0.0001**	**< 0.0001**	0.13
0.14	0.12	0.86	0.59	0.53

Table V
P-VALUES OF SIMULTANEOUS TEST STATISTICS BASED ON $\hat{v}^2_{U,j,j}/\hat{v}^2_{V,j,j}$ SUBSETS. BOLDFACE INDICATES REJECTION FOR $\alpha = 5\%$.

	$j=\{1,2,3,4,5\}$	$j=\{2,3,4,5\}$	$j=\{3,4,5\}$	$j=\{4,5\}$
A	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
B	0.07	0.08	0.20	0.14

for assessing the quality of paper handsheets, where patchiness is particularly of interest. The paper handsheets are made by stirring a slurry of water and wood in a cylindrical vessel. On the bottom of the vessel there is a trap door, which lets the water flow out through a screen collecting the fibers. The wet mats are taken off the screens and squeezed before being stacked and dried on clean metal plates. The images A and B from Fig. 7 are the result of placing the sheets on a film which is sensitive to radiation, and covering them by a plate with an embedded uniform distribution of radioactive elements with low emission rates, followed by developing and digitizing the film negatives. The thickness of the sheets changes from point to point, so each pixel in the images represents a value for the mass per area, also known as basis weight. Visible patches are clumps of fibers, where the fiber length is on the order of a millimeter. Ideally, there should be no directionality in these images; i.e., the mass distribution and thus the thickness of the sheets should be isotropic throughout the image. Hence, if the paper density plot is anisotropic, it might indicate a need to look at the manufacturing process. We use the wavelet-based isotropy test to test this hypothesis.

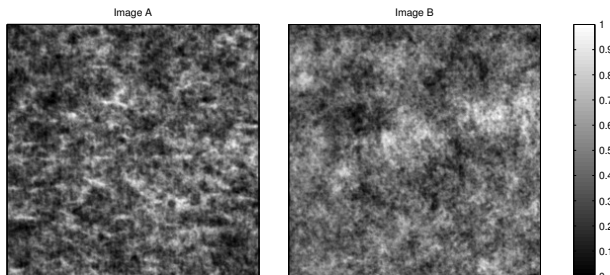Following Sections VII-D and IV-C, for 256×256 images, it suffices to perform tests on the original image and to choose

Figure 7. Images of size 256×256 showing the relative mass density of paper handsheets.

the D(4) wavelet using levels up to and including $J = 5$. The tests require that the wavelet coefficients be zero mean normal random variables. Since the images considered here are well modeled as stationary fields, we choose to histogram equalize them to standard normals to ensure this requirement. This will not effect any directionalities present in the image.

First we apply the single ratio test of Section IV-A. As we consider $J = 5$ levels, there are 35 possible single ratios. The p-values of their respective test statistics are displayed in Table IV and presented in boldface if they are below the significance level of $\alpha = 5\%$. The test rejects the isotropy hypothesis for 23 out of 35 single ratios for image A from Fig. 7, while only 10 out of 35 ratios indicate anisotropy for paper sheet B. The results for ratios involving the unit-level wavelet filters seem to contradict those of other ratios. A wavelet filter of level $j' = 1$ only differences neighboring pixel values so, if the scaling filters also operate on the lower scales, this might simply be an effect of noise. The effect persists for scaling filters up to level three, so it is hard to say if the results point towards there being (an)isotropies present at the smallest scales, or if it is simply the effect of random noise. When disregarding the ratios involving unit level wavelet and/or scaling filters, 18 out of 22 possible ratios lead to rejection for the first image, but only 4 for the second one.

Anisotropy in image A is most significant for tests based on any single ratios $\hat{v}^2_{U,j,j'}/\hat{v}^2_{V,j',j}$ or $\hat{v}^2_{W,j,j'}/\hat{v}^2_{W,j',j}$ with $j' = 3$. Here the differencing operation is performed on rows and/or columns (depending on the chosen ww, ws or sw coefficients) corresponding to scale $\tau_3 = 4$ pixels, on column/row averages on scale $2\tau_j = 2^j$ pixels ($\tau_j = 2^{j-1}$ for the ww case). Image A has a dominance of horizontal patches with thickness close to four pixels, longer ones across the center of the image and many shorter ones throughout the image. These are likely causes for the anisotropy, but are not evident in the second image, where the p-values of the corresponding test statistics do not indicate anisotropy.

Having established (an)isotropy on a scale by scale basis, let us now consider if the images are isotropic over several scales simultaneously. Following the recommendation from Section VI, we look at subsets consisting of diagonal sw/ws ratios, i.e., $\hat{v}^2_{U,j,j}/\hat{v}^2_{V,j,j}$ for a set of levels j. The p-values displayed in Table V indicate that the first image is significantly anisotropic, while evidence for anisotropy in the second is lacking.

The anisotropy of the first handsheet might be due to a directed flow field as the water drains through the screen. If clumps of fibers in the slurry stick to the screen, it may cause the fluid slurry to move sideways to less resistant flow pathways, thereby causing a distorted structure. This might explain the apparent dominance of horizontal patches in image A, though a well controlled experiment is needed to test this further. The relative mass density values of the second

Table VI
EVALUATION OF NULL HYPOTHESIS OF ISOTROPY FOR 106 MAMMOGRAM
IMAGES USING RICHARD–BIERME (RB) [43] AND WAVELET-BASED
TESTS.

		wavelet-based test	
	rejection	no	yes
RB	no	4 (3.8%)	60 (56.6%)
	yes	4 (3.8%)	38 (35.8%)

image appear to be distributed as desired in theory; i.e., our simultaneous tests do not suggest directionality in this image.

IX. MAMMOGRAM EXAMPLE

As a second example we apply our proposed test to mammogram density images. These images have been modeled as isotropic fractional Brownian fields, which are a subclass of intrinsically stationary random fields (see, e.g., [4, 21, 22, 28]). In Section VII we compared the Richard–Bierme (RB) test for isotropy to our wavelet-based test. The paper introducing the RB test [43] applied it to a set of mammograms and concluded that it is more appropriate to model mammograms as *Extended* Fractional Brownian Fields, a more general class of FBFs that allows anisotropic features to be modeled. In this section we compare the performance of our wavelet-based test with the RB test using the same mammogram images as used in [43]; however, it was not possible to use the same regions of interest, i.e., the manually chosen 512×512 regions from the mammograms, as in the original paper. In addition, we have excluded images that are contaminated with large areas consisting of just zeros, thus leaving a total of 106 images from the original 116. The rejection rates reported in [43] for the RB test therefore do not apply here, but we have recalculated this test statistic based on code kindly supplied to us by the authors. For simplicity (and in contrast to the paper density image example), we here concentrate on a single scale combination in the wavelet-based test. As per Section VI, we focus on the diagonal sw/ws ratios using the D(4) wavelet. Since the lowest scales are more sensitive to noise and the highest suffer from a paucity of coefficients, a reasonable choice for the 512×512 images is to use diagonal scales $3 - 6$. For the RB test, we follow the authors' recommendation and use what they call their second test (with sampling factor $v = 2$) [43].

Table VI summarizes the results of the RB and wavelet-based tests. The overall rejection rates for the wavelet-based and RB tests are, respectively, 92.5% and 39.6% (the latter was about 60% in [43]). Our test rejects the isotropic assumption for more mammograms than the RB test does, thus strengthening the overall conclusion in [43] that isotropic FBF models are too simplistic. If the isotropic hypothesis is indeed untenable, the higher rate of rejection by the wavelet-based test is consistent with the simulation study in Section VII-C. Table VI gives the number of cases where the two tests agree or disagree. Since our method rejects more often, the largest category is when our test rejects, but RB test does not. Fig. 8 shows three examples of mammograms falling into each of the four categories (one column for each category). The first (last) two columns show images for which the wavelet-based test failed to reject (rejected) the null hypothesis of isotropy;
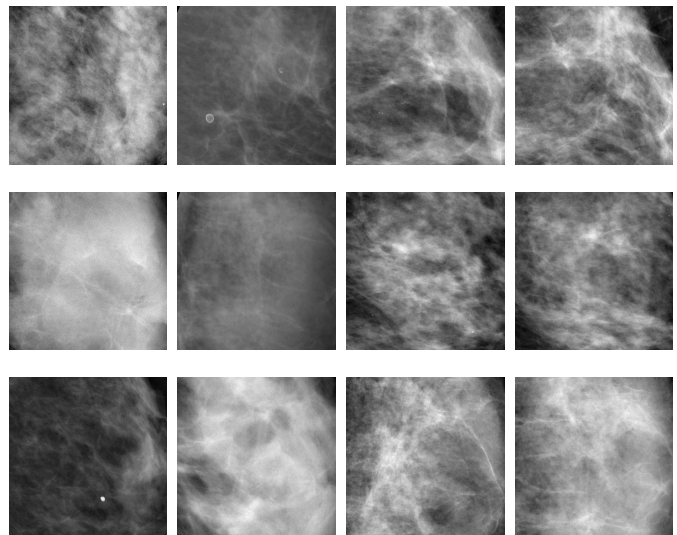

Figure 8. Three examples of mammogram images for each of the four categories from Table VI (one column for each category). First column: RB and wavelet-based tests both do not reject the null hypothesis of isotropy (i.e., rejection: no/no). Second column: yes/no. Third column: no/yes. Fourth column: yes/yes (i.e., both tests reject the null hypothesis and thus agree on anisotropy).

however, it is difficult to ascertain by visual inspection what is causing rejection/non-rejection, and it is also not clear what is causing disagreements between the two tests.

In conclusion, while the results from the RB and wavelet-based tests do not agree in fine detail, both tests reach the same conclusion about the efficacy of isotropic FBF models, namely, that the isotropic assumption is highly questionable. An open research question is whether the scaled-based nature of the anisotropy would be useful in either detecting or classifying tumors.

X. SUMMARY AND DISCUSSION

We have presented a novel method for testing for isotropy of random fields on a regular two-dimensional lattice. We demonstrated its performance on isotropic and anisotropic random fields of different types, parametrizations and sizes. The test consistently rejects close to the nominal level for isotropic fields and has a rejection rate comparable to or exceeding existing methods for anisotropic fields. For wavelets other than the Haar, the method applies, without modification, to intrinsically stationary fields. Since intrinsic models are commonplace (see, e.g., [6, 37]), this is a considerable advantage over existing methods which normally assume stationarity.

The method is computationally efficient, since the wavelet coefficients can be calculated efficiently [18] and the covariance estimates needed for the test can be calculated with the fast Fourier transform. Computational speed depends on the computer in use, image size, the number of ratios included, the length of the unit-level wavelet filter, and on which ratios are to be used in the test. As an example, for a 512×512 image, using the D(4) wavelet and including all permissible ratios up to level four, we need 63 seconds to calculate and perform the test on a Intel Core i7 Q 820 @ 1.73GHz computer with 8GB RAM running on Linux Ubuntu 12.10 (64-bit) with Matlab

version 7.12.0 (R2011a). Using only the diagonal sw/ws ratios along with the ww/ww ratios, this reduces to 25 seconds, and if using only the unit-level sw/ws ratio, 0.22 seconds.

The ability to handle intrinsically stationary fields is a strength of the proposed method, but care must be taken in interpreting the results. A property of intrinsically stationary random fields is that their mean is defined only up to the addition of a plane. Hence, the ability of the method to handle these cases also means that the method will not detect apparent directionality that is caused by a drift or linear trend in otherwise isotropic data when using unit-level wavelet filters longer than two. If it is of interest to capture this, the recommendation would be to use the Haar wavelet and sw/ws ratios. Depending on the slope of the drift, the higher level sw/sw ratios should still detect such a trend in the data. In fact, if using unit-level wavelet filters of length L, the method is essentially invariant to the addition of polynomials of order $d < L/2$ to the random field, so any polynomial trends of order d or lower will not be detected by the proposed methodology, no matter the effect it has on the image in terms of apparent directionalities. When deciding what wavelet type to use in a practical setting, this is yet another factor to consider. As noted in Section IV-C use of longer unit-level wavelet filters yields coefficients well modeled as zero mean normals, but at the price of a reduction in the number of coefficients available for parameter estimation. The additional desire to detect anisotropies due to higher order polynomials suggests avoiding unit-level wavelet filters longer than four, unless the model in question specifically requires it. The existence of a polynomial trend of order higher than unity can violate the assumption of both stationarity and intrinsic stationarity, so it is hardly surprising that the method cannot handle these cases. A full two-dimensional wavelet decomposition will also yield scaling-scaling coefficients which retain the polynomial information that is filtered out by the wavelet filter. These can conceivably be used for capturing polynomial directionalities, though this would require further theory to be developed.

In this work we have exclusively considered intrinsic stationarity of order one, since isotropy is normally defined in terms of the variogram, which is well defined only for intrinsic stationarity up to order one. In the literature on Gaussian Markov random fields it is common also to use models which are second order intrinsically stationary. While isotropy cannot be defined in terms of the variogram in this case, the concept of directionality of the field still makes sense, and the proposed method can capture directionalities that are not an effect of the mean being defined only up to the addition of a second order polynomial, except when using the Haar wavelet.

A strength of the proposed method is that isotropy can be assessed on a scale by scale basis, capturing if there is a significant difference in the amount of predominantly vertical structure compared to predominantly horizontal structure. Unfortunately, the test does not provide information on the direction of anisotropy. If this is of interest, other methods should be considered, e.g., the rose diagram.

The method allows for a large number of tests to be performed, so spurious findings due to the effects of multiple testing is an issue that must be considered. Hence, if many tests are performed, the level of the test should be adjusted by Bonferroni or some other procedure for correcting for multiple testing. One way of minimizing the risk of spurious findings might be to predetermine what scales (preferably only a few) are of interest for the application at hand. Once this has been decided, the first step would be to perform the multiple ratio test including all scales of interest. If there is no rejection here, there is no reason to proceed. If, however, the null hypothesis is rejected, one can perform the single ratio tests for all scales, and ascertain what scales are the most anisotropic.

For simplicity and computational reasons we used the periodogram in Section IV-C for estimating the covariances needed to construct the test. Other spectral estimators might improve the results, as the periodogram is known to be suboptimal. The increase in computational complexity is a drawback, but might be justified if only a few scales are of interest.

So far we have only applied the test globally over an entire image. If applied locally within an image, the test might be a useful texture measure for segmentation and/or classification. A simple approach is to apply the method as is on subimages, doing the wavelet transform on each subimage. A more appealing approach is to calculate the wavelet coefficients over the entire image, and base the local isotropy tests on these, but this would require the development of further theory, as the method developed herein only applies to full images.

XI. Acknowledgments

References

[1] F. Berizzi, G. Bertini, M. Martorella, and M. Bertacca. Two-dimensional variation algorithm for fractal analysis of sea SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2361–2373, 2006.

[2] P. Bloomfield. *Fourier analysis of time series: An introduction*. John Wiley & Sons, Inc., 2nd edition, 2000.

[3] B. Brunet-Imbault, G. Lemineur, C. Chappard, R. Harba, and C.-L. Benhamou. A new anisotropy index on trabecular bone radiographic images using the fast Fourier transform. *BMC Medical Imaging*, 5(1):4, 2005.

[4] A. E. Burgess, F. L. Jacobson, and P. F. Judy. Human observer detection experiments with mammograms and power-law noise. *Medical physics*, 28(4):419–437, 2001.

[5] E. M. Cabaña. Affine processes: A test of isotropy based on level sets. *SIAM Journal on Applied Mathematics*, 47(4):886–891, 1987.

[6] J. P. Chilès and P. Delfiner. *Geostatistics: Modeling spatial uncertainty*. John Wiley & Sons, New Jersey, 2nd edition, 2012.

[7] P. Collaboration. Planck 2013 results. xxiii. isotropy and statistics of the cmb. *arXiv preprint arXiv:1303.5083*, 2013.

[8] N. Cressie. *Statistics for spatial data*. John Wiley & Sons, New York, 1993.

[9] S. Csorgo and J. Faraway. The exact and asymptotic distributions of Cramér-von Mises statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):221–234, 1996.

[10] O. Daniell, Y. Petillot, and S. Reed. Unsupervised sea-floor classification for automatic target recognition. In *Proc. International Conference on Underwater Remote Sensing*, Brest, France, 2012.

[11] I. Daubechies. *Ten Lectures on Wavelets*. Cbms-Nsf Regional Conference Series in Applied Mathematics Series. Society for Industrial and Applied Mathematics (SIAM), 1992.

[12] S. Davies and P. Hall. Fractal analysis of surface roughness by using spatial data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(1):3–37, 1999.

[13] A. Davison. *Statistical models*. Cambridge University Press, 2003.

[14] I. A. Eckley, G. P. Nason, and R. L. Treloar. Locally stationary wavelet fields with application to the modelling and analysis of image texture. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 59(4):595–616, 2010.

[15] E. Fakiris, D. Williams, M. Couillard, and W. Fox. Sea-floor acoustic anisotropy and complexity assessment towards prediction of atr performance. In *Proc. International Conference on Underwater Acoustics*, pages 1277–1284, Corfu, Greece, 2013.

[16] M. Geilhufe. *Statistical Analysis of Lattice Data with Wavelet Variance Methods and Spatiotemporal Modeling of Infectious Disease Spread*. PhD thesis, University of Tromsø, Norway, 2013.

[17] M. Geilhufe and Ø. Midtgaard. Quantifying the complexity in sonar image for MCM performance estimation. In J. S. Papadakis and L. Bjørnø, editors, *Proc. 2nd International Conference and Exhibition on Underwater Acoustics*, pages 1041–1048, Rhodes, Greece, 2014.

[18] M. Geilhufe, D. B. Percival, and H. L. Stern. Two-dimensional wavelet variance estimation with application to sea ice SAR images. *Computers & Geosciences*, 54:351–360, 2013.

[19] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Pearson, 3rd edition, 2008.

[20] Y. Guan, M. Sherman, and J. A. Calvin. A nonparametric test for spatial isotropy using subsampling. *Journal of the American Statistical Association*, 99(467):810–821, 2004.

[21] J. J. Heine, S. R. Deans, R. P. Velthuizen, and L. P. Clarke. On the statistical nature of mammograms. *Medical physics*, 26(11):2254–2265, 1999.

[22] J. J. Heine and R. P. Velthuizen. Spectral analysis of full field digital mammography data. *Medical physics*, 29(5):647–661, 2002.

[23] E. H. Isaaks and R. M. Srivastava. *An Introduction to Applied Geostatistics*. Oxford University Press, 1989.

[24] L. 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.

[25] R. Johnson and D. Wichern. *Applied Multivariate Statistical Analysis, 6th International edition*. Prentice Hall, 2007.

[26] A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, London, 1978.

[27] M. Keim and D. Percival. Assessing characteristic scales using wavelets. *Journal of the Royal Statistical Society: Series C (Applied Statistics), in press*, 2015.

[28] P. Kestener, J. M. Lina, P. Saint-Jean, and A. Arneodo. Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms. *Image Anal. Stereol*, 20:169–174, 2001.

[29] H. R. Künsch. Intrinsic autoregressions and related models on the two-dimensional lattice. *Biometrika*, 74(3):517–524, 1987.

[30] H. R. Künsch. Dependence among observations: Consequences and methods to deal with it. In W. Stahel and S. Weisberg, editors, *Directions in Robust Statistics and Diagnostics, Part I*, volume 33 of *IMA Volumes in Mathematics and its Applications*, pages 131–140. Springer-Verlag, New York, 1991.

[31] R. M. Lark and R. Webster. Changes in variance and correlation of soil properties with scale and location: analysis using an adapted maximal overlap discrete wavelet transform. *European Journal of Soil Science*, 52(4):547–562, 2001.

[32] R. M. Lark and R. Webster. Analysing soil variation in two dimensions with the discrete wavelet transform. *European Journal of Soil Science*, 55(4):777–797, 2004.

[33] N. Lu and D. L. Zimmerman. Testing for directional symmetry in spatial dependence using the periodogram. *Journal of Statistical Planning and Inference*, 129:369–385, 2005.

[34] A. P. Lyons, W. L. Fox, T. Hasiotis, and E. Pouliquen. Characterization of the two-dimensional roughness of wave-rippled sea floors using digital photogrammetry. *IEEE Journal of Oceanic Engineering*, 27(3):515–524, 2002.

[35] A. Maity and M. Sherman. Testing for spatial isotropy under general designs. *Journal of Statistical Planning and Inference*, 142(5):1081 – 1091, 2012.

[36] C. L. Mallows. Linear processes are nearly Gaussian. *Journal of Applied Probability*, 4:313–329, 1967.

[37] G. Matheron. The intrinsic random functions and their applications. *Advances in Applied Probability*, 5(3):439–468, 1973.

[38] A. E. Milne, R. M. Lark, and R. Webster. Spectral and wavelet analysis of gilgai patterns from air photography. *Australian Journal of Soil Research*, 48:309–325, 2010.

[39] A. Molina and F. R. Feito. A method for testing anisotropy and quantifying its direction in digital images. *Computers & Graphics*, 26(5):771 – 784, 2002.

[40] D. Mondal and D. B. Percival. Wavelet variance analysis for random fields on a regular lattice. *IEEE Transactions on Image Processing*, 21(2):537–549, 2012.

[41] D. B. Percival and A. T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, New York, USA, 2000.

[42] M. B. Priestley. *Spectral Analysis and Time Series*. Academic Press, 1981.

[43] F. Richard and H. Bierme. Statistical tests of anisotropy for fractional Brownian textures. Application to full-field digital mammography. *Journal of Mathematical Imaging and Vision*, 36(3):227–240, 2010.

[44] L. Scaccia and R. Martin. Testing axial symmetry and separability of lattice processes. *Journal of Statistical Planning and Inference*, 131(1):19–39, 2005.

[45] M. Schlather. Simulation and analysis of random fields. *R News*, 1(2):18–20, 2001.

[46] M. Sherman. *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*. John Wiley & Sons, Ltd, Chichester, UK, 2010.

[47] E. C. Titchmarsh. *The theory of functions*. Oxford University Press, 1939.

[48] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, 4(11):1549–1560, 1995.

[49] D. L. Zimmerman. Another look at anisotropy in geostatistics. *Mathematical Geology*, 25:453–470, 1993.

Kevin Thon received the M.Sc. degree in physics in 2007, and the Ph.D. degree in statistics in 2013, both from the University of Tromsø, Norway. From 2014 he has been working at the Center for Clinical Documentation and Evaluation (SKDE) in Tromsø. His research interests include spatial statistics and medical image analysis.

Marc Geilhufe received his university diploma degree in business mathematics from Chemnitz University of Technology, Germany, in 2008, and the Ph.D. degree in statistics from the University of Tromsø, Norway, in 2013. His research interests include spatial statistics, image analysis and wavelets.

Donald B. Percival received the B.A. degree in astronomy from the University of Pennsylvania, Philadelphia, in 1968, the M.A. degree in mathematical statistics from George Washington University, Washington, DC, in 1976, and the Ph.D. degree in statistics from the University of Washington, Seattle, in 1982.

From 1968 to 1978, he was an astronomer with the Time Service Division, U.S. Naval Observatory, Washington, DC, where he worked on the generation of atomic clock time scales and on the analysis of the frequency stability of high-performance oscillators. Since 1982, he has been with the Applied Physics Laboratory and the Department of Statistics, University of Washington, where he currently holds the positions of Senior Principal Mathematician and Professor. He is the coauthor (with A.T. Walden) of textbooks "Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques" (Cambridge, U.K.: Cambridge Univ. Press, 1993) and "Wavelet Methods for Time Series Analysis" (Cambridge, U.K.: Cambridge Univ. Press, 2000). His research interests include spectral analysis, wavelets, and the use of statistical methodology in physical sciences.