

The locally stationary dual-tree complex wavelet model

J. D. B. Nelson¹ · A. J. Gibberd² · C. Naornita³ · N. Kingsbury⁴

Received: 15 October 2017 / Accepted: 20 October 2017 / Published online: 26 October 2017
© The Author(s) 2017. This article is an open access publication

Abstract We here harmonise two significant contributions to the field of wavelet analysis in the past two decades, namely the locally stationary wavelet process and the family of dual-tree complex wavelets. By combining these two components, we furnish a statistical model that can simultaneously access benefits from these two constructions. On the one hand, our model borrows the debiased spectrum and auto-covariance estimator from the locally stationary wavelet model. On the other hand, the enhanced directional selectivity is obtained from the dual-tree complex wavelets over the regular lattice. The resulting model allows for the description and identification of wavelet fields with significantly more directional fidelity than was previously possible. The corresponding estimation theory is established for the new model, and some stationarity detection experiments illustrate its practicality.

Keywords Locally stationary wavelet · Random fields · Dual-tree complex wavelets · Stationarity detection

1 Introduction

This paper has been completed by the second, third and fourth authors following the untimely death of the first author, James Nelson, in September 2016. It has been submitted for publication in his memory.

Building on the non-stationary time-series work of [Dahlhaus \(1997\)](#) and first proposed by [Nason et al. \(2000\)](#), the locally stationary wavelet (LSW) model provides an overcomplete, translation-invariant representation of a large class of non-stationary time series. Unlike many other such tools, LSW-based methodology affords a statistically well-principled means to capture the local covariance and local spectrum. That auto-covariance estimation of non-stationary processes by any other means comes entangled with various fundamental difficulties has enabled LSW models to gain traction across a variety of domains such as forecasting for finance ([Fryzlewicz 2005](#)); establishing dependencies in electrophysiological data for neuroscience applications ([Sanderson et al. 2010](#)); and spectral estimation of environmental time series and ECG traces with missing data ([Knight et al. 2012](#)).

Recent years have seen much basic theoretical development of locally stationary processes. For example, work has focussed on: variants of the smoothness constraints placed on the local spectrum function, such as the work by [van Bellegem and von Sachs \(2008\)](#), [Fryzlewicz and Nason \(2006\)](#), and [Nason and Stevens \(2015\)](#); confidence intervals for the empirical local covariance as recently derived by [Nason \(2013\)](#); and changepoint estimation, such as the work of [Killick et al. \(2013\)](#) and [Cho and Fryzlewicz \(2015\)](#).

J. D. B. Nelson and A. J. Gibberd were supported by Dstl.
J. D. B. Nelson was also partially supported by EPSRC Grant EP/N508470/1.

✉ A. J. Gibberd
agibberd@ic.ac.uk
C. Naornita
corina.naornita@upt.ro
N. Kingsbury
ngk@eng.cam.ac.uk

- ¹ Department of Statistical Science, University College, Gower Street, London WC1E 6BT, UK
- ² Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK
- ³ Communications Department, Politehnica University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania
- ⁴ Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK

Of particular interest here is the extension of the LSW theory to two-dimensional processes as recently driven by Eckley et al. (2010) and Eckley and Nason (2011). In particular, they extended the result that the local auto-covariance converges uniformly to auto-covariance as sample size increases and that there is an invertible linear mapping from one to the other. This work was subsequently followed by methodology applied to non-stationarity detection in image textures by Taylor et al. (2014) and segmentation of imagery into stationary regions by Nunes et al. (2014). A regularised smoothing strategy for the two-dimensional LSW process is developed in Gibberd and Nelson (2016). Very recently, Taylor et al. (2017) combined the lattice and multivariate extensions to formulate an LSW model for multivalued image data. As such, this can be exploited to compute multiscale estimates of the local inter-spectral covariance structure of multispectral imagery. The model was then applied to the task of classification of colour image textures.

The locally stationary wavelet model is hence becoming a flexible, extensible means to capture rich behaviour in spatial and image data and to solve practical problems. The overcompleteness of the LSW model, compared to, say, decimated wavelet schemes, is a key property which enables the LSW to elicit the auto-covariance. The increased number of wavelet coefficients permits a multiscale sample auto-covariance of sorts to be computed about any point in space. This is problematic in the decimated case since the periodogram is computed over the dyadic grid, and as Nason et al. (2000) notes, the decimated LSW model fails to accommodate all stationary processes (von Sachs et al. 1998). This contrasts with the non-decimated LSW model, which can describe any stationary process with finite integrated auto-covariance.

Overcompleteness not only affords a feasible means to estimate auto-covariance but also, as a side benefit, results in a translation-invariant periodogram. In effect, it provides the same information as a decimated transform at arbitrary integer shifts of the input data.

The main thrust of our work herein is to posit that further overcompleteness may provide extra value. In particular, when considering random fields such as image data defined over the lattice, directionality can be of significant importance. Often such data present an anisotropic covariance structure. A simple example of an anisotropic field is illustrated in Fig. 1. Here the piecewise non-stationarity is due to a simple relative rotation in the auto-covariance function, between the inner rectangle and outer annulus. In image processing, the ability to distinguish between such directionality of textures greatly aids tasks such as classification and segmentation.

Unfortunately, traditional real-valued (mono) wavelets such as the Daubechies family, favoured by the LSW theory, possess very limited directionality. This constrains their

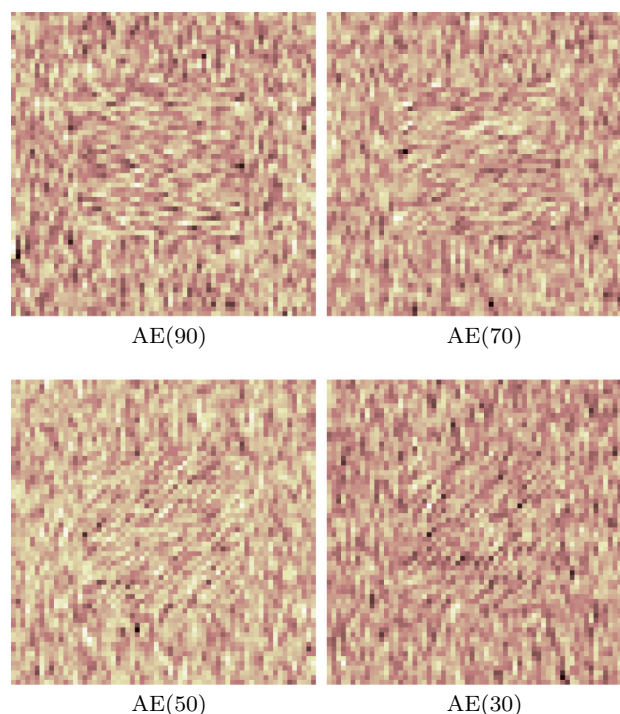


Fig. 1 Anisotropically non-stationary textures drawn from $AE(\theta)$ process, as defined in Sect. 4.2.2. Processes comprise an inner and outer region with different orientations, cf. Nelson and Gibberd (2016)

ability to model and represent highly directional phenomena typical in most modern image processing problems. As such, this has given rise to many so-termed ‘directional wavelets’, including the curvelets of Starck et al. (2002), the contourlets of Do and Vetterli (2002), the steerable pyramids of Portilla and Simoncelli (2000) and the monogenic Riesz–Laplace wavelets of Unser et al. (2009).

We here investigate the incorporation of a special directional wavelet basis, namely the dual-tree complex wavelets (Kingsbury 2001; Selesnick et al. 2005), into the locally stationary wavelet framework. It is shown that, since these complex wavelets are constructed as natural extensions of real-valued wavelets, much of the existing LSW theory can be developed by considering a complex-valued variant of the LSW framework. Indeed, in the sense that the dual-tree complex wavelet framework is a multiwavelet tight frame of order two, it can be considered ‘closer’ to the usual real-valued wavelets than some of the other more exotic and redundant directional wavelets, and as such, it provides a natural means to extend enhanced directionality into the LSW framework.

In Sect. 2, we propose the dual-tree complex locally stationary wavelet processes, denoted hereafter as $LSW(\mathbb{C})$. The processes are real-valued but the definition quite naturally permits complex-valued wavelets and transfer functions—hence extending the usual LSW model.

An additional property, taken straight from dual-tree complex wavelet theory, is placed on the relationship between the real and imaginary parts of the wavelet filters. This stipulates that they form a Hilbert pair and furnishes the wavelets with directional selectivity and polarity. Theorem 1 transfers a central result from the LSW case that the local auto-covariance converges to auto-covariance for LSW(C) processes.

Some estimation theory is covered in Sect. 3. We provide results that show the LSW(C) periodogram is a biased estimator of the spectrum. Furthermore, we establish the result that, under reasonable conditions, the biasing matrix can be inverted; equivalently, the corresponding auto-correlation wavelets retain linear independence; that the spectrum is uniquely defined, given the LSW(C) process; and that the mapping between spectrum and covariance is invertible.

In Sect. 4, simulations and experiments on a recent stationarity detection method confirm the utility of the enhanced directionality afforded by the proposed model.

Thus, for the first time, our work reconciles two significant contributions to the field of wavelet analysis, combining the locally stationary wavelet model with the family of dual-tree complex wavelets.

2 The complex locally stationary wavelet model

The dual-tree complex wavelet transform (Kingsbury 2001; Selesnick et al. 2005) employs a family of carefully constructed wavelet functions with the remarkable property that they jointly maintain approximate translation invariance and admit a full multiresolution analysis in the sense that the resulting two-scale relations are satisfied with finite sequence filters. This enables an implementation of a near-shift-invariant discrete wavelet transform at around the cost of two decimated, fast discrete wavelet transforms which can be computed in parallel.

The two decomposition trees are interpreted as the real and imaginary parts of the complex wavelet coefficients. The near-translation-invariance property is achieved by designing the wavelet filter sequences in such a way that they form an approximate Hilbert transform pair—where the real and imaginary parts are 90° out of phase with each other. That such invariance is achieved in concert with a genuine multiresolution analysis places dual-tree wavelets in a very special and unique category in what many refer to informally as the ‘wavelet zoo’.

In 2-D, the dual-tree scheme also naturally gives rise to improved directionality relative to regular, real-valued wavelets [such as, for example, Daubechies wavelets used by Eckley et al. (2010)]. Figure 2 contrasts the directionality afforded by a regular, real-valued, discrete 2-D wavelet (or impulse response) function and that of a dual-tree com-

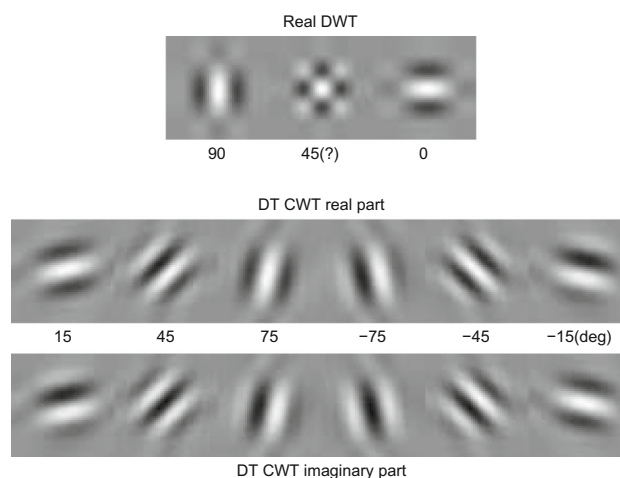


Fig. 2 Dual-tree wavelets provide six directionally selective filters, while real wavelets provide three filters, only two of which have a dominant direction

plex wavelet function. The real-valued wavelet function comprises filters oriented at 0° and 90° to the horizontal in addition to a filter oriented at both $\pm 45^\circ$. This third ‘diagonal’ filter will merge together any content oriented at $\pm 45^\circ$ —a practitioner will therefore be unable to distinguish between these two directions by studying the magnitude of the wavelet coefficients. For this reason, at best, one could perhaps argue that this wavelet function has a directionality of 2.5.

On the other hand, Fig. 2 also illustrates that the dual-tree complex wavelets offer six fully directional filters, oriented at $\{(30\ell - 15)^\circ\}_{\ell=1}^6$. Furthermore, the complex nature of the filters also yields polarity, or phase, information about image edges and singularities.

The resulting dual-tree complex wavelet impulse responses resemble the multiscale directional band-pass filters that the V1 human/mammalian cortical filters use as a front-end process to their visual systems far more faithfully than the directionally limited conventional wavelets—in their experimentally led work Hubel and Wiesel (1962) first remarked on the strong responses to directionality in mammalian vision and, more recently, Ng et al. (2006) provide a review which includes a discussion of the analogous role of wavelets with respect to the underlying biology.

In practice, the directionality offers far richer information than the classical, real-valued wavelet functions. As such, dual-tree complex wavelets and their various extensions continue to find hundreds of applications in image modelling and processing areas such as variational Bayesian enhancement (Zhang and Kingsbury 2015), registration and fusion for atmosphere correction (Anantrasirichai et al. 2013), segmentation via semi-local scaling exponent estimation (Nelson et al. 2016), and many more.

Eckley’s 2-D local stationary wavelet (LSW) model describes a stochastic process or field

$$X = \sum_{j,\ell} \sum_u w_j^\ell(u) \xi_j^\ell(u) \psi_j^\ell(u - \cdot), \tag{1}$$

over the finite lattice $\mathcal{T} := \llbracket 1, T \rrbracket^2$ as a dynamically weighted filtering. This combines a real-valued weighting sequence w_j^ℓ of a centred orthonormal increment sequence ξ_j^ℓ , and an overcomplete basis comprising what Nason et al. (2000) and Eckley et al. (2010) term (undecimated) ‘discrete wavelet functions’ $\psi_j^\ell(\cdot) \in \ell_2(\mathbb{Z}^2)$, indexed over both scale j and orientation ℓ . As usual, the ‘2-D’ wavelet functions are separable and are constructed from ‘1-D’ mother and father wavelets where the discrete mother wavelet functions are defined by the recursions

$$\psi_j(s) = \sum_k h_0(s - 2k) \psi_{j-1}(k), \tag{2}$$

for $j > 0$ and $s \in \llbracket 1, T \rrbracket$, with $\psi_0(s) = h_1(s)$, where the h_0 and h_1 are, respectively, the low- and high-pass filter sequences of the wavelet ψ . Similar recursions hold for the father wavelets, cf. Daubechies (1992) and Eckley et al. (2010). As noted by Nason et al. (2000), these discrete wavelet functions are not (necessarily) the result of discretising a continuous wavelet function. In fact, the recursions described by (2) are equivalent to setting all wavelet coefficients, except one at the j th finest scale level, to zero and then performing an inverse decimated wavelet transform with upsampling. Further discussion on the construction of these sequences, and their numerical generation, can be found in (Daubechies 1992, p. 204).

The LSW can be loosely described as a multiscale, dynamic, moving average model. For only one scale j_0 and direction ℓ_0 , say, the LSW becomes a moving average model with weights that can evolve over space via the sequence $w = \{w(k)\}_k$. If, in addition, the weights are constant over space $w = w_0$, say, then the LSW collapses down to a simple moving average model with parameters $\{w_0 \psi_{j_0}^{\ell_0}(k)\}_k$.

For the sake of controlling notational clutter, we fuse the scale and orientation indexes into one single index $\eta = \eta(j, \ell) \in \mathcal{H} \equiv \{1, \dots, 6J\}$. Without loss of generality, put $\eta = j + (\ell - 1)J$, for $j \in \llbracket 1, J \rrbracket$.

We further note that the LSW can be equivalently stated in terms of convolutions $X = \sum_{\eta \in \mathcal{H}} w_\eta \xi_\eta * \overline{\psi_\eta}$, where the $\overline{\cdot}$ denotes a left-right and up-down spatial flip: $\overline{\varphi} := \varphi(-\cdot)$. Aside from providing an alternative convenient means of conceptualising the model, introducing such operator notation also helps mitigate notational clutter in some of the later proofs.

Definition 1 places further constraints on how quickly the weighting sequence can vary over the spatial support. We

here transfer all of the basic real-valued LSW modelling assumptions in Nason et al. (2000) and Eckley et al. (2010) over to our (dual-tree) complex wavelet case. The differences are that (i) the transfer sequences w and transfer functions W , below, are now complex-valued, and (ii) Property 5 is added in order to exploit the approximate Hilbert transform pair property of the dual-tree complex wavelets and therefore produce wavelets which are directionally selective and which carry phase information.

Unlike the LSW model, the DTCWT, as originally designed, is decimated. Instead, we here incorporate a non-decimated version of the DTCWT into the LSW. Although the resulting extra computation has little affect on translation invariance (the DTCWT is already near-shift invariant), the dual-tree complex wavelets do greatly enhance the poor directional description suffered by the LSW model and the non-decimated construction provides a simple mechanism to introduce them into the LSW framework. Undecimated versions of the DTCWT have recently been proposed by Hill et al. (2015) who, in addition, note the added virtues that the resulting wavelet coefficients are truly co-located over the scale levels rather than lying on the usual dyadic grid. Although we do note that a big attraction of the DTCWT is the ability to perform a decimated transform, the incorporation of decimation, partially or wholly, into the LSW is left as further work for now.

Definition 1 The locally stationary dual-tree complex wavelet stationary processes are defined as the class $\text{LSW}(\mathbb{C})$ of all stochastic processes $X : \mathcal{T} \mapsto \mathbb{R}$ over the finite lattice $\mathcal{T} := \llbracket 1, T \rrbracket^2$ which satisfy the following representation in the mean-squared sense:

$$X = \sum_{\eta \in \mathcal{H}} w_\eta \xi_\eta * \overline{\psi_\eta},$$

with complex-valued sequences $w_\eta \in \ell_2(\mathcal{T})$, centred orthonormal increment processes $\xi_\eta : \mathcal{T} \mapsto \mathbb{R}$, discrete complex wavelet functions $\psi_\eta : \mathcal{T} \mapsto \mathbb{C}$, and where there exist functions $W_\eta : \mathbb{T}^2 := (0, 1)^2 \mapsto \mathbb{C}$, and constants $C_\eta, \lambda_\eta \in \mathbb{R}$ such that all the following properties hold for all $\eta \in \mathcal{H}$.

1. $\sup_{t \in \mathcal{T}} |w_\eta(t) - W_\eta(t/T)| \leq C_\eta/T$, with $\sum_\eta C_\eta < \infty$.
2. W is square-summable over the scale-orientation indexes: $\sum_\eta |W_\eta(z)|^2 < \infty$, uniformly in $z \in \mathbb{T}^2$.
3. W_η is L_1 -Lipschitz, viz.

$$\left\| W_\eta\left(\frac{k+\cdot}{T}\right) - W_\eta\left(\frac{\cdot}{T}\right) \right\|_{L_1(\mathbb{R}^2)} \leq \lambda_\eta \|k/T\|_1$$

4. The Lipschitz constants λ_η are uniformly bounded in η and $\sum_\ell \sum_{j \in \mathbb{N}} 2^{2j} \lambda_{\eta(j,\ell)} < \infty$.

5. The complex wavelet functions satisfy the approximate Hilbert pair property, cf. Kingsbury (2003) and Selesnick et al. (2005), in that its real $\psi_{:,0}$ and imaginary $\psi_{:,1}$ parts satisfy: $\psi = \psi_{:,0} + i\psi_{:,1}$, with $\psi_{:,0}, \psi_{:,1} : \mathcal{T} \mapsto \mathbb{R}$, and where the associated filter sequences h_b^a satisfy the below properties, where $b = 0$ or $b = 1$ signifies the low-pass or high-pass band sequences, respectively, and where $a = 0$ denotes the real part of the tree and $a = 1$ the imaginary part of the tree:

$$h_1^a(k) = (-1)^t h_0^a(n - k), \tag{3}$$

$$h_0^1(k) = h_0^0(n - 1 - k), \tag{4}$$

and where n is the length of the filter sequence.

Note that the approximate Hilbert pair property between $\psi_{:,0}$ and $\psi_{:,1}$ ensures that the negative half of the spectrum of ψ becomes approximately zero. This is the key to obtaining directionally selective filters in 2-D from separable 1-D filters, as explained in Selesnick et al. (2005).

The collection of functions $\{W_\eta\}$ comprise the parameters of the LSW model. They determine how much auto-correlation is present at each scale level. As $T \rightarrow \infty$ and the number of sample points increases, then properties 1 and 2 above associate the collection of transfer sequences $\{w_\eta\}$ with, what become, the bounded family of transfer functions $\{W_\eta\}$. Properties 3 and 4 stipulate that the transfer function should be smooth. Asymptotically, the weights are associated with Lipschitz continuous functions $W_\eta : \mathbb{T}^2 \mapsto \mathbb{C}$, defined over rescaled space $z = t/T, t \in \mathcal{T}$. This condition restricts the variation of the weights over space and thus captures the fact that the local structure becomes more stationary when observed over evermore smaller spatial neighbourhoods.

A central premise, and unique selling point, of the LSW model is that it very naturally offers access to a measure of local covariance via the local spectrum and the auto-correlation wavelet.

Definition 2 (Complex ACW) Let the superscript \star denote complex conjugation. Define the (complex) auto-correlation wavelet ACW associated with the possibly complex-valued, discrete wavelet $\psi : \mathbb{N}^2 \mapsto \mathbb{C}$ by

$$\Psi_\eta = \psi_\eta * \overline{\psi_\eta^\star} = \sum_{t \in \mathcal{T}} \psi_\eta^\star(t) \psi_\eta(\cdot + t). \tag{5}$$

Definition 3 (Local wavelet spectrum) Let $X \in \text{LSW}(\mathbb{C})$ have transfer functions $W_\eta : \mathbb{T}^2 \mapsto \mathbb{C}$. Then, the local wavelet spectrum (LWS) of X is defined by

$$S_\eta(\cdot) := |W_\eta(\cdot)|^2.$$

Definition 4 (Local auto-correlation) Let S be the local wavelet spectrum, defined by Definition 3, of an LSW(\mathbb{C}) process (Definition 1) and let Ψ be the auto-correlation wavelet as described by Definition 2. Then, the local covariance is defined by

$$C(z, t) = \sum_{\eta \in \mathcal{H}} S_\eta(z) \Psi_\eta(t), \quad z \in \mathbb{T}^2, t \in \mathbb{Z}^2.$$

A remarkable property is that this local covariance forms a good estimate for the auto-covariance of the, possibly non-stationary, LSW process. The next result establishes this fact for the locally stationary dual-tree complex wavelet processes.

Theorem 1 (Convergence of auto-covariance) Let C and C_T be, respectively, the local auto-covariance and auto-covariance of an LSW process $X \in \text{LSW}(\mathbb{C})$. Then

$$|C_T(z, t) - C(z, t)| = \mathcal{O}(T^{-1}).$$

as $T \rightarrow \infty$, uniformly in $t \in \mathbb{Z}^2$, and $z \in \mathbb{T}^2$.

Remark 1 (Complex extension of the LSW) For $X \in \text{LSW}(\mathbb{C})$:

1. The auto-correlation wavelet is defined via the natural complex extension of the auto-correlation operator.
2. The local wavelet spectrum quantity works in much the same way as it does in the real LSW case only now it is formed by computing the complex modulus of the transfer function.

3 Estimation theory

The (local wavelet) spectrum S represents: the parameters of the LSW(\mathbb{C}); the amount of ‘energy’ present at a particular scale and orientation; and, via Theorem 1 the spectrum also establishes the second-order behaviour of the field. Two key questions naturally follow. The first is how the spectrum can be estimated and, second, is whether the spectrum is uniquely defined, given a field.

The (undecimated) local wavelet periodogram (LWP) $X_\eta^\sim(\cdot) := |X * \psi_\eta^\star|^2$ offers a straightforward but naive estimate of the spectrum. As Theorem 2 shows, the redundancy in the LWS model spreads content, via a so-termed biasing matrix across the periodogram scales.

Definition 5 (LS-DTCW biasing matrix) Let Ψ be the complex ACW as above. Then, the associated biasing matrix is

$$A_{\eta, \nu} = \langle \Psi_\eta, \Psi_\nu \rangle = \sum_t \Psi_\eta(t) \Psi_\nu^\star(t). \tag{6}$$

Theorem 2 (Periodogram bias) *Let the LSW(\mathbb{C}) process $X : \mathcal{T} \mapsto \mathbb{R}$, defined on the lattice $\mathcal{T} := \llbracket 1, T \rrbracket^2$, have periodogram $X_{\eta}^{\sim} := \left| X * \psi_{\eta}^{\star} \right|^2$, where $\{\psi_{\eta}\}_{\eta \in \mathcal{H}}$ is a discrete, undecimated, dual-tree complex wavelet basis and let A be the associated biasing matrix. Then*

$$\mathbb{E}X_{\eta}^{\sim} = \sum_{\nu} A_{\eta\nu} S_{\nu} \left(\frac{\cdot}{T} \right) + \mathcal{O}(T^{-1}) \quad (7)$$

The biasing matrix is the agent of this spread or mixing of redundancy. Note that Definition 5 is but the complex extension of the inner product used to define the real biasing matrix of the locally stationary, real-valued, wavelet model. In addition to facilitating some subsequent results, it is perhaps of intrinsic interest to note that the biasing matrix is real-valued.

Lemma 1 (Bias is real) *Let Ψ be the ACW (Definition 2). The entries of the biasing matrix (Definition 5) are positive, real, and symmetric.*

Remark 2 (Unbiased spectrum estimator) *If the biasing matrix is invertible, then one can derive an unbiased estimate of the spectrum via:*

$$\widehat{S}_{\eta}(z) = \sum_{\nu} A_{\eta,\nu}^{-1} X_{\nu}^{\sim}(z). \quad (8)$$

The consistency of the smoothed version of the real-valued wavelet periodogram is established and discussed by Nason et al. (2000) in the 1-D case and Eckley et al. (2010) in the 2-D case. A proof of consistency of the Nadaraya–Watson kernel estimator (cf. Nadaraya 2016; Watson 1964) is given by Taylor et al. (2017).

The existence of the debiasing matrix—the inverse of the biasing matrix—rests solely upon the choice of wavelet. Nason et al. (2000) proved existence for the 1-D Daubechies wavelets, and Eckley et al. (2010) extended this to the 2-D Daubechies wavelets. A similar result, for other wavelet families, has remained a conjecture and open problem ever since Nason et al. (2000) carried out their original work.

Instead, we can here report that, in practice, the debiasing matrices exist for 2-D locally stationary dual-tree wavelets with data sizes $2^n \times 2^n$ for $n \in \llbracket 1, 9 \rrbracket$. Furthermore, we present the following result that, given the existence of a debiasing matrix associated with a generic 1-D wavelet function, the corresponding dual-tree wavelet function will also have an associated debiasing matrix.

Theorem 3 (Invertibility of the biasing matrix) *Let the biasing matrix A' associated with an appropriately chosen real-valued wavelet $\psi_{\cdot;0}$ defined over $\llbracket 1, T \rrbracket$ be non-singular. Then, the biasing matrix A (Definition 5) associated with the dual-tree wavelet (cf. Property 5 of Definition 1) $\psi = \psi_{\cdot;0} + i\psi_{\cdot;1}$ is non-singular with $A = 2A'$.*

An equivalent result holds which reveals the connection between the invertibility of the biasing operator and the linear independence of the auto-correlation wavelet.

Theorem 4 (Linear independence of ACW) *Let the auto-correlation wavelet Ψ' associated with an appropriately chosen real discrete wavelet function $\psi_{\cdot;0}$ defined over $\llbracket 1, T \rrbracket$ be linearly independent. Then, the auto-correlation wavelet Ψ associated with the dual-tree wavelet $\psi = \psi_{\cdot;0} + i\psi_{\cdot;1}$ is also linearly independent.*

Remark 3 (Invertibility and ACW independence) *Since the matrix A is a Gramian, the linear independence of Ψ is equivalent to the non-singularity of A .*

Crucially, linear independence naturally gives rise to uniqueness.

Corollary 1 (LWS uniqueness) *The LWS is uniquely defined, given the corresponding LSW(\mathbb{C}) process.*

For completeness, and analogous to the real-valued wavelet case, there is also an inversion property between the spectrum and local auto-covariance function.

Proposition 1 (LS-DTCW covariance inversion, Nelson and Gibberd 2016) *Assume that the debiasing matrix A is non-singular, then*

$$S_{\eta}(z) = \sum_{\nu} A_{\eta,\nu}^{-1} \sum_t c(z, t) \Psi_{\nu}^{\star}(t).$$

4 Experiments

The advantages of the additional directionality afforded by the dual-tree complex wavelets are manifested in various ways according to the task at hand. We here consider just a few numerical experiments, but note that, since the non-LSW form of the dual-tree complex wavelet transform has spawned hundreds of applications in signal and image processing, there are many other such potential applications of this LSW-based model. We focus here on a couple of examples that have been discussed specifically in the LSW literature. Firstly, we qualitatively demonstrate the broader model space provided by the LS-DTCW. Numerical evidence is then offered towards the benefits of greater directionality to the task of non-stationarity detection.

4.1 Simulating anisotropic stationary random fields

An attractive property of the LSW framework is that, given a known local wavelet spectrum S , one can readily draw simulated fields from the prescribed models. In the absence of knowledge about S , a good alternative is to perform estimation of the spectrum. Thus, if one can both estimate S and use

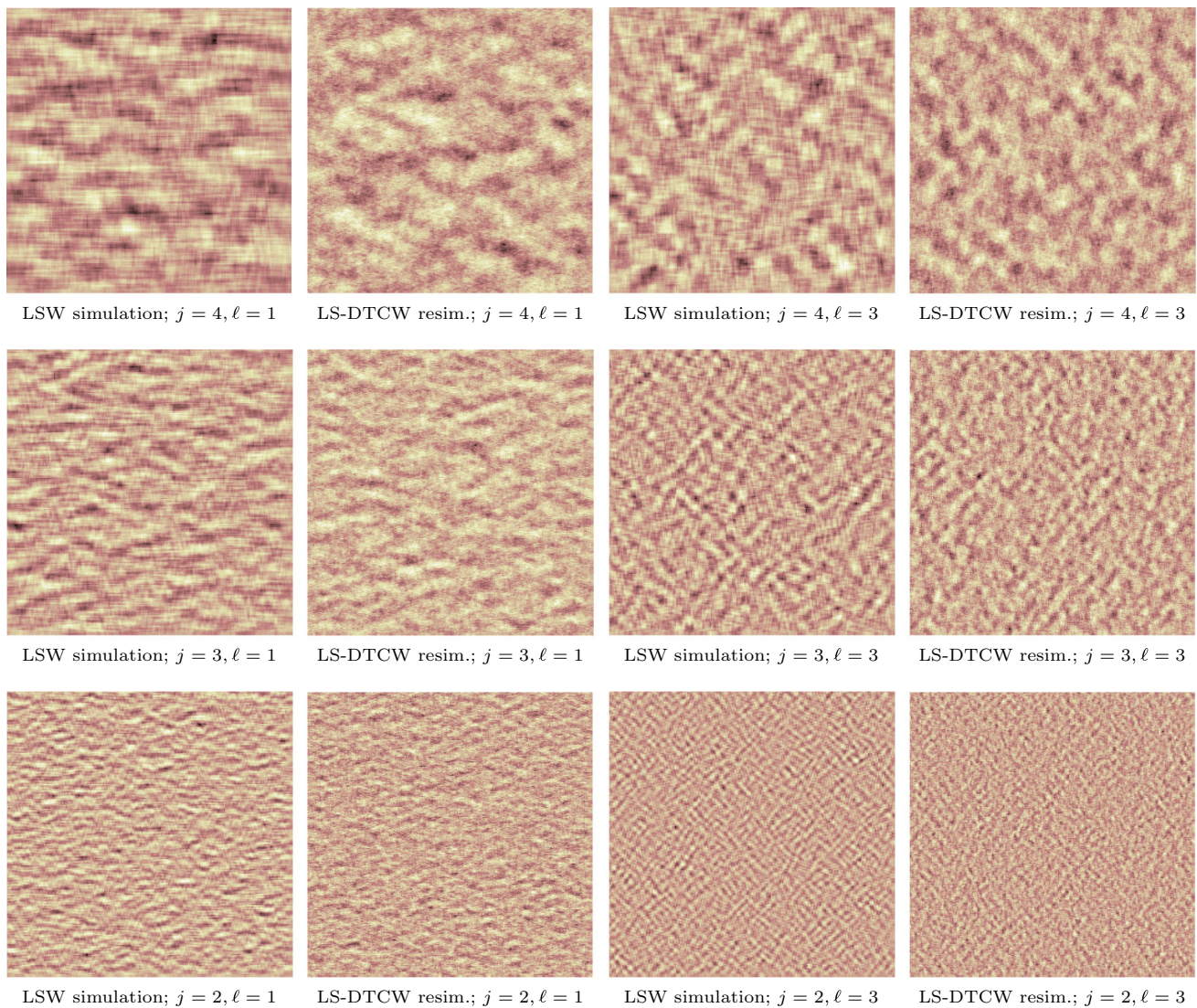


Fig. 3 LS-DTCW simulations of LSW fields. The indexes j refer to scale level, and $\ell = 1, 3$, respectively, correspond to the real filters in the horizontal and diagonal positions, or DTCW filters at $(30\ell - 15)^\circ$

this approximation to simulate a process, then it is possible to informally and qualitatively compare the two model spaces for LSW and LS-DTCW. In the proceeding experiment, a simple texture is simulated using one model (either LSW or LS-DTCW) and then the other model is used to estimate the spectrum of the simulated process. Using the estimate, an attempt to ‘resimulate’ the texture from this spectrum can be made.

As described in Sect. 3, the spectrum can be estimated in a pointwise fashion even when the field in question is non-stationary. Algorithm 1 provides the required simple recipe. First of all, the periodogram $|X * \psi_\eta^*|^2$ of the random field X is computed. This is merely a double convolution over space. Since the wavelets considered here, either Haar in the LSW case or q-shift in the LS-DTCW case, are fairly short—2-taps for Haar, and around 12 or more taps for q-

shift—this is not a significant computational task. For larger filters and, especially, those with infinite support, speed-ups can be obtained by performing the convolution in the Fourier domain. Either way, the periodogram is then bias-corrected by applying the inverse mapping: $\widehat{S}_\eta = \sum_{\eta \in \mathcal{H}} A_{\eta,v}^{-1} X_\eta^\sim(\cdot)$. For each pixel t , this is equivalent to applying a $H \times H$ matrix transform to a vectorised form of the periodogram, namely $[X_1^\sim(t), \dots, X_H^\sim(t)]^\top \in \mathbb{R}^H$, with $H := \#\mathcal{H}$.

In the interests of comparing the extra directionality provided by the LS-DTCW with that of the LSW, it is instructive to, first of all, consider a simple stationary case. In Fig. 3, the sub-figures labelled ‘simulation’, with indexes j and ℓ are fields simulated by setting $S_\eta(\cdot) \equiv 1$ for the particular stated value of $\eta = j + (\ell - 1)J$, where j is the j th coarsest scale level, and ℓ is the directional sub-band such that: $\ell = 1$ refers to a horizontal stripe direction, $\ell = 2$ a vertical stripe direc-

tion, and $\ell = 3$ refers to the ‘diagonal’ sub-band direction. All other values of S_η (for any other (j, ℓ)) are set to zero and the field is simply drawn from the model via $X = \sqrt{S_\eta} \xi_\eta * \psi_\eta$, with $\xi_\eta \sim \mathcal{N}(0, 1)$.

For example, the processes illustrated at the top of the first column of Fig. 3 are simulated using the simple piecewise construction

$$S_{(j,\ell)}(\cdot) \equiv \begin{cases} 1, & \text{if } j = 4 \text{ and } \ell = 1 \\ 0, & \text{otherwise} \end{cases}.$$

Similarly, the processes depicted in the third column of the first row have a spectrum equal to $\delta_{\ell,3} \delta_{j,4}$ across the field.

Of course, strictly speaking, this contravenes the Lipschitz continuity of the spectrum, as per Definition 1, Property 3 but, nonetheless, it serves as a simple, intuitive example. The directional sub-band $\ell = 1$, here refers to a horizontal stripe orientation. As such, horizontally oriented waves of texture should be more perceptible than vertical ones in any of the three ‘LSW simulation’ sub-figures with $\ell = 1$. Similarly, the third column of sub-figures depicts LSW random fields oriented in the ‘diagonal’ direction. Likewise, here, waves of texture oriented at $\pm 45^\circ$ should be evident.

In this simple simulation scenario, the stationarity of the resulting fields means that spectrum estimates can be improved by simply averaging the spectrum estimate over the entire spatial support. Algorithms 1 and 2 describe the recipes required to first estimate the spectrum and then, secondly, resimulate the texture. The second and fourth columns of Fig. 3 illustrate the LS-DTCW resimulations of the LSW textures in columns one and three, respectively.

It can be seen, for example, that the second column reflects the horizontal wave directions of the first column in the appropriate levels of scale fairly well. Likewise, the waves oriented at $\pm 45^\circ$ in the third, LSW simulated, column are also present in the fourth, LS-DTCW resimulated, column.

In contrast, Fig. 4 shows the reverse of this experiment where, firstly, textures are simulated from the LS-DTCW model and then, secondly, they are estimated and resimulated using the LSW model. The processes plotted in the first column have textures oriented at 15° to the horizontal axis. However, owing to its poor directionality and directional selectivity, the LSW is unable to convincingly capture this behaviour. Likewise, the third column depicts LS-DTCW textures orientated at 45° to the horizontal axis. Again, the LSW cannot resimulate textures that resemble anything similar to the original textures. In short, in this example, the LS-DTCW can capture behaviour seen in the LSW model but not vice versa. Since more complex textures can be constructed from linear combinations of these simple building block examples, it could therefore be argued that the LS-DTCW model is, broadly speaking, more general than the LSW model. This is perhaps not surprising since the LS-

DTCW model comprises twice the number of wavelet atoms than the LSW model, but it is interesting to note that the difference manifests quite decidedly when considering the directionality of the texture.

Algorithm 1 Unbiased spectrum estimation, LWSE, cf. Taylor et al. (2014)

$\widehat{S}_\eta = \text{LWSE}(X)$

Input: image X

Output: estimated spectrum \widehat{S}_η

compute periodogram $X_\eta^\sim = |X * \psi_\eta^*|^2$

estimate spectrum $\widehat{S}_\eta = \sum_{v \in \mathcal{H}} A_{\eta,v}^{-1} X_v^\sim(\cdot)$

Algorithm 2 Simulation algorithm, LWSIM, cf. Taylor et al. (2014)

$\widehat{X} = \text{LWSIM}(X)$

Input: image X

Output: simulated image \widehat{X}

for (level, orientation) = η ; **do**

 estimate spectrum $\widehat{S}_\eta = \text{LWSE}(X)$ via Alg. 1

 estimate transfer sequence $\widehat{w}_\eta(\cdot) := \sqrt{T^{-2} \sum_t \widehat{S}_\eta(t)}$

 simulate: $\widehat{X} = \sum_{\eta \in \mathcal{H}} (\widehat{w}_\eta \xi_\eta) * \overline{\psi_\eta}$, with $\xi_\eta \sim \mathcal{N}(0, 1)$

4.2 Stationarity testing

Taylor et al. (2014) explored the role of the LSW to the problem of non-stationarity detection in images. This was applied to the problem of machine-vision detection of pilling effects in fabrics. Since the LSW provides a model which can be simulated from, Taylor et al. (2014) was able to derive a hypothesis testing procedure to determine stationarity. However, the pervasiveness of anisotropic non-stationary random fields in image processing motivates the extension of this test to our locally stationary dual-tree complex wavelet framework. Such an extension recognises the fact that non-stationarity can be caused, at least in part, by a rotation of the auto-covariance function, cf. Fig. 1.

More generally, detection and characterisation of such directional non-stationarity would be of interest to image processing tasks such as segmentation (Nunes et al. 2014; Nelson et al. 2016); denoising or decluttering of highly directional textures such as sand ripples in sonar imagery (Nelson and Kingsbury 2010, 2012); or change detection in remote sensing imagery (Hussain et al. 2013). Non-stationarity detection is also closely aligned with the saliency detection task (Raj et al. 2007). This open problem in image processing is a cru-

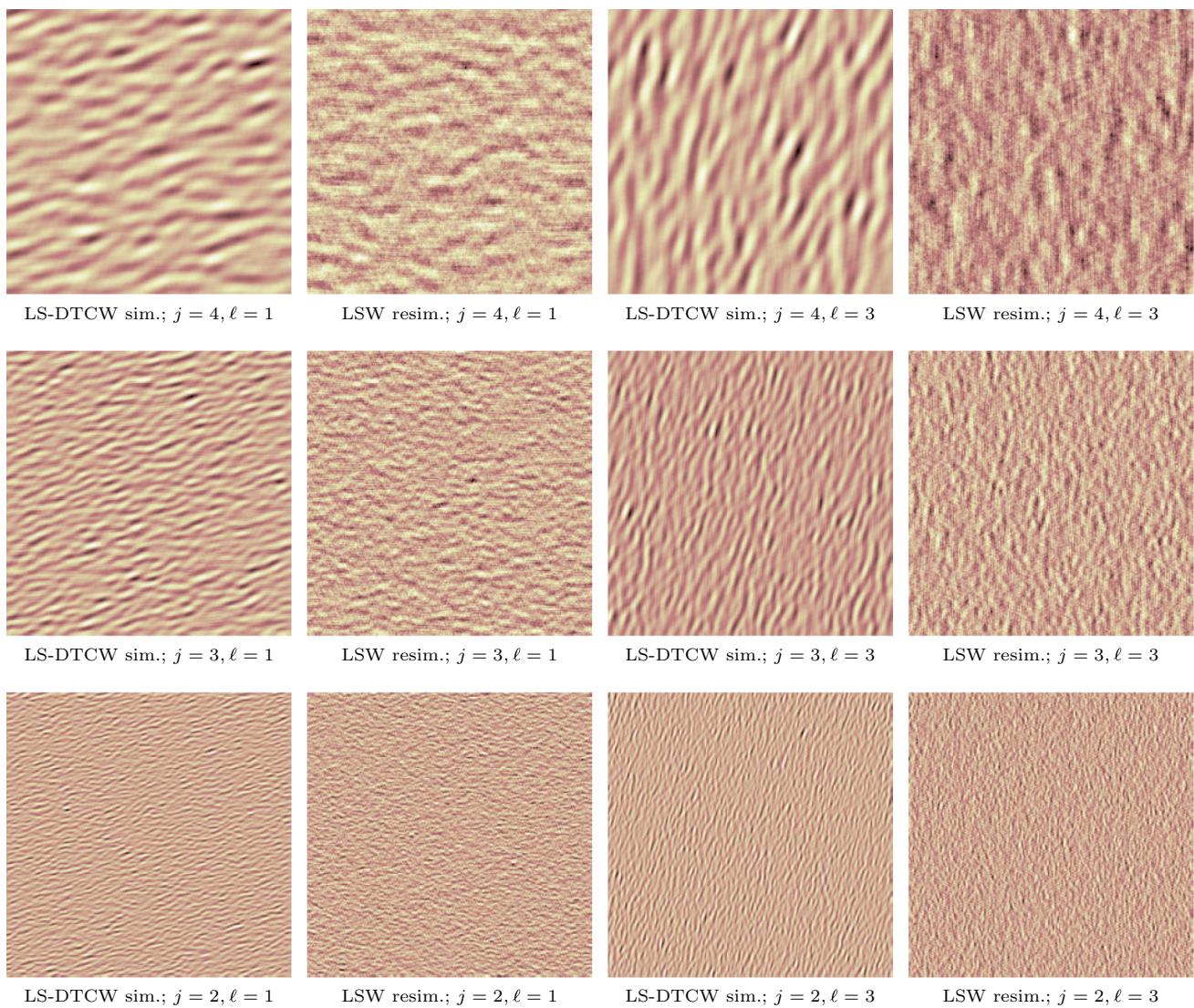


Fig. 4 LSW simulations of LS-DTCW fields

cial and profound challenge towards higher-level automatic semantic understanding of image data.

Taylor et al. (2014) extended ideas from time series to use the LSW model to perform stationarity hypothesis testing on random fields. Bootstrapping forms the central premise. The sample variance of the observed field is compared to that of series of simulated fields, where the simulations are drawn under the assumption that the observed field is stationary. Algorithms 1 and 2 can therefore be employed to perform the necessary spectrum estimation and simulation. The basic intuition is that, if the observation is stationary, then it will have a variance drawn from the same population as the simulations. If it is non-stationary, the observed variance will, on average, tend to be greater than the variances of the (stationary) simulations. Repeatedly drawing simulations from the hypothesised model then yields a p value. The advantage of

this approach is that the distribution of the process need not be known analytically.

We follow the spirit of Taylor et al. (2014) and perform two sets of experiments. The size and power were assessed using a selection of stationary and non-stationary random fields, respectively. Unlike Taylor et al. (2014), and since one of our interests here is the added directionality afforded by the LS-DTCW model, we include some anisotropic models.

4.2.1 Algorithm

Unlike Nelson and Gibberd (2016) who concentrated on the simpler case where the null hypothesis is rejected based on the average spectrum sample variance statistic over all scales and orientations, we follow the multiple hypothesis testing framework. Here, instead, the null hypothesis is tested for all

scale and orientation pairs individually and is rejected if any of the spectra sample variances (over η) deviates significantly from constancy:

$$H_0 : S_\eta(z) = \text{const. } \forall z \in \mathbb{T}^2$$

$$H_A : S_\eta(z) \neq \text{const. } \forall z \in \mathbb{T}^2$$

As in Taylor et al. (2014), we draw on the multiple hypothesis testing bootstrapping framework constructed by Davison and Hinkley (1997) to accommodate any distributional differences and dependencies in the spectrum sample variance statistics over the different scales and orientations. To this end, if π_η is the tail probability of the statistic $\tau(\eta)$ for scale/orientation pair η , then the multiple hypothesis framework uses the test statistic $q = \min_\eta \pi_\eta$. Algorithm 3 lays out the hypothesis testing procedure.

4.2.2 Data

Three stationary processes were used, namely:

1. A simple, white noise process $X \sim \mathcal{N}(0, 1)$.
2. A spatial moving average, order one process

$$\begin{aligned} X \in \text{MA}(1) &\Leftrightarrow X(t_i) \\ &= \epsilon(t_i) + \sum_{k \in \partial(i)} \varphi_k \epsilon(t_{ik}), \quad \varphi_k \equiv 0.9, \end{aligned}$$

where $(t_{ik})_{k \in \partial(i)}$ is the four-neighbourhood of the point t_i .

3. An anisotropic stationary field with exponential covariance $X(t) \in \text{AE}(0) \Leftrightarrow$

$$\text{cov}(X(t), X(t')) = \exp\left(- (t - t')^\top D (t - t')\right),$$

with $D = \text{Diag}(4, 1)$.

Furthermore, three non-stationary fields were considered, viz.:

1. A piecewise constant stationary random field montage of two moving average, order one fields $\text{MA}(1)$, denoted by $\text{MA}(1; \sigma)$, where one half has an innovation process ϵ with $\text{var } \epsilon = 1$ and the other half has $\text{var } \epsilon = \sigma$.
2. An anisotropic non-stationary field with exponential covariance $X(t) \in \text{AE}(\theta) \Leftrightarrow$

$$\text{cov}(X(t), X(t')) = \exp\left(- (t - t')^\top D R_{\theta'} (t - t')\right),$$

with $D = \text{Diag}(4, 1)$ and where the rotation matrix $R_{\theta'}(t)$ is piecewise constant with

$$\theta' := \begin{cases} 0, & \text{for } t \in \mathcal{T}_0 \\ \theta, & \text{otherwise} \end{cases}$$

where \mathcal{T}_0 is a central square region, cf. Fig. 1. The difference in rotation between the inner and outer regions, namely θ , was varied in the experiments to furnish eight non-stationary processes with

$$\theta \in \{5^\circ, 10^\circ, 20^\circ, 30^\circ, 40^\circ, 50^\circ, 70^\circ, 90^\circ\}.$$

3. A montage of Brodatz textures: each half of the image comprised a different texture, selected from the Brodatz texture sub-catalogue: D06, D11, D19, D29, D36, D52, D79, D84, D87, cg. Fig. 5. The images were then rotated by angles $\{0^\circ, 10^\circ, 20^\circ, 30^\circ, 40^\circ\}$.

Remark 4 Note that processes $\text{AE}(\theta)$ are stationary for $\theta = 0$ and non-stationary for any other value. The non-stationarity is due to the angular difference in the orientation of the textures in the inner and outer regions.

Algorithm 3 Multi-hypothesis Stationarity detection, cf. Taylor et al. (2014)

$p = \text{TESTSTAT}(X)$

Input: image X

Output: p-value p

Estimate observed local spectrum $\widehat{S}_\eta = \text{LWSE}(X)$

Compute observed statistic $\tau_*(\eta) = \widehat{\text{var}}(\widehat{S}_\eta)$

for k in $1 : K$ bootstraps **do**

 Simulate $\widehat{X}^k = \text{LWSIM}(X)$ with Alg. 2

 Estimate bootstrapped local spectrum $\widehat{S}_\eta^k = \text{LWSE}(X^k)$

 Compute statistic $\tau_k(\eta) = \widehat{\text{var}}(\widehat{S}_\eta^k)$

for m in $1 : M$ bootstraps **do**

 Simulate $\widehat{X}^{km} = \text{LWSIM}(X^i)$ with Alg. 2

 Est. bootstrapped spectrum $\widehat{S}_\eta^{km} = \text{LWSE}(X^{km})$

 Compute statistic $\tau_{km}(\eta) = \widehat{\text{var}}(\widehat{S}_\eta^{km})$

 Compute test statistic

$$q_k := (M + 1)^{-1} \min\{1 + \#\{\tau_k(\eta) \leq \tau_{km}(\eta)\}_{m=1}^M\}_{\eta \in \mathcal{H}}$$

Compute test statistic

$$q_* := (K + 1)^{-1} \min\{1 + \#\{\tau_*(\eta) \leq \tau_k(\eta)\}_{k=1}^K\}_{\eta \in \mathcal{H}}$$

$$\text{Compute p-value } p := \frac{1 + \#\{q_k \leq q_*\}_{k=1}^K}{K + 1}$$

4.3 Stationarity results

We here compare the locally stationarity (real-valued) wavelet test with that of the locally stationary dual-tree complex wavelet test. For completeness, in addition to the multiple hypothesis tests, denoted by MH in Tables 1, 2, 3 and 4,

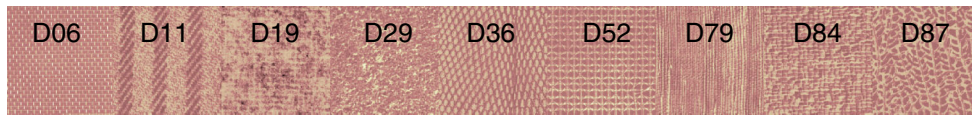


Fig. 5 Brodatz textures used for real texture experiments

Table 1 Proportion (%) of samples with p value below 0.05 for three stationary random fields

Method/process	WN	MA(1)	AE(0)
LS-DTCW (MH)	5	6	5
LS-DTCW (SH)	0	0	6
LSW (MH)	2	2	5
LSW (SH)	0	1	0

Bold values highlight the best performance amongst methods under analysis

Table 2 Proportion (%) of samples with p value below 0.05 for the non-stationary anisotropic random fields, AE(θ)

Method/ θ	5	10	20	30	40	50	70	90
LS-DTCW (MH)	5	26	59	99	100	100	100	100
LS-DTCW (SH)	6	21	39	66	95	98	99	99
LSW (MH)	5	8	8	24	84	98	99	99
LSW (SH)	0	0	0	0	1	5	10	9

Bold values highlight the best performance amongst methods under analysis

Table 3 Proportion (%) of samples with p value below 0.05 for the non-stationary isotropic random fields, MA(1; σ)

Method/ σ	1.2	1.3	1.4	1.5
LS-DTCW (MH)	37	62	91	100
LS-DTCW (SH)	4	30	75	97
LSW (MH)	36	66	92	100
LSW (SH)	1	4	20	53

Bold values highlight the best performance amongst methods under analysis

Table 4 Proportion (%) of samples with p value below 0.05 for the rotated, Brodatz montage data set

Method/ θ	0	10	20	30	40
LS-DTCW (MH)	100	100	100	100	100
LS-DTCW (SH)	60	51	44	37	57
LSW (MH)	56	60	66	60	50
LSW (SH)	7	6	6	7	7

Bold values highlight the best performance amongst methods under analysis

we also include results using the simpler, but less accurate, single-hypothesis test (denoted SH) as described by Taylor et al. (2014) and, later, Nelson and Gibberd (2016).

All multiple hypothesis experiments here below were conducted on 250×250 bootstraps ($K = 250, M = 250$ in Algorithm 3) and 100 instances of each image. The single hypothesis tests were conducted on 250 bootstraps.

Hence, for example, Table 1 records that, two out of 100 white noise processes caused a rejection of the null hypothesis at the 5% level when using the real-valued LSW multiple hypothesis test. The single-hypothesis tests were very conservative, triggering no null-hypotheses at all.

The dual-tree multiple hypothesis test managed to reach a figure closest to the ideal value of 5%. The tests on the MA(1) processes resulted in similar outcomes. Interestingly, unlike the real-valued LSW, the dual-tree single-hypothesis tests resulted in a less conservative figure for the anisotropic stationary process AE(0) experiments than it did for the isotropic processes.

The conservatism of the single-hypothesis test, relative to the multiple hypothesis test, broadly aligns with observations by Cardinali and Nason (2010) and Taylor et al. (2014).

The three non-stationary examples provide a means to assess the power of the stationarity tests. Table 2 provides the number of null-hypothesis rejections for 100 instances of the AE(θ) process, per angle parameter θ . In a sense, this process becomes increasingly stationary as the angle θ is decreased away from 90° towards 0° .

It can be seen that, as expected, all tests improve as the angle increases. However, there is a very clear difference in the rate at which each test improves. The multiple hypothesis LS-DTCW is the most sensitive. The multiple hypothesis tests again obtain better (more sensitive) results than their respective single-hypothesis versions. Interestingly, however, the single-hypothesis version of the LS-DTCW obtains better results than the multiple hypothesis LSW. In this case, the extra directionality provided by the dual-tree wavelet model is able to elicit smaller changes in the anisotropic non-stationarity.

Table 3 provides results on the isotropic and non-stationary process MA(1, σ). One hundred realisations were used per value of variance $\sigma \in \{1.2, 1.3, 1.4, 1.5\}$. The results suggest that, even though the single hypothesis version of the LS-DTCW appears more sensitive to the single-hypothesis LSW, the multihypothesis LS-DTCW is only roughly comparable to the multihypothesis LSW test. Intuitively, since the ‘information’ about non-stationarity is spread over all directions there appears to be no advantage in splitting up the directional content of the processes any

further than the three bands of the usual real-valued wavelet decomposition.

Finally, Table 4 shows the results from the Brodatz data. The percentage of null-hypothesis rejections are presented over the data set of 72 montage images. Again, the multiple hypothesis tests detect more non-stationarities than the single hypothesis tests. The LS-DTCW correctly declared all images to be non-stationary, irrespective of what angle the data were rotated. The single-hypothesis LS-DTCW performed a little worse and much more inconsistently with respect to angle than the multiple hypothesis LSW test. The single-hypothesis LSW test performed very poorly and only managed to pick up 6 or 7% of the non-stationarities. This experiment lends some credence to the view that, as long as some of the non-stationarity is distinguishable using directional information, then a wavelet decomposition with greater directionality may provide a more sensitive test.

5 Conclusions and further work

This work has focused on the case where the LSW spectrum is smoothly varying in the Lipschitz sense. However, for complicated image scenes with multiple objects, and distinct boundaries between those objects, a more realistic scenario would be to permit piecewise jumps in the underlying spectrum. This alternative formalism, whereby the transfer function W is assumed to have piecewise bounded variation, has been explored for univariate, time-series LSW models in several other works such as Fryzlewicz and Nason (2006) and Nason and Stevens (2015) via the Haar–Fisz device. Extending this concept to the spatial and complex LSW model would provide a natural target for further work.

The main thrust of this work was to incorporate a wavelet basis into the existing LSW model which was even more overcomplete than the existing undecimated basis. This idea can be pushed further. Other bases are possible such as the M-band extension of the dual-tree wavelets constructed by Chaux et al. (2006), wavelet packets schemes such as Bayram and Selesnick (2008), and mixtures of wavelets such as Nelson (2015). Although these constructions offer useful additional functionality, their statistical properties have yet to be explored. Placing such constructions in the LSW framework could provide an interesting means to extend the current theory and functionality.

Further work towards the non-stationarity detection application could also consider a more complete range of anisotropy testing machinery such as that proposed by Mondal and Percival (2012) and Thon et al. (2015). It could also consider incorporating the dual-tree complex framework into the very recent multibasis approach of Cardinali and Nason (2016) which discovers a broader range of non-stationary behaviours that is possible with only one basis on its own.

Acknowledgements This paper was primarily written by Dr. James D. B. Nelson who sadly passed away September 2016. At various points throughout his life all the co-authors of this paper were fortunate enough to have worked with James. It is an honour and privilege to complete this work on behalf of, and in the memory of such a humble and intelligent member of our community.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

A Proofs

Lemma 2 Let w , W , and C be the transfer sequence, function, and coefficients, respectively, from Definition 3. Then

$$|w_\eta(t)|^2 = |W_\eta(t/T)|^2 + \mathcal{O}(C_\eta/T). \quad (9)$$

Proof The result is easily obtained from Definition 1, property (1), viz.

$$\begin{aligned} \sup_t |w_\eta(t) - W_\eta(t/T)| &\leq \frac{C_\eta}{T} \\ \sup_t |w_\eta(t)| - |W_\eta(t/T)| &\leq \frac{C_\eta}{T} \\ |w_\eta(t)| &= |W_\eta(t/T)| + \mathcal{O}(C_\eta/T) \\ |w_\eta(t)|^2 &= |W_\eta(t/T)|^2 + \mathcal{O}(C_\eta/T) \\ &= S_\eta(t/T) + \mathcal{O}(C_\eta/T). \end{aligned}$$

□

Lemma 3 Let S be the local wavelet spectrum of the LSW process $X \in \text{LSW}(\mathbb{C})$. Then

$$S_\eta\left(\frac{k+\cdot}{T}\right) = S_\eta\left(\frac{\cdot}{T}\right) + \mathcal{O}(\lambda_\eta \|k/T\|_1).$$

Proof The proof follows that of Eckley et al. (2010) without any necessary adjustments. We include it here for completeness. Using Definition 1, property (3), the triangular inequality, the square summability of the transfer function (Definition 1, property 2) and the boundedness of the Lipschitz constants (Definition 1, property 4), respectively, we have that

$$\begin{aligned} |W_\eta\left(\frac{k+\cdot}{T}\right) - W_\eta\left(\frac{\cdot}{T}\right)| &\leq \lambda_\eta \|k/T\|_1 \\ |W_\eta\left(\frac{k+\cdot}{T}\right)| - |W_\eta\left(\frac{\cdot}{T}\right)| &\leq \lambda_\eta \|k/T\|_1 \end{aligned}$$

$$|W_\eta\left(\frac{k+\cdot}{T}\right)| = |W_\eta\left(\frac{\cdot}{T}\right)| + \mathcal{O}(\lambda_\eta\|k/T\|_1)$$

$$|W_\eta\left(\frac{k+\cdot}{T}\right)|^2 = |W_\eta\left(\frac{\cdot}{T}\right)|^2 + \mathcal{O}(\lambda_\eta\|k/T\|_1).$$

□

Theorem 1 (Convergence of auto-covariance) *Let C and C_T be, respectively, the local auto-covariance and auto-covariance of an LSW process $\{X(t)\}_{t \in \mathcal{T}} \in \text{LSW}(\mathbb{C})$. Then*

$$|C_T(z, t) - C(z, t)| = \mathcal{O}(T^{-1}).$$

as $T \rightarrow \infty$, uniformly in $t \in \mathbb{Z}^2$, and $z \in \mathbb{T}^2$.

Proof Since X is zero mean,

$$C_T(z, t) = \mathbb{E}\{X([zT])X([zT] + t)\}$$

$$= \mathbb{E}\left\{\sum_{\eta, \nu \in \mathcal{H}} (w_\eta \xi_\eta * \psi_\eta)([zT])(w_\nu \xi_\nu * \psi_\nu^*)([zT] + t)\right\},$$

where $[\cdot]$ denotes the integer part of \cdot . The orthonormality of the increments gives

$$C_T(z, t) = \sum_{\eta \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |w_\eta(k + [zT])|^2 \Psi_\eta(k, t),$$

with $\Psi_\eta(k, t) := \psi_\eta(k) \psi_\eta^*(k - t)$. The remainder of the proof now follows the same arguments as Eckley et al. (2010), namely:

$$|C_T(z, t) - C(z, t)|$$

$$= \left| \sum_{\eta \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |w_\eta(k + [zT])|^2 \Psi_\eta(k, t) - \sum_{\eta \in \mathcal{H}} S_\eta(z) \Psi_\eta(t) \right|$$

$$= \left| \sum_{\eta \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} (|w_\eta(k + [zT])|^2 - S_\eta(k/T + z)) \Psi_\eta(k, t) \right.$$

$$\left. - \sum_{\eta \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} (S_\eta(k/T + z) - S_\eta(z)) \Psi_\eta(k, t) \right|$$

$$\leq \sum_{\eta \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |\Psi_\eta(k, t)| \left(\left| |w_\eta(k + [zT])|^2 - S_\eta(k/T + z) \right| \right.$$

$$\left. + |S_\eta(k/T + z) - S_\eta(z)| \right)$$

Lemmas 2 and 3 now yield

$$|C_T(z, t) - C(z, t)| \leq \sum_{\eta \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |\Psi_\eta(k, t)| \frac{C_\eta + \lambda_\eta \|k\|_1}{T},$$

and the result is obtained by invoking the boundedness properties (1 and 4 from Definition 1) of the constants C and λ . □

Theorem 2 (Periodogram bias) *Let the LSW(\mathbb{C}) process $X : \mathcal{T} \mapsto \mathbb{R}$, defined on the lattice $\mathcal{T} := \llbracket 1, T \rrbracket^2$, have periodogram $X_\eta^\sim := |X * \psi_\eta^*|^2$, where $\{\psi_\eta\}_{\eta \in \mathcal{H}}$ is a discrete, undecimated, dual-tree complex wavelet basis and let A be the associated biasing matrix. Then*

$$\mathbb{E}X_\eta^\sim = \sum_v A_{\eta\nu} S_\nu\left(\frac{\cdot}{T}\right) + \mathcal{O}(T^{-1}) \tag{10}$$

Proof The proof follows the spirit of Eckley et al. (2010), but, here, special attention is required to take care of the fact that both the wavelet functions ψ and the transfer sequences w are complex-valued, rather than real-valued. In addition, a slightly more ‘operator-theory’ flavour is applied to the final elements of the calculations whereby convolution properties are exploited to circumvent the need for some tedious changes of variables.

$$\mathbb{E}X_\eta^\sim = \mathbb{E}|X * \psi_\eta^*|^2$$

$$= \mathbb{E}\left| \sum_{t \in \mathcal{T}} X(t) \psi_\eta^*(\cdot - t) \right|^2$$

$$= \mathbb{E}\left| \sum_{t \in \mathcal{T}} \sum_{\nu \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} w_\nu(k) \xi_\nu(k) \psi_\nu(k - t) \psi_\eta^*(\cdot - t) \right|^2$$

$$= \sum_{\nu \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |w_\nu(k)|^2 \left| \sum_{t \in \mathcal{T}} \psi_\nu(k - t) \psi_\eta^*(\cdot - t) \right|^2$$

$$= \sum_{\nu \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |w_\nu(k + \cdot)|^2 \left| \sum_{t \in \mathcal{T}} \psi_\nu(k - t) \psi_\eta^*(-t) \right|^2$$

$$= \sum_{\nu \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} |w_\nu(k + \cdot)|^2 |(\psi_\nu * \overline{\psi_\eta^*})(k)|^2,$$

where the fourth equality results from the orthonormality of the increment processes ξ . From Lemma 2, we now use the fact that

$$|w_\nu(k + \cdot)|^2 = S_\nu\left(\frac{k + \cdot}{T}\right) + \mathcal{O}\left(\frac{C_\nu}{T}\right),$$

to write

$$\mathbb{E}X_\eta^\sim = \sum_{\nu \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} \left(S_\nu\left(\frac{k + \cdot}{T}\right) + \mathcal{O}\left(\frac{C_\nu}{T}\right) \right) |(\psi_\nu * \overline{\psi_\eta^*})(k)|^2.$$

Now, since $\sum_{\nu \in \mathcal{H}} C_\nu < \infty$ and, since the $|(\psi_\nu * \overline{\psi_\eta^*})(k)|^2$ is finite and bounded as a function of k (recall ψ has finite support), then

$$\mathbb{E}X_\eta^\sim = \sum_{\nu \in \mathcal{H}} \sum_{k \in \mathbb{Z}^2} S\left(\frac{k + \cdot}{T}\right) |(\psi_\nu * \overline{\psi_\eta^*})(k)|^2 + \mathcal{O}(T^{-1}).$$

From Lemma 3, we similarly use

$$S\left(\frac{k+\cdot}{T}\right) = S\left(\frac{\cdot}{T}\right) + \mathcal{O}\left(\frac{\lambda_v \|k\|_1}{T}\right),$$

to get that

$$\mathbb{E}X_{\eta}^{\sim} = \sum_{v \in \mathcal{H}} S\left(\frac{\cdot}{T}\right) \sum_{k \in \mathbb{Z}^2} |(\psi_v * \overline{\psi_{\eta}^*})(k)|^2 + \mathcal{O}(T^{-1}).$$

Finally, the following convolution property completes the proof.

$$\begin{aligned} \sum_{k \in \mathbb{Z}^2} |(\psi_v * \overline{\psi_{\eta}^*})(k)|^2 &= \sum_{k \in \mathbb{Z}^2} (\psi_v * \overline{\psi_{\eta}^*})(k) (\overline{\psi_v} * \psi_{\eta}^*)(k) \\ &= \sum_{k \in \mathbb{Z}^2} (\overline{\psi_{\eta}^*} * \psi_v)(k) (\overline{\psi_v} * \psi_{\eta}^*)(k) \\ &= \sum_{k \in \mathbb{Z}^2} \Psi_{\eta}(k) \Psi_v^*(k) \\ &= \langle \Psi_{\eta}, \Psi_v \rangle \\ &= A_{\eta v}. \end{aligned}$$

□

Lemma 1 (Bias is real) *Let Ψ be the ACW (Definition 2). The entries of the biasing matrix (Definition 5) are positive, real, and symmetric.*

Proof By construction, A is a Gram matrix. Hence, symmetry follows immediately. We then invoke Plancherel to write $A_{\eta, v} = \langle \psi_{\eta} * \overline{\psi_{\eta}^*}, \psi_v * \overline{\psi_v^*} \rangle = \langle |\psi_{\eta}^{\wedge}|^2, |\psi_v^{\wedge}|^2 \rangle$. □

Theorem 3 (Invertibility of the biasing matrix) *Let the biasing matrix A' associated with an appropriately chosen real-valued wavelet $\psi_{\cdot;0}$ defined over $\llbracket 1, T \rrbracket$ be non-singular. Then, the biasing matrix A (Definition 5) associated with the dual-tree wavelet (cf. Property 5 of Definition 1) $\psi = \psi_{\cdot;0} + i\psi_{\cdot;1}$ is non-singular with $A = 2A'$.*

Proof We use the same arguments as the proof of Lemma 1 and then note that the Fourier magnitudes of the real and imaginary parts of the dual-tree wavelets are identical by design:

$$\begin{aligned} A_{\eta, v} &= \langle \psi_{\eta} * \overline{\psi_{\eta}^*}, \psi_v * \overline{\psi_v^*} \rangle \\ &= \langle |\psi_{\eta}^{\wedge}|^2, |\psi_v^{\wedge}|^2 \rangle \\ &= \langle |\psi_{\eta;0}^{\wedge}|^2 + |\psi_{\eta;1}^{\wedge}|^2, |\psi_{v;0}^{\wedge}|^2 + |\psi_{v;1}^{\wedge}|^2 \rangle \\ &= 2 \langle |\psi_{\eta;0}^{\wedge}|^2, |\psi_{v;0}^{\wedge}|^2 \rangle. \end{aligned}$$

Unlike the brief proof in Nelson and Gibberd (2016), we here justify why the real and imaginary parts have the same filter characteristic, namely $|\psi_{\eta;1}^{\wedge}|^2 = |\psi_{\eta;0}^{\wedge}|^2$. The strategy is to

write $\psi_{\eta;1}^{\wedge}$ in terms of the filter sequence frequency responses $h_1^{1\wedge}$ and $h_0^{1\wedge}$. The filter relations in Eqs. (3) and (4) are then used to write these in terms of $h_1^{0\wedge}$ and $h_0^{0\wedge}$ and hence strike an explicit comparison with respect to the filter response of the real wavelet $\psi_{\eta;0}^{\wedge}$.

Firstly, from the dual-tree wavelet filter properties (3) and (4) and the dual-tree version of the recursions in Equation (2), namely:

$$\psi_{j;a} = \sum_k h_0^a(\cdot - 2k) \psi_{j-1}(k), \quad j > 0 \tag{11}$$

$$\psi_{0;a} = h_1^a, \tag{12}$$

we have that:

$$\begin{aligned} \psi_{0;1}^{\wedge}(\omega) &= h_1^{1\wedge}(\omega) = \frac{1}{\sqrt{2}} \sum_t h_1^1(t) e^{-i\omega t} \\ &= (-1)^{n-1} \frac{1}{\sqrt{2}} \sum_t h_1^0(n-t-1) e^{-i\omega t} \\ &= (-1)^{n-1} e^{-i(n-1)\omega} h_1^{0\wedge}(-\omega). \end{aligned} \tag{13}$$

Likewise, we have that

$$\begin{aligned} h_0^{1\wedge}(\omega) &= \frac{1}{\sqrt{2}} \sum_t h_0^1(t) e^{-i\omega t} \\ &= \frac{1}{\sqrt{2}} \sum_t h_0^0(n-t-1) e^{-i\omega t} \\ &= e^{-i(n-1)\omega} h_0^0(-\omega). \end{aligned} \tag{14}$$

Now, from Equation (11), we have

$$\psi_{j;0}^{\wedge}(\omega) = 2^{(j-1)/2} h_1^{0\wedge}(2^j \omega) \prod_{\ell=0}^{j-1} h_0^{0\wedge}(2^{\ell} \omega).$$

Similarly,

$$\psi_{j;1}^{\wedge}(\omega) = 2^{(j-1)/2} h_1^{1\wedge}(2^j \omega) \prod_{\ell=0}^{j-1} h_0^{1\wedge}(2^{\ell} \omega).$$

Now, using (13) and (14) gives

$$\begin{aligned} \psi_{j;1}^{\wedge}(\omega) &= 2^{(j-1)/2} (-1)^{n-1} e^{-i(n-1)2^j \omega} h_1^{0\wedge}(-2^j \omega) \times \\ &\quad \prod_{\ell=0}^{j-1} e^{-i(n-1)2^{\ell} \omega} h_0^0(-2^{\ell} \omega) \\ &= z_j 2^{(j-1)/2} h_1^{0\wedge}(-2^j \omega) \prod_{\ell=0}^{j-1} h_0^0(-2^{\ell} \omega), \end{aligned}$$

with $z_j = (-1)^{n-1} e^{-i(n-1)(2^{j+1}-1)\omega}$. But the right-hand side can be expressed in terms of $\psi_{j;1}^\wedge(\omega)$, viz.

$$\begin{aligned} \psi_{j;1}^\wedge(\omega) &= z_j \psi_{j;0}^\wedge(-\omega) \\ &= z_j \psi_{j;0}^{\wedge*}(\omega). \end{aligned}$$

Noting that $|z_j| = 1$ completes the proof. \square

Theorem 4 (Linear independence of ACW) *Let the auto-correlation wavelet Ψ' associated with an appropriately chosen real discrete wavelet function $\psi_{\cdot;0}$ defined over $\llbracket 1, T \rrbracket$ be linearly independent. Then, the auto-correlation wavelet Ψ associated with the dual-tree wavelet $\psi = \psi_{\cdot;0} + i\psi_{\cdot;1}$ is also linearly independent.*

Proof This result holds courtesy of the equivalent result in Theorem 3. However, we here provide an extra proof restricted to the real scalars as an alternative means to establish Corollary 1.

We have

$$\begin{aligned} \sum_{\eta \in \mathcal{H}} \Delta_\eta(z) \Psi_\eta(t) &= 0 \\ \Rightarrow \Re \sum_{\eta \in \mathcal{H}} \Delta_\eta(z) \Psi_\eta(t) &= 0 \\ \Rightarrow \sum_{\eta \in \mathcal{H}} \Delta_\eta(z) \Re(\psi_\eta * \overline{\psi_\eta^*})(t) &= 0 \\ \Rightarrow \sum_{\eta \in \mathcal{H}} \Delta_\eta(z) (\psi_{\eta;0} * \overline{\psi_{\eta;0}} + \psi_{\eta;1} * \overline{\psi_{\eta;1}}) &= 0 \end{aligned}$$

Now we note that $\psi_{\eta;1} * \overline{\psi_{\eta;1}} = \psi_{\eta;0} * \overline{\psi_{\eta;0}}$ because convolution is commutative and $\psi_{\cdot;1}$ (and $\psi_{\cdot;1}$) are just translated versions of $\psi_{\cdot;0}$ (and $\overline{\psi_{\cdot;0}}$), respectively. Hence, we have established that

$$\sum_{\eta \in \mathcal{H}} \Delta_\eta(z) \Psi_\eta(t) = 0 \Rightarrow \sum_{\eta \in \mathcal{H}} \Delta_\eta(z) \Psi'_\eta(t) = 0,$$

for all $t \in \mathbb{Z}^2, z \in \mathbb{T}^2$. But we know that, since Ψ' is linearly independent we must therefore have that $\Delta_\eta(z) = 0$. Hence, Ψ is also linearly independent. \square

Corollary 1 (LWS uniqueness) *The LWS is uniquely defined, given the corresponding LSW(C) process.*

Proof Similar to Nason et al. (2000) and Eckley et al. (2010), we follow a proof by contradiction and assume that the LWS is not uniquely defined. In which case, there exist two different LWS sequences $\{S_\eta^{(1)}(z)\}_{z \in \mathbb{T}^2}$ and $\{S_\eta^{(2)}(z)\}_{z \in \mathbb{T}^2}$, say, which are associated with the same process and hence give rise to the same local auto-correlation, namely

$$C(z, t) = \sum_{\eta \in \mathcal{H}} S_\eta^{(1)}(z) \Psi_\eta(t) = \sum_{\eta \in \mathcal{H}} S_\eta^{(2)}(z) \Psi_\eta(t),$$

for $z \in \mathbb{T}^2, t \in \mathbb{Z}^2$. Defining

$$\Delta_\eta(z) := S_\eta^{(1)}(z) - S_\eta^{(2)}(z) \in \mathbb{R},$$

this would mean that

$$\sum_{\eta \in \mathcal{H}} \Delta_\eta(z) \Psi_\eta(t) = 0. \tag{15}$$

However, by Theorem 4, Ψ is linearly independent with respect to real scalars and the only way that (15) could hold, therefore, is if $\Delta_\eta(z) = 0$ which contradicts the original assumption. \square

Proposition 1 (LS-DTCW covariance inversion) *Assume that the debiasing matrix A is non-singular, then*

$$S_\eta(z) = \sum_v A_{\eta,v}^{-1} \sum_t C(z, t) \Psi_v^*(t).$$

Proof As in Nelson and Gibberd (2016) and Nason et al. (2000), we follow a proof by verification and then note that A is both symmetric and, by Lemma 1, real.:

$$\begin{aligned} \sum_v A_{\eta,v}^{-1} \sum_t \sum_{\eta'} S_{\eta'}(z) \Psi_{\eta'}(t) \Psi_v^*(t) &= \sum_{\eta'} S_{\eta'}(z) \sum_v A_{\eta,v}^{-1} \sum_t \Psi_{\eta'}(t) \Psi_v^*(t) \\ &= \sum_{\eta'} S_{\eta'}(z) \sum_v A_{\eta,v}^{-1} A_{v,\eta'} = \sum_{\eta'} S_{\eta'}(z) \delta_{\eta,\eta'} = S_\eta(z). \end{aligned}$$

\square

References

Anantrasirichai, N., Achim, A., Bull, D.R., Kingsbury, N.: Atmospheric turbulence mitigation using complex wavelet-based fusion. *IEEE Trans. Image Process.* **22**(6), 2398–2408 (2013)

Bayram, I., Selesnick, I.W.: On the dual-tree complex wavelet packet and M-band transforms. *IEEE Trans. Signal Process.* **56**(6), 2298–2310 (2008)

Cardinali, A., Nason, G.P.: Costationarity of locally stationary time series. *J. Time Ser. Econom.* **2**(2), 1–33 (2010)

Cardinali, A., Nason, G.P.: Practical powerful wavelet packet tests for second-order stationarity. *Appl. Comput. Harmon. Anal.* (2016). <https://doi.org/10.1016/j.acha.2016.06.006>

Chaux, C., Duval, L., Pesquet, J.-C.: Image analysis using a dual-tree M-band wavelet transform. *IEEE Trans. Image Process.* **15**(8), 2397–2412 (2006)

Cho, H., Fryzlewicz, P.: Multiple change-point detection for high-dimensional time series via sparsified binary segmentation. *J. R. Stat. Soc. B* **77**, 475–507 (2015)

Dahlhaus, R.: Fitting time series models to nonstationary processes. *Ann. Stat.* **25**, 1–37 (1997)

Daubechies, I.: *Ten Lectures on Wavelets*. SIAM, Philadelphia (1992)

Davison, A., Hinkley, D.: *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge (1997)

- Do, M.N., Vetterli, M.: Contourlets: a directional multiresolution image representation. In: IEEE International Conference on Image Processing, pp. 357–360 (2002)
- Eckley, I.A., Nason, G.P.: LS2W: locally stationary wavelet fields in R. *J. Stat. Softw.* **43**(3), 1–23 (2011)
- Eckley, I.A., Nason, G.P., Treloar, R.L.: Locally stationary wavelet fields with application to the modeling and analysis of image texture. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **59**(4), 595–616 (2010). <https://doi.org/10.1111/j.1467-9876.2009.00721.x>
- Fryzlewicz, P.: Modelling and forecasting financial log-returns as locally stationary wavelet processes. *J. Appl. Stat.* **32**, 503–528 (2005)
- Fryzlewicz, P., Nason, G.P.: Haar-fisz estimation of evolutionary wavelet spectra. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **68**, 611–634 (2006)
- Gibberd, A.J., Nelson, J.D.B.: Regularised estimation of 2d-locally stationary wavelet processes. In: 2016 IEEE Statistical Signal Processing Workshop (SSP) (2016). <https://doi.org/10.1109/ssp.2016.7551838>
- Hill, P.R., Anantrasirichai, N., Achim, A., Al-Mualla, M.E., Bull, D.R.: Undecimated dual-tree complex wavelet transforms. *Signal Process. Image Commun.* **35**, 61–70 (2015)
- Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *J. Physiol.* **160**(1), 106–154.2 (1962)
- Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D.: Change detection from remotely sensed images: from pixel-based to object-based approaches. *ISPRS J. Photogram. Remote Sens.* **80**, 91–106 (2013)
- Killick, R., Eckley, I.A., Jonathan, P.: A wavelet-based approach for detecting changes in second order structure within nonstationary time series. *Electron. J. Stat.* **7**, 1167–1183 (2013)
- Kingsbury, N.G.: Complex wavelets for shift invariant analysis and filtering of signals. *J. Appl. Comput. Harmon. Anal.* **10**(3), 234–253 (2001)
- Kingsbury, N.G.: Design of q-shift complex wavelets for image processing using frequency domain energy minimisation. In: IEEE International Conference on Image Processing (2003)
- Knight, M.I., Nunes, M.A., Nason, G.P.: Spectral estimation for locally stationary time series with missing observations. *Stat. Comput.* **22**(4), 877–895 (2012)
- Mondal, D., Percival, D.B.: Wavelet variance analysis for random fields on a regular lattice. *IEEE Trans. Image Process.* **21**(2), 537–549 (2012)
- Nadaraya, E.: On estimating regression. *Theory Probab. Appl.* **9**(1), 141–142 (2016)
- Nason, G.P.: A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **75**, 879–904 (2013)
- Nason, G., Stevens, K.: Bayesian wavelet shrinkage of the Haar-Fisz transformed wavelet periodogram. *PLoS One* **10**(9), e0137662 (2015). <https://doi.org/10.1371/journal.pone.0137662>
- Nason, G.P., von Sachs, R., Kroisandt, G.: Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **62**, 271–292 (2000)
- Nelson, J.D.B.: Enhanced B-wavelets via mixed, composite packets. *IEEE Trans. Signal Process.* **63**(12), 3191–3203 (2015)
- Nelson, J.D.B., Gibberd, A.J.: Introducing the locally stationary dual-tree complex wavelet model. In: IEEE International Conference on Image Processing (2016)
- Nelson, J.D.B., Kingsbury, N.G.: Fractal dimension based sand ripple suppression for mine hunting with sidescan sonar. In: International Conference on Synthetic Aperture Sonar and Synthetic Aperture Radar (2010)
- Nelson, J.D.B., Kingsbury, N.G.: Fractal dimension, wavelet shrinkage and anomaly detection for mine hunting. *IET Signal Process.* **6**(5), 484–493 (2012)
- Nelson, J.D.B., Naornita, C., Isar, A.: Semi-local scaling exponent estimation with box-penalty constraints and total-variation regularization. *IEEE Trans. Image Process.* **25**(7), 3167–3181 (2016)
- Ng, J., Bharath, A.A., Zhaoping, L.: A survey of architecture and function of the primary visual cortex (V1). *EURASIP J. Adv. Signal Process.* **2007**(1), 097961 (2006). <https://doi.org/10.1155/2007/97961>
- Nunes, M.A., Taylor, S., Eckley, I.A.: A multiscale test of spatial stationarity for textured images in R. *R J.* **6**, 20–30 (2014)
- Portilla, J., Simoncelli, E.P.: A parametric texture model based on joint statistics of complex wavelet coefficients. *Int. J. Comput. Vis.* **40**, 49–70 (2000)
- Raj, R.G., Bovik, A.C., Geisler, W.S.: Non-stationarity detection in natural images. In: IEEE International Conference on Image Processing (2007)
- Sanderson, J., Fryzlewicz, P., Jones, M.W.: Estimating linear dependence between nonstationary time series using the locally stationary wavelet model. *Biometrika* **97**(2), 435–446 (2010)
- Selesnick, I.W., Baraniuk, R.G., Kingsbury, N.G.: The dual-tree complex wavelet transform. *IEEE Signal Process. Mag.* **22**(6), 123–151 (2005)
- Starck, J.-L., Cands, E.J., Donoho, D.L.: The curvelet transform for image denoising. *IEEE Trans. Image Process.* **11**(6), 670–684 (2002)
- Taylor, S., Eckley, I.A., Nunes, M.A.: A test of stationarity for textured images. *Technometrics* **56**, 291–301 (2014)
- Taylor, S., Eckley, I.A., Nunes, M.A.: Multivariate locally stationary 2d wavelet processes with application to colour texture analysis. *Stat. Comput.* (2017)
- Thon, K., Geilhufe, M., Percival, D.B.: A multiscale wavelet-based test for isotropy of random fields on a regular lattice. *IEEE Trans. Image Process.* **24**(2), 694–708 (2015)
- Unser, M., Sage, D., Van De Ville, D.: Multiresolution monogenic signal analysis using the Riesz–Laplace wavelet transform. *IEEE Trans. Image Process.* **18**(11), 2402–2418 (2009)
- van Bellegem, S., von Sachs, R.: Locally adaptive estimation of evolutionary wavelet spectra. *Ann. Stat.* **36**, 1879–1024 (2008)
- von Sachs, R., Nason, G. P., Kroisandt, G.: Spectral representation and estimation for locally-stationary wavelet processes. In: Dubuc, S. (eds.) CRM Proceedings and Lecture Notes, vol. 18, pp. 381–397 (1998)
- Watson, G.: Smooth regression analysis. *Indian J. Stat.* **26**, 359–372 (1964)
- Zhang, G., Kingsbury, N.: Variational Bayesian image restoration with group-sparse modeling of wavelet coefficients. *Digit. Signal Process. (Special Issue in Honour of William J. (Bill) Fitzgerald)* **47**, 157–168 (2015)