The Annals of Statistics
2003, Vol. 31, No. 6, 1772–1821

ESTIMATING DEFORMATIONS OF STATIONARY PROCESSES

BY MAUREEN CLERC[1] AND STÉPHANE MALLAT[2]

Ecole Nationale des Ponts et Chaussées, and Ecole Polytechnique and Courant Institute

This paper studies classes of nonstationary processes, such as warped processes and frequency-modulated processes, that result from the deformation of stationary processes. Estimating deformations can often provide important information about an underlying physical phenomenon. A computational harmonic analysis viewpoint shows that the deformed autocovariance satisfies a transport equation at small scales, with a velocity proportional to a deformation gradient. We derive an estimator of the deformation from a single realization of the deformed process, with a proof of consistency under appropriate assumptions.

1. Introduction. When a nonstationary process X results from the deformation of a stationary process Y, estimating the deformation can provide important information about an underlying physical process of interest. From one realization of $X = DY$, we wish to recover the deformation operator D, which is assumed to belong to a specific transformation group $\mathcal{D}$. For example, a Doppler effect produces a warping deformation in time $X(x) = Y(\theta(x))$, where $\theta'(x)$ depends on velocity. The deformation of a stationary texture by perspective in an image also produces a warping, where $x \in \mathbb{R}^2$ is a spatial variable; recovering the Jacobian matrix of $\theta(x)$ characterizes the shape of the three-dimensional surface that is being viewed [9]. The frequency modulation of a stationary process $X(x) = Y(x)\exp(i\theta(x))$ corresponds to another class of deformations encountered in signal processing, in transmissions by frequency modulation, where the message is carried by $\theta'(x)$.

Estimating the deformation $D \in \mathcal{D}$ from $X = DY$ is an inverse problem. As we suppose no prior knowledge about the stationary process Y, the deformation D can only be recovered up to the subgroup $\mathcal{G}$ of $\mathcal{D}$ which leaves the set of stationary processes globally invariant. Rather than the deformation itself, we therefore seek to estimate the equivalence class of D in $\mathcal{D}/\mathcal{G}$. We consider cases where $\mathcal{G}$ is a finite-dimensional Lie group. Under appropriate assumptions, this equivalence class can be represented by a vector field on $\mathcal{G}$, which corresponds to a deformation gradient. A local analysis of the deformation is performed by decomposing the

Received June 2000; revised December 2002.

[1]Part of this work was completed while the author was visiting the Statistics Department at Stanford University.

[2]Supported in part by MURI Grant F49620-96-1-0028 and NSF Grant IIS-0114391.

AMS 2000 subject classifications. Primary 62M10, 60G12; secondary 60G35.

Key words and phrases. Nonstationary processes, inverse problem, wavelets, warping, frequency modulation, scalogram, spectrogram.

autocovariance of X over an appropriate family of localized functions, which are called atoms in the harmonic analysis literature. The deformation gradient is shown to appear as a velocity vector in a transport equation satisfied by a localized autocovariance. This general result is applied to one-dimensional warping and frequency modulation, where the atoms are classical wavelets, and multidimensional warping, where the atoms are called warplets.

Computing the deformation gradient requires estimating the autocovariance of X projected over a family of localized atoms, from a single realization. Under certain conditions on the autocovariance of the stationary process Y, one can obtain consistent estimators for one-dimensional warping and frequency modulation. Numerical examples illustrate these results. Let us mention that the stationarity hypothesis on Y can be relaxed by supposing only that Y has stationary increments, in which case our estimation of the deformation gradient remains consistent.

The paper is organized in three main sections: after discussing the well-posedness of the inverse problem in Section 2, we establish in Section 3 a transport equation for the localized autocovariance of a deformed process; Section 4 introduces estimators and proves their consistency.

2. Inverse problem. We want to estimate a deformation operator D, which belongs to a known group $\mathcal{D}$, from a single realization of $X = DY$. The process Y, which is not known a priori, is assumed wide sense stationary. Since we are limited to a single realization, we concentrate on second-order moments. For this reason, stationarity will always be understood in the wide sense, meaning that

$$\mathbb{E}\{Y(x)\} = \mathbb{E}\{Y(0)\}$$

and

$$\mathbb{E}\{Y(x)Y^*(y)\} = c_Y(x-y) \qquad \text{with } c_Y(0) < +\infty,$$

where z^* denotes the complex conjugate of $z \in \mathbb{C}$. We shall further suppose that $Y(x)$ is stochastically continuous, which means that its covariance $c_Y(x)$ is continuous at $x = 0$. The deformation operators D that we shall consider are defined over distributions, but, for simplicity, we restrict their domain to functions of $\mathbb{R}^d$. An operator D acts on a stochastic process Y realization by realization. We shall thus consider processes Y whose realizations are functions of $\mathbb{R}^d$. For example, if Y is a Gaussian process and $|c_Y(x) - c_Y(0)| = O(|\log|x||^{-1-\varepsilon})$ for some $\varepsilon > 0$, then one can prove that its realizations are continuous with probability 1 [2]. This hypothesis will be satisfied by our estimation theorems.

2.1. *Class of solutions.* Since we only know that the process Y is stationary, the set of solutions to the inverse problem is the set of all operators $\tilde{D} \in \mathcal{D}$ such that $\tilde{D}^{-1}X$ is stationary. In general, this set is larger than $\{D\}$. Let $\mathcal{G}$ be the set of all operators $G \in \mathcal{D}$ such that if Y is a wide-sense stationary process, then GY

is also wide-sense stationary. One can verify that $\mathcal{G}$ is a subgroup of $\mathcal{D}$, which we call the *stationarity-invariant group*. Clearly, if D is a solution of the inverse problem, any operator $\tilde{D} = DG$ with $G \in \mathcal{G}$ is also a solution. The set of solutions of the inverse problem therefore contains the equivalence class of D in the quotient group $\mathcal{D}/\mathcal{G}$. In order for the set of solutions to the inverse problem to be exactly equal to the equivalence class of D in $\mathcal{D}/\mathcal{G}$, we need to impose a condition on the stationary process Y, so that any deformation $\tilde{D} \in \mathcal{D}$ such that $\tilde{D}Y$ is wide-sense stationary necessarily belongs to $\mathcal{G}$. This is not true for all stationary processes Y, but we give sufficient conditions on the covariance of Y to guarantee this form of uniqueness. In this paper, we concentrate on three deformation groups.

EXAMPLE 1. The frequency modulation group modifies the signal frequency:

$$\mathcal{D} = \{D : Df(x) = e^{i\theta(x)} f(x), \text{ where } \theta(x) \text{ is real and } \mathbf{C}^4\}. \tag{1}$$

In transmissions with frequency modulation, $\theta'(x)$ is proportional to the signal to be transmitted, and the stationary process Y, which is in this case assumed to have zero mean, is the carrier. The stationarity-invariant group is

$$\mathcal{G} = \{G_{(\phi,\xi)} : G_{(\phi,\xi)} f(x) = e^{i(\phi+\xi x)} f(x), \text{ with } (\phi, \xi) \in \mathbb{R}^2\}.$$

Two operators D_1 and D_2 such that $D_1 f(x) = e^{i\theta_1(x)} f(x)$ and $D_2 f(x) = e^{i\theta_2(x)} f(x)$ are in the same equivalence class in $\mathcal{D}/\mathcal{G}$ if and only if $\theta_1(x) = \phi + \xi x + \theta_2(x)$ and hence

$$\theta_1''(x) = \theta_2''(x). \tag{2}$$

The following proposition gives a sufficient condition on the covariance $c_Y(x)$ to identify $\theta''(x)$ from the covariance of $X = DY$. The proof is given in Appendix A.

PROPOSITION 2.1. *Let $X = DY$, where Y is a stationary process and D belongs to the frequency modulation group $\mathcal{D}$ in* (1). *If there exists an $\varepsilon > 0$ such that*

$$\forall x \in [-\varepsilon, \varepsilon], \qquad c_Y(x) > 0,$$

then the equivalence class of D in $\mathcal{D}/\mathcal{G}$ is uniquely characterized by the covariance of X.

EXAMPLE 2. The one-dimensional warping group is defined by

$$\mathcal{D} = \{D : D\, f(x) = f(\theta(x)), \text{ where } \theta(x) \text{ is } \mathbf{C}^3 \text{ and } \theta'(x) > 0\}. \tag{3}$$

Such time warping appears in many physical phenomena, such as the Doppler effect. We easily verify that the stationarity-invariant group is the affine group

$$\mathcal{G} = \{G_{(u,s)} : G_{(u,s)} f(x) = f(u + sx), \text{ with } (u, s) \in \mathbb{R} \times \mathbb{R}^{+*}\}.$$

Note that operators in the stationarity invariant group $\mathcal{G}$ are not time invariant: they do not commute with translations.

Two warping operators D_1 and D_2 are in the same equivalence class in $\mathcal{D}/\mathcal{G}$ if and only if there exists (u, s) such that $\theta_1(x) = u + s\,\theta_2(x)$ or, equivalently,

$$\frac{\theta_1''(x)}{\theta_1'(x)} = \frac{\theta_2''(x)}{\theta_2'(x)}. \tag{4}$$

The following proposition, whose proof is given in Appendix A, gives a sufficient condition on Y to characterize the equivalence class of D uniquely. Perrin and Senoussi [13] provide a similar result.

PROPOSITION 2.2. *Let $X = DY$, where Y is stationary, and $D \in \mathcal{D}$, where $\mathcal{D}$ is the warping group* (3). *If there exists an $\varepsilon > 0$ such that c_Y is $\mathbf{C}^1$ on $(0, \varepsilon]$ and*

$$\forall x \in (0, \varepsilon], \qquad c_Y'(x) < 0, \tag{5}$$

then the equivalence class of D in $\mathcal{D}/\mathcal{G}$ is uniquely characterized by the covariance of X.

Condition (5) is met by a very wide range of stationary processes, including Poisson pulse processes, and Ornstein–Uhlenbeck processes [16]. White noise, however, violates (5) since $c_Y(x) = 0$ for $x \neq 0$.

EXAMPLE 3. The warping problem in two dimensions has an important application in image analysis, particularly in recovering a three-dimensional surface shape by analyzing texture deformations. Warping deformations are also used in geostatistics (see [12] and [15]) to model nonstationary phenomena. Stationarizing the data is suggested as an initial step before applying classical geostatistical methods such as kriging. We study a d-dimensional warping problem, specified by an invertible function $\theta(x)$ from $\mathbb{R}^d$ to $\mathbb{R}^d$ with

$$\theta(x_1, \ldots, x_d) = \big(\theta_1(x_1, \ldots, x_d), \ldots, \theta_d(x_1, \ldots, x_d)\big).$$

The Jacobian matrix of θ at position $x \in \mathbb{R}^d$ is written as

$$J_\theta(x) = \left(\frac{\partial \theta_i(x)}{\partial x_j}\right)_{1 \le i, j \le d}. \tag{6}$$

If the Jacobian determinant $\det J_\theta(x)$ does not vanish, it corresponds to a change of metric and $\theta(x)$ is invertible. We consider a group of regular warping operators

$$\mathcal{D} = \big\{D : Df(x) = f(\theta(x)),\ \text{where } \theta(x) \text{ is in } \mathbf{C}^3(\mathbb{R}^d) \text{ and } \det J_\theta(x) > 0\big\}. \tag{7}$$

Let $GL^+(\mathbb{R}^d)$ be the group of linear operators in $\mathbb{R}^d$ with a strictly positive determinant. We easily verify that the stationarity-invariant group is the affine group

$$\mathcal{G} = \big\{G_{(u,S)} : G_{(u,S)} f(x) = f(u + Sx),\ \text{with } (u, S) \in \mathbb{R}^d \times GL^+(\mathbb{R}^d)\big\}.$$

Two operators D and $\tilde{D}$ such that $Df(x) = f(\theta(x))$ and $\tilde{D}f(x) = f(\tilde{\theta}(x))$ are in the same equivalence class in $\mathcal{D}/\mathcal{G}$ if and only if

$$\exists\,(u, S) \in \mathbb{R}^d \times GL^+(\mathbb{R}^d), \qquad \theta(x) = u + S\tilde{\theta}(x). \tag{8}$$

The partial derivative of the Jacobian matrix in a fixed direction x_k is again a matrix:

$$\frac{\partial J_\theta(x)}{\partial x_k} = \left(\frac{\partial^2 \theta_i(x)}{\partial x_k\,\partial x_j}\right)_{1\le i,j\le d}.$$

We will use the notation $\vec{\nabla} J_\theta(x)$ to denote the set of matrices $\{\partial J_\theta(x)/\partial x_k\}_{k=1,\dots,d}$. Condition (8) is equivalent to the following matrix equalities, which generalize (4):

$$\forall\,k \in \{1,\dots,d\}, \qquad J_\theta^{-1}(x)\frac{\partial J_\theta(x)}{\partial x_k} = J_{\tilde{\theta}}^{-1}(x)\frac{\partial J_{\tilde{\theta}}(x)}{\partial x_k}. \tag{9}$$

Proof of this equivalence can be found at the end of the proof of Proposition 2.3 in Appendix A.

There are cases for which the inverse warping problem cannot be solved. For example, consider a stationary process $Y(x) = Y_1(x_1)$ that only depends on the first variable and a warping deformation that leaves x_1 invariant: $\theta(x_1,\dots,x_d) = (x_1, \theta_1(x_2,\dots,x_d))$. In this case,

$$X(x) = Y\big(x_1, \theta_1(x_2,\dots,x_d)\big) = Y_1(x_1) = Y(x). \tag{10}$$

Hence, θ cannot be recovered. The following proposition, whose proof is given in Appendix A, gives a sufficient condition on $c_Y(x)$ to guarantee that the inverse warping problem has a unique solution in $\mathcal{D}/\mathcal{G}$.

PROPOSITION 2.3. *Let $X = DY$, where Y is stationary, and $D \in \mathcal{D}$, where $\mathcal{D}$ is the multidimensional warping group* (7). *If the covariance of Y satisfies*

$$c_Y(0) - c_Y(x) = |x|^h \eta(x) \qquad \text{with } h > 0 \text{ and } \eta(0) > 0, \tag{11}$$

where $\eta(x)$ is $\mathbf{C}^2$ in a neighborhood of 0, *then the equivalence class of D in $\mathcal{D}/\mathcal{G}$ is uniquely characterized by the autocovariance of X.*

The inverse warping problem has been applied to the reconstruction of three-dimensional surfaces from deformations of textures in images [6]. One can model the image of a textured three-dimensional surface as

$$X(x) = Y(\theta(x)).$$

The stationary process Y depends on the textured reflectance of the surface, and $\theta(x)$ is the two-dimensional warping due to the imaging process, which projects the surface onto the image plane [5]. We showed in (9) that solving the inverse warping problem is equivalent to computing normalized partial derivatives of the Jacobian matrix J_θ:

$$J_\theta^{-1}(x)\frac{\partial J_\theta(x)}{\partial x_1} \quad \text{and} \quad J_\theta^{-1}(x)\frac{\partial J_\theta(x)}{\partial x_2}. \tag{12}$$

Differential geometry derivations by Gårding [9] have proved that these matrices specify the local orientation and curvature of the three-dimensional surface in the scene. Knowing these surface parameters, it is then possible to recover the three-dimensional coordinates of the surface, up to a constant scaling factor. We will see in Section 3.4 that the Jacobian matrices (12) appear as velocity vectors in a transport equation satisfied by the autocovariance of X.

2.2. *Stationarity-invariant group.* The stationarity-invariant group $\mathcal{G}$ specifies the class of solutions of the inverse problem $X = DY$, and Section 3 will show that it is also an important tool to identify the equivalence class of D in $\mathcal{D}/\mathcal{G}$. This section examines the properties of operators that belong to such a group. Recall that an operator G is said to be *stationarity invariant* if, for any wide-sense stationary and stochastically continuous process Y, the process $X = GY$ is also wide-sense stationary.

The following theorem characterizes this class of operators. We denote by $x \cdot y$ the inner product between two vectors x and y of $\mathbb{R}^d$.

THEOREM 2.1. *An operator G is stationarity invariant if and only if there exists $\hat{\rho}(\omega)$ from $\mathbb{R}^d$ to $\mathbb{C}$ with $\operatorname{ess\,sup}_{\omega\in\mathbb{R}^d} |\hat{\rho}(\omega)| < \infty$, and $\lambda(\omega)$ from $\mathbb{R}^d$ to $\mathbb{R}^d$, such that*

$$Ge^{i\omega\cdot x} = \hat{\rho}(\omega)e^{i\lambda(\omega)\cdot x}. \tag{13}$$

The proof is given in Appendix A. This theorem proves that a stationarity-invariant operator acts on a sinusoid by transposing its frequency and modifying its amplitude. The examples given in the previous section correspond to specific classes of such operators, where $\lambda(\omega)$ is affine in ω. Suppose that $\lambda(\omega) = \overline{S}\omega + \xi$, with $\xi \in \mathbb{R}^d$ and where S is an invertible linear operator in $\mathbb{R}^d$ whose adjoint is denoted $\overline{S}$. In this case, the operator G in (13) satisfies

$$Gf(x) = e^{i\xi\cdot x} f \star \rho(Sx), \tag{14}$$

where $\rho(x)$ is the inverse Fourier transform of $\hat{\rho}(\omega)$. If $\rho(x) = e^{i\phi}\delta(x - v)$, then the operator G defined in (14) consists of both a frequency modulation and a warping.

Let us define the translation operator T_v for $v \in \mathbb{R}^d$ by

$$T_v f(x) = f(x - v).$$

The following proposition proves that linear operators of the form (14) are characterized by a weak form of commutativity with T_v.

PROPOSITION 2.4. *A linear operator G that is bounded in $\mathbf{L}^2(\mathbb{R}^d)$ is stationarity invariant and satisfies* (14) *if and only if it satisfies*

$$\exists \xi \in \mathbb{R}^d,\ \exists S \in GL^+(\mathbb{R}^d),\ \forall v \in \mathbb{R}^d, \qquad G\,T_{Sv} = e^{i\xi \cdot v} T_v G. \tag{15}$$

This result, which can be viewed as a transport property, is proved in Appendix A. In the rest of the paper, we concentrate on deformations for which the stationarity-invariant operators satisfy (15).

3. Conservation and transport. The stationarity of a random process Y is a conservation property of its autocovariance through translation. After deforming Y, one obtains a process $X(x) = DY(x)$, which is no longer stationary and whose autocovariance thus does not satisfy the same conservation property. Yet, we show that the stationarity of Y implies a conservation of the autocovariance of X, along characteristic curves in an appropriate parameter space. These characteristic curves, which identify the equivalence class of D in $\mathcal{D}/\mathcal{G}$, are computed by locally approximating D^{-1} by a "tangential" operator $G_{\beta(v)} \in \mathcal{G}$. Assuming that the stationarity-invariant operators satisfy the transport property (15), the conservation equation can be rewritten as a transport equation whose velocity term, called the *deformation gradient*, is related to $\vec{\nabla}_v \beta(v)$. This deformation gradient characterizes the equivalence class of D in $\mathcal{D}/\mathcal{G}$. Section 3.1 gives the general transport equation and Sections 3.2–3.4 apply this result to one-dimensional warping, frequency modulation and multidimensional warping.

3.1. *Transport in groups.* We consider a stationarity-invariant group $\mathcal{G}$ whose elements satisfy the transport property (15) and can be written under a parametric form

$$G_\beta f(x) = G_{(\phi,\xi,S,v)} f(x) = e^{i(\xi\cdot x+\phi)} f \star \rho(Sx - v),$$

where ρ is a tempered distribution. With this assumption, the stationarity-invariant group $\mathcal{G}$ is a finite-dimensional Lie group. The translation parameter v is isolated because of its particular role, and since the phase ϕ has no influence on the autocovariance, we also set it apart and write

$$G_\beta = e^{i\phi} F_\alpha T_v,$$

with

$$F_\alpha f(x) = e^{i\xi\cdot x} f \star \rho(Sx) \quad \text{and} \quad \alpha = (\xi, S).$$

The group product and inverse are denoted by

$$F_{\alpha_1} F_{\alpha_2} = F_{\alpha_1 * \alpha_2} \quad \text{and} \quad F_\alpha^{-1} = F_{\alpha^{-1}}.$$

To identify the tangential deformation $G_{\beta(v)} \in \mathcal{G}$, which approximates D^{-1} for functions supported in a neighborhood of $v \in \mathbb{R}^d$, we use a family of test functions constructed from a single function $\psi(x)$ whose support is in $[-1,1]^d$. For $\sigma>0$, $\psi_\sigma(x)=\psi(x/\sigma)$ has its support in $[-\sigma,\sigma]^d$. Let $\overline{F}_\alpha$ be the adjoint of F_α. An *atomic decomposition* of a process $X(x)$ is obtained by applying its autocovariance operator to a family of deformed and translated test functions, which are called *atoms*:

$$A_X^\sigma(u,\alpha)=\mathbb{E}\{|\langle X, T_u\overline{F}_\alpha\psi_\sigma\rangle|^2\}.$$

This atomic decomposition only depends on X through its autocovariance.

Let us now explain how to identify the tangential deformation $G_{\beta(v)}$ from a conservation property of atomic decompositions. If Y is a stationary process, then $A_Y^\sigma(u,\alpha)$ does not depend on u; therefore, $\vec{\nabla}_u A_Y^\sigma(u,\alpha)=0$. This is not the case for the atomic decomposition of the deformed process $X=DY$:

$$A_X^\sigma(u,\alpha)=\mathbb{E}\{|\langle X, T_u\overline{F}_\alpha\psi_\sigma\rangle|^2\}=\mathbb{E}\{|\langle Y, \overline{D}T_u\overline{F}_\alpha\psi_\sigma\rangle|^2\}.$$

However, we now show that this atomic decomposition satisfies a conservation property along characteristic lines that depend on D. The following proposition proves that if $\overline{D^{-1}}$ can be approximated by a certain $\overline{G}_{\beta(v)}$, for functions having support in a neighborhood of v, then there exists a function γ such that, for all u and α,

$$\vec{\nabla}_u A_X^\sigma\big(u,\alpha*\gamma(u)\big)\approx 0 \qquad \text{for } \sigma \text{ small}.$$

Before stating the proposition, let us set some notation. If $f(x)$ and $g(x)$ are two functions defined in $\mathbb{R}^d$, then $\vec{\nabla}_x g$ is a vector with d components and $\langle f,\vec{\nabla}_x g\rangle$ is also a vector, whose d components are the inner products $\langle f,\partial g/\partial x_k\rangle$. We denote by $\mathrm{Re}\langle f,\vec{\nabla}_x g\rangle$ the real part of this vector. We write $c(\sigma)=O(\sigma)$ if there exists a constant C such that, for σ small, $|c(\sigma)|\le C\sigma$, without specifying the sign. We define the covariance operator of X by

$$K_X f(x)=\int \mathbb{E}\{X(x)X^*(y)\}f(y)\,dy.$$

PROPOSITION 3.1. *Let $X=DY$, with Y stationary. Suppose that, for each $v\in\mathbb{R}^d$, there exists $\beta(v)$ such that, for each α, the function $\psi_{v,\alpha,\sigma}=\overline{G}_{\beta(v)}T_v\overline{F}_\alpha\psi_\sigma$ satisfies*

$$\begin{aligned}&\big|\mathrm{Re}\big\langle K_X\psi_{v,\alpha,\sigma},\overline{D^{-1}}(\vec{\nabla}_v+\vec{\nabla}_x)\overline{D}\psi_{v,\alpha,\sigma}\big\rangle\big|\\&\qquad=O(\sigma)\big|\mathrm{Re}\langle K_X\psi_{v,\alpha,\sigma},\vec{\nabla}_x\psi_{v,\alpha,\sigma}\rangle\big|.\end{aligned}\tag{16}$$

If there exist a differentiable invertible map $u(v)$ and two functions $\phi(u)$ and $\gamma(u)$ such that

$$\overline{G}_{\beta(v)}\,T_v=e^{i\phi(u(v))}\,T_{u(v)}\overline{F}_{\gamma(u(v))},\tag{17}$$

then, for each (u,α), we have, at $t=u$,

$$\big|\vec{\nabla}_u A_X^\sigma\big(u,\alpha*\gamma(t)\big)+\vec{\nabla}_t A_X^\sigma\big(u,\alpha*\gamma(t)\big)\big|=O(\sigma)\big|\vec{\nabla}_u A_X^\sigma\big(u,\alpha*\gamma(t)\big)\big|.\tag{18}$$

The norms in (18) are Euclidean norms of d-dimensional vectors. The proof of Proposition 3.1 can be found in Appendix B.

We have assumed that the stationarity-invariant group $\mathcal{G}$ is composed of elements of the form $G_\beta = e^{i\phi} F_\alpha T_v$.

To shed some light on the meaning of (16) and the role of the function $\beta(v)$, we examine the case when the deformation operator D itself is a stationarity-invariant operator. In this case, a function β such that $\overline{G}_{\beta(v)} T_v \overline{F}_\alpha \psi_\sigma$ satisfies (16) can be chosen independently of v. Indeed, $\overline{G}_{\beta(v)} = \overline{D^{-1}}$ yields $\overline{D}\psi_{v,\alpha,\sigma} = T_v \overline{F}_\alpha \psi_\sigma$. For any function $f(x)$, $(\vec{\nabla}_v + \vec{\nabla}_x) T_v f(x) = 0$. Therefore,

$$(\vec{\nabla}_v + \vec{\nabla}_x)\overline{D}\psi_{v,\alpha,\sigma} = (\vec{\nabla}_v + \vec{\nabla}_x) T_v \overline{F}_\alpha \psi_\sigma = 0,$$

and the left-hand side of (16) vanishes. When D is not stationarity invariant, $\beta(v)$ must be chosen so that (16) holds when $\sigma \to 0$. This imposes a form of tangency between $\overline{G}_{\beta(v)}$ and $\overline{D^{-1}}$, when applied to functions supported in a neighborhood of v. In the following three sections, we will identify the β which are appropriate for each type of deformation considered.

The partial differential equation (18) that results from the above proposition can be written as a transport equation in the $(u; \alpha)$ domain by expanding the gradient with respect to t:

$$\vec{\nabla}_t A_X^\sigma\big(u, \alpha * \gamma(t)\big) = \vec{\nabla}_t\big(\alpha * \gamma(t)\big) \cdot \vec{\nabla}_\alpha A_X^\sigma\big(u, \alpha * \gamma(t)\big),$$

where $\vec{\nabla}_\alpha A_X^\sigma(u, \alpha)$ is a vector of partial derivatives with respect to each component of parameter α. Replacing the free variable α by $\alpha * \gamma^{-1}(u)$ in (18) gives, at $t = u$,

$$\big|\vec{\nabla}_u A_X^\sigma(u, \alpha) + \vec{\nabla}_t\big(\alpha * \gamma^{-1}(u) * \gamma(t)\big) \cdot \vec{\nabla}_\alpha A_X^\sigma(u, \alpha)\big| = O(\sigma)|\vec{\nabla}_u A_X^\sigma(u, \alpha)|. \tag{19}$$

When σ is sufficiently small, the right-hand side can be neglected, yielding a transport partial differential equation. This is illustrated in the next three sections, in which we apply this proposition to the warping deformation and the frequency modulation problems. Section 4 will afterwards show how, from a single realization of X, we can estimate the partial derivatives of $A_X^\sigma(u, \alpha)$ and compute the deformation gradient.

3.2. *Scale transport.* If D is a one-dimensional warping deformation $Df(x) = f(\theta(x))$ with $x \in \mathbb{R}$, then $\overline{D^{-1}} f(x) = \theta'(x) f(\theta(x))$. The stationarity-invariant subgroup is the affine group, whose elements are $G_\beta f(x) = f(u + sx)$ with $\beta = (u, s)$. The adjoint of G_β is

$$\overline{G}_\beta f(x) = s^{-1} f\big((x - u)/s\big) = T_u \overline{F}_s f(x) \qquad \text{with } \overline{F}_s f(x) = s^{-1} f(x/s).$$

Let ψ be a function whose integral vanishes: $\int \psi(x)\, dx = 0$. The function ψ is called a *wavelet* [11]. Using the above expression of the adjoint operator $\overline{F}_s$, the atomic decomposition $A_X^\sigma(u, s) = \mathbb{E}\{|\langle X, T_u \overline{F}_s \psi_\sigma\rangle|^2\}$ can be written as

$$A_X^\sigma(u, s) = \mathbb{E}\big\{\big|\big\langle X(x), s^{-1}\psi\big((s\sigma)^{-1}(x - u)\big)\big\rangle\big|^2\big\}.$$

We reduce the number of parameters by dividing $A_X^\sigma(u,s)$ by σ^2 and replacing the product $s\sigma$ by a single scale parameter s. The resulting atomic decomposition

$$A_X(u,s) = \mathbb{E}\{|\langle X(x), s^{-1}\psi(s^{-1}(x-u))\rangle|^2\} \tag{20}$$

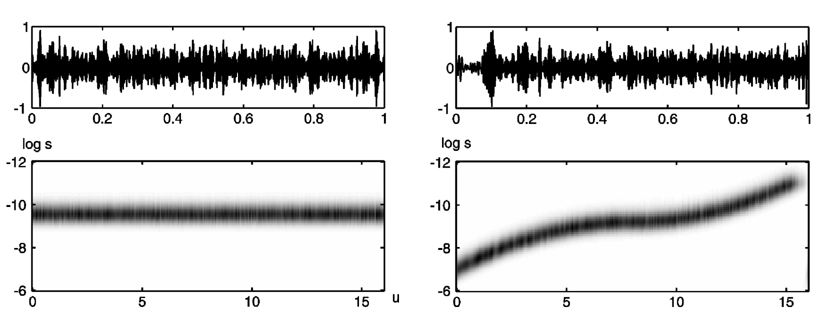
is called a *scalogram* and can be interpreted as the expected value of a squared wavelet transform. Figure 1(a) shows the scalogram $A_Y(u,s)$ of a stationary process Y. As expected, its value does not depend on u. Figure 1(b) gives $A_X(u,s)$ for a warped process $X(x) = DY(x) = Y(\theta(x))$. The warping causes the values of the scalogram of Y to migrate in the $(u;\log s)$ plane.

Let us now give the expression of $\beta(u)$ corresponding to the tangential approximation of Proposition 3.1. A tangential approximation of $\overline{D^{-1}}$ can be found by noting that, for a regular function f supported in a neighborhood of $v = \theta(u)$,

$$\overline{D^{-1}} f(x) \approx \theta'(u) f(v + \theta'(u)(x-u)). \tag{21}$$

The right-hand side of (21) can be written as $\overline{G}_{\beta(v)} f(x)$, and operators $\overline{D^{-1}}$ and $\overline{G}_{\beta(v)}$ both translate the support of f from a neighborhood of v to a neighborhood of $u(v)$.

To derive a transport equation from Proposition 3.1, we must make some assumptions on the autocovariance of Y, which will guarantee the uniqueness of the inverse warping problem at the same time. Proposition 2.2 shows that it is necessary to specify the behavior of the autocovariance kernel $c_Y(x)$ in a neighborhood of 0. The following theorem supposes that $c_Y(x)$ is nearly h-homogeneous in a neighborhood of 0. We denote $\partial f/\partial a = \partial_a f$ and $\partial_{\log s} f = \partial f/\partial \log s = s\,\partial f/\partial s$.



(a) (b)

FIG. 1. (a) *Scalogram $A_Y(u,s)$ of a stationary process Y. The horizontal and vertical axes represent u and $\log s$, respectively. The darkness of a point is proportional to the value of $A_Y(u,s)$.* (b) *Scalogram $A_X(u,s)$ of a warped process X.*

THEOREM 3.1 (Scale transport). *Let Y be a stationary process whose covariance satisfies*

$$c_Y(0) - c_Y(x) = |x|^h \eta(x), \qquad \text{with } h > 0,\ \eta(0) > 0, \tag{22}$$

and where η is $\mathbf{C}^1$ in a neighborhood of 0. *Let $\psi(x)$ be a $\mathbf{C}^1$ function supported in* $[-1, 1]$ *such that*

$$\int \psi(x)\,dx = 0 \quad \text{and} \quad \operatorname{Re} \iint |x-y|^h \psi^*(x)\psi(y)\,dx\,dy \neq 0. \tag{23}$$

If

$$X(x) = Y(\theta(x)),$$

where $\theta(x)$ is $\mathbf{C}^3$ and $\theta'(x) > 0$, then, for each $u \in \mathbb{R}$ such that $\theta''(u) \neq 0$, when s tends to 0

$$\partial_u A_X(u,s) - (\log\theta')'(u)\,\partial_{\log s} A_X(u,s) = O(s)\,\partial_u A_X(u,s). \tag{24}$$

The proof is given in Appendix B. The conditions imposed on c_Y and ψ in this theorem guarantee that $\partial_{\log s} A_X(u,s)$ does not vanish for $s > 0$. The deformation gradient $(\log\theta')'(u)$, which specifies the equivalence class of D in $\mathcal{D}/\mathcal{G}$, can thus be computed from (24) by letting s go to 0. It is therefore not surprising that (22) imposes a stronger condition on c_Y than the uniqueness condition (5) of Proposition 2.2. The estimation of $(\log\theta')'(u)$ from a single realization of X will be studied in Section 4.1.

REMARK 3.1. Since $\psi(x)$ has a zero integral, one can verify that $A_X(u,s)$ can be expressed from the covariance of the increments of $X(x)$, which itself depends on the covariance of the increments of $Y(x)$. It is therefore possible to extend this theorem by supposing only that $Y(x)$ has stationary increments and by replacing (22) by a similar condition on the autocovariance of the increments. Fractional Brownian motions (see [1] and [7]) are examples of processes $Y(x)$ with stationary increments whose deformations by warping satisfy (24).

3.3. *Frequency transport.* If the deformation operator D is a frequency modulation, $Df(x) = e^{i\theta(x)} f(x)$, the stationarity-invariant subgroup $\mathcal{G}$ is composed of operators G_β such that

$$G_\beta f(x) = e^{i(\phi + \xi x)} f(x).$$

In this case, $F_\xi f(x) = e^{i\xi x} f(x)$. Let us choose an even, positive window function $\psi(x) \geq 0$, with a support equal to $[-1, 1]$. The atomic decomposition of process X is the well-known *spectrogram*:

$$\begin{aligned} A_X^\sigma(u,\xi) &= \mathbb{E}\{|\langle X(x), \psi_\sigma(x-u)e^{-i\xi(x-u)}\rangle|^2\} \\ &= \mathbb{E}\{|\langle X(x), \psi_\sigma(x-u)e^{-i\xi x}\rangle|^2\}. \end{aligned}$$

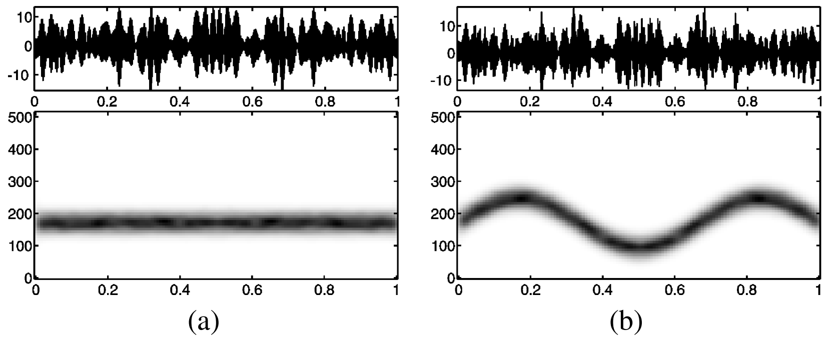


FIG. 2. (a) *Spectrogram $A_Y^\sigma(u,\xi)$ of a stationary process Y. The horizontal and vertical axes represent position u and frequency ξ, respectively. The darkness of a point is proportional to the value of $A_Y^\sigma(u,\xi)$.* (b) *Spectrogram $A_X^\sigma(u,\xi)$ of a frequency-modulated process X.*

Figure 2(a) shows a spectrogram $A_Y^\sigma(u,\xi)$, whose values do not depend on u because Y is stationary. Figure 2(b) depicts $A_X^\sigma(u,\xi)$ for $X(x) = DY(x) = e^{i\theta(x)}Y(x)$, with $\theta(x) = \lambda_1\cos(\lambda_2 x)$, where λ_1 and λ_2 are two constants. The frequency modulation translates the spectrogram of Y nonuniformly along the frequency axis.

Let us now give the expression of $\beta(v)$ corresponding to the tangential approximation of Proposition 3.1, when D is a frequency modulation. A tangential approximation of $\overline{D^{-1}}$ can be found by noting that if f is supported in a neighborhood of v, then

$$\overline{D^{-1}}f(x) = e^{i\theta(x)}f(x) \approx e^{i(\theta(v)+\theta'(v)(x-v))}f(x),$$

and one can define a stationarity-invariant approximation of $\overline{D^{-1}}$ by $G_{\beta(v)}$ such that

$$\overline{G}_{\beta(v)}f(x) = e^{i(\theta(v)+\theta'(v)(x-v))}f(x). \tag{25}$$

The following theorem uses this tangential approximation to derive from Proposition 3.1 a transport equation, satisfied by the spectrogram $A_X^\sigma(u,\xi)$ in the $(u;\xi)$ plane, when the window width σ decreases to 0. The frequency ξ is chosen large enough so that the period of $e^{i\xi x}$ is smaller than the support size σ of ψ_σ. We set $\xi = \xi_0/\sigma$ and select ξ_0 so that $\hat{\psi}(\omega)$ and its first $\lceil h\rceil + 2$ derivatives vanish at $\omega = \xi_0$, where $\lceil h\rceil$ denotes the smallest integer greater than or equal to h.

THEOREM 3.2 (Frequency transport). *Let Y be a stationary process such that there exists $h > 0$ with*

$$c_Y(0) - c_Y(x) = |x|^h\eta(x), \tag{26}$$

where η is continuous in a neighborhood of 0 *and* $\eta(0) > 0$. *Let ψ be an even, positive $\mathbf{C}^1$ function supported in* $[-1, 1]$ *and let ξ_0 be such that $\hat{\psi}(\omega)$ and its first*

$\lceil h \rceil + 2$ *derivatives vanish at* $\omega = \xi_0$ *but*

$$\iint |x-y|^h (x-y) \sin[\xi_0(x-y)]\psi(x)\psi(y)\,dx\,dy \neq 0.$$

If

$$X(x) = e^{i\theta(x)} Y(x), \qquad \textit{where } \theta(x) \textit{ is } \mathbf{C}^{\lceil h \rceil + 4},$$

then, for each $u \in \mathbb{R}$ *such that* $\theta''(u) \neq 0$ *and for* $\xi = \xi_0/\sigma$, *when* $\sigma \to 0$,

$$\partial_u A_X^\sigma(u,\xi) - \theta''(u)\,\partial_\xi A_X^\sigma(u,\xi) = O(\sigma^2)\,\partial_u A_X^\sigma(u,\xi). \tag{27}$$

The proof is given in Appendix B. To satisfy the theorem hypotheses, one may choose $\psi(x)$ to be a box spline obtained by convolving the indicator function $\mathbb{1}_{[-1/2m,1/2m]}$ with itself m times: let $m \geq \lceil h \rceil + 3$ and

$$\hat{\psi}(\omega) = \left(\frac{\sin(\omega/(2m))}{\omega/(2m)}\right)^m \exp\left(\frac{-i\varepsilon\omega}{4m}\right), \tag{28}$$

where $\varepsilon = 1$ if m is odd and $\varepsilon = 0$ if m is even. With $\xi_0 = 2m\pi$, ψ satisfies the theorem hypotheses. The deformation gradient $\theta''(u)$ can be characterized from (27) by letting σ go to 0, and we proved in (2) that $\theta''(u)$ specifies the equivalence class of D in $\mathcal{D}/\mathcal{G}$. Section 4.2 will impose additional conditions on c_Y and θ to obtain a consistent estimation of $\theta''(u)$ from one realization of the frequency-modulated process X.

3.4. *Multidimensional scale transport.* For a multidimensional warping, where $Df(x) = f(\theta(x))$ with $x \in \mathbb{R}^d$, the adjoint of D^{-1} is $\overline{D^{-1}} f(x) = \det J_\theta(x)\, f(\theta(x))$. The matrix $J_\theta(x)$ is the Jacobian matrix of θ at position x, as defined in (6). The stationarity-invariant group $\mathcal{G}$ is the affine group, composed of operators G_β with $\beta = (u, S) \in \mathbb{R}^d \times GL^+(\mathbb{R}^d)$, such that

$$G_\beta f(x) = f(u + Sx).$$

The adjoint of G_β is

$$\overline{G}_\beta f(x) = \det S^{-1} f\bigl(S^{-1}(x-u)\bigr) = T_u \overline{F}_S f(x),$$

where $\overline{F}_S f(x) = \det S^{-1} f(S^{-1}x)$ and

$$S = (s_{l,m})_{1 \leq l,m \leq d}.$$

Similarly to (21), we find a tangential approximation to $\overline{D^{-1}}$ by noting that, for a regular function f supported in a neighborhood of $v = \theta(u)$,

$$\overline{D^{-1}} f(x) \approx \det J_\theta(u) f\bigl(\theta(u) + J_\theta(u)\,(x-u)\bigr) = \overline{G}_{\beta(v)} f(x). \tag{29}$$

The operators $\overline{D^{-1}}$ and $\overline{G}_{\beta(v)}$ both translate the support of f from a neighborhood of v to a neighborhood of $u(v) = \theta^{-1}(v)$.

Let ψ be a function such that $\int_{\mathbb{R}^d} \psi(x)\,dx = 0$. A multidimensional extension of the scalogram is given by

$$\begin{aligned} A_X^\sigma(u,S) &= \mathbb{E}\{|\langle X(x), \det S^{-1}\psi_\sigma(S^{-1}(x-u))\rangle|^2\} \\ &= \mathbb{E}\{|\langle X(x), \det S^{-1}\,\psi(\sigma^{-1}S^{-1}(x-u))\rangle|^2\}. \end{aligned}$$

As in the one-dimensional case, we divide $A_X^\sigma(u,s)$ by σ^{2d} and replace the product σS by a matrix, which we still denote S. The resulting atomic decomposition

$$A_X(u,S) = \mathbb{E}\{|\langle X(x), \det S^{-1}\psi(S^{-1}(x-u))\rangle|^2\} \tag{30}$$

is similar to the scalogram (20) but since the scale parameter s is replaced by a warping matrix S, we call it a *warpogram*.

For a one-dimensional warping, the velocity term of transport equation (24) is $(\log\theta')'(u) = \theta''(u)/\theta'(u)$. In two dimensions, it becomes a set of matrices, indexed by the direction k of spatial differentiation:

$$\text{for } 1 \le k \le d, \qquad J_\theta^{-1}(u)\frac{\partial J_\theta(u)}{\partial u_k} = (\gamma_{l,m}^k(u))_{1\le l,m\le d}. \tag{31}$$

This set of matrices has been shown in (9) to specify the equivalence class of D in $\mathcal{D}/\mathcal{G}$. For each (l,m), we introduce the vector

$$\vec{\gamma}_{l,m}(u) = (\gamma_{l,m}^k(u))_{1\le k\le d}.$$

The partial derivative $\partial_{\log s}A_X(u,s) = s\,\partial_s A_X(u,s)$, which appears in the one-dimensional transport equation (24), now becomes a matrix product between a matrix composed of partial derivatives with respect to the scale parameters and the transpose S^t of S:

$$\left(\frac{\partial A_X(u,S)}{\partial s_{i,j}}\right)_{1\le i,j\le d} S^t = (a_{i,j}(u,S))_{1\le i,j\le d}. \tag{32}$$

The following theorem isolates $\sigma = (\det S)^{1/d}$ by writing $S = \sigma\tilde{S}$ with $\det\tilde{S} = 1$ and gives a d-dimensional transport equation when σ goes to 0.

THEOREM 3.3. *Suppose that $X(x) = Y(\theta(x))$, where Y is stationary, $\theta(x)$ is $\mathbf{C}^3$ and $\det J_\theta(x) > 0$. Suppose that the autocovariance kernel c_Y of Y satisfies*

$$c_Y(0) - c_Y(x) = |x|^h\,\eta(x), \tag{33}$$

with $\eta(0) > 0$ and $\eta \in \mathbf{C}^2$ in a neighborhood of 0. For each $u \in \mathbb{R}^d$ and for each $\tilde{S}$ with $\det\tilde{S} = 1$, if there exists $C(u,\tilde{S}) > 0$ such that, for $S = \sigma\tilde{S}$ and σ small enough,

$$\begin{aligned} &\left|\mathrm{Re}\iint \vec{\nabla}c_Y(S(x-y))\vec{\nabla}J_\theta(u)J_\theta^{-1}(u)S(x-y)\psi^*(x)\psi(y)\,dx\,dy\right| \\ &\qquad \ge C(u,\tilde{S})\sigma^h, \end{aligned} \tag{34}$$

then, *when* σ *goes to* 0,

$$\left|\vec{\nabla}_u A_X(u,S) - \sum_{l,m=1}^{d} \vec{\gamma}_{l,m}(u)\, a_{l,m}(u,S)\right| = O(\sigma)|\vec{\nabla}_u A_X(u,S)|. \tag{35}$$

The proof of this theorem is given in Appendix B.

The purpose of condition (34) is to ensure that $|\vec{\nabla}_u A_X(u,S)|$ on the right-hand side of (35) is not too small. If $\theta(x)$ is a separable warping function of the form $\theta(x_1,\dots,x_d) = (\theta_1(x_1),\dots,\theta_d(x_d))$, then one can verify (see Appendix B) that (34) holds for all functions θ_i such that θ_i'' does not vanish if

$$\mathrm{Re} \iint |\tilde{S}(x-y)|^h \psi^*(x)\psi(y)\, dx\, dy \neq 0. \tag{36}$$

The above condition is similar to the second part of condition (23) in Theorem 3.1. Condition (34) is more involved, however, because, in general, coupling occurs between different directions.

For σ sufficiently small, neglecting the right-hand side of (35) yields d scalar equations:

$$\text{for } 1 \le k \le d, \qquad \partial_{u_k} A_X(u,S) - \sum_{l,m=1}^{d} \gamma_{l,m}^k(u) a_{l,m}(u,S) = 0.$$

For any (u,S), the values $\partial_{u_k} A_X(u,S)$ and $a_{l,m}(u,S)$ defined in (32) depend on the autocovariance of X and have to be estimated. For each direction k, there are a total of d^2 unknown coefficients $\gamma_{l,m}^k(u)$, equal to the d^2 matrix components of $J_\theta^{-1}(u)\, \partial_{u_k} J_\theta(u)$. To compute them, we need to select d^2 warping matrices $\{S_i\}_{i=1,\dots,d^2}$ and invert the linear system:

$$\begin{pmatrix} a_{1,1}(u,S_1) & a_{1,2}(u,S_1) & \cdots & a_{d,d}(u,S_1) \\ \vdots & \vdots & \vdots & \vdots \\ a_{1,1}(u,S_{d^2}) & a_{1,2}(u,S_{d^2}) & \cdots & a_{d,d}(u,S_{d^2}) \end{pmatrix} \begin{pmatrix} \gamma_{1,1}^k(u) \\ \gamma_{1,2}^k(u) \\ \vdots \\ \gamma_{d,d}^k(u) \end{pmatrix} = \begin{pmatrix} \partial_{u_k} A_X(u,S_1) \\ \vdots \\ \partial_{u_k} A_X(u,S_{d^2}) \end{pmatrix}. \tag{37}$$

Changing the direction index k only modifies the right-hand side of (37). Note that, in order for the system to be invertible, the matrix on the left-hand side of (37) must have full rank. The matrices S_k must therefore be appropriately chosen, and the inverse warping problem must have a unique solution. This is not always the case, as shown by the example in (10).

The system (37) has been used in Computer Vision for relief reconstruction from photographs of surfaces with regular patterns, a problem called shape-from-texture [6]. Examples of such "textured surfaces" are provided in Figure 3.

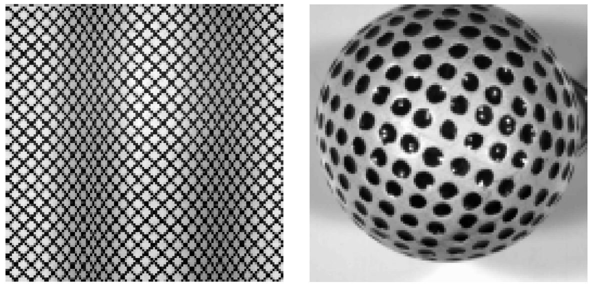


FIG. 3. *Shape from texture*: *examples of objects whose shape is inferred by analyzing texture* (*i.e.*, *pattern*) *variations*.

Perceptually, by analyzing the variations of size, shape and density of the patterns within the two-dimensional images, we can infer the three-dimensional shape of the objects. This cue to three-dimensional shape is monocular (i.e., it uses a single image), unlike stereo vision, in which two images of a scene taken from different viewing points are compared. Mathematically, one can model the pattern variations across the image by the two deformation gradient vectors

$$\begin{pmatrix} \gamma_{1,1}^1(u) \\ \gamma_{1,2}^1(u) \\ \gamma_{2,1}^1(u) \\ \gamma_{2,2}^1(u) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \gamma_{1,1}^2(u) \\ \gamma_{1,2}^2(u) \\ \gamma_{2,1}^2(u) \\ \gamma_{2,2}^2(u) \end{pmatrix}.$$

These two vectors represent the elements of the two matrices defined in (12), and as mentioned in Section 2.1, one can recover from them the normal vector field of the surface being viewed, and hence its shape.

4. Estimation of deformations. The deformation gradient appears as a velocity vector in the transport (19). To recover it from a single realization of X, the derivatives $\vec{\nabla}_u A_X^\sigma(u,\alpha)$ and $\vec{\nabla}_\alpha A_X^\sigma(u,\alpha)$ of the atomic decomposition of X have to be estimated. With a single realization, a sample mean estimator has a variance of the same order of magnitude as the term it estimates. This variance can be reduced with a spatial smoothing, while the bias, which is proportional to the width of the smoothing kernel, is controlled. The next two sections study the consistency of such smoothed estimators for the one-dimensional warping problem and the frequency modulation problem. Finally, Section 4.3 discusses multidimensional warping estimation.

4.1. *Warping in one dimension.* The scalogram of X is an expected value

$$A_X(u,s) = \mathbb{E}\{|\langle X, \psi_{u,s}\rangle|^2\},$$

with $\psi_{u,s}(x) = s^{-1}\psi((x-u)/s)$. If $X(x) = Y(\theta(x))$, then Theorem 3.1 proves that

$$\partial_u A_X(u,s) - (\log\theta')'(u)\, \partial_{\log s} A_X(u,s) = O(s)\, \partial_u A_X(u,s). \tag{38}$$

From $X(x)$ approximated at a resolution N, one can compute the *empirical scalogram* $|\langle X, \psi_{u,s}\rangle|^2$ at scales $s \geq N^{-1}$ and locations $u = k/N$ with $k \in \mathbb{Z}$ [11]. We introduce a kernel estimator $A_X(u,s)$, using the averaging kernel

$$g(x) = \begin{cases} \Delta^{-1}(1 - |x/\Delta|), & \text{if } |x| \leq \Delta, \\ 0, & \text{if } |x| > \Delta. \end{cases} \tag{39}$$

Let

$$\begin{aligned} &\widehat{\partial_u A_X}(u,s) \\ &\quad = 2N^{-1} \sum_{|k/N-u| \leq \Delta} g(u - k/N)\, \mathrm{Re}\big[\langle X, \psi_{k/N,s}\rangle \langle X, \partial_u \psi_{k/N,s}\rangle^*\big], \end{aligned} \tag{40}$$

$$\begin{aligned} &\widehat{\partial_{\log s} A_X}(u,s) \\ &\quad = 2N^{-1} \sum_{|k/N-u| \leq \Delta} g(u - k/N)\, \mathrm{Re}\big[\langle X, \psi_{k/N,s}\rangle \langle X, \partial_{\log s} \psi_{k/N,s}\rangle^*\big]. \end{aligned} \tag{41}$$

In view of (38), we suggest the following estimator for $(\log\theta')'(u)$:

$$(\widehat{\log\theta'})'(u) = \frac{\widehat{\partial_u A_X}(u, N^{-1})}{\widehat{\partial_{\log s} A_X}(u, N^{-1})}.$$

One must guarantee that, when $s = N^{-1}$ and N increases, $\widehat{\partial_u A_X}(u,s)$ and $\widehat{\partial_{\log s} A_X}(u,s)$ are close to $\partial_u A_X(u,s)$ and $\partial_{\log s} A_X(u,s)$, respectively. For this, we introduce a Gaussian assumption on the underlying stationary process Y. This assumption allows us to derive the consistency of the estimators from the fast spatial decorrelation (in k) of $\langle X, \psi_{k/N,s}\rangle$ and $\langle X, \partial_{\log s} \psi_{k/N,s}\rangle$. Fast decorrelation of these two random variables will be ensured by adjusting the choice of wavelet ψ to the behavior of the autocovariance function c_Y in a neighborhood of 0.

The following theorem proves the weak consistency of the proposed estimator $(\widehat{\log\theta'})'(u)$ of $(\log\theta')'(u)$ by selecting the averaging interval Δ according to the scale $s = N^{-1}$. A wavelet $\psi(x)$ is said to have p vanishing moments if

$$\int x^k \psi(x)\, dx = 0 \qquad \text{for } 0 \leq k < p.$$

THEOREM 4.1 (Consistency, warping). *Let $X(x) = Y(\theta(x))$, where Y is a stationary Gaussian process whose covariance satisfies*

$$c_Y(0) - c_Y(x) = |x|^h \eta(x) \qquad \text{with } h > 0 \text{ and } \eta(0) > 0. \tag{42}$$

Let ψ be a $\mathbf{C}^2$ wavelet supported in $[-1, 1]$ *with p vanishing moments such that*

$$2p - h > 1/2 \quad \textit{and} \quad \iint |x-y|^h \psi^*(x)\,\psi(y)\,dx\,dy \neq 0.$$

Let $\Delta = N^{-1/5}$. If $\eta(x)$ is $\mathbf{C}^{2p}$ in a neighborhood of 0 *and if $\theta(x) \in \mathbf{C}^3 \cap \mathbf{C}^{2p}$, then, for each $u \in \mathbb{R}$ such that $\theta''(u) \neq 0$,*

$$\text{Prob}\Big\{\big|(\widehat{\log\theta'})'(u) - (\log\theta')'(u)\big| \le 2\,(\log N)\,N^{-1/5}\Big\} \underset{N\to\infty}{\longrightarrow} 1. \tag{43}$$

The proof is given in Appendix C. Since all estimations are based on wavelet coefficients, one can easily verify that the results still hold if Y is not stationary but has stationary increments. In particular, it applies to fractional Brownian motion (see [1] and [7]), for which $\eta(x) = 1$.

Figure 4 displays a numerical experiment conducted on a single realization of a warped process. The signal X in Figure 4(b) is obtained by warping a stationary signal Y, depicted in Figure 4(a). Figure 4(c) shows (dotted line) the estimate $\widehat{\log\theta'}$

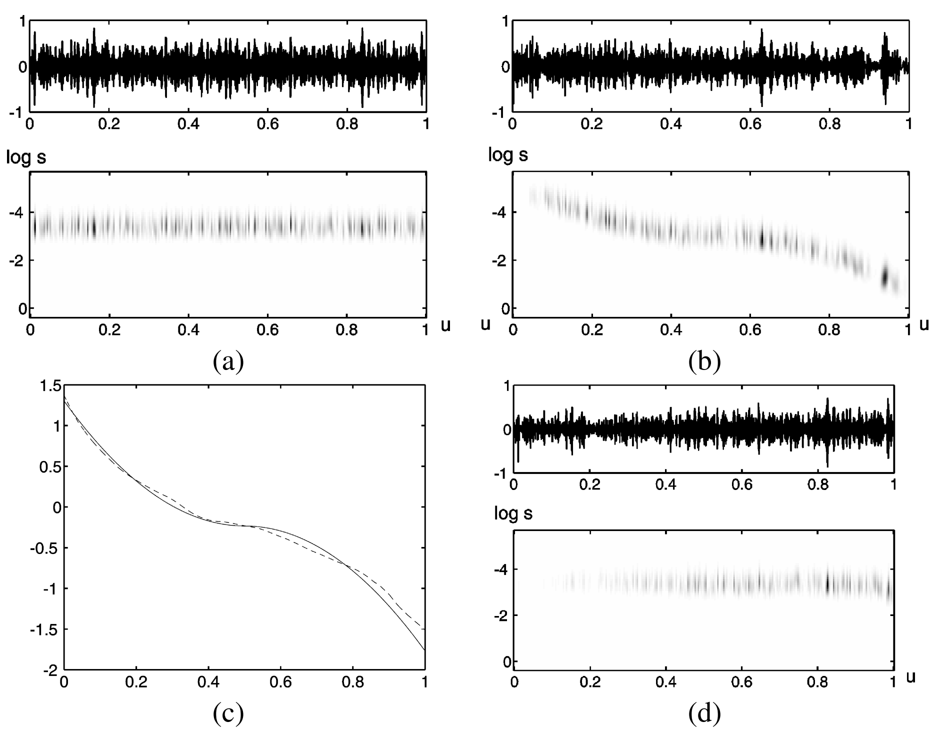


FIG. 4. (a) *Stationary signal Y and its empirical scalogram* $|\langle Y, \psi_{u,s}\rangle|^2$. (b) *Warped signal* $X(x) = Y(\theta(x))$ *and its empirical scalogram.* (c) $\log\theta'(x) = \lambda_1 + \lambda_2\,\text{sign}(1/2 - x)\,|x - 1/2|^2$, *where λ_1 and λ_2 are two constants (solid line), and its estimation from X (dashed line).* (d) *Stationarized signal and its empirical scalogram.*

of $\log\theta'$ obtained by integrating the estimate $(\widehat{\log\theta'})'(u)$ and choosing the additive integration constant so that $\int_0^1 \exp(\widehat{\log\theta'}) = \int_0^1 \theta'$. An estimate $\widehat{\theta}$ for the warping function can be obtained up to an additive constant by integrating $\exp\widehat{\log\theta'}$. It is then possible to stationarize the deformed signal X by computing $X \circ (\widehat{\theta})^{-1}$. Figure 4(d) displays such a stationarized signal. We refer the reader to [6] for details on the numerical implementation of this method.

4.2. *Frequency modulation.* For a frequency-modulated process, $X(x) = Y(x)e^{i\theta(x)}$, Theorem 3.2 shows that the deformation gradient $\theta''(u)$ can be computed from the spectrogram

$$A_X^\sigma(u,\xi) = \mathbb{E}\{|\langle X(x), \psi_\sigma(x-u)e^{i\xi(x-u)}\rangle|^2\},$$

with the equation

$$\partial_u A_X^\sigma(u,\xi) - \theta''(u)\,\partial_\xi A_X^\sigma(u,\xi) = O(\sigma^2)\,\partial_u A_X^\sigma(u,\xi) \tag{44}$$

evaluated at a frequency $\xi = \xi_0/\sigma$.

To compute an estimator of the smoothed partial derivatives of the spectrogram, we relate the spectrogram coefficients to a particular wavelet transform. Observe that

$$\psi_\sigma(x-u)\exp\left(i\xi_0\,\frac{x-u}{\sigma}\right) = \psi^1\left(\frac{x-u}{\sigma}\right), \tag{45}$$

where

$$\psi^1(x) = \psi(x)e^{i\xi_0 x}. \tag{46}$$

Since ψ is real, $\hat{\psi}(\omega)$ is even, and if $\hat{\psi}(\omega)$ has a 0 of order $\lceil h\rceil + 3$ at $\omega = \xi_0$, then

$$\int x^k\,\psi^1(x)\,dx = (-i)^k \frac{d^k\hat{\psi}^1}{d\omega^k}(-\xi_0) = 0 \qquad \text{for } k \le \lceil h\rceil + 2.$$

This means that ψ^1 is a wavelet with $\lceil h\rceil + 3$ vanishing moments [11]. We write $\psi^1_{u,\sigma}(x) = \sigma^{-1}\,\psi^1(\sigma^{-1}(x-u))$. The scalogram associated to this wavelet is defined by $A_X(u,\sigma) = \mathbb{E}\{|\langle X, \psi^1_{u,\sigma}\rangle|^2\}$. From (45),

$$A_X^\sigma(u,\xi_0/\sigma) = \mathbb{E}\{|\langle X(x), \psi^1(\sigma^{-1}(x-u))\rangle|^2\} = \sigma^2\,A_X(u,\sigma),$$

and hence

$$\partial_u A_X^\sigma(u,\xi_0/\sigma) = \sigma^2\,\partial_u A_X(u,\sigma).$$

Let $\widehat{\partial_u A_X}(u,\sigma)$ be the estimator calculated in (40) at a scale $\sigma = N^{-1}$ for a certain averaging width Δ from a single realization of $X(x)$ sampled at a resolution N.

We introduce the estimator

$$\widehat{\partial_u A_X^\sigma}(u, \xi_0/\sigma) = \sigma^2\, \widehat{\partial_u A_X}(u, \sigma).$$

To compute an empirical estimator of the other partial derivative, $\partial_\xi A_X^\sigma(u, \xi_0/\sigma)$, observe that

$$\partial_\xi A_X^\sigma(u, \xi) = 2\,\mathrm{Re}\big[\mathbb{E}\big\{\big\langle X(x), \psi_\sigma(x-u)e^{i\xi(x-u)}\big\rangle\big\langle X(x), \partial_\xi\big[\psi_\sigma(x-u)e^{i\xi(x-u)}\big]\big\rangle^*\big\}\big].$$

Introducing a new wavelet

$$\psi^2(x) = x\psi^1(x) = x\psi(x)e^{i\xi_0 x}, \tag{47}$$

and $\psi^2_{u,\sigma}(x) = \sigma^{-1}\psi^2(\sigma^{-1}(x-u))$, this partial derivative can be rewritten, for $\xi = \xi_0/\sigma$, as

$$\partial_\xi A_X^\sigma(u, \xi_0/\sigma) = 2\sigma^3\,\mathrm{Im}\big[\mathbb{E}\big\{\langle X, \psi^1_{u,\sigma}\rangle\langle X, \psi^2_{u,\sigma}\rangle^*\big\}\big].$$

Similarly to (41), for $\sigma = N^{-1}$, we define the averaging kernel estimator

$$\begin{aligned}\widehat{\partial_\xi A_X^\sigma}&(u, \xi_0/\sigma)\\ &= 2\sigma^3 N^{-1} \sum_{|k/N-u|\le\Delta} g(u-k/N)\,\mathrm{Im}\big[\langle X, \psi^1_{k/N,\sigma}\rangle\langle X, \psi^2_{k/N,\sigma}\rangle^*\big].\end{aligned} \tag{48}$$

The following theorem proves that, for $\sigma = N^{-1}$ and an appropriate choice of the averaging interval Δ,

$$\widehat{\theta''}(u) = \frac{\widehat{\partial_u A_X^\sigma}(u, N\xi_0)}{\widehat{\partial_\xi A_X^\sigma}(u, N\xi_0)}$$

is a weakly consistent estimator of $\theta''(u)$ when $N \to \infty$. The proof resides in the spatial decorrelation of wavelet coefficients of X, and for the same reasons as in Theorem 4.1, we introduce a Gaussian assumption on the underlying stationary process Y.

THEOREM 4.2 (Consistency, frequency modulation). *Let $X(x) = Y(x)e^{i\theta(x)}$, where Y is a stationary Gaussian process such that there exists $h > 0$ with*

$$c_Y(0) - c_Y(x) = |x|^h\eta(x) \quad \textit{and} \quad \eta(0) > 0. \tag{49}$$

Suppose that $\psi^1(x) = \psi(x)e^{i\xi_0 x}$ is a compactly supported wavelet with $p \ge \lceil h\rceil + 3$ vanishing moments such that

$$\iint |x-y|^h(x-y)\sin[\xi_0(x-y)]\psi(x)\psi(y)\,dx\,dy \neq 0.$$

Let $\Delta = N^{-1/5}$. If $\eta \in \mathbf{C}^{2p}$ in a neighborhood of 0 *and if $\theta \in \mathbf{C}^{2p}$, then, for each $u \in \mathbb{R}$,*

$$\mathrm{Prob}\big\{\big|\widehat{\theta''}(u) - \theta''(u)\big| \le 2(\log N)N^{-1/5}\big\} \underset{N\to\infty}{\longrightarrow} 1. \tag{50}$$

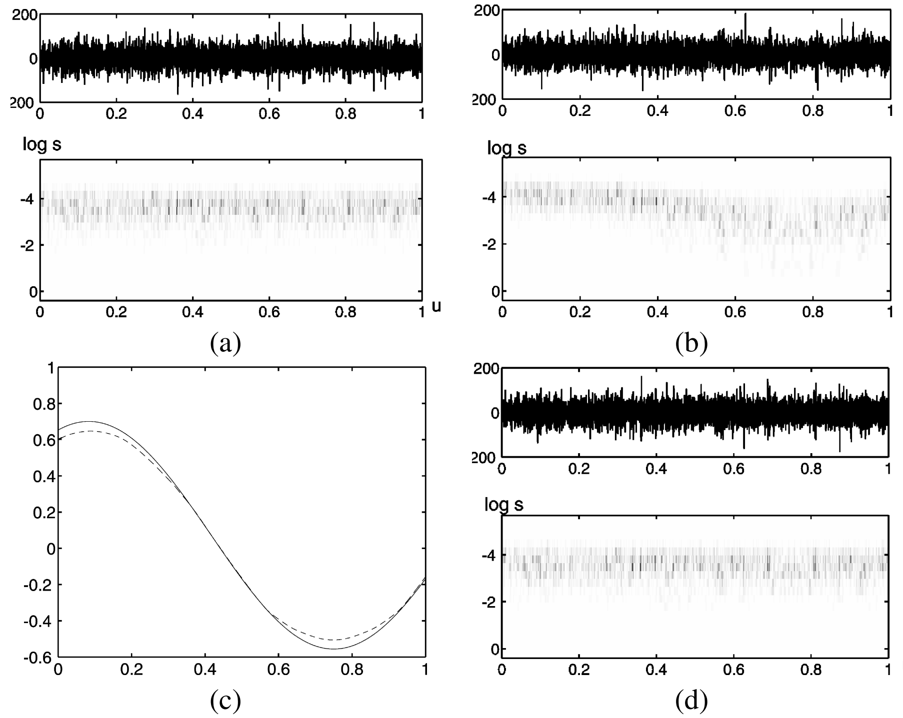


FIG. 5. (a) *Stationary signal Y and its empirical scalogram* $|\langle Y, \psi^1_{u,s}\rangle|^2$. (b) *Frequency-modulated signal* $X(x) = Y(x)\exp(i\theta(x))$ *and its empirical scalogram.* (c) *Frequency modulation* $\theta'(x)$ (*solid line*) *and its estimation from* X (*dashed line*). (d) *Stationarized signal and its empirical scalogram.*

The proof is given in Appendix C. The numerical example in Figure 5 shows the estimation of a frequency modulation, with a box spline window $\psi(x)$ defined in (28). We explained that the empirical estimator $\widehat{\theta''}(u)$ is, in fact, computed from wavelet coefficients associated to the two wavelets ψ^1 and ψ^2 defined in (46) and (47). Figure 5(a) shows a realization of a stationary signal $Y(x)$ and the corresponding empirical scalogram $\widehat{A_Y}(u,s) = |\langle Y, \psi^1_{u,s}\rangle|^2$. The frequency-modulated signal $X(x) = Y(x)\exp(i\theta(x))$ and its empirical scalogram are shown in Figure 5(b). The derivative θ' of the frequency modulation is plotted in Figure 5(c) (solid line). An estimate $\widehat{\theta'}$ of θ' is obtained by integrating $\widehat{\theta''}$ and choosing the additive integration constant so that $\int_0^1 \widehat{\theta'} = \int_0^1 \theta'$. Figure 5(c) plots $\widehat{\theta'}$ (dashed line) superposed on the theoretical function θ' (solid line). Finally, Figure 5(d) represents the stationarized process $X(x)\exp(-i\widehat{\theta}(x))$ and its empirical scalogram.

4.3. *Warping in higher dimensions.* For a multidimensional warping, at each position u, the deformation gradient corresponds to a set of d matrices $\vec{\gamma}_{l,m}(u) = (\gamma^k_{l,m}(u))_{1\le k\le d}$ defined in (31). Theorem 3.3 shows that these coefficients appear in the velocity term of the transport equation (35) satisfied by the warpogram of X:

$$A_X(u, S) = \mathbb{E}\{|\langle X, \psi_{u,S}\rangle|^2\},$$

with

$$\psi_{u,S}(x) = (\det S^{-1})\psi\big(S^{-1}(x-u)\big).$$

At a sufficiently small scale σ, the error on the right-hand side of the transport equation (35) can be neglected. The vector transport equation can then be written as a linear system

$$\begin{aligned}
&\begin{pmatrix} a_{1,1}(u, S_1) & a_{1,2}(u, S_1) & \cdots & a_{d,d}(u, S_1) \\ \vdots & \vdots & \vdots & \vdots \\ a_{1,1}(u, S_{d^2}) & a_{1,2}(u, S_{d^2}) & \cdots & a_{d,d}(u, S_{d^2}) \end{pmatrix} \begin{pmatrix} \gamma^k_{1,1}(u) \\ \vdots \\ \gamma^k_{d,d}(u) \end{pmatrix} \\
&\quad = \begin{pmatrix} \partial_{u_k} A_X(u, S_1) \\ \vdots \\ \partial_{u_k} A_X(u, S_{d^2}) \end{pmatrix},
\end{aligned} \tag{51}$$

where

$$\big(a_{l,m}(u,S)\big)_{1\le l,m\le d} = \big(\partial_{s_{i,j}} A_X(u,S)\big)_{1\le i,j\le d} S^t.$$

If the process X is measured at a resolution N, we compute the warpogram with functions $\psi_{u,S}$ whose support in any direction is larger than N^{-1}. We therefore require that $S = \sigma\tilde{S}$, where $\sigma \ge KN^{-1}$ and all the eigenvalues of $\tilde{S}$ are greater than K^{-1}. The position parameter is also restricted to a uniform grid $u = N^{-1}k$, with $k \in \mathbb{Z}^d$. For $a = u$ or $s_{i,j}$, we have $\partial_a A_X(u,S) = 2\operatorname{Re}\mathbb{E}\{\langle X, \psi_{u,S}\rangle\langle X, \partial_a\psi_{u,S}\rangle^*\}$, and we introduce the kernel estimator

$$\begin{aligned}
&\widehat{\partial_a A_X}(u,S) \\
&\quad = 2N^{-d} \sum_{|N^{-1}k-u|\le\Delta} g_2(u - N^{-1}k)\operatorname{Re}\{\langle X, \psi_{N^{-1}k,S}\rangle\langle X, \partial_a\psi_{N^{-1}k,S}\rangle^*\},
\end{aligned}$$

where $g_2(u_1, u_2) = g(u_1)\,g(u_2)$ is the separable product of two window functions defined in (39).

We also define

$$\big(\widehat{a_{l,m}}(u,S)\big)_{1\le l,m\le d} = \big(\widehat{\partial_{s_{i,j}} A_X}(u,S)\big)_{1\le i,j\le d} S^t.$$

The kth component of the deformation gradient, $(\gamma^k_{l,m}(u))$, can then be estimated by

$$\begin{pmatrix} \widehat{\gamma^k_{1,1}}(u) \\ \vdots \\ \widehat{\gamma^k_{d,d}}(u) \end{pmatrix} = \begin{pmatrix} \widehat{a_{1,1}}(u, S_1) & \cdots & \widehat{a_{d,d}}(u, S_1) \\ \vdots & \vdots & \vdots \\ \widehat{a_{1,1}}(u, S_{d^2}) & \cdots & \widehat{a_{d,d}}(u, S_{d^2}) \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\partial_{u_k} A_X}(u, S_1) \\ \vdots \\ \widehat{\partial_{u_k} A_X}(u, S_{d^2}) \end{pmatrix}. \tag{52}$$

Extending the consistency theorem (Theorem 4.1) to more than one dimension is possible, but requires technical hypotheses that are not yet well understood.

In Section 2.1, we mentioned that the warping of textures in images specifies the three-dimensional shapes of the textured objects appearing in the scene. The estimator defined in (52) is used in [5] and [6] to compute shape from texture and provides a good estimation of three-dimensional surfaces from two-dimensional images.

APPENDIX A

Proofs of Section 2.

PROOF OF PROPOSITION 2.1. Let Y be a stationary process and suppose that there exists an $\varepsilon > 0$ such that $c_Y(x) > 0$ for $|x| \leq \varepsilon$. Let $\tilde{Y}$ be another stationary process. We want to show that if the autocovariance kernels $Y(x)\exp[i\theta(x)]$ and of $\tilde{Y}(x)\exp[i\tilde{\theta}(x)]$ are equal, that is, if

$$c_Y(x-y)\exp\bigl(i[\theta(x)-\theta(y)]\bigr) = c_{\tilde{Y}}(x-y)\exp\bigl(i[\tilde{\theta}(x)-\tilde{\theta}(y)]\bigr), \tag{53}$$

then $\theta''(x) = \tilde{\theta}''(x)$. The functions θ and $\tilde{\theta}$ are assumed $\mathbf{C}^4$; therefore, $\mu = \theta - \tilde{\theta}$ is also $\mathbf{C}^4$. Let us fix $x \in \mathbb{R}$; our goal is to prove that $\mu''(x) = 0$. We choose $y \in \mathbb{R}$ such that $|x-y| < \varepsilon$. After dividing both sides of (53) by $c_Y(x-y) > 0$, it appears that $e^{i[\mu(x)-\mu(y)]}$ is a function of $x-y$. Therefore, $\mu(x)-\mu(y)$ is also a function of $x-y$, and, in particular, for all z,

$$\mu(x)-\mu(y) = \mu(x+z)-\mu(y+z).$$

Differentiating this expression with respect to x shows that $\mu(x) = \mu'(x+z)$ and thus $\mu''(x) = 0$. $\square$

PROOF OF PROPOSITION 2.2. Let Y be a stationary process and let $\varepsilon > 0$ such that $c_Y(x)$ is $\mathbf{C}^1$ for $0 < |x| \leq \varepsilon$, with $c'_Y(x) < 0$. Let $\tilde{Y}$ denote another stationary process and let us suppose that the autocovariance kernels of $Y(\theta(x))$ and $\tilde{Y}(\tilde{\theta}(x))$ are equal. The functions θ and $\tilde{\theta}$ are assumed $\mathbf{C}^3$; therefore, $\mu = \theta \circ \tilde{\theta}^{-1}$ is also $\mathbf{C}^3$. Proving the proposition amounts to proving that μ is linear or, equivalently, that μ'' vanishes everywhere. By the definition of μ,

$$c_{\tilde{Y}}(x-y) = c_Y\bigl(\mu(x)-\mu(y)\bigr). \tag{54}$$

Let us fix $x \in \mathbb{R}$ and choose $y \neq x$, but sufficiently close to x so that $|\mu(x) - \mu(y)| < \varepsilon$. Differentiating (54) with respect to x and y shows that

$$c'_Y\bigl(\mu(x)-\mu(y)\bigr)\mu'(y) = c'_Y\bigl(\mu(x)-\mu(y)\bigr)\mu'(x).$$

Since $c'_Y(\mu(x)-\mu(y)) < 0$, we obtain $\mu'(x) = \mu'(y)$ and therefore $\mu''(x) = 0$. $\square$

PROOF OF PROPOSITION 2.3. Let Y be a stationary process such that c_Y satisfies (11). Let $\tilde{Y}$ denote another stationary process and suppose that the autocovariance kernels of $Y(\theta(x))$ and $\tilde{Y}(\tilde{\theta}(x))$ are equal. Let $\mu = \theta \circ \tilde{\theta}^{-1}$. By the definition of μ,

$$c_Y\big(\mu(x) - \mu(y)\big) = c_{\tilde{Y}}(x - y).$$

Differentiating this expression with respect to x and y, for $x \neq y$, shows that

$$\vec{\nabla} c_Y\big(\mu(x) - \mu(y)\big) J_\mu(y) = \vec{\nabla} c_Y\big(\mu(x) - \mu(y)\big) J_\mu(x). \tag{55}$$

Let us fix $x \in \mathbb{R}^d$ and prove that $\vec{\nabla} J_\mu(x) = 0$. Let $\varepsilon > 0$ such that $\eta(z)$ is $\mathbf{C}^2$ for $|z| < \varepsilon$. Let us choose $y \in \mathbb{R}^d$ such that $0 < |\mu(x) - \mu(y)| < \varepsilon$ and let $z = \mu(x) - \mu(y)$:

$$\vec{\nabla} c_Y(z) = -|z|^{h-2}\big(h\eta(z)z + |z|^2 \vec{\nabla}\eta(z)\big).$$

Replacing this expression in (55) and dividing both sides by $-h|z|^{h-2}\eta(z)$ proves that

$$\big(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z)\big) J_\mu(y) = \big(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z)\big) J_\mu\big(\mu^{-1}\big(z + \mu(y)\big)\big),$$

so

$$\big(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z)\big) J_\mu(y) J_\mu^{-1}\big(\mu^{-1}\big(z + \mu(y)\big)\big) = z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z).$$

Introducing a function $\tilde{\mu}$ such that

$$\tilde{\mu}(z) = J_\mu(y)\mu^{-1}\big(z + \mu(y)\big), \tag{56}$$

this can be rewritten as

$$\big(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z)\big) J_{\tilde{\mu}}(z) = \big(z + h^{-1}|z|^2 \vec{\nabla} \log \eta(z)\big).$$

Noticing that $z\, J_{\tilde{\mu}}(\lambda z) = (d/d\lambda)\tilde{\mu}(\lambda z)$, we have, for $\lambda \geq 0$,

$$\frac{d}{d\lambda}\tilde{\mu}(\lambda z) = z + h^{-1}|z|^2 \lambda \vec{\nabla} \log \eta(\lambda z)\big(Id - J_{\tilde{\mu}}(\lambda z)\big),$$

which, when integrated between $\lambda = 0$ and $\lambda = 1$, gives

$$\tilde{\mu}(z) - \tilde{\mu}(0) = z + h^{-1}|z|^2 \int_0^1 \lambda \vec{\nabla} \log \eta(\lambda z)\big(Id - J_{\tilde{\mu}}(\lambda z)\big)\, d\lambda.$$

After replacing $\tilde{\mu}$ with (56) and noticing that $\tilde{\mu}(0) = J_\mu(y)\, y$, we obtain

$$\begin{aligned} &\mu^{-1}\big(z + \mu(y)\big) \\ &\quad = J_\mu^{-1}(y)z + y + J_\mu^{-1}(y)h^{-1}|z|^2 \int_0^1 \lambda \vec{\nabla} \log \eta(\lambda z)\big(Id - J_{\tilde{\mu}}(\lambda z)\big)\, d\lambda. \end{aligned} \tag{57}$$

Since c_Y is even, $\vec{\nabla}\eta(0) = 0$ and so $\vec{\nabla} \log \eta(0) = 0$. Let us denote $\vec{\nabla} \log \eta(z) = |z|\, \vec{a}(z)$. Recalling that η is twice continuously differentiable in a neighborhood of z for $0 \le |z| < \varepsilon$, the function $\vec{a}(z)$ is differentiable for $0 < |z| < \varepsilon$, and its gradient [the matrix $(\partial_j a_i)_{ij}$] is uniformly bounded for $0 < |z| < \varepsilon$. Differentiating (57) with respect to z shows that

$$J_\mu^{-1}\big(\mu^{-1}(z+\mu(y))\big) = J_\mu^{-1}(y)\big(Id + |z|^2 A(z)\big), \tag{58}$$

where $A(z)$ is a matrix defined by

$$A(z) = |z|^{-2}\frac{\partial}{\partial z}\left(h^{-1}|z|^2 \int_0^1 \lambda|\lambda z|\vec{a}(\lambda z)\big(Id - J_{\tilde{\mu}}(\lambda z)\big)\, d\lambda\right). \tag{59}$$

Let us calculate $A(z)$ explicitly: $|z|^2 A(z)$ is the sum of three terms:

$$\frac{\partial}{\partial z}\left(h^{-1}|z|^2 \int_0^1 \lambda|\lambda z|\vec{a}(\lambda z)\big(Id - J_{\tilde{\mu}}(\lambda z)\big)\, d\lambda\right) = \text{(I)} + \text{(II)} + \text{(III)}.$$

If a and b are two vectors, we denote by $a \otimes b$ the matrix whose elements are $(a_i b_j)_{ij}$. Then

$$\begin{aligned} \text{(I)} &= 2h^{-1}\left(\int_0^1 \lambda|\lambda z|\vec{a}(\lambda z)\big(Id - J_{\tilde{\mu}}(\lambda z)\big)\, d\lambda\right) \otimes z, \\ \text{(II)} &= h^{-1}|z|^2 \int_0^1 |\lambda|^2\left(\vec{a}(\lambda z) \otimes \frac{z}{|z|} + \lambda|z|\vec{\nabla}\vec{a}(\lambda z)\right)\big(Id - J_{\tilde{\mu}}(\lambda z)\big)\, d\lambda, \\ \text{(III)} &= -h^{-1}|z|^3 \int_0^1 |\lambda|^3 \vec{a}(\lambda z)\vec{\nabla}_z J_{\tilde{\mu}}(\lambda z)\, d\lambda. \end{aligned} \tag{60}$$

More precisely for (III), the i, j element of $\vec{a}(\lambda z)\vec{\nabla}_z J_{\tilde{\mu}}(\lambda z)$ is $\sum_k a_k(\lambda z) \times \frac{\partial}{\partial z_j}(J_{\tilde{\mu}})_{ki}(\lambda z)$. Because $\tilde{\mu}$ is in $\mathbf{C}^2$ and because $\vec{a}(\lambda z)$ as well as $\vec{\nabla}\vec{a}(\lambda z)$ is uniformly bounded for $0 < |\lambda z| < \varepsilon$, the matrix $A(z)$ resulting from the division of (I) + (II) + (III) by $|z|^2$ is uniformly bounded for $0 < |z| < \varepsilon$. Replacing z by $\mu(x) - \mu(y)$ in (58) gives

$$J_\mu^{-1}(x) = J_\mu^{-1}(y)\big(Id + |\mu(x) - \mu(y)|^2 A\big(\mu(x) - \mu(y)\big)\big).$$

Therefore, for any unit-length vector $x_k \in \mathbb{R}^d$,

$$\begin{aligned} \frac{\partial}{\partial x_k} J_\mu^{-1}(x) &= \lim_{\lambda \to 0} \frac{J_\mu^{-1}(x + \lambda x_k) - J_\mu^{-1}(x)}{\lambda} \\ &= \lim_{\lambda \to 0} \frac{J_\mu^{-1}(x + \lambda x_k)|\mu(x) - \mu(x + \lambda x_k)|^2}{\lambda} A\big(\mu(x) - \mu(x + \lambda x_k)\big) \\ &= 0. \end{aligned}$$

This proves that $\vec{\nabla} J_\mu^{-1}(x) = 0$ and therefore $\vec{\nabla} J_\mu(x) = 0$. As a consequence, for each direction x_k, $(\partial/\partial x_k) J_\mu(\tilde{\theta}(x)) = 0$. Since $J_\mu(x) = J_\theta(\tilde{\theta}^{-1}(x)) J_{\tilde{\theta}}^{-1}(\tilde{\theta}^{-1}(x))$, we obtain

$$\frac{\partial}{\partial x_k}\big(J_\theta(x) J_{\tilde{\theta}}^{-1}(x)\big) = 0,$$

and expanding the above differential expression then gives

$$\frac{\partial J_\theta(x)}{\partial x_k} J_{\tilde\theta}^{-1}(x) - J_\theta(x) J_{\tilde\theta}^{-1}(x) \frac{\partial J_{\tilde\theta}(x)}{\partial x_k} J_{\tilde\theta}^{-1}(x) = 0,$$

which is equivalent to (9).

Let us prove the equivalence between (8) and (9). If θ and $\tilde\theta$ satisfy (8), then $J_\theta(x) = S J_{\tilde\theta}(x)$, which implies that $J_\theta(x) J_{\tilde\theta}^{-1}(x)$ is independent of x. The preceding calculations then prove (9). Conversely, assuming that θ and $\tilde\theta$ satisfy (9), then, for each direction x_k, $\frac{\partial}{\partial x_k}(J_\theta(x) J_{\tilde\theta}^{-1}(x)) = 0$, which proves that $J_\theta(x) J_{\tilde\theta}^{-1}(x)$ is a constant matrix, belonging to $GL^+(\mathbb{R}^d)$ as the product of two elements of $GL^+(\mathbb{R}^d)$. The partial differential system $J_\theta(x) = S\, J_{\tilde\theta}(x)$ can be integrated to prove that $\theta(x) = u + S\tilde\theta(x)$. $\square$

PROOF OF THEOREM 2.1. Let us consider a specific family of zero-mean wide-sense stationary processes defined by

$$Y_\omega(x) = Y e^{i\omega\cdot x},$$

where Y is a zero-mean random variable with variance σ^2. Then

$$\mathbb{E}\{Y_\omega(x) Y_\omega^*(y)\} = \sigma^2 \exp\bigl(i\omega\cdot(x-y)\bigr) = c_{Y_\omega}(x-y).$$

Let G be a stationarity-invariant operator. If $X_\omega(x) = GY_\omega(x)$, then

$$\mathbb{E}\{X_\omega(x) X_\omega^*(y)\} = \sigma^2 f_\omega(x) f_\omega^*(y),$$

where $f_\omega(x) = Ge^{i\omega\cdot x}$. However, since G is stationarity invariant, X_ω is stationary; therefore, $\mathbb{E}\{X_\omega(x) X_\omega^*(y)\}$ is a function of $x-y$. This implies that, for any (x, y), the product $f_\omega(x) f_\omega(y)^*$ is a function of $x-y$. Therefore, there exist $\hat\rho(\omega) \in \mathbb{C}$ and $\lambda(\omega) \in \mathbb{R}^d$ such that

$$f_\omega(x) = Ge^{i\omega\cdot x} = \hat\rho(\omega) e^{i\lambda(\omega)\cdot x}. \tag{61}$$

Let us now prove that an operator that satisfies (61) is indeed stationarity invariant if and only if $\operatorname{ess\,sup}_{\omega\in\mathbb{R}^d} |\hat\rho(\omega)| < \infty$. Let Y be a zero-mean, stochastically continuous wide-sense stationary process. It therefore admits a spectral representation

$$Y(x) = \int_{\mathbb{R}^d} e^{i\omega\cdot x}\, dZ(\omega),$$

where $Z(\omega)$ is an orthogonal process [14]. Let $dH(\omega) = \mathbb{E}\{|dZ(\omega)|^2\}$. We have

$$c_Y(0) = \int_{\mathbb{R}^d} dH(\omega) < +\infty.$$

Since $\operatorname{ess\,sup}_{\omega\in\mathbb{R}^d} |\hat\rho(\omega)| < \infty$,

$$\int_{\mathbb{R}^d} e^{i\lambda(\omega)\cdot x} \hat\rho(\omega)\, dZ(\omega)$$

is convergent in the mean-squared sense. Therefore, $GY(x) = G\int_{\mathbb{R}^d} e^{i\omega\cdot x}\,dZ(\omega)$ is equal to $\int_{\mathbb{R}^d} e^{i\lambda(\omega)\cdot x}\hat{\rho}(\omega)\,dZ(\omega)$. This shows that GY is wide-sense stationary, since

$$\mathbb{E}\{GY(x)GY^*(y)\} = \int_{\mathbb{R}^d} e^{i\lambda(\omega)\cdot(x-y)}|\hat{\rho}(\omega)|^2\,dH(\omega)$$

is a function of $x - y$ and $\mathbb{E}\{|GY(x)|^2\} < \infty$. For any wide-sense stationary process Y, one can write

$$GY(x) = G\mathbb{E}\{Y(0)\} + G\big(Y(x) - \mathbb{E}\{Y(0)\}\big).$$

Since $Y(x) - \mathbb{E}\{Y(0)\}$ is zero-mean and wide-sense stationary, $G(Y(x) - \mathbb{E}\{Y(0)\})$ is wide-sense stationary, therefore so is $GY(x)$. $\square$

In order for G to be stationarity invariant, for any positive integrable measure $dH(\omega)$ one must have $\int_{\mathbb{R}^d}|\hat{\rho}(\omega)|^2\,dH(\omega) < +\infty$. One can verify that a necessary and sufficient condition is that $\operatorname{ess\,sup}_{\omega\in\mathbb{R}^d}|\hat{\rho}(\omega)| < \infty$.

PROOF OF PROPOSITION 2.4. The autocovariance operator of a process Z is defined by

$$K_Z f(x) = \int \mathbb{E}\{Z(x)Z^*(y)\} f(y)\,dy.$$

Let Y be a stationary process, let G be a bounded linear operator satisfying (15) and let $X = GY$. The autocovariance operators of X and Y satisfy

$$K_X = GK_Y\bar{G}.$$

Since Y is stationary, K_Y commutes with the translation operator T_v for any $v \in \mathbb{R}^d$. We derive from (15) that K_X also commutes with T_v and hence X is wide-sense stationary. The operator G is therefore stationarity invariant and Theorem 2.1 proves that

$$Ge^{i\omega\cdot x} = \hat{\rho}(\omega)e^{i\lambda(\omega)\cdot x}. \tag{62}$$

Inserting this expression in the equality $GT_{Sv}f(x) = e^{i\xi\cdot v}T_v Gf(x)$ for $f(x) = e^{i\omega\cdot x}$ implies that

$$\hat{\rho}(\omega)e^{i\lambda(\omega)\cdot x}e^{-iSv\cdot\omega} = \hat{\rho}(\omega)e^{i\lambda(\omega)\cdot(x-v)}e^{i\xi\cdot v},$$

from which we derive that $\lambda(\omega) = \overline{S}\omega + \xi$ for all ω, where $\hat{\rho}(\omega) \neq 0$. For ω such that $\hat{\rho}(\omega) = 0$, (62) clearly holds with $\lambda(\omega) = \overline{S}\omega + \xi$. So G can indeed be written as in (14).

Conversely, if G satisfies (14) then a direct calculation shows that (15) holds. $\square$

APPENDIX B

Proofs of Section 3.

PROOF OF PROPOSITION 3.1 (Transport). The autocovariance operator of $X = DY$ satisfies $K_X = DK_Y\overline{D}$. Therefore,

$$\langle K_X\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle = \langle K_Y\overline{D}\psi_{v,\alpha,\sigma}, \overline{D}\psi_{v,\alpha,\sigma}\rangle.$$

Let us compute

$$\begin{aligned}\vec{\nabla}_v\langle K_X\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle &= 2\,\mathrm{Re}\big\langle K_Y\ \overline{D}\psi_{v,\alpha,\sigma}, \vec{\nabla}_v\overline{D}\psi_{v,\alpha,\sigma}\big\rangle \\ &= 2\,\mathrm{Re}\big\langle K_Y\ \overline{D}\psi_{v,\alpha,\sigma}, (\vec{\nabla}_v + \vec{\nabla}_x)\overline{D}\psi_{v,\alpha,\sigma}\big\rangle \\ &\quad - 2\,\mathrm{Re}\big\langle K_Y\overline{D}\psi_{v,\alpha,\sigma}, \vec{\nabla}_x\overline{D}\psi_{v,\alpha,\sigma}\big\rangle.\end{aligned}$$

Since Y is stationary, for any g we have $\langle K_Y g, \vec{\nabla}_x g\rangle = 0$, so

$$\vec{\nabla}_v\langle K_X\,\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle = 2\,\mathrm{Re}\big\langle K_X\psi_{v,\alpha,\sigma}, \overline{D^{-1}}(\vec{\nabla}_v + \vec{\nabla}_x)\overline{D}\psi_{v,\alpha,\sigma}\big\rangle.$$

Hypothesis (16) thus implies that

$$|\vec{\nabla}_v\langle K_X\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle| = O(\sigma)|\,\mathrm{Re}\langle K_X\psi_{v,\alpha,\sigma}, \vec{\nabla}_x\psi_{v,\alpha,\sigma}\rangle|. \tag{63}$$

Since $\psi_{v,\alpha,\sigma} = \overline{G}_{\beta(v)}T_v\overline{F}_\alpha\psi_\sigma$, the transport property (17) shows that

$$\psi_{v,\alpha,\sigma} = e^{i\phi(u(v))}T_{u(v)}\overline{F}_{\gamma(u(v))}\overline{F}_\alpha\psi_\sigma = e^{i\phi(u(v))}T_{u(v)}\overline{F}_{\alpha*\gamma(u(v))}\psi_\sigma.$$

The phase $e^{i\phi(u(v))}$ disappears from $\langle K_X\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle$. Indeed,

$$\begin{aligned}\langle K_X\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle &= \big\langle K_X T_{u(v)}\overline{F}_{\alpha*\gamma(u(v))}, T_{u(v)}\overline{F}_{\alpha*\gamma(u(v))}\big\rangle \\ &= A_X^\sigma\big(u(v), \alpha*\gamma(u(v))\big)\end{aligned}$$

by the definition of A_X^σ. Since $\vec{\nabla}_v f = \vec{\nabla}_u f\, J_v^{-1}(u)$,

$$\begin{aligned}\vec{\nabla}_v\langle K_X\,\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle &= \vec{\nabla}_v A_X^\sigma\big(u(v), \alpha*\gamma(u(v))\big) \\ &= \vec{\nabla}_u A_X^\sigma\big(u(v), \alpha*\gamma(u(v))\big)J_v^{-1}(u).\end{aligned}$$

This implies that

$$\big|\vec{\nabla}_u A_X^\sigma\big(u(v), \alpha*\gamma(u(v))\big)\big| \le \|J_v(u)\|\,|\vec{\nabla}_v\langle K_X\psi_{v,\alpha,\sigma}, \psi_{v,\alpha,\sigma}\rangle|,$$

where $\|J_v(u)\|$ is the operator sup norm of $J_v(u)$. Using (63) shows that, for u fixed,

$$\big|\vec{\nabla}_u A_X^\sigma\big(u, \alpha*\gamma(u)\big)\big| = O(\sigma)|\,\mathrm{Re}\langle K_X\,\psi_{v,\alpha,\sigma}, \vec{\nabla}_x\psi_{v,\alpha,\sigma}\rangle|. \tag{64}$$

Note that the gradient with respect to u on the left-hand side of the above expression involves partial derivatives of A_X^σ with respect to u and α since the variable u appears in $\alpha*\gamma(u)$.

Since $\vec{\nabla}_u T_u f(x) = -\vec{\nabla}_x T_u f(x)$, using the symmetry of K_X, we get

$$2\,\mathrm{Re}\langle K_X \psi_{v,\alpha,\sigma}, \vec{\nabla}_x \psi_{v,\alpha,\sigma}\rangle = -\vec{\nabla}_u A_X^\sigma\big(u, \alpha * \gamma(t)\big) \qquad \text{at } t = u. \tag{65}$$

Inserting (65) in (64) finally proves that

$$\big|\vec{\nabla}_u A_X^\sigma\big(u, \alpha * \gamma(u)\big)\big| = O(\sigma)\big|\vec{\nabla}_u A_X^\sigma\big(u, \alpha * \gamma(t)\big)\big| \qquad \text{at } t = u,$$

which implies (18). □

PROOF OF THEOREM 3.1 (Scale transport). This theorem is proved as a consequence of Proposition 3.1. The operator $G_{\beta(v)}$ is given by (21),

$$\overline{G}_{\beta(v)} f(x) = \theta'(u) f\big(v + \theta'(u)(x-u)\big),$$

where $u = \theta^{-1}(v)$ is a differentiable invertible map. Notice that

$$\overline{G}_{\beta(v)} T_v = T_{u(v)} \overline{F}_{\alpha(u(v))},$$

with $\alpha(u) = 1/\theta'(u)$; therefore, transport property (17) holds.

Let us now verify hypothesis (16) concerning

$$\psi_{v,s,\sigma} = \overline{G}_{\beta(v)} T_v \overline{F}_s \psi_\sigma,$$

with $\overline{F}_s f(x) = 1/s f(x/s)$. The scalogram renormalization (20) is equivalent to dividing $\psi_\sigma(x)$ by σ, which yields $\psi_\sigma(x) = 1/\sigma\, \psi(x/\sigma)$, and replacing σs by s, which gives

$$\psi_{v,s,\sigma}(x) = \varphi_{v,s}(x) = \frac{\theta'(u)}{s}\psi\bigg(\frac{\theta'(u)}{s}(x-u)\bigg).$$

If we can prove that

$$\big|\mathrm{Re}\big\langle K_X \varphi_{v,s}, \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\varphi_{v,s}\big\rangle\big| = O(s)\,|\,\mathrm{Re}\langle K_X \varphi_{v,s}, \partial_x \varphi_{v,s}\rangle|, \tag{66}$$

then Proposition 3.1 can be applied: we obtain a transport equation (19) with $\alpha = s$, $\gamma(u) = 1/\theta'(u)$ and $\alpha * \gamma(t) = s/\theta'(t)$:

$$\bigg|\partial_u A_X(u,s) - s\theta'(u)\frac{\theta''(t)}{(\theta'(t))^2}\,\partial_s A_X(u,s)\bigg| = O(s)|\partial_u A_X(u,s)| \qquad \text{at } t = u,$$

which proves (24). It now only remains to prove (66).

Let us compute

$$\phi_{v,s} = \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\varphi_{v,s}.$$

Since $\overline{D} f(x) = (\theta'(\theta^{-1}(x)))^{-1} f(\theta^{-1}(x))$ and $\overline{D^{-1}} f(x) = \theta'(x) f(\theta(x))$, a direct calculation gives

$$\begin{aligned}
&\overline{D^{-1}}\partial_x \overline{D}\varphi_{v,s}(x) \\
&\quad = -\frac{\theta'(u)\theta''(x)}{s\theta'(x)^2}\psi\bigg(\frac{\theta'(u)}{s}(x-u)\bigg) + \frac{\theta'(u)^2}{s^2\theta'(x)}\psi'\bigg(\frac{\theta'(u)}{s}(x-u)\bigg)
\end{aligned}$$

and

$$\overline{D^{-1}}\partial_v \overline{D}\varphi_{v,s}(x) = \frac{\theta''(u)}{s\theta'(u)}\psi\left(\frac{\theta'(u)}{s}(x-u)\right) + \frac{1}{s^2}\big((x-u)\theta''(u) - \theta'(u)\big)\psi'\left(\frac{\theta'(u)}{s}(x-u)\right).$$

Therefore,

$$\phi_{v,s}(x) = \frac{1}{s\theta'(u)}\left(\theta''(u) - \left(\frac{\theta'(u)}{\theta'(x)}\right)^2\theta''(x)\right)\psi\left(\frac{\theta'(u)}{s}(x-u)\right) + \frac{1}{s^2}\left(\frac{\theta'(u)}{\theta'(x)}\big(\theta'(u)-\theta'(x)\big) - (u-x)\theta''(u)\right)\psi'\left(\frac{\theta'(u)}{s}(x-u)\right).$$

Since ψ is supported in $[-1,1]$, $\phi_{v,s}$ is supported in $[u - s/\theta'(u), u + s/\theta'(u)]$. Since $\theta \in \mathbf{C}^3$, Taylor series expansions of $\theta'(x)$ and of $\theta''(x)$ around position u prove that, for x close to position u,

$$\left|\phi_{v,s}(x) - \left(2\frac{\theta''(u)^2}{\theta'(u)^3} - \frac{\theta'''(u)}{\theta'(u)^2}\right)\left[\frac{\theta'(u)(x-u)}{s}\psi\left(\frac{\theta'(u)}{s}(x-u)\right) + \frac{1}{2}\left(\frac{\theta'(u)(x-u)}{s}\right)^2\psi'\left(\frac{\theta'(u)}{s}(x-u)\right)\right]\right| = O(x-u).$$

The autocovariance kernel of $X(x)$ is $c_X(x,y) = c_Y(\theta(x)-\theta(y))$. Hence,

$$\langle K_X\varphi_{v,s}, \phi_{v,s}\rangle = \iint c_Y\big(\theta(x)-\theta(y)\big)\varphi_{v,s}^*(x)\phi_{v,s}(y)\,dx\,dy.$$

Since $\int \varphi_{v,s}(x)\,dx = \int \psi(x)\,dx = 0$,

$$\langle K_X\varphi_{v,s}, \phi_{v,s}\rangle = -\iint \big(c_Y(0) - c_Y\big(\theta(x)-\theta(y)\big)\big)\varphi_{v,s}^*(x)\phi_{v,s}(y)\,dx\,dy.$$

The supports of $\varphi_{v,s}$ and $\phi_{v,s}$ are in $[u - s/\theta'(u), u + s/\theta'(u)]$ and for z in a neighborhood of 0, the continuity of η implies that $c_Y(0) - c_Y(z) = \eta(0)|z|^h + o(|z|^h)$. Since θ' is continuous at u, a Taylor series expansion of θ around u combined with a change of variables $x' = (x-u)\theta'(u)/s$ and $y' = (y-u)\theta'(u)/s$ yield, for s sufficiently small,

$$\langle K_X\varphi_{v,s}, \phi_{v,s}\rangle + \iint \eta(0)|s(x'-y')|^h\psi^*(x')\left(2\frac{\theta''(u)^2}{\theta'(u)^3} - \frac{\theta'''(u)}{\theta'(u)^2}\right) \times \left[y'\psi(y') + \frac{y'^2}{2}\psi'(y')\right]\frac{s}{\theta'(u)}\,dx'\,dy' = o(s^{h+1}). \tag{67}$$

Therefore,

$$|\mathrm{Re}\langle K_X\varphi_{v,s}\phi_{v,s}\rangle| = O(s^{h+1}). \tag{68}$$

Let us now compute

$$\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle = \iint c_Y\big(\theta(x)-\theta(y)\big)\varphi^*_{v,s}(x)\frac{d}{dy}\varphi_{v,s}(y)\,dx\,dy.$$

With an integration by parts,

$$\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle = \iint \theta'(y)c'_Y\big(\theta(x)-\theta(y)\big)\varphi^*_{v,s}(x)\varphi_{v,s}(y)\,dx\,dy,$$

and since $c'_Y(z)$ is antisymmetric,

$$\begin{aligned}&\mathrm{Re}\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle\\ &\quad= -\tfrac{1}{2}\iint\big(\theta'(x)-\theta'(y)\big)c'_Y\big(\theta(x)-\theta(y)\big)\,\mathrm{Re}\big(\varphi^*_{v,s}(x)\,\varphi_{v,s}(y)\big)\,dx\,dy.\end{aligned}$$

A change of variables $x'=(x-u)\theta'(u)/s$ and $y'=(y-u)\theta'(u)/s$ gives

$$\begin{aligned}\mathrm{Re}\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle = -\tfrac{1}{2}\iint&\big(\theta'\big(u+sx'/\theta'(u)\big)-\theta'\big(u+sy'/\theta'(u)\big)\big)\\ &\times c'_Y\big(\theta\big(u+sx'/\theta'(u)\big)-\theta\big(u+sy'/\theta'(u)\big)\big)\\ &\times \mathrm{Re}\big(\psi^*(x')\psi(y')\big)\,dx'\,dy'.\end{aligned}$$

Because of assumption (22), since η is $\mathbf{C}^1$ in a neighborhood of 0, $c'_Y(z) = h\eta(0)\,\mathrm{sign}(z)|z|^{h-1}+o(|z|^{h-1})$. With a Taylor expansion for θ, we get, for s small enough,

$$\begin{aligned}&\mathrm{Re}\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle\\ &\qquad+\frac{1}{2}\frac{\theta''(u)}{\theta'(u)}\iint h\eta(0)s^h|x-y|^h\,\mathrm{Re}\big(\psi^*(x)\,\psi(y)\big)\,dx\,dy = o(s^h).\end{aligned}\tag{69}$$

Since $\mathrm{Re}\iint|x-y|^h\psi^*(x)\psi(y)\,dx\,dy \neq 0$ and $\theta''(u)\neq 0$, there exists $a(u)>0$ such that

$$|\mathrm{Re}\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle| \geq a(u)s^h + o(s^h),$$

and (68) implies that

$$|\mathrm{Re}\langle K_X\varphi_{v,s}\phi_{v,s}\rangle| = O(s)|\mathrm{Re}\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle|.$$

We have therefore proved (66). □

PROOF OF THEOREM 3.2 (Frequency transport). This theorem is proved as a consequence of Proposition 3.1. The operator $G_{\beta(v)}$ is defined in (25) by

$$\overline{G}_{\beta(v)}f(x) = e^{i(\theta(v)+\theta'(v)(x-v))}f(x).\tag{70}$$

Let $u(v)=v$, which is clearly a differentiable and invertible map. We have

$$\overline{G}_{\beta(v)}T_v = e^{i\theta(u(v))}T_{u(v)}\overline{F}_{\alpha(u(v))},$$

with $\overline{F}_\alpha f(x) = e^{-i\alpha x} f(x)$ and $\alpha(u) = -\theta'(u)$. Therefore, transport property (17) holds.

Let us now verify hypothesis (16):

$$\begin{aligned}
&\bigl|\mathrm{Re}\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\psi_{v,\xi,\sigma}\rangle\bigr| \\
&\qquad = O(\sigma)\left|\mathrm{Re}\left\langle K_X\psi_{v,\xi,\sigma}, \frac{\partial}{\partial x}\psi_{v,\xi,\sigma}\right\rangle\right|,
\end{aligned}\tag{71}$$

with

$$\begin{aligned}
\psi_{v,\xi,\sigma}(x) &= \overline{G}_{\beta(v)} T_v \overline{F}_\xi \psi_\sigma(x) \\
&= \exp\bigl[i\bigl(\theta(v) + \theta'(v)(x-v)\bigr)\bigr]\exp[-i\xi(x-v)]\psi\left(\frac{x-v}{\sigma}\right).
\end{aligned}\tag{72}$$

A direct calculation shows that

$$\begin{aligned}
&\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\psi_{v,\xi,\sigma}\rangle \\
&\quad = \iint c_Y(x-y)\exp\bigl[i\bigl(\theta(x) - \theta(y) - \theta'(v)(x-y)\bigr)\bigr]\exp[i\xi(x-y)] \\
&\qquad\quad \times i\bigl(-\theta''(v)(y-v) + \theta'(y) - \theta'(v)\bigr)\psi\left(\frac{x-v}{\sigma}\right)\psi\left(\frac{y-v}{\sigma}\right)dx\,dy,
\end{aligned}$$

and with a change of variables $x' = (x-v)/\sigma$ and $y' = (y-v)/\sigma$, introducing $\xi_0 = \sigma\xi$,

$$\begin{aligned}
&\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\psi_{v,\xi,\sigma}\rangle \\
&\quad = \sigma^2 \iint c_Y\bigl(\sigma(x'-y')\bigr) \\
&\qquad\qquad \times \exp\bigl[i\bigl(\theta(v+\sigma x') - \theta(v+\sigma y') - \sigma\theta'(v)(x'-y')\bigr)\bigr] \\
&\qquad\qquad \times i\bigl(\theta'(v+\sigma y') - \theta'(v) - \theta''(v)\sigma y'\bigr) \\
&\qquad\qquad \times e^{i\xi_0(x'-y')}\psi(x')\psi(y')\,dx'\,dy'.
\end{aligned}\tag{73}$$

The function ψ is real; therefore $\hat{\psi}$ is even, and since it vanishes at ξ_0, it also vanishes at $-\xi_0$. Therefore

$$\begin{aligned}
&\sigma^2 \iint c_Y(0) i\bigl(\theta'(v+\sigma y) - \theta'(v) - \theta''(v)\sigma y\bigr) \\
&\qquad \times e^{i\xi_0(x-y)}\,\psi(x)\psi(y)\,dx\,dy = 0.
\end{aligned}\tag{74}$$

After subtracting (74) from (73),

$$\begin{aligned}
&\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v+\partial_x)\overline{D}\psi_{v,\xi,\sigma}\rangle \\
&\quad = \sigma^2 \iint c_Y(\sigma(x-y)) \\
&\qquad\quad \times \big(\exp[i(\theta(v+\sigma x)-\theta(v+\sigma y)-\sigma\theta'(v)(x-y))]-1\big) \\
&\qquad\quad \times i\big(\theta'(v+\sigma y)-\theta'(v)-\theta''(v)\sigma y\big)e^{i\xi_0(x-y)}\psi(x)\psi(y)\,dx\,dy \\
&\qquad + \sigma^2 \iint \big(c_Y(\sigma(x-y))-c_Y(0)\big) \\
&\qquad\quad \times i\big(\theta'(v+\sigma y)-\theta'(v)-\theta''(v)\sigma y\big)e^{i\xi_0(x-y)}\psi(x)\psi(y)\,dx\,dy.
\end{aligned} \tag{75}$$

Since $\theta \in \mathbf{C}^{4+\lceil h\rceil}$, we can perform the following Taylor expansions, where a_k, b_k and c_k are real parameters that depend on the derivatives $\theta^{(k)}(v)$, for $(x, y) \in [0,1]^2$:

$$\begin{aligned}
&\exp[i(\theta(v+\sigma x)-\theta(v+\sigma y)-\sigma\theta'(v)(x-y))] \\
&\quad = 1 + i\sum_{k=2}^{2+\lceil h\rceil} a_k\sigma^k(x-y)^k + \sum_{k=4}^{2+\lceil h\rceil} b_{k-2}\sigma^k(x-y)^k + O(\sigma^{3+\lceil h\rceil}),
\end{aligned} \tag{76}$$

$$\theta'(v+\sigma y)-\theta'(v) = \sum_{k=1}^{2+\lceil h\rceil} c_{k+1}\sigma^k y^k + O(\sigma^{3+\lceil h\rceil}). \tag{77}$$

In particular, $a_2 = \theta''(v)/2$ and $c_k = \theta^{(k)}(v)/(k-1)!$.

Replacing these Taylor expansions in (75), we obtain

$$\begin{aligned}
&\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v+\partial_x)\overline{D}\psi_{v,\xi,\sigma}\rangle \\
&\quad = -\sigma^2 \iint c_Y(\sigma(x-y))\left[i\sum_{k=2}^{2+\lceil h\rceil} a_k\sigma^k(x-y)^k + \sum_{k=4}^{2+\lceil h\rceil} b_{k-2}\sigma^k(x-y)^k\right] \\
&\qquad\qquad \times \sum_{k=2}^{2+\lceil h\rceil} c_{k+1}\sigma^k y^k e^{i\xi_0(x-y)}\psi(x)\psi(y)\,dx\,dy \\
&\qquad - i\sigma^2 \iint \sigma^h|x-y|^h\eta(\sigma(x-y)) \\
&\qquad\qquad \times \left[\sum_{k=2}^{2+\lceil h\rceil} c_{k+1}\sigma^k y^k\right] e^{i\xi_0(x-y)}\psi(x)\psi(y)\,dx\,dy \\
&\qquad + o(\sigma^{5+\lceil h\rceil}).
\end{aligned}$$

In the first of these two integrals, one can replace $c_Y(\sigma(x-y))$ by $c_Y(0) - \sigma^h|x-y|^h\eta(\sigma(x-y))$. Since $\hat{\psi}$ and its first $\lceil h\rceil + 2$ derivatives vanish at ξ_0, we derive that $e^{i\xi_0 t}\psi(t)$ is a function with $\lceil h\rceil + 3$ vanishing moments [11], so the first integral is on the order of $O(\sigma^{6+h})$.

Consider the real part of the second integral: because η is even, exchanging x and y shows that

$$\iint |x-y|^h \eta\bigl(\sigma(x-y)\bigr) y^2 \sin\bigl(\xi_0(y-x)\bigr)\psi(x)\psi(y)\,dx\,dy = 0.$$

Since $2 + \lceil h\rceil \geq 3$, the real part of the second integral is equal to

$$\sigma^{5+h}\eta(0)\frac{\theta^{(4)}(v)}{6}\iint \sin\bigl(\xi_0(x-y)\bigr)|x-y|^h y^3 \psi(x)\psi(y)\,dx\,dy + o(\sigma^{5+h}).$$

As a consequence,

$$\begin{aligned} &\operatorname{Re}\bigl\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\psi_{v,\xi,\sigma}\bigr\rangle \\ &\quad = \sigma^{5+h}\eta(0)\frac{\theta^{(4)}(v)}{6}\iint \sin\bigl(\xi_0(x-y)\bigr)|x-y|^h y^3\psi(x)\psi(y)\,dx\,dy \\ &\qquad + o(\sigma^{5+h}). \end{aligned} \tag{78}$$

Let us now compute the leading term of $|\operatorname{Re}\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle|$. After a change of variables,

$$\begin{aligned} &\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle \\ &\quad = \sigma^2\iint c_Y\bigl(\sigma(x-y)\bigr)\exp\bigl[i\bigl(\theta(v+\sigma x) - \theta(v+\sigma y) - \sigma\theta'(v)(x-y)\bigr)\bigr] \\ &\qquad\qquad \times \exp\bigl[i\xi_0(x-y)\bigr] i\bigl(\theta'(v) + \xi_0/\sigma\bigr)\psi(x)\psi(y)\,dx\,dy \\ &\qquad + \sigma^2\iint c_Y\bigl(\sigma(x-y)\bigr)\exp\bigl[i\bigl(\theta(v+\sigma x) - \theta(v+\sigma y) - \sigma\theta'(v)(x-y)\bigr)\bigr] \\ &\qquad\qquad \times \exp[i\xi_0(x-y)]\psi(x)\frac{1}{\sigma}\psi'(y)\,dx\,dy. \end{aligned}$$

Using Taylor expansions (76) and (77), we obtain

$$\begin{aligned} &\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle \\ &\quad = \sigma^2\iint c_Y\bigl(\sigma(x-y)\bigr) \\ &\qquad\qquad \times \left[1 + i\sum_{k=2}^{2+\lceil h\rceil} a_k\sigma^k(x-y)^k + \sum_{k=4}^{2+\lceil h\rceil} b_{k-2}\sigma^k(x-y)^k\right] \\ &\qquad\qquad \times \exp[i\xi_0(x-y)]i(\theta'(v) + \xi_0/\sigma)\psi(x)\psi(y)\,dx\,dy \end{aligned} \tag{79}$$

$$
\begin{aligned}
&+\sigma^2 \iint c_Y\big(\sigma(x-y)\big) \\
&\qquad \times \Bigg[1 + i \sum_{k=2}^{2+\lceil h\rceil} a_k \sigma^k (x-y)^k + \sum_{k=4}^{2+\lceil h\rceil} b_{k-2}\sigma^k (x-y)^k\Bigg] \\
&\qquad \times \exp[i\xi_0(x-y)]\psi(x)\frac{1}{\sigma}\psi'(y)\,dx\,dy + O\big(\sigma^{4+\lceil h\rceil}\big).
\end{aligned}
$$

Exchanging x and y shows that

$$
\iint c_Y\big(\sigma(x-y)\big)\sin[\xi_0(y-x)]\psi(x)\psi(y)\,dx\,dy = 0,
$$

and since ψ is even and ψ' is odd, changing x to $-x$ and y to $-y$ shows that

$$
\iint c_Y\big(\sigma(x-y)\big)\cos[\xi_0(y-x)]\psi(x)\,\psi'(y)\,dx\,dy = 0.
$$

Writing $c_Y(\sigma(x-y)) = c_Y(0) - \sigma^h|x-y|^h\eta(\sigma(x-y))$ and noticing that $e^{i\xi_0 t}\psi(t)$ is a function with $\lceil h\rceil + 3$ vanishing moments, the first integral in (79) has a real part equal to

$$
\sigma^{3+h}\frac{\theta''(v)}{2}\eta(0)\iint |x-y|^{2+h}\xi_0\cos[\xi_0(y-x)]\psi(x)\psi(y)\,dx\,dy + o(\sigma^{3+h}).
$$

Because ψ is even, the second integral in (79) has a real part equal to

$$
-\sigma^{3+h}\frac{\theta''(v)}{2}\eta(0)\iint |x-y|^{2+h}\sin[\xi_0(y-x)]\psi(x)\psi'(y)\,dx\,dy + o(\sigma^{1+h}).
$$

An integration by parts with respect to y shows that

$$
\begin{aligned}
&\iint |x-y|^{2+h}\sin[\xi_0(y-x)]\psi(x)\psi'(y)\,dx\,dy \\
&\quad = -\xi_0\iint |x-y|^{2+h}\cos[\xi_0(y-x)]\psi(x)\psi(y)\,dx\,dy \\
&\qquad + (2+h)\iint |x-y|^{1+h}\operatorname{sign}(x-y)\sin[\xi_0(y-x)]\psi(x)\psi(y)\,dx\,dy.
\end{aligned}
$$

Summing the two contributions, we see that

$$
\begin{aligned}
&\operatorname{Re}\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle \\
&\quad = \sigma^{3+h}\eta(0)(1+h/2)\theta''(v) \\
&\qquad\qquad \times \iint |x-y|^h(x-y)\sin[\xi_0(x-y)]\psi(x)\psi(y)\,dx\,dy \\
&\qquad + o(\sigma^{3+h}).
\end{aligned} \tag{80}
$$

Because of the hypothesis that

$$\iint |x-y|^h (x-y) \sin[\xi_0(x-y)]\psi(x)\psi(y)\,dx\,dy \neq 0,$$

comparing (80) and (78) proves a result that is stronger than (71), because the right-hand side has order $O(\sigma^2)$ instead of $O(\sigma)$:

$$\big|\operatorname{Re}\big\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v+\partial_x)\overline{D}\psi_{v,\xi,\sigma}\big\rangle\big| = O(\sigma^2)|\operatorname{Re}\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle|.$$

Adapting Proposition 3.1 to account for the $O(\sigma^2)$ term, we obtain a transport equation (19) with $\alpha=\xi$, $\gamma(u)=-\theta'(u)$ and $\alpha_1 * \alpha_2 = \alpha_1+\alpha_2$: for u such that $\theta''(u)\neq 0$,

$$|\partial_u A_X^\sigma(u,\xi) - \theta''(u)\,\partial_\xi A_X^\sigma(u,\xi)| = O(\sigma^2)|\partial_u A_X^\sigma(u,\xi)|,$$

which proves (27). $\square$

PROOF OF THEOREM 3.3 (Multidimensional scale transport). The proof of this theorem follows the same lines as the proof of Theorem 3.1. The hypotheses of Proposition 3.1 are verified in order to apply (19) in d dimensions.

Let us verify hypothesis (16) concerning

$$\psi_{v,S,\sigma} = \overline{G}_{\beta(v)} T_v \overline{F}_S \psi_\sigma,$$

with $\overline{F}_S f(x) = \det S^{-1} f(S^{-1}x)$ and where $\overline{G}_{\beta(v)}$ has been defined in (29). Note that the transport property (17) clearly holds. The warpogram renormalization (30) is equivalent to dividing $\psi_\sigma(x)$ by σ^d and replacing σS by S. We replace $\psi_{v,S,\sigma}(x)$ by

$$\varphi_{v,S}(x) = \det\big(S^{-1}J_\theta(u)\big)\psi\big(S^{-1}J_\theta(u)(x-u)\big). \tag{81}$$

Let us define the vector of functions

$$\vec{\phi}_{v,S} = \overline{D^{-1}}(\vec{\nabla}_v + \vec{\nabla}_x)\overline{D}\varphi_{v,S}.$$

We now prove that, for any fixed u and $\tilde{S}$ such that $\det\tilde{S}=1$, if $S=\sigma\tilde{S}$, then

$$|\operatorname{Re}\langle K_X\varphi_{v,S}, \vec{\phi}_{v,S}\rangle| = O(\sigma)|\operatorname{Re}\langle K_X\varphi_{v,S}, \vec{\nabla}_x\varphi_{v,S}\rangle|. \tag{82}$$

Let us first compute an upper bound for $|\operatorname{Re}\langle K_X\varphi_{v,S}, \vec{\phi}_{v,S}\rangle|$. Since

$$\overline{D^{-1}} f(x) = \det(J_\theta(x)) f(\theta(x))$$

and

$$\overline{D} f(x) = \det\big(J_\theta^{-1}(\theta^{-1}(x))\big) f\big(\theta^{-1}(x)\big),$$

we have

$$\begin{aligned}
&\overline{D^{-1}}\vec{\nabla}_x \overline{D}\varphi_{v,S}(x)\\
&\quad = \sigma^{-d}\biggl[-\frac{\det J_\theta(u)}{\det J_\theta(x)}\vec{\nabla}\det J_\theta(x) J_\theta^{-1}(x)\psi\bigl(S^{-1}J_\theta(u)(x-u)\bigr)\\
&\qquad\qquad + \det J_\theta(u)\vec{\nabla}\psi\bigl(S^{-1}J_\theta(u)(x-u)\bigr)S^{-1}J_\theta(u)J_\theta^{-1}(x)\biggr]
\end{aligned}$$

and

$$\begin{aligned}
&\overline{D^{-1}}\vec{\nabla}_v \overline{D}\varphi_{v,S}(x)\\
&\quad = \sigma^{-d}\bigl[\vec{\nabla}\det J_\theta(u)\psi\bigl(S^{-1}J_\theta(u)(x-u)\bigr)J_\theta^{-1}(u)\\
&\qquad\qquad + \det J_\theta(u)\vec{\nabla}\psi\bigl(S^{-1}J_\theta(u)(x-u)\bigr)\\
&\qquad\qquad\quad \times S^{-1}\bigl(\vec{\nabla}J_\theta(u)(x-u) - J_\theta(u)\bigr)J_\theta^{-1}(u)\bigr].
\end{aligned}$$

After summing these two expressions, a Taylor expansion of $\det J_\theta$, J_θ^{-1} and $\vec{\nabla}\det J_\theta$ in the vicinity of position u shows that, for $S = \sigma\tilde{S}$ and σ small, there exists $C(u,\tilde{S})$ such that

$$|\vec{\phi}_{v,S}| \le C(u,\tilde{S})\sigma^{1-d}. \tag{83}$$

By the definition of K_X,

$$\langle K_X\varphi_{v,S}, \vec{\phi}_{v,S}\rangle = \iint c_Y\bigl(\theta(x)-\theta(y)\bigr)\varphi^*_{v,S}(x)\vec{\phi}_{v,S}(y)\,dx\,dy.$$

The wavelet ψ has one vanishing moment, so $\int \varphi_{v,S}(x)\,dx = 0$, and therefore

$$\langle K_X\varphi_{v,S}, \vec{\phi}_{v,S}\rangle = \iint \bigl[c_Y\bigl(\theta(x)-\theta(y)\bigr) - c_Y(0)\bigr]\varphi^*_{v,S}(x)\vec{\phi}_{v,S}(y)\,dx\,dy,$$

which implies that

$$|\langle K_X\varphi_{v,S}, \vec{\phi}_{v,S}\rangle| \le \iint \bigl|c_Y\bigl(\theta(x)-\theta(y)\bigr) - c_Y(0)\bigr|\,|\varphi_{v,S}(x)|\,|\vec{\phi}_{v,S}(y)|\,dx\,dy.$$

Substituting (83) and (81) in the above inequality and using condition (33) on c_Y, after a change of variables and a Taylor expansion of θ around u, we obtain

$$|\langle K_X\varphi_{v,S}, \vec{\phi}_{v,S}\rangle| = O(\sigma^{h+1}).$$

To prove (82), we now show that there exists $C'(u,\tilde{S}) > 0$ such that

$$|\operatorname{Re}\langle K_X\varphi_{v,S}, \vec{\nabla}_x\varphi_{v,S}\rangle| \ge C'(u,\tilde{S})\sigma^h. \tag{84}$$

With an integration by parts,

$$\langle K_X\varphi_{v,S}, \vec{\nabla}_x\varphi_{v,S}\rangle = \iint \vec{\nabla}c_Y\bigl(\theta(x)-\theta(y)\bigr)J_\theta(y)\varphi^*_{v,S}(x)\varphi_{v,S}(y)\,dx\,dy,$$

and using the fact that $\vec{\nabla} c_Y(x)$ is antisymmetric,

$$\langle K_X \varphi_{v,S}, \vec{\nabla}_x \varphi_{v,S} \rangle = -\frac{1}{2} \iint \vec{\nabla} c_Y\big(\theta(x) - \theta(y)\big)\big(J_\theta(x) - J_\theta(y)\big) \operatorname{Re}\big(\varphi_{v,S}^*(x)\varphi_{v,S}(y)\big)\, dx\, dy.$$

Therefore,

$$\begin{aligned}
&\operatorname{Re}\langle K_X \varphi_{v,S}, \vec{\nabla}_x \varphi_{v,S} \rangle \\
&\quad + \frac{1}{2} \iint \vec{\nabla} c_Y\big(S(x-y)\big) \vec{\nabla} J_\theta(u) J_\theta^{-1}(u) S(x-y) \operatorname{Re}\big(\psi^*(x)\psi(y)\big)\, dx\, dy \\
&\qquad = -\frac{1}{2} \iint \Big(\vec{\nabla} c_Y\big(\theta\big(u + J_\theta^{-1}(u)Sx\big) - \theta\big(u + J_\theta^{-1}(u)Sy\big)\big) - \vec{\nabla} c_Y\big(S(x-y)\big)\Big) \\
&\qquad\qquad \times \big(J_\theta\big(u + J_\theta^{-1}(u)Sx\big) - J_\theta\big(u + J_\theta^{-1}(u)Sy\big)\big) \\
&\qquad\qquad \times \operatorname{Re}\big(\psi^*(x)\psi(y)\big)\, dx\, dy \\
&\qquad\quad - \frac{1}{2} \iint \vec{\nabla} c_Y\big(S(x-y)\big)\Big(J_\theta\big(u + J_\theta^{-1}(u)Sx\big) - J_\theta\big(u + J_\theta^{-1}(u)Sy\big) \\
&\qquad\qquad\qquad - \vec{\nabla} J_\theta(u) J_\theta^{-1}(u) S(x-y)\Big) \\
&\qquad\qquad \times \operatorname{Re}\big(\psi^*(x)\, \psi(y)\big)\, dx\, dy.
\end{aligned}$$

Because $\vec{\nabla} c_Y$ is $\mathbf{C}^1$ in a neighborhood of 0 excluding 0, for small σ, second-order Taylor series expansions for θ and J_θ around position u prove that

$$\begin{aligned}
&\operatorname{Re}\langle K_X \varphi_{v,S}, \vec{\nabla}_x \varphi_{v,S} \rangle \\
&\quad + \frac{1}{2} \iint \vec{\nabla} c_Y\big(S(x-y)\big) \vec{\nabla} J_\theta(u) J_\theta^{-1}(u) S(x-y) \operatorname{Re}\big(\psi^*(x)\psi(y)\big)\, dx\, dy \\
&\quad = o(\sigma^h).
\end{aligned}$$

Hypothesis (34) guarantees that (84) holds, and therefore (82) is satisfied. Now that conditions (16) and (17) of Proposition 3.1 have been verified, the resulting transport equation (19) can be applied, with $\alpha = S$, $\gamma(u) = J_\theta^{-1}(u)$ and $S_1 * S_2 = S_2 S_1$. This yields

$$|\vec{\nabla}_u A_X(u,S) + [J_\theta(u)^{-1} \vec{\nabla}_u J_\theta(u) S] \cdot \vec{\nabla}_S A_X(u,S)| = O(\sigma) |\vec{\nabla}_u A_X(u,S)|.$$

The final result (35) is derived from this equation by noting that

$$[J_\theta^{-1}(u) \vec{\nabla}_u J_\theta(u) S] \cdot \vec{\nabla}_S A_X(u,S) = [J_\theta^{-1}(u) \vec{\nabla}_u J_\theta(u)] \cdot [\vec{\nabla}_S A_X(u,S) S^t]. \qquad \square$$

PROOF THAT (36) IMPLIES (34) FOR A SEPARABLE WARPING FUNCTION. Using the fact that θ is assumed separable, we obtain that $\vec{\nabla}_u J_\theta(u) J_\theta(u)^{-1} S(x-y)$ is a diagonal matrix whose ith element along the diagonal is

$$\sigma \frac{\theta_i''(u_i)}{\theta_i'(u_i)} \big(\tilde{S}(x-y)\big)_i.$$

Using (33), the leading term of $\vec{\nabla} c_Y(S(x-y))$, when $\sigma \to 0$, is

$$\eta(0) h \sigma^{h-1} \tilde{S}(x-y) |\tilde{S}(x-y)|^{h-2}.$$

Therefore, when $\sigma \to 0$, the ith component of

$$\mathrm{Re} \iint \vec{\nabla} c_Y\big(S(x-y)\big) \vec{\nabla} J_\theta(u) J_\theta^{-1}(u) S(x-y) \psi^*(x)\psi(y)\, dx\, dy$$

is equivalent to

$$\eta(0) h \sigma^h \frac{\theta_i''(u_i)}{\theta_i'(u_i)} \mathrm{Re} \iint |\tilde{S}(x-y)|^{h-2} \big(\tilde{S}(x-y)\big)_i^2 \psi^*(x)\psi(y)\, dx\, dy.$$

Using the fact that $(\sum_{i=1}^d |k_i|^2)^{1/2} \geq (1/\sqrt{d}) \sum_{i=1}^d |k_i|$, we obtain

$$\begin{aligned} &\left| \mathrm{Re} \iint \vec{\nabla} c_Y\big(S(x-y)\big) \vec{\nabla} J_\theta(u) J_\theta^{-1}(u) S(x-y) \psi^*(x)\psi(y)\, dx\, dy \right| \\ &\qquad \geq \frac{\eta(0)h}{\sqrt{d}} \sigma^h \min_i \left| \frac{\theta_i''(u_i)}{\theta_i'(u_i)} \right| \left| \mathrm{Re} \iint |\tilde{S}(x-y)|^h \psi^*(x)\psi(y)\, dx\, dy \right|. \end{aligned}$$

This shows that if

$$\mathrm{Re} \iint |\tilde{S}(x-y)|^h \psi^*(x)\psi(y)\, dx\, dy \neq 0$$

and if none of the θ_i'' vanish, then (34) is satisfied with $C(u, \tilde{S}) > 0$. □

APPENDIX C

Proofs of Section 4.

PROOF OF THEOREM 4.1. With a slight modification of the proof of Theorem 3.1, one can prove a stronger result than (24), which is stated in the following lemma.

LEMMA C.1. *Under the hypotheses of Theorem* 3.1,

$$\partial_u A_X(u,s) - (\log \theta')'(u)\, \partial_{\log s} A_X(u,s) = s\big(C(u) + o(1)\big)\, \partial_u A_X(u,s), \tag{85}$$

where C is continuous.

Let a be a generic variable denoting either u or $\log s$ and let us introduce

$$\overline{\partial_a A_X}(u,s) = \int g(u-v)\, \partial_a A_X(v,s)\, dv. \tag{86}$$

If u is such that $\theta''(u) \neq 0$ and Δ is small enough, then $\partial_u A_X(v,s)$ keeps a constant sign over $[u-\Delta, u+\Delta]$. Therefore, by the continuity of C, convolving the right-hand side of (85) with g gives

$$O(s)\, \overline{\partial_u A_X}(u,s).$$

Convolving the left-hand side of (85) with g gives

$$\overline{\partial_u A_X}(u,s) - \int g(u-v)(\log\theta')'(v)\,\partial_{\log s} A_X(v,s)\,dv.$$

The hypotheses of Theorem 3.1 imply that $\partial_{\log s} A_X(u,s)$ does not vanish. By continuity, $\partial_{\log s} A_X(v,s)$ therefore keeps a constant sign for v in $[u-\Delta, u+\Delta]$. Because $(\log\theta')''$ is bounded over $[u-\Delta,\ u+\Delta]$,

$$\begin{aligned}&\int g(u-v)(\log\theta')'(v)\,\partial_{\log s} A_X(v,s)\,dv\\ &\qquad = (\log\theta')'(u)\,\overline{\partial_{\log s} A_X}(u,s) + O(\Delta)\,\overline{\partial_{\log s} A_X}(u,s).\end{aligned}$$

Regrouping the left- and right-hand sides, we obtain

$$\begin{aligned}&\overline{\partial_u A_X}(u,s) - (\log\theta')'(u)\overline{\partial_{\log s} A_X}(u,s)\\ &\qquad = O(s)\,\overline{\partial_u A_X}(u,s) + O(\Delta)\,\overline{\partial_{\log s} A_X}(u,s),\end{aligned}\tag{87}$$

which can be viewed as an averaged transport equation.

The following lemma, whose proof is given later, shows that the two estimators $\widehat{\partial_u A_X}(u,s)$ and $\widehat{\partial_{\log s} A_X}(u,s)$ are consistent.

LEMMA C.2. *Let X, Y and ψ satisfy the hypotheses of Theorem* 4.1. *For each u, for s small enough,*

$$\operatorname{Prob}\Bigl\{\bigl|\widehat{\partial_{\log s} A_X}(u,s) - \overline{\partial_{\log s} A_X}(u,s)\bigr| \ge C\bigl|\overline{\partial_{\log s} A_X}(u,s)\bigr|\Bigr\} \le \varepsilon_1,\tag{88}$$

$$\operatorname{Prob}\Bigl\{\bigl|\widehat{\partial_u A_X}(u,s) - \overline{\partial_u A_X}(u,s)\bigr| \ge C\bigl|\overline{\partial_u A_X}(u,s)\bigr|\Bigr\} \le \varepsilon_2,\tag{89}$$

where

$$C = \frac{\log(N\Delta)}{\Delta\sqrt{N\Delta}}, \qquad \varepsilon_1 = \frac{C_1(u)\Delta^2}{(\log(N\Delta))^2}, \qquad \varepsilon_2 = 6(N\Delta)^{-1/(2C_2(u))}.$$

The parameters $C_1(u)$ and $C_2(u)$, which are defined in the proof of the lemma, are both positive. The weak consistency of

$$\frac{\widehat{\partial_u A_X}(u, N^{-1})}{\widehat{\partial_{\log s} A_X}(u, N^{-1})}$$

as an estimator of $(\log\theta')'(u)$ then results from the following lemma, whose proof is straightforward.

LEMMA C.3. *If X_1 and X_2 are two random variables and $C < 1$ is a constant such that*

$$\operatorname{Prob}\bigl\{|X_1 - \mathbb{E}\{X_1\}| \le C\,|\mathbb{E}\{X_1\}|\bigr\} \ge 1 - \varepsilon_1,$$

$$\operatorname{Prob}\bigl\{|X_2 - \mathbb{E}\{X_2\}| \le C\,|\mathbb{E}\{X_2\}|\bigr\} \ge 1 - \varepsilon_2,$$

then

$$\operatorname{Prob}\left(\left|\frac{X_2}{X_1}-\frac{\mathbb{E}\{X_2\}}{\mathbb{E}\{X_1\}}\right|\le\frac{2C}{1-C}\left|\frac{\mathbb{E}\{X_2\}}{\mathbb{E}\{X_1\}}\right|\right)\ge 1-\varepsilon_1-\varepsilon_2.$$

In view of Lemma C.2, one can apply Lemma C.3 to $X_1=\widehat{\partial_{\log s}A_X}(u,N^{-1})$ and $X_2=\widehat{\partial_u A_X}(u,N^{-1})$ with $C=\log(N\Delta)/\Delta\sqrt{N\Delta}$, yielding

$$\operatorname{Prob}\left\{\left|\frac{\widehat{\partial_u A_X}(u,N^{-1})}{\widehat{\partial_{\log s}A_X}(u,N^{-1})}-\frac{\overline{\partial_u A_X}(u,N^{-1})}{\overline{\partial_{\log s}A_X}(u,N^{-1})}\right|\right.$$
$$\left.\le\frac{2\log(N\Delta)}{\Delta\sqrt{N\Delta}-\log(N\Delta)}\left|\frac{\overline{\partial_u A_X}(u,N^{-1})}{\overline{\partial_{\log s}A_X}(u,N^{-1})}\right|\right\}\ge 1-\varepsilon_1-\varepsilon_2.$$

Because of the averaged transport equation (87),

$$(\log\theta')'(u)=O(\Delta)+\frac{\overline{\partial_u A_X}(u,N^{-1})}{\overline{\partial_{\log s}A_X}(u,N^{-1})}\bigl(1+O(N^{-1})\bigr).$$

Since $\Delta>N^{-1}$ and $(\log\theta')'(u)$ is bounded, we derive

$$(\log\theta')'(u)=O(\Delta)+\frac{\overline{\partial_u A_X}(u,N^{-1})}{\overline{\partial_{\log s}A_X}(u,N^{-1})}.$$

Therefore,

$$\operatorname{Prob}\left(\left|\frac{\widehat{\partial_u A_X}(u,N^{-1})}{\widehat{\partial_{\log s}A_X}(u,N^{-1})}-(\log\theta')'(u)\right|\right.$$
$$\left.\le\frac{2\log(N\Delta)}{\Delta\sqrt{N\Delta}-\log(N\Delta)}|(\log\theta')'(u)|+O(\Delta)\right)\ge 1-\varepsilon_1-\varepsilon_2.$$

We choose Δ such that $\Delta^{-1}(N\Delta)^{-1/2}=\Delta$, that is, $\Delta=N^{-1/5}$. When $N\to\infty$, ε_1 and ε_2, whose expressions are given in Lemma C.2, both tend to 0. Moreover, for N large enough,

$$|(\log\theta')'(u)|\frac{2\log(N\Delta)}{\Delta\sqrt{N\Delta}-\log(N\Delta)}+O(\Delta)\le 2(\log N)N^{-1/5}.$$

Therefore,

$$\lim_{N\to\infty}\operatorname{Prob}\left(\left|\frac{\widehat{\partial_u A_X}(u,N^{-1})}{\widehat{\partial_{\log s}A_X}(u,N^{-1})}-(\log\theta')'(u)\right|\le 2(\log N)N^{-1/5}\right)=1. \qquad \square$$

PROOF OF THEOREM 4.2.

LEMMA C.4. *Under the hypotheses of Theorem* 3.2,

$$\begin{aligned}\partial_u A_X^\sigma(u,\xi_0/\sigma) - \theta''(u)\,\partial_\xi A_X^\sigma(u,\xi_0/\sigma)\\ = \sigma^2\big(C(u)+o(1)\big)\,\partial_u A_X^\sigma(u,\xi_0/\sigma),\end{aligned} \tag{90}$$

where C is continuous. □

PROOF. The proof mimicks the proof of Lemma C.1. In the proof of Theorem 3.2, we showed in (78) that

$$\operatorname{Re}\big\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v+\partial_x)\overline{D}\psi_{v,\xi,\sigma}\big\rangle = \sigma^{3+h}\big(A(v)+o(1)\big) \tag{91}$$

and in (80) that

$$\operatorname{Re}\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle = \sigma^{1+h}\big(B(v)+o(1)\big), \tag{92}$$

with $B(v)$ continuous. Comparing (91) and (92) shows that

$$\begin{aligned}\operatorname{Re}\big\langle K_X\psi_{v,\xi,\sigma}, \overline{D^{-1}}(\partial_v+\partial_x)\overline{D}\psi_{v,\xi,\sigma}\big\rangle\\ = \sigma^2\big(C(v)+o(1)\big)\operatorname{Re}\langle K_X\psi_{v,\xi,\sigma}, \partial_x\psi_{v,\xi,\sigma}\rangle,\end{aligned}$$

with $C(v)$ continuous. This implies, by repeating the argument of Lemma C.1, that (90) is satisfied. □

Using Lemma C.4, it is easy to see that

$$\begin{aligned}\overline{\partial_u A_X^\sigma}(u,\xi_0/\sigma) - \overline{\partial_\xi A_X^\sigma}(u,\xi_0/\sigma)\\ = O(\sigma^2)\,\overline{\partial_u A_X^\sigma}(u,\xi_0/\sigma) + O(\Delta)\,\overline{\partial_\xi A_X^\sigma}(u,\xi_0/\sigma).\end{aligned} \tag{93}$$

As in the proof of Theorem 4.1, one can combine the following lemma with Lemma C.3 to prove the weak consistency result (50).

LEMMA C.5. *Let X, Y and ψ satisfy the hypotheses of Theorem* 4.2. *Then, for each u, for N large enough*,

$$\operatorname{Prob}\Big\{\Big|\widehat{\partial_\xi A_X^\sigma}(u,N\xi_0) - \overline{\partial_\xi A_X^\sigma}(u,N\xi_0)\Big| \ge C\big|\overline{\partial_\xi A_X^\sigma}(u,N\xi_0)\big|\Big\} \le \varepsilon_1,$$

$$\operatorname{Prob}\Big\{\Big|\widehat{\partial_u A_X^\sigma}(u,N\xi_0) - \overline{\partial_u A_X^\sigma}(u,N\xi_0)\Big| \ge C\big|\overline{\partial_u A_X^\sigma}(u,N\xi_0)\big|\Big\} \le \varepsilon_2,$$

where ε_1 and ε_2 are defined in Lemma C.2.

The proof of Lemma C.5 is almost identical to that of Lemma C.2; the only difference is that ψ^2 has $p-1$ vanishing moments instead of p, so that Lemma C.7 must be replaced with the following lemma, which is proved by using the same method.

LEMMA C.6. *Let $Y(x) = X(x)e^{i\theta(x)}$, let $W_k = \langle X, \psi^1_{k/N,\sigma}\rangle$ and let $Z_k = \langle X, \psi^2_{k/N,\sigma}\rangle^*$. Under the hypotheses of Lemma* C.5, *for σ small enough, there exist two continuous functions M_1 and M_2 such that, for $|k-l| \le 2$,*

$$|\mathbb{E}\{W_k W_l^*\}| \le M_1(\sigma k)\sigma^h,$$
$$|\mathbb{E}\{W_k Z_l^*\}| \le M_1(\sigma k)\sigma^h,$$
$$|\mathbb{E}\{Z_k Z_l^*\}| \le M_1(\sigma k)\sigma^h,$$

and, for $|k-l| > 2$,

$$|\mathbb{E}\{W_k W_l^*\}| \le M_2(\sigma k)\frac{\sigma^{2p}}{(\sigma(|k-l|-2))^{2p-h}},$$
$$|\mathbb{E}\{W_k Z_l^*\}| \le M_2(\sigma k)\frac{\sigma^{2p-1}}{(\sigma(|k-l|-2))^{2p-1-h}},$$
$$|\mathbb{E}\{Z_k Z_l^*\}| \le M_2(\sigma k)\frac{\sigma^{2p-2}}{(\sigma(|k-l|-2))^{2p-2-h}}.$$

Since $p \ge \lceil h \rceil + 3$, we have $2(2p-2-h) > 1$. Therefore, the variance term

$$\mathbb{E}\left\{\left|\widehat{\partial_\xi A_X^\sigma}(u,\xi) - \overline{\partial_\xi A_X^\sigma}(u,\xi)\right|^2\right\}$$

can be controlled as in the proof of Lemma C.2.

PROOF OF LEMMA C.1. In one dimension, the proof of Proposition 3.1 can be adapted to show that, if (16) is replaced by

$$\operatorname{Re}\langle K_X\psi_{v,\tilde\beta,\sigma}, \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\psi_{v,\tilde\beta,\sigma}\rangle = c(u,\sigma)\operatorname{Re}\langle K_X\psi_{v,\tilde\beta,\sigma}, \partial_x\psi_{v,\tilde\beta,\sigma}\rangle \tag{94}$$

and if (17) holds, then the resulting transport equation (19) is replaced by

$$\partial_u A_X^\sigma(u,\alpha) + \partial_t\big(\alpha * \gamma^{-1}(u) * \gamma(t)\big)\,\partial_\alpha A_X^\sigma(u,\alpha) = c(u,\sigma)\,\partial_u A_X^\sigma(u,\alpha).$$

Recall that, in the proof of Theorem 3.1, (67) proves that

$$\operatorname{Re}\langle K_X\varphi_{v,s}, \phi_{v,s}\rangle = s^{h+1}\big(B(u) + o(1)\big), \tag{95}$$

where B is continuous. On the other hand, (69) proves that

$$\begin{aligned}&\operatorname{Re}\langle K_X\varphi_{v,s}, \partial_x\varphi_{v,s}\rangle\\ &\quad= \frac{1}{2}\frac{\theta''(u)}{\theta'(u)}s^h\big(1+o(1)\big)h\eta(0)\iint |x-y|^h\psi^*(x)\psi(y)\,dx\,dy,\end{aligned} \tag{96}$$

where $\theta''(u)/\theta'(u)$ is continuous in u.

Comparing (95) and (96) and recalling that

$$\varphi_{v,s} = \psi_{v,\tilde{\beta},\sigma},$$
$$\phi_{v,s} = \overline{D^{-1}}(\partial_v + \partial_x)\overline{D}\psi_{v,\tilde{\beta},\sigma},$$

we see that (94) holds, with

$$c(u,\sigma) = s\bigl(C(u) + o(1)\bigr)$$

and C continuous. This proves that (85) is indeed satisfied. $\square$

PROOF OF LEMMA C.2. We start by proving (88). Let $n = N\Delta$ and let us choose $u = 0$ without loss of generality. We seek an upper bound for the variance of $\widehat{\partial_{\log s} A_X}(0,s)$,

$$V_{\log s} = \mathbb{E}\Bigl\{\Bigl|\widehat{\partial_{\log s} A_X}(0,s) - \overline{\partial_{\log s} A_X}(0,s)\Bigr|^2\Bigr\}.$$

One can see that

$$\left|\frac{\partial^2}{\partial u\,\partial \log s} A_X(u,s)\right| = O(s^h),$$

and a Riemann series approximation shows that

$$\int g(v)\,\partial_{\log s} A_X(v,s)\,dv - N^{-1}\sum_{k=-n}^{n} g\Bigl(\frac{k}{N}\Bigr)\,\partial_{\log s} A_X\Bigl(\frac{k}{N,s}\Bigr) = O\Bigl(\frac{s^h}{N}\Bigr).$$

Replacing $\widehat{\partial_{\log s} A_X}(0,s)$ by its expression (41) and noticing that the real part is smaller than the modulus, we obtain

$$V_{\log s} \le \frac{4}{N^2}\mathbb{E}\Biggl\{\Biggl|\sum_{|k|\le n} g_k W_k Z_k - g_k \mathbb{E}\{W_k Z_k\}\Biggr|^2\Biggr\} + O\Bigl(\frac{s^{2h}}{N^2}\Bigr),$$

where g_k, W_k and Z_k, respectively, denote $g(k/n)$, $\langle X, \psi_{k/N,s}\rangle$ and $\langle X, \partial_{\log s}\psi_{k/N,s}\rangle^*$. Expanding $|\sum_{|k|\le n}|^2$ into $(\sum_{|k|\le n})\cdot(\sum_{|l|\le n})^*$,

$$\begin{aligned}
V_{\log s} &\le \frac{4}{N^2}\mathbb{E}\Biggl\{\sum_{|k|\le n}[g_k W_k Z_k - g_k\mathbb{E}\{W_k Z_k\}]\sum_{|l|\le n}[g_l W_l Z_l - g_l \mathbb{E}\{W_l Z_l\}]^*\Biggr\}\\
&\quad + O\Bigl(\frac{s^{2h}}{N^2}\Bigr)\\
&\le \frac{4}{N^2}\sum_{|k|\le n, |l|\le n}\bigl[g_k g_l \mathbb{E}\{W_k Z_k W_l^* Z_l^*\} - g_k g_l \mathbb{E}\{W_k Z_k\}\mathbb{E}\{W_l^* Z_l^*\}\bigr]\\
&\quad + O\Bigl(\frac{s^{2h}}{N^2}\Bigr).
\end{aligned}$$

Since Y is Gaussian, so is X, as well as the random variables W_k and Z_k. Hence,

$$\mathbb{E}\{W_k Z_k W_l^* Z_l^*\}$$
$$= \mathbb{E}\{W_k Z_k\}\mathbb{E}\{W_l^* Z_l^*\} + \mathbb{E}\{W_k W_l^*\}\mathbb{E}\{Z_k Z_l^*\} + \mathbb{E}\{W_k Z_l^*\}\mathbb{E}\{Z_k W_l^*\}.$$

Therefore,

$$\begin{aligned} V_{\log s} &\le \frac{4}{N^2} \sum_{|k|\le n, |l|\le n} g_k g_l \big[\mathbb{E}\{W_k W_l^*\}\mathbb{E}\{Z_k Z_l^*\} + \mathbb{E}\{W_k Z_l^*\}\mathbb{E}\{Z_k W_l^*\}\big] \\ &\quad + O\biggl(\frac{s^{2h}}{N^2}\biggr) \\ &\le \frac{4}{n^2} \sum_{|k|\le n, |l|\le n} \big[|\mathbb{E}\{W_k W_l^*\}|\,|\mathbb{E}\{Z_k Z_l^*\}| + |\mathbb{E}\{W_k Z_l^*\}|\,|\mathbb{E}\{Z_k W_l^*\}|\big] \\ &\quad + O\biggl(\frac{s^{2h}}{N^2}\biggr). \end{aligned} \tag{97}$$

Each of the terms appearing in the sum above is now bounded due to the following decorrelation lemma.

LEMMA C.7. *Let $X(x) = Y(\theta(x))$, let $W_k = \langle X, \psi_{k/N,s}\rangle$ and let $Z_k = \langle X, \partial_{\log s}\psi_{k/N,s}\rangle^*$. Under the hypotheses of Theorem* 4.1, *for s small enough*, *there exist continuous functions M_1 and M_2 such that*, *for $|k-l| \le 2$*,

$$|\mathbb{E}\{W_k W_l^*\}| \le M_1(sk)s^h, \tag{98a}$$

$$|\mathbb{E}\{W_k Z_l^*\}| \le M_1(sk)s^h, \tag{98b}$$

$$|\mathbb{E}\{Z_k Z_l^*\}| \le M_1(sk)s^h, \tag{98c}$$

and, *for $|k-l| > 2$*,

$$|\mathbb{E}\{W_k W_l^*\}| \le M_2(sk)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}, \tag{99a}$$

$$|\mathbb{E}\{W_k Z_l^*\}| \le M_2(sk)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}, \tag{99b}$$

$$|\mathbb{E}\{Z_k Z_l^*\}| \le M_2(sk)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}}. \tag{99c}$$

The proof of the above lemma is given later.

Replacing (98) and (99) in (97), we see that, since M_1 and M_2 are continuous and since $k/N = \Delta \to 0$ when $N \to \infty$,

$$\begin{aligned}
V_{\log s} \leq{}& \frac{4}{n^2} \sum_{|k-l|\leq 2, |k|, |l| \leq n} 2\big(M_1(0)^2 + o(1)\big) s^{2h} \\
&+ \frac{4}{n^2} \sum_{|k-l|>2, |k|, |l| \leq n} \frac{2(M_2(0)^2 + o(1))\, s^{4p}}{(s(|k-l|-2))^{4p-2h}} + O\left(\frac{s^{2h}}{N^2}\right).
\end{aligned} \tag{100}$$

Since $4p - 2h > 1$,

$$\sum_{|k-l|>2, |k|, |l| \leq n} (|k-l|-2)^{2h-4p} = K_p n. \tag{101}$$

Replacing (101) in (100), we obtain

$$V_{\log s} \leq 8C^2 \frac{s^{2h}}{n} \big(3\, M_1(0)^2 + K_p M_2(0)^2\big) + o\left(\frac{s^{2h}}{n}\right). \tag{102}$$

In the proof of Theorem 3.1, (69) proves that there exists $a(u) > 0$ such that

$$|\partial_u A_X(u, s)| \geq a(u) s^h + o(s^h).$$

For Δ small enough, $\partial_u A_X(v, s)$ does not change sign for $|v| \leq \Delta$. Thus, after convolution with g,

$$|\overline{\partial_u A_X}(0, s)| \geq A s^h + o(s^h).$$

Because of (87), the same applies to $\overline{\partial_{\log s} A_X}(0, s)$. Therefore, there exists a constant C_1 such that

$$V_{\log s} \leq C_1 \left[\frac{|\overline{\partial_{\log s} A_X}(0, s)|}{\sqrt{n}}\right]^2.$$

Applying Chebyshev's lemma [4] then proves that, for all $\varepsilon > 0$,

$$\operatorname{Prob}\left\{\left|\widehat{\partial_{\log s} A_X}(0, s) - \overline{\partial_{\log s} A_X}(0, s)\right| \geq \frac{\sqrt{C_1}\,|\overline{\partial_{\log s} A_X}(0, s)|}{\varepsilon\sqrt{n}}\right\} \leq \varepsilon^2,$$

and (88) follows by choosing

$$\varepsilon = \frac{\sqrt{C_1}\,\Delta}{\log n} \quad \text{and} \quad \varepsilon_1 = \frac{C_1 \Delta^2}{(\log n)^2}.$$

Let us now prove (89). We denote $D_u = |\widehat{\partial_u A_X}(0, s) - \overline{\partial_u A_X}(0, s)|$. Recall that

$$\overline{\partial_u A_X}(0, s) = \int_{-\Delta}^{\Delta} g(-v)\, \partial_u A_X(v, s)\, ds.$$

An integration by parts shows that

$$\overline{\partial_u A_X}(0, s) = \frac{1}{\Delta^2}\left[\int_0^{\Delta} A_X(v, s)\, dv - \int_{-\Delta}^{0} A_X(v, s)\, dv\right]$$

and since $|\partial_u A_X(u,s)| = O(s^h)$, with a Riemann series approximation,

$$\Delta^{-2}\int_0^{\Delta} A_X(v,s)\,dv - \Delta^{-1}n^{-1}\sum_{k=1}^{n} A_X\left(\frac{k}{N,s}\right) = O\left(\frac{s^h}{n}\right).$$

On the other hand, an integration by parts also shows that $\widehat{\partial_u A_X}(0,s)$, defined in (40), satisfies

$$\widehat{\partial_u A_X}(0,s) = \frac{1}{N\Delta^2}\sum_{0<k/N-u\le\Delta} |\langle X, \psi_{k/N,s}\rangle|^2 - \frac{1}{N\Delta^2}\sum_{-\Delta\le k/N-u\le 0} |\langle X, \psi_{k/N,s}\rangle|^2 + O\left(\frac{s^h}{N}\right).$$

Therefore, using once again the notation $W_k = \langle X, \psi_{k/N,s}\rangle$,

$$D_u = \frac{1}{n\Delta}\left|\sum_{k=1}^{n}(|W_k|^2 - \mathbb{E}\{|W_k|^2\}) - \sum_{k=-n+1}^{0}(|W_k|^2 - \mathbb{E}\{|W_k|^2\})\right| + O\left(\frac{s^h}{n}\right).$$

Denoting $\widetilde{W}^- = \sum_{k=-n+1}^{0}|W_k|^2$ and $\widetilde{W}^+ = \sum_{k=1}^{n}|W_k|^2$, we have

$$D_u \le \frac{1}{n\Delta}\left(|\widetilde{W}^+ - \mathbb{E}\{\widetilde{W}^+\}| + |\widetilde{W}^- \mathbb{E}\{\widetilde{W}^-\}|\right) + O\left(\frac{s^h}{n}\right). \tag{103}$$

We are now going to prove that there exists a strictly positive constant C_2 such that

$$\forall y,\qquad \mathrm{Prob}\left\{D_u > yC_2\frac{s^h}{\Delta\sqrt{n}}\right\} \le 6e^{-y/2} \tag{104}$$

and since $|\overline{\partial_u A_X}(0,s)| \ge A\,s^h + o(s^h)$ with $A>0$, choosing $y = \log n/C_2$ will then imply (89).

Let us consider the random vector $W = (W_1, W_2, \dots, W_n)$, let K_W denote the covariance operator of W and let $(e_j)_{j=1,\dots,n}$ be its Karhunen–Loève basis. If $(\alpha_j)_{j=1,\dots,n}$ are the eigenvalues of K_W corresponding to the eigenvectors $(e_j)_{j=1,\dots,n}$, then

$$W = \sum_{j=1}^{n}\sqrt{\alpha_j}\,\bar{W}_j e_j,$$

where $\bar{W}_j$ are independent random variables with variance 1. As a consequence,

$$\widetilde{W}^+ = \|W\|^2 = \sum_{j=1}^{n}\alpha_j\bar{W}_j^2.$$

The following lemma, which is proved in [8], relies on a theorem by Bakirov [3].

LEMMA C.8. *If $\widehat{W} = \sum_j \beta_j \bar{W}_j^2$, where $\bar{W}_j$ are independent Gaussian random variables with variance* 1, *and $\sum_j \beta_j^2 = 1$, then*

$$\forall y, \qquad \operatorname{Prob}\{|\widehat{W} - \mathbb{E}\{\widehat{W}\}| > y\} \le 6e^{-y/2}.$$

The random variable $\widehat{W}^+ = (\sum_j \alpha_j^2)^{-1/2} \widetilde{W}^+$ satisfies the requirements of Lemma C.8. Therefore,

$$\forall y, \qquad \operatorname{Prob}\left\{|\widetilde{W}^+ - \mathbb{E}\{\widetilde{W}^+\}| > y\left(\sum_j \alpha_j^2\right)^{1/2}\right\} \le 6e^{-y/2}$$

but $\sum_j \alpha_j^2$ is equal to the Hilbert–Schmidt norm of K_W,

$$\sum_j \alpha_j^2 = \sum_{j,k} \mathbb{E}\{W_j W_k^*\},$$

which is bounded by $Bs^{2h}n$ because of (98a) and (99a). Hence,

$$\forall y, \qquad \operatorname{Prob}\{|\widetilde{W}^+ - \mathbb{E}\{\widetilde{W}^+\}| > y\sqrt{B}s^h\sqrt{n}\} \le 6e^{-y/2}.$$

The same applies to $\widetilde{W}^-$, and by combining the two and using (103), we obtain (104).

PROOF OF LEMMA C.7. The three terms $\mathbb{E}\{W_k W_l^*\}$, $\mathbb{E}\{W_k Z_l^*\}$ and $\mathbb{E}\{Z_k Z_l^*\}$ can be written as

$$I = \iint c_Y\big(\theta(u+sx) - \theta(v+sy)\big)\psi(x)\tilde{\psi}(y)\,dx\,dy,$$

where $(u, v) = (sk, sl)$ and ψ and $\tilde{\psi}$ are two wavelets with p vanishing moments. Clearly,

$$I = \iint \big[c_Y\big(\theta(u+sx) - \theta(v+sy)\big) - c_Y(0)\big]\psi(x)\tilde{\psi}(y)\,dx\,dy. \tag{105}$$

For $|u - v| \le \Delta$, $|x| \le 1$ and $|y| \le 1$, we have

$$|\theta(u+sx) - \theta(v+sy)| \le (\Delta + 2s) \sup_{|x-u| \le \Delta+2s} [\theta'(x)] \le (\Delta + 2s)C_u,$$

because θ is continuously differentiable. For Δ small enough, $|\theta(u+sx) - \theta(v+sy)|$ is therefore in a neighborhood of 0. Since η is assumed continuous in a neighborhood of 0,

$$\big|\eta\big(\theta(u+sx) - \theta(v+sy)\big)\big| \le B$$

for $|u - v| \le \Delta$, $|x| \le 1$ and $|y| \le 1$.

Hence,

$$|I| \leq \iint |\theta(u+sx) - \theta(v+sy)|^h B|\psi(x)|\,|\tilde{\psi}(y)|\,dx\,dy$$
$$\leq C(u)s^h,$$

where $C(u)$ is continuous. This proves (98a), (98b) and (98c).

Let us now prove (99). Since ψ and $\tilde{\psi}$ in (105) are compactly supported and have p vanishing moments, there exist two compactly supported functions β and $\tilde{\beta}$ such that $\psi(x) = \beta^{(p)}(x)$ and $\tilde{\psi}(y) = \tilde{\beta}^{(p)}(y)$. Integrating (105) by parts p times with respect to x and to y gives

$$I = \iint \frac{\partial^p}{\partial x^p}\frac{\partial^p}{\partial y^p}\{|\theta(u+sx) - \theta(v+sy)|^h \eta(\theta(u+sx) - \theta(v+sy))\}$$
$$\times \beta(x)\tilde{\beta}(y)\,dx\,dy.$$

For $|u - v| > 2s$, however, one can show that

$$\left|\frac{\partial^p}{\partial x^p}\frac{\partial^p}{\partial y^p}\{|\theta(u+sx) - \theta(v+sy)|^h \eta(\theta(u+sx) - \theta(v+sy))\}\right|$$
$$\leq \frac{M(u)\,s^{2p}}{(|u-v| - 2s)^{2p-h}},$$

where $M(u)$ depends on h, on derivatives of θ up to order $2p$ in a neighborhood of u and on derivatives of η up to order $2p$ in a neighborhood of 0. Therefore, there exists a continuous $M_2(u)$ such that

$$|I| \leq M_2(u)\frac{s^{2p}}{(s(|k-l|-2))^{2p-h}},$$

which proves (99a), (99b) and (99c). □

REFERENCES

[1] ABRY, P., GONÇALVES, P. and FLANDRIN, P. (1995). Wavelets, spectrum analysis and $1/f$ processes. *Wavelets and Statistics. Lecture Notes in Statist.* **103** 15–29. Springer, New York.

[2] ADLER, R. J. (1981). *The Geometry of Random Fields.* Wiley, New York.

[3] BAKIROV, N. K. (1989). Extrema of the distributions of quadratic forms of Gaussian variables. *Theory Probab. Appl.* **34** 207–215.

[4] BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd ed. Wiley, New York.

[5] CLERC, M. and MALLAT, S. (1999). Shape from texture through deformations. In *Proc. Seventh International Conference on Computer Vision* **1** 405–410.

[6] CLERC, M. and MALLAT, S. (2002). The texture gradient equation for recovering shape from texture. *IEEE Trans. Pattern Anal. Machine Intelligence* **24** 536–549.

[7] FLANDRIN, P. (1989). On the spectrum of fractional Brownian motions. *IEEE Trans. Inform. Theory* **35** 197–199.

[8] GAO, H.-Y. (1993). Wavelet estimation of spectral densities in times series analysis. Ph.D. dissertation, Dept. Statistics, Univ. California, Berkeley.

[9] GÅRDING, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *J. Math. Imaging Vision* **2** 327–350.

[10] MALIK, J. and ROSENHOLTZ, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *Internat. J. Comput. Vision* **23** 149–168.

[11] MALLAT, S. (1999). *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, New York.

[12] PATIL, G. P. and RAO, C. R. (1994). *Handbook of Statistics* **12**. *Environmental Statistics*. 661–689. North-Holland, Amsterdam.

[13] PERRIN, O. and SENOUSSI, R. (1999). Reducing non-stationary stochastic processes to stationarity by a time deformation. *Statist. Probab. Lett.* **43** 393–397.

[14] PRIESTLEY, M. B. (1965). Evolutionary spectra and non-stationary processes (with discussion). *J. Roy. Statist. Soc. Ser. B* **27** 204–237.

[15] SAMPSON, P. D. and GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Statist. Assoc.* **87** 108–119.

[16] YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions* **1**. Springer, Berlin.

CERMICS
ECOLE NATIONALE DES PONTS ET CHAUSSÉES
INRIA
06902 SOPHIA-ANTIPOLIS
FRANCE
E-MAIL: Maureen.Clerc@cermics.enpc.fr

CMAP
ECOLE POLYTECHNIQUE
PALAISEAU
FRANCE
AND
COURANT INSTITUTE OF
MATHEMATICAL SCIENCES
NEW YORK UNIVERSITY
NEW YORK, NEW YORK 10012
E-MAIL: Stephane.Mallat@polytechnique.fr