Polyharmonic Smoothing Splines and the Multidimensional Wiener Filtering of Fractal-Like Signals

Shai Tirosh, *Student Member, IEEE*, Dimitri Van De Ville, *Member, IEEE*, and Michael Unser, *Fellow, IEEE*

Abstract—Motivated by the fractal-like behavior of natural images, we develop a smoothing technique that uses a regularization functional which is a fractional iterate of the Laplacian. This type of functional was initially introduced by Duchon for the approximation of nonuniformly sampled, multidimensional data. He proved that the general solution is a smoothing spline that is represented by a linear combination of radial basis functions (RBFs). Unfortunately, this is tedious to implement for images because of the poor conditioning of RBFs and their lack of decay. Here, we present a much more efficient method for the special case of a uniform grid. The key idea is to express Duchon's solution in a fractional polyharmonic B-spline basis that spans the same space as the RBFs. This allows us to derive an algorithm where the smoothing is performed by filtering in the Fourier domain. Next, we prove that the above smoothing spline can be optimally tuned to provide the MMSE estimation of a fractional Brownian field corrupted by white noise. This is a strong result that not only yields the best linear filter (Wiener solution), but also the optimal interpolation space, which is not bandlimited. It also suggests a way of using the noisy data to identify the optimal parameters (order of the spline and smoothing strength), which yields a fully automatic smoothing procedure. We evaluate the performance of our algorithm by comparing it against an oracle Wiener filter, which requires the knowledge of the true noiseless power spectrum of the signal. We find that our approach performs almost as well as the oracle solution over a wide range of conditions.

Index Terms—Fractal image model, multidimensional Wiener filter, optimal multidimensional linear filtering, polyharmonic smoothing splines.

I. INTRODUCTION

DENOISING algorithms are part of many image processing systems. The methods available are quite diverse and range from simple linear filtering (image smoothing) [1], to more complex nonlinear schemes such as total variation image regularization [2] or anisotropic diffusion [3], which are iterative processes. There are also many intermediate approaches such as wavelet denoising [4] that are nonlinear but have a direct implementation.

From a conceptual point of view, it is always better when an algorithm can be justified based on some optimality principle. In this respect, the three dominant paradigms are as follows:

1) the Wiener formulation, where one minimizes the mean square estimation error (MMSE) for a given class of stochastic processes [5], [6];
2) the Bayesian framework, where one searches for the maximum *a posteriori* solution given some prior knowledge of the statistical distribution of its parameters [7], [8];
3) the variational approach, where one minimizes an energy functional that favors solutions with some desirable features (eg., nonoscillating, and/or with sharp edges) [9], [10]; formulations have been proposed for both discrete [11] and continuously defined signals [12], [13].

The Wiener formulation typically produces a linear solution (Wiener filter) which is simple to implement; however, it requires complete knowledge of the second-order statistics of the signal which is often not available. The Bayesian and variational formulations are more versatile and, therefore, often favored by researchers and practitioners [14], [15]. When the underlying distributions (respectively, energy functionals) are non-Gaussian (respectively, nonquadratic), they yield nonlinear estimators that are typically implemented iteratively. Note that except for the context, which is either statistical or deterministic, both types of formulations are essentially equivalent: An energy function can be converted into a probability density by defining an appropriate Gibbs distribution, while the reverse equivalence is achieved by considering the log-likelihood function of the data. The most basic variational algorithm is the Tikhonov filter (or regularized least-squares estimator) which constrains the solution to be smooth via a quadratic regularization functional (e.g., energy of the gradient or Laplacian) [1], [9]. This approach has been extended by considering a variety of nonquadratic regularization terms; the most prominent are total variation [2], half-quadratic regularization [16], as well as penalized maximum-likelihood and Bayesian estimation under the generalized Gaussian assumption [17]–[19], which also fit the framework. Interestingly, there is also a variational formulation of wavelet denoising that yields a most effective algorithm which amounts to applying a suitable nonlinearity to the wavelet coefficients of the signal [20].

The approach that is proposed in this paper also falls within the variational framework. However, it departs from the traditional formulation in that we seek a function $s(\boldsymbol{x})$ of the continuous space variable $\boldsymbol{x} \in \mathbb{R}^d$ that best fits the noisy samples of an image (or multidimensional signal). This means that we are not only attempting to recover the noise-free pixel values but also an optimal interpolation model that can provide an image prediction (or MMSE estimate) at any arbitrary location $\boldsymbol{x}_0 \in \mathbb{R}^d$. Since the present continuous-space estimation problem is more involved mathematically than the classical one where the input

and output signals are both discrete, it is justifiable in a first stage to concentrate on the case of a quadratic cost function. Our choice of regularization functional, however, is not arbitrary. First, we justify it on deterministic grounds: We want a criterion that is universal in the sense of being invariant with respect to the viewing geometry; in particular, translations, rotations, and scaling. This narrows it down to the energy of a fractional[1] iterate of the Laplacian whose transfer function is $\|\boldsymbol{\omega}\|^s$ [22], [23]. Second, we want to take advantage of the fractal-like behavior of many natural images [24]. Indeed, several researchers have observed that the power spectrum of images often follows an inverse power-law distribution, $1/\|\boldsymbol{\omega}\|^{2H+d}$, which is characteristic of fractal processes [25], [26]. Now, if we recall the classical equivalence between Wiener and regularization filters in the case of additive white noise, this observation suggests using a regularization operator that whitens the signal; this points, once again, towards a fractional derivative. Of course, this argument needs to be made more rigorous because self-similar processes are not stationary, meaning that their Wiener filter cannot be defined in the conventional way. The processes can still be estimated but one has to resort to alternative techniques (in particular, kriging [27]) that were developed in geostatistics for the reconstruction of functions from their nonuniform samples. Interestingly, these methods happen to be closely related to splines and radial basis functions [28], [29].

The mathematical foundation for solving our variational estimation problem is provided by Duchon's general theory of (m, s)-splines [22]. Duchon considered the most general situation where the location of the measurements is arbitrary (i.e., nonuniform samples) and proved that the solution could be expressed as a weighted sum of radial basis functions (RBF). He also showed how to determine the optimal weights through the solution of a linear system of equations. However, the RBFs are poorly conditioned, thus making this approach very delicate to implement when there are many data points. In addition, the basic method (thin-plate splines [30]) is computationally extremely expensive[2] and, therefore, out of reach for image processing applications.

Our contribution in this paper is three-fold. Our first result is a fast computational solution to Duchon's smoothing spline problem that takes advantage of the fact that we are dealing with data on a uniform grid. The key idea is to express Duchon's solution in a fractional polyharmonic B-spline basis. These functions, which were first introduced by Rabut [31], are localized versions of RBFs and, thus, span the same space when the grid is uniform. The second is a proof of the optimality of polyharmonic smoothing splines for the estimation of fractional Brownian fields (which are multidimensional fractal stochastic processes). The third contribution is a practical denoising algorithm that estimates the optimal spline parameters from the noisy data, and the demonstration of its suitability for image processing.

The paper is organized as follows. In Section II, we briefly review the essential properties of polyharmonic B-splines. In Section III, we use these to derive the optimal solution of Duchon's

smoothing spline problem which enforces a penalty on the energy of the fractional derivative of the solution. The fact that we are working on a uniform grid yields a fast Fourier-based filtering algorithm which is applicable in any number of dimensions. In Section IV, we present an alternative stochastic derivation and show that the proposed smoothing estimator is optimal for the class of fractal-like signals. In particular, we show that the optimal order of the spline is directly related to the fractal dimension of the signal (or its Hurst exponent). We also get a closed form expression for the optimal regularization parameter. In Section IV, we briefly review fractal models of images and show how these can be identified in the frequency domain. In particular, we present a robust method that estimates the fractal parameters of our measurement model for the radially averaged periodiogram of the noisy image. The task is formulated as a nonlinear weighted least-squares fit using a cost function that takes into account the bias and variance inherent to sampling and averaging along radial annuli. Finally, in Section VI, we discuss some practical implementation issues and present experimental results for a wide range of signal-to-noise ratios. In particular, we compare our method against the oracle Wiener filter, which requires the knowledge of the true uncorrupted signal spectrum. The remarkable fact is that our method which is parameterized by two numbers only—the order of the spline and the regularization parameter λ—essentially keeps up (within less than 2 dB) with the oracle Wiener filter over the whole range of testing conditions. Since the oracle solution is not available in a practical denoising problem, our approach presents a viable way to get the best linear filtering solution.

II. MATHEMATICAL PRELIMINARIES

A. Notation

- Multidimensional functions: A continuously defined signal is a function of $\boldsymbol{x} \in \mathbb{R}^d$ and is usually denoted as $f(\boldsymbol{x})$; its Fourier transform is given by $\hat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) e^{-j\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle} d\boldsymbol{x}$.
- Multidimensional sequences: A discrete signal is a sequence indexed by $\boldsymbol{k} \in \mathbb{Z}^d$ and is typically denoted as $c[\boldsymbol{k}]$; its discrete Fourier transform is given by $C(e^{j\boldsymbol{\omega}}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} c[\boldsymbol{k}] e^{-j\langle \boldsymbol{\omega}, \boldsymbol{k} \rangle}$.
- The inner product for discrete domain sequences is defined as: $\langle c_1, c_2 \rangle_{\ell_2} = \sum_{\boldsymbol{l} \in \mathbb{Z}^d} c_1[\boldsymbol{l}] c_2^*[\boldsymbol{l}]$, and the corresponding ℓ_2-norm is: $\|c\|_{\ell_2} = \sqrt{\langle c, c \rangle_{\ell_2}}$.

B. Fractional Differentiation

Following Liouville's definition of a fractional derivative [32], the differentiation operator can be easily extended to noninteger exponents in the Fourier domain

$$\partial^s f(x) \longleftrightarrow (j\omega)^s \hat{f}(\omega) \qquad (1)$$

which is defined in the sense of distributions. Similarly, we can define a symmetric differentiation operator as

$$\partial_*^s f(x) \longleftrightarrow |\omega|^s \hat{f}(\omega). \qquad (2)$$

---

[1] The present meaning of fractional is real-valued (as opposed to integer only) in accordance with the terminology used in the fractal literature [21].

[2] The complexity of the thin-plate spline method is $O(N^3)$ for computing the optimal weights (N is the total number of data points). It then requires an additional $O(N)$ for each evaluation of the RBF model at a given location $\boldsymbol{x}_0$.
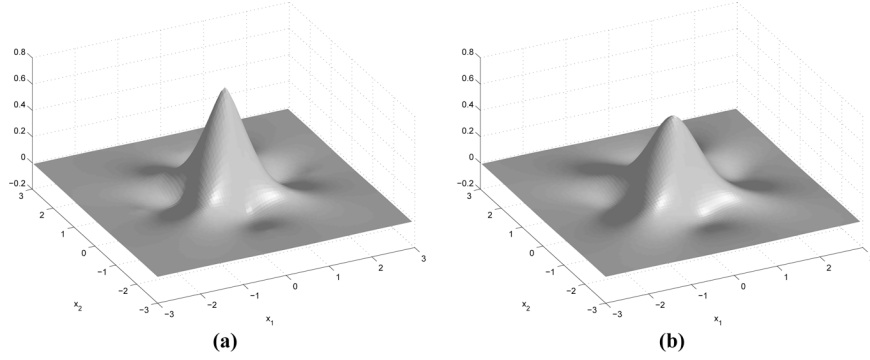
Fig. 1. Example of 2-D polyharmonic B-splines. Note that they are not compactly supported but decay like $O\left(\|\boldsymbol{x}\|^{-d-\min(2,\gamma)}\right)$ as $\|\boldsymbol{x}\| \to \infty$. These functions also show (slightly) negative sidelobes. (a) Order $\gamma = 1.5$. (b) Order $\gamma = 2.6$.

In order to extend fractional differentiation to multiple dimensions, it is convenient to consider the isotropic fractional derivative, which can be interpreted as the $s/2$-iterated Laplacian operator:

$$\partial_*^s f(\boldsymbol{x}) \longleftrightarrow \|\boldsymbol{\omega}\|^s \hat{f}(\boldsymbol{\omega}). \tag{3}$$

Its discrete counterpart is the fractional finite difference operator

$$\Delta_*^s f(\boldsymbol{x}) \longleftrightarrow 2^s \left(\sum_{i=1}^d \sin^2\left(\frac{\omega_i}{2}\right)\right)^{s/2} \hat{f}(\boldsymbol{\omega})$$
$$= \left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^s \hat{f}(\boldsymbol{\omega})$$

with the convention that $\sin(\boldsymbol{\omega}/2) = (\sin(\omega_1/2),\ldots,\sin(\omega_d/2))$. This operator can also be represented using the sequence d_s

$$\Delta_*^s f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} d_s[\boldsymbol{k}] f(\boldsymbol{x}-\boldsymbol{k}) \tag{4}$$

whose Fourier transform is $D_s(e^{j\boldsymbol{\omega}}) = \|2\sin(\boldsymbol{\omega}/2)\|^s$.

C. Polyharmonic B-Splines

The multidimensional basis functions that are of interest to us are the polyharmonic B-splines [31], [33], [34]. We denote the function of order γ, which corresponds to a spline degree $\alpha = \gamma - d$, as $\beta_\gamma(\boldsymbol{x})$, and specify its Fourier transform

$$\mathcal{F}\{\beta_\gamma\}(\boldsymbol{\omega}) = \hat{\beta}_\gamma(\boldsymbol{\omega}) = \frac{\left\|\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^\gamma}{\left\|\frac{\boldsymbol{\omega}}{2}\right\|^\gamma}, \text{ for } \gamma > \frac{d}{2} \tag{5}$$

with $\boldsymbol{\omega} = (\omega_1,\ldots,\omega_d) \in \mathbb{R}^d$. Note that the order of those functions can be fractional (see also [35] and [36]). In Fig. 1, we show two examples of two-dimensional (2-D) polyharmonic B-splines.

We denote the shift-invariant signal space generated by β_γ as

$$\mathcal{V}(\beta_\gamma) = \left\{\sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\beta_\gamma(\boldsymbol{x}-\boldsymbol{k}); c \in \ell_2(\mathbb{Z}^d)\right\}. \tag{6}$$

This means that any function $f(\boldsymbol{x})$ that lies in this spline space can be expressed in terms of a polyharmonic B-spline expansion

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} c[\boldsymbol{k}]\beta_\gamma(\boldsymbol{x}-\boldsymbol{k}). \tag{7}$$

A useful notion for our purpose is the sampled or discrete polyharmonic B-spline, given by

$$b_\gamma[\boldsymbol{k}] = \beta_\gamma(\boldsymbol{k}). \tag{8}$$

These functions naturally appear when we are calculating the spline representation at the integers

$$f(\boldsymbol{k}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} c[\boldsymbol{i}]\beta_\gamma(\boldsymbol{k}-\boldsymbol{i}) = \sum_{\boldsymbol{i}\in\mathbb{Z}^d} c[\boldsymbol{i}]b_\gamma[\boldsymbol{k}-\boldsymbol{i}] = (c*b_\gamma)[\boldsymbol{k}]. \tag{9}$$

The polyharmonic B-splines are designed so that they satisfy $\partial_*^\gamma \beta_\gamma(\boldsymbol{x}) = 0$, for $\boldsymbol{x} \in \mathbb{R}^d\backslash\mathbb{Z}^d$, where ∂_*^γ denotes the $\gamma/2$-iterated Laplacian operator of (3). These splines can be represented as linear combinations of shifted radial basis functions (RBFs) called thin plate splines (for integer orders)

$$\beta_\gamma(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} d_\gamma[\boldsymbol{k}]\rho_\gamma(\boldsymbol{x}-\boldsymbol{k}) \tag{10}$$

where these RBFs are, in fact, the Green functions (inverse filter) of the γth-order symmetric differentiation operator ∂_*^γ; they are given by

$$\rho_\gamma(\boldsymbol{x}) = \|\boldsymbol{x}\|^{\gamma-d}\left(c_{d,\gamma} + c'_{d,\gamma}\ln\|\boldsymbol{x}\|\right) \longleftrightarrow \hat{\rho}_\gamma(\boldsymbol{\omega}) = \frac{1}{\|\boldsymbol{\omega}\|^\gamma} \tag{11}$$

where $c_{d,\gamma}$ and $c'_{d,\gamma}$ are appropriate constants [37]

$$\begin{cases} c_{d,\gamma} = 0, c'_{d,\gamma} = \dfrac{(-1)^{(\gamma-d)/2+1}}{2^{\gamma-1}\sqrt{\pi^d}\Gamma\left(\frac{\gamma}{2}\right)\left(\frac{\gamma-d}{2}\right)!}, & (\gamma-d) \text{ even} \\ c'_{d,\gamma} = 0, c_{d,\gamma} = \dfrac{\Gamma\left(-\frac{\gamma-d}{2}\right)}{2^\gamma\sqrt{\pi^d}\Gamma\left(\frac{\gamma}{2}\right)}, & \text{otherwise.} \end{cases} \tag{12}$$

Here,[3] $\Gamma(x)$ is Euler's gamma function. In the case where $(\gamma - d)$ is not even, $c_{d,\gamma}$ can be also written as

$$c_{d,\gamma} = \frac{-\pi}{2^\gamma \sqrt{\pi^d} \sin\left(\frac{\pi(\gamma-d)}{2}\right) \Gamma\left(\frac{\gamma}{2}\right) \Gamma\left(\frac{\gamma-d}{2}+1\right)}. \quad (13)$$

This implies that the closure of the harmonic B-splines is spanning the same signal space as a set of RBFs positioned at the multiintegers [31]; that is, when the grid is uniform which is mostly the case in signal processing applications.

The polyharmonic B-splines have the following useful properties.

1) *Convolution property*: The polyharmonic B-splines satisfy the convolution relation $\beta_{\gamma_1} * \beta_{\gamma_2} = \beta_{\gamma_1+\gamma_2}$. This important property follows directly from their definition in the Fourier domain.

2) *Differentiation property*: The fractional derivative of a polyharmonic B-splines is a spline of reduced order that is simply given by

$$\partial_*^s \beta_\gamma(\boldsymbol{x}) = \Delta_*^s \beta_{\gamma-s}(\boldsymbol{x}). \quad (14)$$

This formula is readily established in the Fourier domain.

3) *Partition of unity*: The polyharmonic B-splines satisfy the partition of unity property, which guarantees that they reproduce the constant

$$\sum_{\boldsymbol{k} \in \mathbb{Z}^d} \beta_\gamma(\boldsymbol{x} - \boldsymbol{k}) = 1. \quad (15)$$

4) *Decay*: Unlike traditional B-splines of integer order, Rabut's polyharmonic B-splines are not compactly supported. However, they decay like $O(1/\|\boldsymbol{x}\|^{d+\min(2,\gamma)})$ when $\|\boldsymbol{x}\| \to \infty$. The proof is given in [31], [36].

5) *Order of approximation*: A polyharmonic B-spline, β_γ, has an order of approximation γ. This means that they can approximate a well-behaved function $f(\boldsymbol{x})$ to any required accuracy by projecting it into a rescaled spline space with step size h

$$\inf_{c[\boldsymbol{k}]} \left\| f(\boldsymbol{x}) - \sum_{\boldsymbol{k} \in \mathbb{Z}^d} c[\boldsymbol{k}] \beta_\gamma\left(\frac{\boldsymbol{x}}{h} - \boldsymbol{k}\right) \right\|_{\mathbf{L}_2} \leq \text{Const} \times h^\gamma \quad (16)$$

where the constant does not depend on h. It is also easy to see that

$$\hat{\beta}_\gamma(2\pi\boldsymbol{k} + \boldsymbol{\omega}) = O\left(\|\boldsymbol{\omega}\|^\gamma\right), \text{ for } \boldsymbol{\omega} \to \mathbf{0}, \boldsymbol{k} \in \mathbb{Z}^d \backslash \mathbf{0}. \quad (17)$$

6) *Riesz basis*: These B-splines also generate a Riesz basis, which is equivalent to the condition

$$\forall c \in \ell_2(\mathbb{Z}^d) : \exists C_0, C_1 \in (0, \infty) :$$

$$C_0 \|c\|_{\ell_2} \leq \left\| \sum_{\boldsymbol{k} \in \mathbb{Z}^d} c[\boldsymbol{k}] \beta_\gamma(\boldsymbol{x} - \boldsymbol{k}) \right\|_{\mathbf{L}_2} \leq C_1 \|c\|_{\ell_2}. \quad (18)$$

---

[3]Technically, the constants for the case where $(\gamma - d)$ is even depend on the interpretation given to the principal value of $1/\|\boldsymbol{\omega}\|^\gamma$. Here, we selected the form that does not add any Dirac distributions to its Fourier transform.

III. SMOOTHING SPLINE FORMULATION

We now present a deterministic algorithm for fitting noisy image data. This algorithm solves Duchon's smoothing problem. In the next section, we will prove that this solution also yields the optimal Wiener estimate for a multidimensional fractional Brownian motion process with Hurst exponent $0 < H < 1$.

The problem is to find a continuous-space function $f(\boldsymbol{x})$ that approximately fits a series of noisy measurements $\{g[\boldsymbol{x_i}] = f(\boldsymbol{x_i}) + n[\boldsymbol{x_i}]\}_{\boldsymbol{x_i} \in \text{S}}$, where $n[\boldsymbol{x_i}]$ is some additive noise; $\text{S} = \{\boldsymbol{x_i}\}$ is the set of sampling locations which can be essentially arbitrary. Duchon's smoothing formulation is variational with a Tikhonov-like regularization [11], [22]. The solution of the smoothing problem is specified as

$$f = \arg\min_{\tilde{f} \in \mathbf{L}_2} \left\{ \epsilon_s^2(\tilde{f}) \right\} \quad (19)$$

where

$$\epsilon_s^2(\tilde{f}) = \sum_{\boldsymbol{x_i} \in \text{S}} \left(g[\boldsymbol{x_i}] - \tilde{f}(\boldsymbol{x_i}) \right)^2 + \lambda \left\| \partial_*^s \tilde{f}(\boldsymbol{x}) \right\|_{\mathbf{L}_2}^2. \quad (20)$$

The first component of ϵ_s^2, denoted as the *data term*, quantifies the distance between our solution and the given measurements $g[\boldsymbol{x_i}]$ on S—the set of points at which we are observing the signal. The second *regularization term* penalizes the lack of smoothness of the solution. Here, $\lambda \geq 0$ is a regularization parameter, making a balance between the two terms: A higher λ means more smoothing and less fidelity to the data, and vice versa.

Duchon[4] has shown that the solution of (19) in the general case can be written as

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{x_i} \in \text{S}} a_{\boldsymbol{i}} \rho_{2s}(\boldsymbol{x} - \boldsymbol{x_i}) + p(\boldsymbol{x}) \quad (21)$$

where $\rho_{2s}(\boldsymbol{x})$ is the RBF specified by (11), $\{a_{\boldsymbol{i}}\}$ is a suitable set of coefficients, and $p(\boldsymbol{x})$ is a polynomial of degree $\lceil s \rceil - 1$. The main problem with this representation is the lack of decay of $\rho_{2s}(\boldsymbol{x})$ due to the singularity of its Fourier transform at $\boldsymbol{\omega} = \mathbf{0}$. Note that the only difference between $\hat{\rho}_{2s}(\boldsymbol{\omega}) = \|\boldsymbol{\omega}\|^{-\gamma}$ and the Fourier transform (5) of the polyharmonic B-spline is the missing localization operator in the nominator. The primary role of this factor is to temper $\hat{\beta}_{2s}(\boldsymbol{\omega})$ at the origin, making it differentiable and essentially localized in the space domain.

Here, we concentrate on the special case $\text{S} = \mathbb{Z}^d$, which will allow us to express (21) in a polyharmonic B-spline basis that is much better conditioned than the RBFs [cf. (18)]. Since we are dealing with an infinite number of data points, it is important to make sure that the criterion (20) is well defined. The regularization term is obviously finite for any function $f \in \mathbf{W}_2^s$ (the

---

[4]In his paper, Duchon uses a more elaborate formulation of the regularizing functional that does not require the condition $f \in \mathbf{L}_2$. Here, we need not worry about the fact that the null space of ∂_*^s defined by (2) is potentially larger than the polynomials, since we are excluding all harmonic solutions [23].

Sobolev space[5] of order s). Moreover, we can invoke the following proposition which ensures that the data term is bounded as well for $s > d/2$.

Proposition 1: For any function $f \in \mathbf{W}_2^s$ with $s > d/2$, we have $f\left(T\mathbb{Z}^d\right) \in \ell_2\left(\mathbb{Z}^d\right)$, for every $0 < T \in \mathbb{R}$.

This can be established by adapting the proof of an earlier one-dimensional (1-D) result [38, Appendix C-A].

Thanks to the uniform setting, we can now apply Fourier techniques to find the solution. Furthermore, the polyharmonic B-splines of order $2s$ generate polynomials of degree $\lceil 2s \rceil - 1$, which also takes care of the second polynomial term in (21).

Theorem 1: For $s > d/2$, the solution of the variational problem (19), with discrete input data $g[\boldsymbol{k}]$ on a uniform grid $(\mathrm{S} = \mathbb{Z}^d)$, is given by

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} (h_\lambda * g)[\boldsymbol{k}]\, \beta_{2s}(\boldsymbol{x} - \boldsymbol{k}) \qquad (22)$$

where h_λ is a digital filter defined by its Fourier transform

$$H_\lambda\left(e^{j\boldsymbol{\omega}}\right) = \frac{1}{B_{2s}\left(e^{j\boldsymbol{\omega}}\right) + \lambda\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2s}}. \qquad (23)$$

Here, $B_{2s}(e^{j\boldsymbol{\omega}}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} b_{2s}[\boldsymbol{k}] e^{-j\langle \boldsymbol{\omega}, \boldsymbol{k}\rangle}$ is the Fourier transform of the discrete polyharmonic B-spline $b_{2s}[\boldsymbol{k}] := \beta_{2s}(\boldsymbol{k})$.

Proof: First, we need to prove that the global solution of the variational problem (19), among all possible functions, lies in

$$\mathcal{V}(\beta_{2s}) = \mathrm{span}_{\boldsymbol{k} \in \mathbb{Z}^d}\{\beta_{2s}(\boldsymbol{x} - \boldsymbol{k})\} \qquad (24)$$

the space of polyharmonic splines of order $2s$. To this end, we rely on the following spline energy decomposition property.

Proposition 2: Let $s > d/2$. Then, for all $f \in \mathbf{W}_2^s$, we have

$$\|\partial_*^s f(\boldsymbol{x})\|_{\mathbf{L}_2}^2 = \|\partial_*^s f_{\mathrm{int}}(\boldsymbol{x})\|_{\mathbf{L}_2}^2 + \|\partial_*^s (f - f_{\mathrm{int}})(\boldsymbol{x})\|_{\mathbf{L}_2}^2 \quad (25)$$

where the polyharmonic spline interpolator $f_{\mathrm{int}}(\boldsymbol{x})$ is the unique function of $\mathcal{V}(\beta_{2s})$ that satisfies $f_{\mathrm{int}}(\boldsymbol{k}) = f(\boldsymbol{k})$ for all $\boldsymbol{k} \in \mathbb{Z}^d$.

Proposition 2 is a variation on a standard theme in spline theory that goes back to the work of Schoenberg [39] and Duchon [22] (with a finite number of data points). There are also results for an infinite number of data points on a uniform grid; e.g., the work of Madych and Nelson [34] which covers the polyharmonic splines of integer order s, and that of Unser and Blu [35] for fractional orders in the 1-D B-spline case. The proofs given in these two latter papers can be extended to our situation.

As the cost function is quadratic, the problem (19) is guaranteed to have a solution. Using Proposition 2, we rewrite (20) as

$$\epsilon_s^2(\tilde{f}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \left(g[\boldsymbol{k}] - \tilde{f}(\boldsymbol{k})\right)^2$$
$$+ \lambda \left\|\partial_*^s \tilde{f}_{\mathrm{int}}(\boldsymbol{x})\right\|_{\mathbf{L}_2}^2 + \lambda \left\|\partial_*^s \left(\tilde{f}(\boldsymbol{x}) - \tilde{f}_{\mathrm{int}}(\boldsymbol{x})\right)\right\|_{\mathbf{L}_2}^2$$

[5]The Sobolev space with $\mathbf{L}_2$-norm and order of differentiability s, denoted as $\mathbf{W}_2^s$, is the collection of all functions $f \in \mathbf{L}_2$ such that $\|\partial_*^s f\|_{\mathbf{L}_2}$ is bounded.

where $\tilde{f}_{\mathrm{int}}(\boldsymbol{x})$ is the polyharmonic spline interpolator of the discrete signal $\tilde{f}(\boldsymbol{k})$. The left most data term is entirely determined by the integer samples $\tilde{f}(\boldsymbol{k})$. Now, if we assume that the $\tilde{f}(\boldsymbol{k})$s are the samples of the optimal solution, we can further minimize the regularization part of the criterion by selecting the solution $\tilde{f}(\boldsymbol{x}) = \tilde{f}_{\mathrm{int}}(\boldsymbol{x})$ that sets the third positive term to zero, whereas the first two terms remain unchanged. This implies that the globally optimal solution $f(\boldsymbol{x})$ of Problem (19) lies in $\mathcal{V}(\beta_{2s})$ and, therefore, can be represented by its polyharmonic B-spline expansion (7).

We can now compute the solution explicitly. By substituting (7) in (20) and using the B-spline differentiation property (14), we obtain

$$\epsilon_s^2(f) = \|g - c * b_{2s}\|_{\ell_2}^2 + \lambda \|\partial_*^s f(\boldsymbol{x})\|_{\mathbf{L}_2}^2 \qquad (26)$$
$$= \langle g, g\rangle_{\ell_2} - 2\langle b_{2s}^T * g, c\rangle_{\ell_2} + \langle b_{2s}^T * b_{2s} * c, c\rangle_{\ell_2}$$
$$+ \lambda\langle d_s * c, d_s * b_{2s} * c\rangle_{\ell_2}. \qquad (27)$$

where $b^T[\boldsymbol{k}] = b[-\boldsymbol{k}]$. Taking the partial derivative with respect to $c[\boldsymbol{k}]$ and equating it to zero, we get

$$\frac{\partial \epsilon_s^2}{\partial c[\boldsymbol{k}]} = \left(-2\left(b_{2s}\right)^T * g + 2\left(b_{2s}\right)^T * b_{2s} * c\right.$$
$$\left. + 2\lambda\left(d_s\right)^T * d_s * \left(b_{2s}\right)^T * c\right)[\boldsymbol{k}]$$
$$= 0.$$

Due to the invertibility of b_{2s} (since the B-splines generate a Riesz basis), we find

$$g[\boldsymbol{k}] = \left(\left(b_{2s} + \lambda\left(d_s\right)^T * d_s\right) * c\right)[\boldsymbol{k}]. \qquad (28)$$

Going to the discrete Fourier domain, this yields

$$C\left(e^{j\boldsymbol{\omega}}\right) = \frac{1}{B_{2s}\left(e^{j\boldsymbol{\omega}}\right) + \lambda\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2s}}\, G\left(e^{j\boldsymbol{\omega}}\right)$$
$$= H_\lambda\left(e^{j\boldsymbol{\omega}}\right) G\left(e^{j\boldsymbol{\omega}}\right)$$

where $B_{2s}(e^{j\boldsymbol{\omega}})$ is the Fourier transform of $b_{2s}[\boldsymbol{k}]$. $\blacksquare$

The Fourier domain version of (22), which will be the basis of our implementation, is

$$\hat{f}(\boldsymbol{\omega}) = H_\lambda\left(e^{j\boldsymbol{\omega}}\right) G\left(e^{j\boldsymbol{\omega}}\right) \hat{\beta}_{2s}(\boldsymbol{\omega})$$
$$= \frac{\hat{\beta}_{2s}(\boldsymbol{\omega})}{B_{2s}\left(e^{j\boldsymbol{\omega}}\right) + \lambda\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2s}}\, G\left(e^{j\boldsymbol{\omega}}\right). \quad (29)$$

For the case without regularization ($\lambda = 0$), we obtain the (fractional) interpolating polyharmonic B-splines, often called Lagrangean polyharmonic spline [31], [33].

The solution of (29) is a continuous one. This is useful, for example, for getting a higher resolution version of our signal. Note, however, that when we want to actually implement our algorithm, we should use a discrete version of it, obtaining $f[\boldsymbol{k}]$

instead of $f(\boldsymbol{x})$. In this case, our solution in the Fourier domain becomes

$$
\begin{aligned}
F\left(e^{j\boldsymbol{\omega}}\right) &= H_\lambda\left(e^{j\boldsymbol{\omega}}\right) G\left(e^{j\boldsymbol{\omega}}\right) B_{2s}\left(e^{j\boldsymbol{\omega}}\right) \\
&= \frac{B_{2s}\left(e^{j\boldsymbol{\omega}}\right)}{B_{2s}\left(e^{j\boldsymbol{\omega}}\right) + \lambda\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2s}} G\left(e^{j\boldsymbol{\omega}}\right). \quad (30)
\end{aligned}
$$

For computational purposes, we also express $B_{2s}(e^{j\boldsymbol{\omega}})$ as the periodized version of $\hat{\beta}_{2s}(\boldsymbol{\omega})$ (using Poisson summation formula)

$$
B_{2s}\left(e^{j\boldsymbol{\omega}}\right) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \hat{\beta}_{2s}\left(\boldsymbol{\omega} + 2\pi\boldsymbol{k}\right) = \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \frac{\left\|\sin\left(\frac{\boldsymbol{\omega}}{2} + \pi\boldsymbol{k}\right)\right\|^{2s}}{\left\|\frac{\boldsymbol{\omega}}{2} + \pi\boldsymbol{k}\right\|^{2s}}.
$$
(31)

This filter may also be regarded as the Fourier transform of the autocorrelation sequence of β_s. We have found an efficient way to calculate it, by applying a two-scale relation in the Fourier domain. The details of this calculation and complete mathematical analysis of this procedure will be published elsewhere [40].

IV. MMSE ESTIMATION

We now present a stochastic interpretation of our algorithm. For that, we model our signal as a fractional Brownian field, and prove that our smoothing spline algorithm yields the MMSE solution. We are also able to find an expression for the regularization parameter that achieves MMSE.

A. Stochastic Modeling of Self-Similar Processes

We model our noiseless signal, $f(\boldsymbol{x})$, as a multidimensional fractional Brownian motion (fBm) [21], [41]–[43]; the multidimensional extension of the classical prototype of a self-similar process [44].

There are a number of technical difficulties with the modeling of fBms. The main reason is that these processes are nonstationary. Therefore, their spectral power density cannot be defined in the conventional sense [45], [46]. However, the $\boldsymbol{\tau}$-lag increment derived process, $y_{\boldsymbol{\tau}}(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{x} - \boldsymbol{\tau})$, where $f(\boldsymbol{x})$ denotes our fBm process, is zero-mean and second-order stationary.

As a result, we can characterize such a process by its variogram (see Appendix I). The variogram of a multidimensional fBm process with parameter Hurst exponent H is given by

$$
\gamma_f(\boldsymbol{\tau}) = \sigma_0^2 \cdot \|\boldsymbol{\tau}\|^{2H}, 0 < H < 1 \quad (32)
$$

where σ_0^2 is a global energy factor. It is self-similar of order $2H$ in the sense that

$$
\gamma_f(a\boldsymbol{\tau}) = a^{2H}\gamma_f(\boldsymbol{\tau}). \quad (33)
$$

The variance of the fBm is nonstationary and has the same functional form

$$
\sigma_f^2(\boldsymbol{x}) \propto \sigma_0^2 \cdot \|\boldsymbol{x}\|^{2H}. \quad (34)
$$

Proposition 3: The variogram of a multidimensional fBm can be expressed using a polyharmonic B-spline representation of order $2H + d$

$$
\gamma_f(\boldsymbol{\tau}) = \sigma_0^2 \cdot \sum_{\boldsymbol{k}\in\mathbb{Z}^d} p[\boldsymbol{k}]\beta_{2H+d}(\boldsymbol{\tau} - \boldsymbol{k}) \quad (35)
$$

where the filter p is defined by its Fourier transform

$$
P\left(e^{j\boldsymbol{\omega}}\right) = \frac{1}{c_{d,2H+d} \cdot \left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2H+d}} \quad (36)
$$

with the constant in the denominator given by (12).

Proof: As the variogram of a multidimensional fBm is a radial basis function of degree $2H$ [see (32)], we can directly use the definition of a polyharmonic B-spline (see Section II-C).

Taking the Fourier transform of $\gamma_f(\boldsymbol{\tau})$ from (32), and using (11), we get

$$
\begin{aligned}
\gamma_f(\boldsymbol{\tau}) &\longleftrightarrow \sigma_0^2 \cdot \frac{1}{c_{d,2H+d}} \cdot \frac{1}{\|\boldsymbol{\omega}\|^{2H+d}} \\
&= \sigma_0^2 \cdot \frac{\left[\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2H+d}\right]^{-1}}{c_{d,2H+d}} \cdot \frac{\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2H+d}}{\|\boldsymbol{\omega}\|^{2H+d}} \\
&= \sigma_0^2 \cdot \frac{\left[\left\|2\sin\left(\frac{\boldsymbol{\omega}}{2}\right)\right\|^{2H+d}\right]^{-1}}{c_{d,2H+d}} \cdot \hat{\beta}_{2H+d}(\boldsymbol{\omega}). \quad (37)
\end{aligned}
$$

We then obtain (35) by taking the inverse Fourier transform: The central factor, which is 2π-periodic, represents the discrete Fourier transform of $p[\boldsymbol{k}]$. ∎

If we evaluate (35) on a uniform grid, we can write it as a discrete convolution

$$
\gamma_f(\boldsymbol{k}) = \sigma_0^2 \cdot (p * b_{2H+d})[\boldsymbol{k}]. \quad (38)
$$

It is interesting to note that a fBm process, even though it is nonstationary, has the property that $\partial_*^r f(\boldsymbol{x}) = n_0(\boldsymbol{x})$ (in the sense of distributions), where $n_0(\boldsymbol{x})$ is white Gaussian noise; i.e., ∂_*^r is the whitening operator of the process. The whitened process has spectral power density σ_0^2.

B. MMSE Estimation

We now consider the problem of estimating a realization f of a multidimensional fBm process given its noisy measurements at the integers: $\{g[\boldsymbol{k}] = f(\boldsymbol{k}) + n[\boldsymbol{k}]\}_{\boldsymbol{k}\in\mathbb{Z}^d}$, where $n[\boldsymbol{k}]$ is additive white gaussian noise. The noise is zero-mean with variance

σ^2 and is assumed to be independent from the signal f. The general form of the linear estimator of $f(\boldsymbol{x})$ at a fixed location $\boldsymbol{x}$ is

$$\tilde{f}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] \, g[\boldsymbol{k}]. \tag{39}$$

We also impose the constraint

$$\sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] = 1 \tag{40}$$

which ensures the perfect reconstruction of a constant background. Since the condition is imposed for all $\boldsymbol{x} \in \mathbb{R}^d$, it is equivalent to the partition of unity.

Proposition 4: The best linear constrained estimation of $f(\boldsymbol{x})$—an fBm process with parameter Hurst exponent $0 < H = s - d/2 < 1$ and signal energy σ_0^2—given its noisy samples $\{g[\boldsymbol{k}] = f(\boldsymbol{k}) + n[\boldsymbol{k}]\}_{\boldsymbol{k} \in \mathbb{Z}^d}$, where $n[\boldsymbol{k}]$ is white noise with variance σ^2, is the smoothing spline, $\tilde{f}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} (h_\lambda * g) [\boldsymbol{k}] \beta_{2s}(\boldsymbol{x} - \boldsymbol{k})$, of order $2s = 2H + d$, where the filter h_λ is given by (23) with $\lambda = (-c_{d,2s}) \cdot \sigma^2 / \sigma_0^2$.

Proof: We need to find the regression coefficients $a_{\boldsymbol{x}}[\boldsymbol{k}]$ in (39) so that the estimator $\tilde{f}(\boldsymbol{x})$ minimizes the estimation error subject to (40). Using this constraint and the fact that the noise is independent from the signal, we write the estimation error and manipulate it as shown in the equation at the bottom of the page. Observing that

$$(f(\boldsymbol{x}) - f[\boldsymbol{k}]) \cdot (f(\boldsymbol{x}) - f[\boldsymbol{l}]) = \frac{1}{2}(f(\boldsymbol{x}) - f[\boldsymbol{k}])^2$$
$$+ \frac{1}{2}(f(\boldsymbol{x}) - f[\boldsymbol{l}])^2 - \frac{1}{2}(f(\boldsymbol{k}) - f[\boldsymbol{l}])^2$$

we then rewrite the error in terms of the variogram (cf. Appendix I) as shown in the equation at the bottom of the page.

To minimize this expression, we take the partial derivative with respect to $a_{\boldsymbol{x}}[\boldsymbol{l}]$ and set it to zero, which yields

$$\gamma_f(\boldsymbol{x} - \boldsymbol{l}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{x_0}[\boldsymbol{k}] \left(\gamma_f(\boldsymbol{k} - \boldsymbol{l}) - \sigma^2 \delta_{\boldsymbol{k}-\boldsymbol{l}} \right). \tag{41}$$

The key now is to recognize that this system can be rewritten as a discrete convolution equation

$$\sigma_0^2 \cdot p * \beta_{2s}(\boldsymbol{x} - \boldsymbol{l}) = (\sigma_0^2 \cdot p * b_{2s} - \sigma^2 \delta_{\boldsymbol{k}-\boldsymbol{l}}) * a_{\boldsymbol{x}}[\boldsymbol{l}] \tag{42}$$

where we have made use of the spline representation of the variogram (38) with $H + d/2 = s$. We then convolve this identity with p^{-1}—the convolution inverse of p—whose Fourier transform is $P\left(e^{j\boldsymbol{\omega}}\right)^{-1} = c_{d,2s} \cdot \|2 \sin(\boldsymbol{\omega}/2)\|^{2s}$. Dividing by σ_0^2, we get

$$\left(b_{2s} - \frac{\sigma^2}{\sigma_0^2} \cdot p^{-1} \right) * a_{\boldsymbol{x}_0}[\boldsymbol{l}] = \beta_{2s}(\boldsymbol{x} - \boldsymbol{l}). \tag{43}$$

We then solve this equation by applying the convolution inverse of the filter on the left-hand side, which yields

$$a_{\boldsymbol{x}}[\boldsymbol{l}] = h_\lambda * \beta_{2s}(\boldsymbol{x} - \boldsymbol{l}) \tag{44}$$

where the frequency response of the inverse filter h_λ coincides with the smoothing spline filter formula (23) provided that we set $\lambda = -c_{d,2s} \cdot \sigma^2 / \sigma_0^2$. Observe that λ is necessarily positive because $c_{d,2H+d} < 0$ when $0 < H < 1$ [cf. (13)]. ∎

Note that the first part of the above derivation [until (41)] is a transposition to our problem of the kriging formalism pioneered by Matheron [47], which relies on the variogram to specify the best linear unbiased estimator of a spatial process at location $\boldsymbol{x}$ given N samples at (arbitrary) positions $\{\boldsymbol{x}_i\}_{(i=1,...,N)}$. The

$$E\left\{ \left| f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x}) \right|^2 \right\} = E\left\{ \left| f(\boldsymbol{x}) - \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] g[\boldsymbol{k}] \right|^2 \right\}$$

$$= E\left\{ \left| \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}](f(\boldsymbol{x}) - g[\boldsymbol{k}]) \right|^2 \right\}$$

$$= \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \sum_{\boldsymbol{l} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] a_{\boldsymbol{x}}[\boldsymbol{l}] E\left\{ (f(\boldsymbol{x}) - g[\boldsymbol{k}]) \cdot (f(\boldsymbol{x}) - g[\boldsymbol{l}]) \right\}$$

$$= \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \sum_{\boldsymbol{l} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] a_{\boldsymbol{x}}[\boldsymbol{l}] \left(E\left\{ (f(\boldsymbol{x}) - f[\boldsymbol{k}]) \cdot (f(\boldsymbol{x}) - f[\boldsymbol{l}]) \right\} + \sigma^2 \delta_{\boldsymbol{k}-\boldsymbol{l}} \right)$$

$$E\left\{ \left| f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x}) \right|^2 \right\} = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \sum_{\boldsymbol{l} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] a_{\boldsymbol{x}}[\boldsymbol{l}] \left(\gamma_f(\boldsymbol{x} - \boldsymbol{k}) + \gamma_f(\boldsymbol{x} - \boldsymbol{l}) - \gamma_f(\boldsymbol{k} - \boldsymbol{l}) + \sigma^2 \delta_{\boldsymbol{k}-\boldsymbol{l}} \right)$$

$$= 2 \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] \gamma_f(\boldsymbol{x} - \boldsymbol{k}) + \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \sum_{\boldsymbol{l} \in \mathbb{Z}^d} a_{\boldsymbol{x}}[\boldsymbol{k}] a_{\boldsymbol{x}}[\boldsymbol{l}] \left(\sigma^2 \delta_{\boldsymbol{k}-\boldsymbol{l}} - \gamma_f(\boldsymbol{k} - \boldsymbol{l}) \right)$$

specificity of our approach is that we are considering an infinite number of measurements on a uniform grid and that we are taking advantage of the "shift-invariant" structure of the problem and of the existence of the polyharmonic B-splines to derive an efficient digital filtering algorithm that has not been reported before.

Since the MMSE property in Proposition 4 is valid for any $\boldsymbol{x} \in \mathbb{R}^d$, our result shows that the deterministic smoothing spline algorithm described in Section III does, in fact, provide the best possible estimator when the underlying signal satisfies the fBm model. The smoothing spline may, therefore, be interpreted as a hybrid version of the classical Wiener filter[6] for which the input is discrete and the output continuous. The important point here is that the formulation not only yields an optimal digital restoration filter h_λ, but also the optimal interpolation space (polyharmonic splines) which is a powerful theoretical result. Of course, as in the deterministic case, we can also obtain a purely discrete version of the algorithm by sampling the output signal at the integers, which is achieved by digital postfiltering with b_{2s} [cf. (36)].

Another important aspect is that Proposition 4 gives the value of the optimal regularization factor λ. This is very useful and will allow us to develop a completely automatic smoothing process where the optimal smoothing spline parameters (s and λ) are estimated from the noisy data.

V. FRACTAL MODELING OF IMAGES AND IDENTIFICATION

The polyharmonic smoothing spline algorithm proposed here is a noise filtration technique that has two parameters: the order of the spline ($\gamma = 2s$), and the regularization factor λ, which controls the smoothing strength. We have also shown in the previous section that this method is optimal for the processing of multidimensional fBm-like processes; these are isotropic and are fully characterized by two corresponding parameters: 1) the signal energy σ_0^2, and 2) the Hurst exponent $0 < H < 1$, which is a measure of fractality.[7] Even though these processes are nonstationary, we will now argue that they provide a good model of the correlation structure of a large class of images. We will also indicate how the fractal-like behavior of natural images can be identified in the frequency domain. We will then propose a practical method for identifying the model parameters from the noisy input data.

A. Fractal-Like Behavior of Natural Images

Natural phenomena are widely presumed to be self-similar from a statistical perspective. As a consequence, natural images tend to be scale invariant—seeing an object from ten yards or one yard will result in very similar images transmitted by our visual system [26], [48].

The notion of self-similarity is reminiscent of fractals, which possess this property. This leads to a fractal-like model for natural images—a spectral power density that behaves like

---

[6]Note, however, that the classical Wiener filter (either continuous or discrete) is not defined for fBms since these processes are nonstationary.

[7]The fractal dimension is given by $D = d + 1 - H$ where d is the topological dimension (cf. [21]).

$(1/\|\boldsymbol{\omega}\|^{2s})$, where $2s$ is the fractal order [25] and is linked to the Hurst exponent as $s = H + d/2$.

Here, we choose to model such signals as fBms, as described in Section IV-A. Even though the power spectrum of a fBm is not defined mathematically (because the process in nonstationary), the fBm has a characteristic, inverse-power-law signature in the frequency domain. Specifically, a multidimensional fBm has the following generalized spectral representation (cf. [46], [49], and [50])

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{e^{j\langle \boldsymbol{\omega}, \boldsymbol{x} \rangle} - 1}{\|\boldsymbol{\omega}\|^{H+d/2}} dW(\boldsymbol{\omega}) \qquad (45)$$

which involves the inverse Fourier transform (in the sense of distributions) of the random measure $dW(\boldsymbol{\omega})$ where $W(\boldsymbol{\omega})$ is a complex multidimensional Wiener process (or Levy field). Intuitively, the random measure represents the Fourier domain equivalent of a white Gaussian noise process; i.e., $dW(\boldsymbol{\omega}) = N_0(\boldsymbol{\omega})d\boldsymbol{\omega}$. Thus, if one excludes the -1 in the denominator of (45), which is there to neutralize the singularity of $\|\boldsymbol{\omega}\|^{-H-d/2}$ and to enforce the condition $f(\boldsymbol{x}) = 0$, we see that the Fourier components of the fBm are decorrelated with a standard deviation that decays like $\|\boldsymbol{\omega}\|^{-s}$ with $s = H + d/2$. Another equivalent way of expressing this characteristic spectral behavior is that the fBm process can be whitened by application of an isotropic fractional derivative of order s whose frequency response is precisely $\|\boldsymbol{\omega}\|^s$. Concretely, this means that one can still perform a standard spectral analysis provided that one excludes the origin $\boldsymbol{\omega} = (0, \ldots, 0)$.

Example of signals that tend to behave in this way are functional MRI images [51]. Note that this kind of behavior is not limited to 2-D images, and can occur in multidimensional signals as well. Our experimental results, as well as those of others [24], [25], indicate that many natural images fit this model well, with s being typically fractional between 1 to 1.5. This is a good inducement for using fractional polyharmonic splines.

B. Estimation of Fractal Parameters

The final important issue we need to deal with is how to correctly identify the fractal model parameters from the input data. There has been much research on estimating the fractal order of a signal, e.g., [52]–[54]. However, in our case, we also need to estimate the variance of the additive noise. Therefore, we have to revert to a more general estimation technique.

We choose to obtain the parameters from the signal's periodogram, which is an estimation of the spectral power density of our signal. We use an iterative optimization technique to solve a nonlinear weighted least squares fit of our parameters. The power spectral density of our noisy signal should essentially behave like

$$S(\boldsymbol{\omega}) = \frac{\sigma_0^2}{\|\boldsymbol{\omega}\|^{2s}} + \sigma^2, \quad \boldsymbol{\omega} \neq (0, 0, \ldots, 0) \qquad (46)$$

which leaves us with three parameters to estimate. Note that we sometimes abuse the notation and use $S(\|\boldsymbol{\omega}\|)$ instead of $S(\boldsymbol{\omega})$, since we have $S(\boldsymbol{\omega}_1) = S(\boldsymbol{\omega}_2)$ when $\|\boldsymbol{\omega}_1\| = \|\boldsymbol{\omega}_2\|$.

For simplicity, we assume that we are given measurements of a noisy signal, $g[\boldsymbol{k}], \boldsymbol{k} \in \mathrm{S}$, on a square domain with

$$\mathrm{S} = \left\{ \boldsymbol{k} \in \mathbb{Z}^d : 0 \leq k_1, k_2, \ldots, k_d \leq M - 1 \right\}. \tag{47}$$

We then compute the periodogram of this signal, $I_M[\boldsymbol{n}]$, which is defined as the square modulus of the unitary discrete Fourier transform over the domain

$$I_M[\boldsymbol{n}] = \frac{1}{\mathrm{card}(\mathrm{S})} \left| G\left(e^{j\boldsymbol{n}2\pi/M} \right) \right|^2 \tag{48}$$

$$= \frac{1}{\mathrm{card}(\mathrm{S})} \left| \sum_{\boldsymbol{k} \in \mathrm{S}} g[\boldsymbol{k}] e^{-j\langle 2\pi \boldsymbol{n}/M, \boldsymbol{k} \rangle} \right|^2 \tag{49}$$

where
$\boldsymbol{n} \in \mathrm{N} = \left\{ \boldsymbol{n} \in \mathbb{Z}^d : -M/2 + 1 \leq n_1, n_2, \ldots, n_d \leq M/2 \right\}$ (assuming that M is even).

The periodogram is a noisy and wiggly estimation of the power spectral density $S(\boldsymbol{\omega})$. For a stationary Gaussian process, we have asymptotically for large M (almost surely)

$$I_M[\boldsymbol{n}] = S\left(\frac{2\pi}{M} \boldsymbol{n} \right) u_{\boldsymbol{n}}, \quad \boldsymbol{n} \neq [0, 0, \ldots, 0] \quad \text{a.s.} \tag{50}$$

where the random variables $\{u_{\boldsymbol{n}}\}$ are i.i.d as $(1/2)\chi_2^2$ (except the estimates on the envelope which are independent and distributed like χ_1^2), and $E[u_{\boldsymbol{n}}] = 1$ [55], [56] (a more accurate analysis is given in [57]). Note that this result is also valid for an fBm process (except for the zero-frequency component), even though the fBm is not second order stationary (cf. Section V-A). The multiplicative model (50) makes the task of estimating our parameters directly from the periodogram somewhat tricky.

To exploit the isotropic behavior of our spectral power density, we use the radial periodogram, which is a 1-D function, and offers a great computational advantage. We compute the radial periodogram by averaging the periodogram along annuli. Specifically, we define R uniformly spaced bins, their center denoted as the radial frequency, θ. Given a periodogram, $I_M[\boldsymbol{n}]$, the radial periodogram is computed by

$$I_R[\theta] = \frac{1}{|V_\theta|} \sum_{\boldsymbol{n} \in V_\theta} I_M[\boldsymbol{n}],$$
$$\text{with } \theta = \frac{\pi\sqrt{d}}{R}, \frac{3\pi\sqrt{d}}{R}, \ldots, \frac{(R-1)\pi\sqrt{d}}{R} \tag{51}$$

where the set of discrete frequencies in each bin, V_θ, is defined as

$$V_\theta = \left\{ \boldsymbol{n} \in \mathrm{N} : \theta - \frac{\pi\sqrt{d}}{R} \leq \frac{2\pi\|\boldsymbol{n}\|}{M} < \theta + \frac{\pi\sqrt{d}}{R} \right\}. \tag{52}$$

Following Wahba [58], we transform the multiplicative model of (50) into an additive one by taking the log, thus achieving a

less wiggly behavior of our samples. The next proposition shows how we can build an unbiased estimate of the power spectrum from the radial periodogram. We also make the assumption that M is large enough for (50) to hold. The proof is given in Appendix II.

Proposition 5: The radial periodogram $I_R[\theta]$ defined for the periodogram $I_M[\boldsymbol{n}]$ of a signal with an isotropic spectral power density $S(\theta)$ satisfies

$$E[\log I_R[\theta]] = \log S(\theta) + \psi\left(\frac{n_\theta}{2} \right) - \log\left(\frac{|V_\theta|}{2} \right), \text{ and} \tag{53}$$

$$\mathrm{Var}[\log I_R[\theta]] = \psi'\left(\frac{n_\theta}{2} \right) \tag{54}$$

for some n_θ (which is given in the proof). ψ is the Psi (digamma) function and ψ' (the derivative of ψ) is the trigamma function [59, Chapter 6].

From Proposition 4, we see that averaging along annuli decreases the bias of the estimation. Furthermore, the variance of each annulus depends on the number of samples in it. Therefore, we estimate the spectral parameters of our stochastic model by a weighted least squares approach with weights that are inversely proportional to the variance of the component. This yields the cost function shown in the equation at the bottom of the page, where the parameters σ_0^2, s, and σ^2 represent the signal power, its fractal order, and the noise variance, respectively. Robust estimates of these quantities are obtained by minimizing this cost function iteratively using either Levenberg–Marquardt or Powell's method in three dimensions. We are then ready to apply Proposition 3, which tells us that the corresponding optimal signal estimator is a polyharmonic smoothing spline of order $\gamma = 2s$ and regularization parameter $\lambda = (-c_{d,2H+d}) \cdot \sigma^2/\sigma_0^2$.

VI. RESULTS

We have implemented two versions of the polyharmonic spline denoising algorithm: one in Matlab for testing purposes, and a second as a JAVA plugin for ImageJ [60]. A demonstration applet is available at the site: http://bigwww.epfl.ch/demo/fractaldenoising/.

We tried our approach on many images, and demonstrate it here on four representative ones (cf. Fig. 2). The first, a biomedical image, is a slice of an MRI T2* volume, as used in functional analysis of brain activity. The second image is a texture image (Brodatz D59 [61]). The third one is a natural scenery image which is often used as a standard test image (Goldhill). The last is the well-known Lena image.

We first check the validity of our fractal model on both noisy and noiseless images. In the noiseless case, the model reduces to $S(\boldsymbol{\omega}) = \sigma_0^2/\|\boldsymbol{\omega}\|^{2s}$, for which the parameters σ_0^2 and s can be found by a simple linear regression in a log-log scale. In Fig. 3, we show the estimated radial spectral power density

$$\epsilon^2\left(\sigma_0^2, \sigma^2, s \right) = \sum_{\theta \neq 0} \frac{1}{\psi'\left(\frac{n_\theta}{2} \right)} \left[\log\left(\frac{\sigma_0^2}{|\theta|^{2s}} + \sigma^2 \right) - \log\left(I_R[\theta] \right) + \psi\left(\frac{n_\theta}{2} \right) - \log\left(\frac{|V_\theta|}{2} \right) \right]^2$$
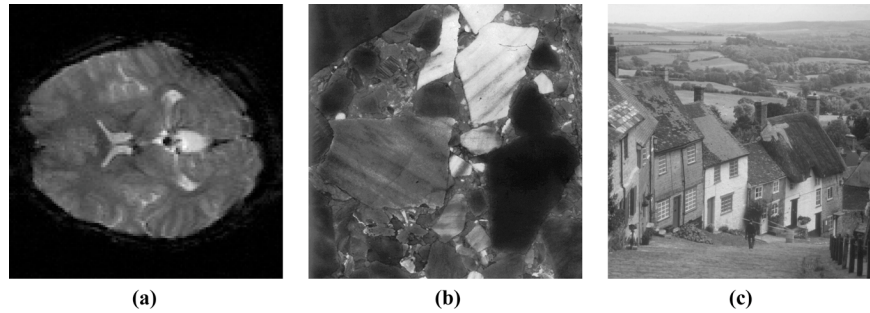
Fig. 2. Noiseless test images. We included images of a different nature: A medical image, a Brodatz texture, a natural scenery, and the classical "Lena" image which we omitted from this figure. (a) MRI T2* image. (b) Brodatz D59 image. (c) Goldhill image.
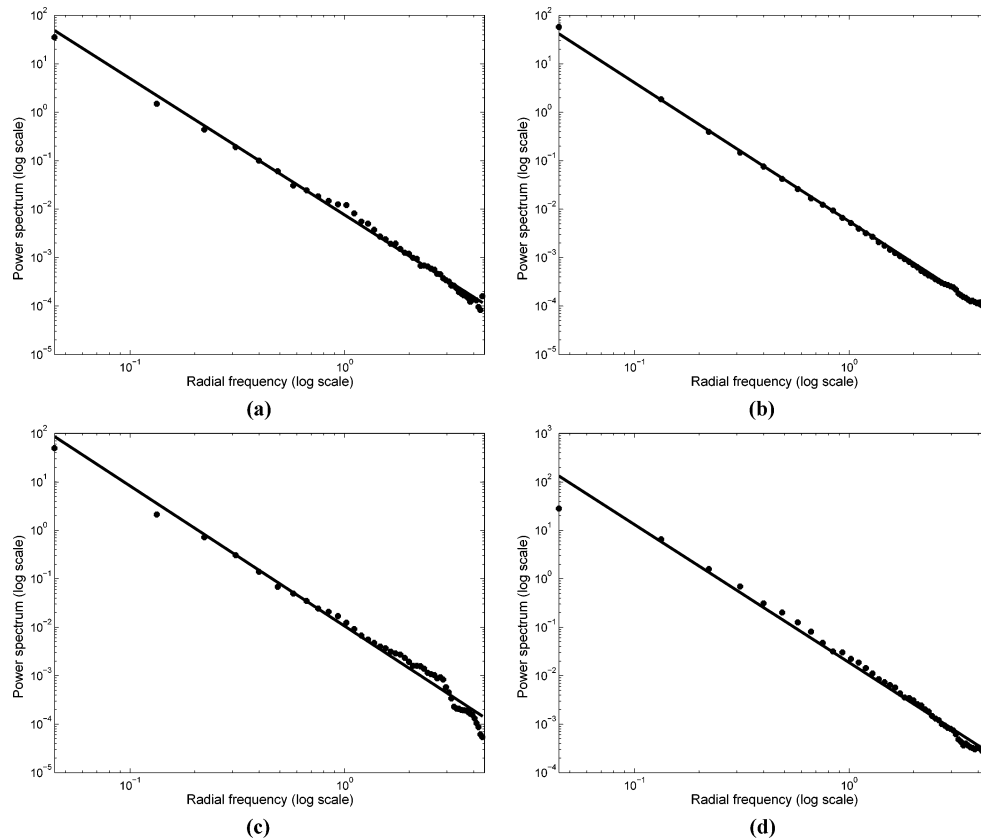

Fig. 3. Fractal behavior of the noiseless test images. The estimated radial spectral power density of the original, noiseless image, with the corresponding regressed line (on a log-log scale). (a) $s = 1.407, \sigma_0 = 0.017$; (b) $s = 1.430, \sigma_0 = 0.0145$; (c) $s = 1.446, \sigma_0 = 0.02$; (d) $s = 1.423, \sigma_0 = 0.027$. (a) MRI T2 image. (b) Brodatz D59 image. (c) Goldhill image. (d) Lena image.

of the noiseless images, and the estimated fractal model (on a log-log scale). Note that the estimated values of s are within the range of admissibility of the fBm model; i.e., $H \in (0, 1) \Rightarrow s \in (1, 2)$ for $d = 2$.

We now check the model for noisy images. In Fig. 4, we show the estimated radial spectral power density of a noisy version of the test images, and the estimated fractal model. The noisy images used for Fig. 4 were created by adding white Gaussian noise with standard deviation of around 22% of the intensity range of the original image. We omit the results for the noisy versions for the Goldhill and Lena images, as the results are very similar. As discussed in Section V-A, the "optimal" value of s

is chosen according to this analysis. We can see that the model fits the spectral power density quite well in both the noiseless and noisy cases.

We compare two different settings of our algorithm. The first one, denoted as the theoretical solution, is the solution obtained when estimating s and σ_0^2 from the noiseless image, and using the known true value of σ^2. The second one, denoted as the estimated solution, is the solution obtained when estimating the parameters from the noisy image, and is the solution which is feasible in practice. Finally, we also compare our results to the best linear solution, which is denoted as the Wiener oracle. This algorithm requires the knowledge of the power spectrum of the
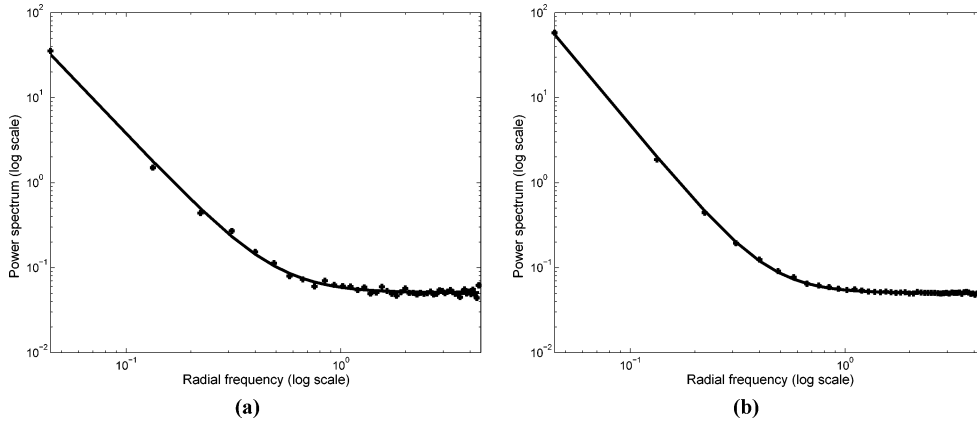
Fig. 4. Fractal model fit for the noisy test images. The estimated radial spectral power density of a noisy version of the test images, with noise standard deviation of 22% of the range of the test image, and the fit of the fractal model (on a log-log scale). (a) Input SNR: 0.07 dB ; $s = 1.351, \sigma_0 = 0.018, \sigma = 0.224$ (true σ is 0.224). (b) Input SNR: 4.23 dB; $s = 1.51, \sigma_0 = 0.012, \sigma = 0.225$ (true σ is 0.224). (a) MRI T2 image. (b) Brodatz D59 image.
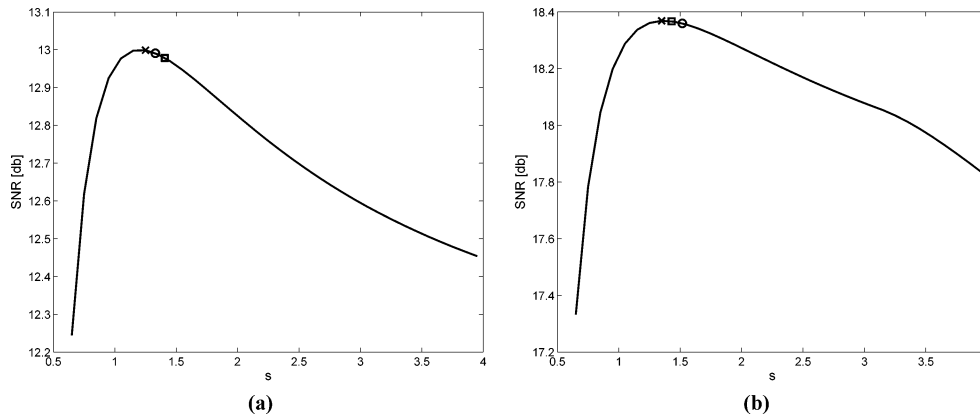

Fig. 5. Best possible SNR value vs. value of s. We use the empirically optimal value of λ. We note the theoretical value with a square, the optimal empirical value with an "x" and the estimated value from the noisy image with a circle. (a) Input SNR: 0.07 dB. (b) Input SNR: 4.23 dB. (a) MRI T2 image. (b) Brodatz D59 image.

original noiseless image, $\left| F \left(e^{j\boldsymbol{\omega}} \right) \right|^2$, and the variance of the noise, and is computed by filtering with

$$H_{\mathrm{oracle}} \left(e^{j\boldsymbol{\omega}} \right) = \frac{\left| F \left(e^{j\boldsymbol{\omega}} \right) \right|^2}{\left| F \left(e^{j\boldsymbol{\omega}} \right) \right|^2 + \sigma^2 \cdot \mathrm{card}(\mathrm{S})}. \qquad (55)$$

We now check the results using the theoretical and estimated prediction of the parameter s. In Fig. 5, we show the best SNR values possible vs. the values of s (using brute force optimization of the λ values). This is in fact an upper bound for the performance of our algorithm when we use a specific s value. We mark the location of s which corresponds to the fractal order of the image that we estimate from the noiseless images (the theoretical solution), as well as the estimated value from the noisy image, and the optimal empirical value. The noisy images are the same as the ones used in Fig. 4. We can conclude that choosing s in the proposed way is indeed close to optimal (it is within 0.05 dB of the best value in all cases), and we can also conclude that our estimation algorithm works quite well for s. It is also worth mentioning that the optimal value of s is fractional in all cases—a motivation for using fractional (polyharmonic) splines. Here, too, we do not show the results for the last two test images because the graphs are very similar.

Finally, we compare the overall performance for varying values of input SNR. In Fig. 6, we show the resulting SNR after smoothing the noisy images when using the theoretical and estimated parameters of our algorithm. We also compare the performance with the optimal Wiener oracle solution. We observe that the theoretical and estimated solutions perform almost identically. We can also see that our algorithm is always close in performance to the Wiener oracle solution (always within 2 dB of it), which is the optimal linear technique, but impossible to use in practice because the noise-free spectrum is generally unknown. In fact, it is quite surprising that an algorithm like ours, which has only two free parameters (i.e., s and λ), can essentially keep up with the best oracle estimator that requires the knowledge of as many parameters as there are pixels in the image.

VII. CONCLUSION

We proposed an efficient way to solve Duchon's multidimensional smoothing spline problem when the data is located on a uniform grid, as is mostly the case in image processing applications. The corresponding estimator is essentially a Tikhonov-like smoothing filter that is parametrized by two quantities: the
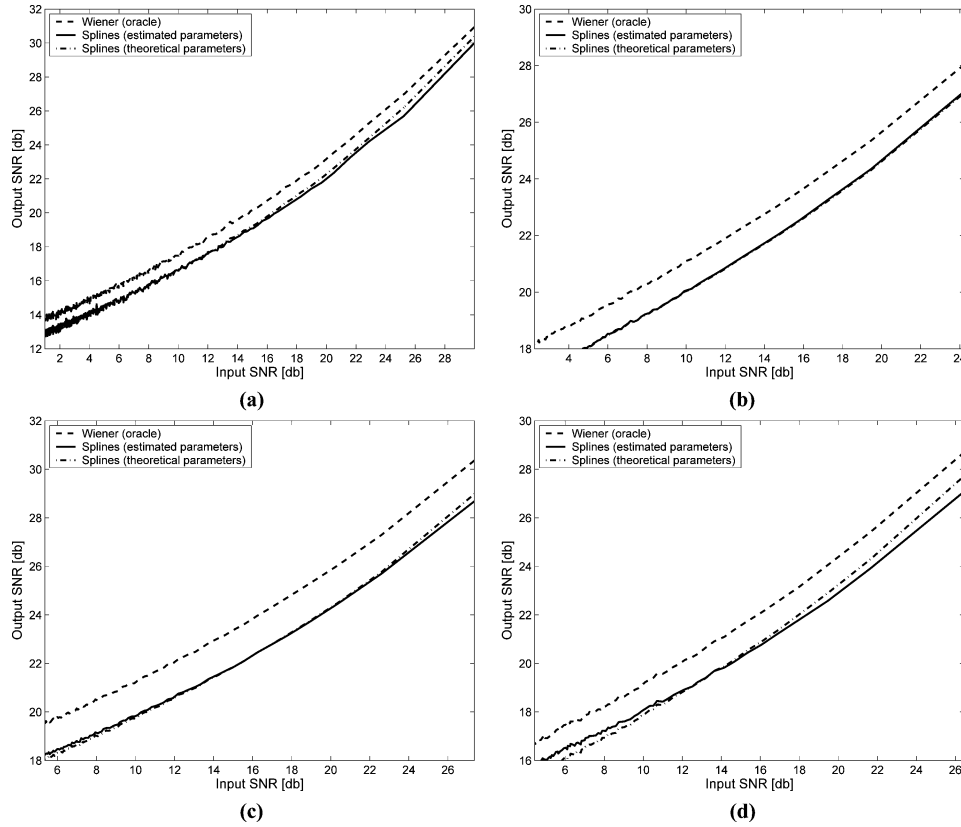
Fig. 6. Smoothing results of the test images. Results of smoothing for the theoretical and estimated values of parameters, and a comparison with the optimal oracle Wiener solution for varying input SNR values. Maximum difference between the SNR obtained with the oracle Wiener solution and the smoothing splines with estimated parameters solution: (a) 1.28 dB; (b) 1.14 dB; (c) 1.678 dB; (d) 1.616 dB. (a) MRI T2 image. (b) Brodatz D59 image. (c) Goldhill image. (d) Lena image.

order of the spline and the regularization factor λ that controls the smoothing strength.

We justified the procedure statistically by proving that it provides the MMSE estimation for fractal-like signals corrupted by additive white noise. This result is interesting conceptually because it establishes an equivalence between multidimensional smoothing splines and Wiener filtering. It also yields the optimal (MMSE) parameters for the denoising algorithm, including the optimal reconstruction space which is not bandlimited.

We proposed a practical FFT-based implementation of the method, and introduced a viable way of estimating the best smoothing parameters from the noisy data. We demonstrated that the approach performs quite well for the processing of a variety of images, which is probably due to the fact that the underlying fractal model is quite successful in capturing their primary correlation structure. This is remarkable given the extreme conciseness of the representation.

The final point that we want to make is that the proposed method is a linear one, albeit not a conventional one because it reconstructs a continuous-space version of the signal from a discrete set of input data. Except for the special case of fractional Brownian fields for which the procedure is provably MMSE optimal, we do not want to oversell the technique by claiming that it will compete with the best nonlinear denoising methods that were mentioned in the introduction. However, we do believe that our algorithm can serve as a gold standard of what can realistically be achieved with a linear algorithm that is optimized for best performance.

APPENDIX I
VARIOGRAM AND AUTOCORRELATION

The variogram is one way of characterizing the second order statistics of a multidimensional stochastic process $f(\boldsymbol{x})$ [47], [62]. It is defined as

$$\gamma_f(\boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{2} E\left\{|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)|^2\right\}. \quad (56)$$

where $E\{\cdot\}$ is the expectation operator. It is easy to see that γ_f is symmetric in $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and vanishes when $\boldsymbol{x}_1 = \boldsymbol{x}_2$. The variogram is related to the autocorrelation function of a process through the relation

$$\begin{aligned} c_{ff}(\boldsymbol{x}_1, \boldsymbol{x}_2) &\equiv E\{f(\boldsymbol{x}_1) f(\boldsymbol{x}_2)\} \\ &= \frac{1}{2}\sigma_f^2(\boldsymbol{x}_1) + \frac{1}{2}\sigma_f^2(\boldsymbol{x}_2) - \gamma_f(\boldsymbol{x}_1, \boldsymbol{x}_2) \quad (57) \end{aligned}$$

where $\sigma_f^2(\boldsymbol{x}) \equiv E\left\{f(\boldsymbol{x})^2\right\} = c_{ff}(\boldsymbol{x}, \boldsymbol{x})$ is the variance of the process at location $\boldsymbol{x}$.

A process is called *intrinsically stationary* if

$$E\{f(\boldsymbol{x}) - f(\boldsymbol{x} - \boldsymbol{\tau})\} = 0, \forall \boldsymbol{x} \in \mathbb{R}^d \quad (58)$$

$$\gamma_f(\boldsymbol{x}, \boldsymbol{x} - \boldsymbol{\tau}) = \gamma_f(\boldsymbol{\tau}) \quad (59)$$

i.e., when the variogram is a function of the space lag $\tau = x_2 - x_1$ only.

A process is called *second order stationary* if

$$E\{f(x)\} = \mu, \forall x \in \mathbb{R}^d \tag{60}$$

$$c_{ff}(x, x - \tau) = c_{ff}(\tau) \tag{61}$$

i.e., when the autocorrelation is a function of the space lag only. In this case (57) simplifies to

$$c_{ff}(\tau) = c_{ff}(0) - \gamma_f(\tau) = \sigma_f^2 - \gamma_f(\tau) \tag{62}$$

and the variance is constant over space. Note that the second order stationarity implies intrinsic stationarity, but not the reverse.

Finally, a process is called *isotropic* if $\gamma_f(\tau) = \gamma_f(\|\tau\|)$. Otherwise, if there is a dependence on the direction, the process is anisotropic.

APPENDIX II
PROOF OF PROPOSITION 4

Starting from (51) and using (50), we get (under the assumption that M is large)

$$I_R[\theta] = \frac{1}{|V_\theta|} \sum_{n \in V_\theta} S\left(\frac{2\pi}{M}n\right) u_n, \quad \text{a.s.} \tag{63}$$

Here, we have a linear combination of χ^2 variables. We refer to [63] for a numerically accurate way to solve this kind of problem, and to [64] for an overview of most known approximation methods. Using the fact that the sum of $\chi_{p_k}^2$ random variables is a $\chi_{\sum p_k}^2$ random variable, we regroup the u_n with the same factor in the previous equation as

$$I_R[\theta] = \frac{S(\theta)}{|V_\theta|} \sum_k \lambda_k u_k, \quad \text{a.s.} \tag{64}$$

where the u_k are independent and distributed as $\chi_{n_k}^2$; the weights λ_k are all different and are either equal to $S((2\pi/M)n)/S(\theta)$ or $S((2\pi/M)n)/2S(\theta)$ for some $n \in V_\theta$, depending on whether this element was on the envelope or not. We now use the following (very good) approximation: given a random variable $U \sim \chi_n^2$, then

$$aU \sim \chi_{n+2J}^2, \text{where } J \sim \text{NB}\left(\frac{a-1}{a}, \frac{n}{2}\right) \tag{65}$$

and NB stands for negative-binomial distribution [65]. Combining the last two equations, we have

$$I_R[\theta] = \frac{S(\theta)}{|V_\theta|} U_\theta, \ U_\theta \sim \chi_{\sum_k n_k + 2J_k}^2 \quad \text{a.s.} \tag{66}$$

where the J_k's are independent and distributed as $\text{NB}((\lambda_k - 1)/\lambda_k, n_k/2)$. Since $E[J_k] =$

$(\lambda_k - 1)(2 - n_k)/\lambda_k n_k$, and we are only interested in finding the mean and the variance, we can further approximate U_θ as

$$U_\theta \sim \chi_{n_\theta}^2, \ n_\theta = \sum_k \left[n_k + 2\frac{(\lambda_k - 1)(2 - n_k)}{\lambda_k n_k}\right]. \tag{67}$$

Now, it is easy to see that

$$E[\log I_R[\theta]] = \log S(\theta) + E\left[\log\left(\frac{U_\theta}{|V_\theta|}\right)\right]. \tag{68}$$

Computing the second term of (68), we get

$$E\left[\log\left(\frac{U_\theta}{|V_\theta|}\right)\right] = \int_0^\infty \log\left(\frac{\theta}{|V_\theta|}\right) \frac{\theta^{n_\theta/2-1}e^{-\theta/2}}{\Gamma\left(\frac{n_\theta}{2}\right)2^{n_\theta/2}}d\theta$$

$$= \frac{\Gamma'\left(\frac{n_\theta}{2}\right) - \log\left(\frac{|V_\theta|}{2}\right)\Gamma\left(\frac{n_\theta}{2}\right)}{\Gamma\left(\frac{n_\theta}{2}\right)}$$

$$= \psi\left(\frac{n_\theta}{2}\right) - \log\left(\frac{|V_\theta|}{2}\right) \tag{69}$$

where we have used properties from [59, Chapter 6]. Finally, plugging (69) into (68), we get (53).

We now calculate the variance. Using (67), we have

$$\text{Var}[\log I_R[\theta]] = \text{Var}[\log S(\theta)] + \text{Var}\left[\log\left(\frac{U_\theta}{|V_\theta|}\right)\right]$$

$$= \text{Var}\left[\log\left(\frac{U_\theta}{|V_\theta|}\right)\right]. \tag{70}$$

Computing the squared expected value we have

$$E\left[\log^2\left(\frac{U_\theta}{|V_\theta|}\right)\right]$$

$$= \int_0^\infty \log^2\left(\frac{\theta}{|V_\theta|}\right) \frac{\theta^{n_\theta/2-1}e^{-\theta/2}}{\Gamma\left(\frac{n_\theta}{2}\right)2^{n_\theta/2}}d\theta$$

$$= \frac{\Gamma''\left(\frac{n_\theta}{2}\right) - 2\log\left(\frac{|V_\theta|}{2}\right)\Gamma'\left(\frac{n_\theta}{2}\right) + \log^2\left(\frac{|V_\theta|}{2}\right)\Gamma\left(\frac{n_\theta}{2}\right)}{\Gamma\left(\frac{n_\theta}{2}\right)}$$

$$= \frac{\Gamma''\left(\frac{n_\theta}{2}\right)}{\Gamma\left(\frac{n_\theta}{2}\right)} - 2\log\left(\frac{|V_\theta|}{2}\right)\psi\left(\frac{n_\theta}{2}\right) + \log^2\left(\frac{|V_\theta|}{2}\right). \tag{71}$$

We can now calculate the variance using (69) and (71)

$$\text{Var}\left[\log\left(\frac{U_\theta}{|V_\theta|}\right)\right] = E\left[\log^2\left(\frac{U_\theta}{|V_\theta|}\right)\right] - E^2\left[\log\left(\frac{U_\theta}{|V_\theta|}\right)\right]$$

$$= \frac{\Gamma''\left(\frac{n_\theta}{2}\right)}{\Gamma\left(\frac{n_\theta}{2}\right)} - \psi^2\left(\frac{n_\theta}{2}\right)$$

$$= \psi'\left(\frac{n_\theta}{2}\right). \tag{72}$$

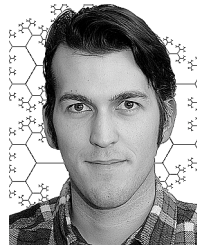Finally, by plugging (72) into (70), we get (54).

REFERENCES

[1] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

[3] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[4] D. L. Donoho, "De-noising by soft thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[5] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[6] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge, MA: MIT Press, 1949.

[7] G. T. Herman, H. Hurwitz, A. Lent, and H.-P. Lung, "On the Bayesian approach to image reconstruction," *Inf. Control*, vol. 42, pp. 60–71, 1979.

[8] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[9] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. New York: Wiley, 1977.

[10] T. Poggio, V. Torre, and C. Koch, "Computational vision and regularization theory," *Nature*, vol. 317, pp. 314–319, 1985.

[11] N. B. Karyiannis and A. N. Venetsanopoulos, "Regularization theory in image restoration — The stabilizing functional approach," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1155–1179, Jul. 1990.

[12] R. Ramlau and G. Teschke, "Regularization of Sobolev embedding operators and applications—Part I: Fourier and wavelet based methods," *Sampling Theory Signal Image Process.*, vol. 3, no. 2, pp. 175–196, 2004.

[13] ——, "Regularization of Sobolev embedding operators and applications—Part II: Data driven regularization and applications," *Sampling Theory Signal Image Process.*, vol. 3, no. 3, pp. 225–246, 2004.

[14] G. Demoment, "Image reconstruction and restoration—Overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2024–2036, Dec. 1989.

[15] L. Tenorio, "Statistical regularization of inverse problems," *SIAM Rev.*, vol. 43, no. 2, pp. 347–366, 2001.

[16] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 298–311, Feb. 1997.

[17] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.

[18] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving map estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, Jul. 1993.

[19] J. A. Fessler and A. O. Hero, III, "Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms," *IEEE Trans. Image Process.*, vol. 4, no. 10, pp. 1417–1429, Oct. 1995.

[20] A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 319–335, Mar. 1998.

[21] B. B. Mandelbrot, *Gaussian Self-Affinity and Fractals*. New York: Springer-Verlag, 2002.

[22] J. Duchon, , W. Schempp and K. Zeller, Eds., "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Multivariate Approximation Theory*. Basel, Switzerland: Birkhäuser-Verlag, 1979, pp. 85–100.

[23] J. Kybic, T. Blu, and M. Unser, "Generalized sampling: A variational approach—Part I: Theory," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 1965–1976, Aug. 2002.

[24] B. Pesquet-Popescu and J. L. Véhel, "Stochastic fractal models for image processing," *IEEE Signal Process. Mag.*, vol. 19, no. 5, pp. 48–62, May 2002.

[25] A. P. Pentland, "Fractal–based description of natural scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 661–674, Nov. 1984.

[26] D. Mumford and B. Gidas, "Stochastic models for generic images," *Quart. Appl. Math.*, vol. 59, pp. 85–111, 2001.

[27] G. Matheron, "The intrinsic random functions and their applications," *Appl. Prob.*, vol. 5, no. 12, pp. 439–468, 1973.

[28] D. E. Myers, "Kriging, cokriging, radial basis functions and the role of positive definiteness," *Comput. Math. Appl.*, vol. 24, no. 12, pp. 139–148, 1992.

[29] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990, vol. 59.

[30] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," *Zeit. Ang. Math. Phys.*, vol. 30, pp. 292–304, 1979.

[31] C. Rabut, "Elementary m-harmonic cardinal B-splines," *Numer. Alg.*, vol. 2, pp. 39–62, 1992.

[32] J. Liouville, "Sur une formule pour les différentielles à indices quelconques à l'occasion d'un mémoire de M. Tortolini," (in French) *J. Math. Pure Appl.*, vol. 20.

[33] W. R. Madych and S. A. Nelson, "Polyharmonic cardinal splines," *J. Approx. Theory*, vol. 60, pp. 141–156, 1990.

[34] ——, "Polyharmonic cardinal splines: A minimization property," *J. Approx. Theory*, vol. 63, pp. 303–320, 1990.

[35] M. Unser and T. Blu, "Fractional splines and wavelets," *SIAM Rev.*, vol. 42, no. 1, pp. 43–67, Mar. 2000.

[36] D. Van De Ville, T. Blu, and M. Unser, "Isotropic polyharmonic B-splines: Scaling functions and wavelets," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1798–1813, Nov. 2005.

[37] L. Schwartz, *Théorie des Distributions* (in French). Paris, France: Hermann, 1966.

[38] T. Blu and M. Unser, "Approximation error for quasi-interpolators and (multi-) wavelet expansions," *Appl. Comput. Harmon. Anal.*, vol. 6, no. 2, pp. 219–251, Mar. 1999.

[39] I. J. Schoenberg, "Spline functions and the problem of graduation," *Proc. Nat. Acad. Sci.*, vol. 52, pp. 947–950, 1964.

[40] T. Blu, D. Van De Ville, and M. Unser, "Numerical methods for the computation of wavelet correlation sequences," *SIAM J. Math. Anal.*, to be published.

[41] J. P. Kahane, *Some Random Series of Functions*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1985, vol. 5, Cambridge studies in advanced mathematics.

[42] H. O. Peitgen and D. Saupe, Eds., *The Science of Fractal Images*. New York: Springer-Verlag, 1988.

[43] R. J. Adler, *The Geometry of Random Fields*. New York: Wiley, 1985.

[44] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, no. 4, pp. 422–437, Oct. 1968.

[45] P. Flandrin, "On the spectrum of fractional Brownian motions," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 197–199, Jan. 1989.

[46] I. S. Reed, P. C. Lee, and T. K. Truong, "Spectral representation of fractional Brownian motion in n dimensions and its properties," *IEEE Trans. Inf. Theory*, vol. 41, no. 9, pp. 1439–1451, Sep. 1995.

[47] G. Matheron, *Les Variables Régionalisées et Leur Estimation, une Application de la Théorie de Fonctions Aléatoires aux Sciences de la Nature* (in French). Paris, France: Masson et Cie, 1965.

[48] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Phys. Rev. Lett.*, vol. 73, no. 6, pp. 814–818, Aug. 1994.

[49] A. M. Yaglom, *Correlation Theory of Stationary and Related Random Functions I: Basic Results*. New York: Springer, 1986.

[50] R. L. Dobrushin, "Gaussian and their subordinated self-similar random generalized fields," *Ann. Prob.*, vol. 7, no. 1, pp. 1–28, 1979.

[51] E. Zarahn, G. Aguirre, and M. D'Esposito, "Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions," *NeuroImage*, vol. 5, pp. 179–197, 1997.

[52] T. Lundahl, W. J. Ohley, S. M. Kay, and R. Stiffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. Med. Imag. IEEE Trans. Inf. Theory*, vol. MI-5, pp. 152–161, 1986.

[53] S. Soltania, P. Simarda, and D. Boichub, "Estimation of the self-similarity parameter using the wavelet transform," *Signal Process.*, vol. 84, pp. 117–123, 2004.

[54] M. S. Taqqu, V. Teverovsky, and W. Willinger, "Estimators for long-range dependence: An empirical study," *Fractals*, vol. 3, no. 4, pp. 785–798, 1995.

[55] A. M. Walker, "Some asymptotic results for the periodogram of a stationary time series," *J. Austral. Math. Soc.*, vol. 5, pp. 107–128, 1965.

[56] D. R. Brillinger, *Time Series Data Analysis and Theory*. New York: Holt Rinehart and Winston, 1975.

[57] G. Wittwer, "On the distribution of the periodogram for stationary random sequences (with discussion)," *Statistics*, vol. 17, pp. 201–219, 1986.

[58] G. Wahba, "Automatic smoothing of the log periodogram," *J. Amer. Statist. Assoc.*, vol. 75, pp. 122–132, Mar. 1980.

[59] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Washington, DC: U.S. Government Printing Office, 1964.

[60] D. Sage and M. Unser, "Teaching image-processing programming in Java," *IEEE Signal Process. Mag.*, vol. 20, no. 6, pp. 43–52, Nov. 2003.

[61] P. Brodatz, *Textures*. New York: Dover, 1966.

[62] N. A. Cressie, *Statistics for Spatial Data*. New York: Wiley, 1993.

[63] J. P. Imhof, "Computing the distribution of quadratic forms in normal variables," *Biometrika*, vol. 48, pp. 419–426, 1961.

[64] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1994, vol. 1.

[65] S. D. Oman and S. Zacks, "A mixture approximation to the distribution of a weighted sum of chi-squared variables," *J. Statist. Comput. Simul.*, vol. 13, pp. 215–224, 1981.

Dimitri Van De Ville (M'02) was born in Dendermonde, Belgium, in 1975. He received the Engineering and Ph.D. degrees in computer sciences from Ghent University, Ghent, Belgium, in July 1998 and January 2002, respectively.

While studying for the Ph.D. degree, he was a Research Assistant with the Fund for Scientific Research, Flanders, Belgium, at the Medical Image and Signal Processing Group within the Department of Electronics and Information Systems. He is currently with the Biomedical Imaging Group, Swiss Federal Institute of Technology Lausanne, Lausanne, Switzerland. He is a Technical Editor of the *Wavelet Digest*, the electronic newsletter of the wavelet community. His current research interests include splines, wavelets, approximation and sampling theory, and biomedical signal and imaging applications, such as fMRI and microscopy imaging.

Dr. Van De Ville is an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS.

Michael Unser (M'89–SM'94–F'99) received the M.S. (*summa cum laude*) and Ph.D. degrees in electrical engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1981 and 1984, respectively.

From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, where he headed the Image Processing Group. He is now Professor and Director of the Biomedical Imaging Group, EPFL. He is the Editor-in-Chief of the *Wavelet Digest*, the electronic newsletter of the wavelet community. He also served as regular chair for SPIE's Conference on Wavelets from 1993 to 2005. His research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, and the use of splines for image processing, and he is the author of over 100 published journal papers in these areas.

Dr. Unser is the Associate Editor-in-Chief of the IEEE TRANSACTIONS ON MEDICAL IMAGING. He has acted as Associate Editor or member of the editorial boards for the *IEEE Signal Processing Magazine*, the IEEE TRANSACTIONS ON IMAGE PROCESSING (1992 to 1995), and the IEEE SIGNAL PROCESSING LETTERS (1994 to 1998). He was general Co-Chair for the first IEEE International Symposium on Biomedical Imaging (ISBI 2002), Washington, DC, July 7–10, 2002. He received the IEEE Signal Processing Society's 1995 Best Paper Award and the IEEE Signal Processing Society's 2000 Magazine Award.

Shai Tirosh (S'03) received the B.Sc. degrees in computer science and electrical engineering from the Technion–Israel Institute of Technology, Haifa, and the postgraduate degree in telecommunications from the Swiss Federal Institute of Technology, Lausanne (EPFL).

He was with the Biomedical Imaging Group, EPFL. His current areas of interest include signal and image processing and mathematical finance and economics.