Sampling and (sparse) stochastic processes: A tale of splines and innovation

Michael Unser
Biomedical Imaging Group
Ecole polytechnique fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland
Email: michael.unser@epfl.ch

***Abstract*—The commonality between splines and Gaussian or sparse stochastic processes is that they are ruled by the same type of differential equations. Our purpose here is to demonstrate that this has profound implications for the three primary forms of sampling: uniform, nonuniform, and compressed sensing.**

The connection with classical sampling is that there is a one-to-one correspondence between spline interpolation and the minimum-mean-square-error reconstruction of a Gaussian process from its uniform or nonuniform samples. The caveat, of course, is that the spline type has to be matched to the operator that whitens the process.

The connection with compressed sensing is that the non-Gaussian processes that are ruled by linear differential equations generally admit a parsimonious representation in a wavelet-like basis. There is also a construction based on splines that yields a wavelet-like basis that is matched to the underlying differential operator. It has been observed that expansions in such bases provide excellent M-term approximations of sparse processes. This property is backed by recent estimates of the local Besov regularity of sparse processes.

I. Introduction

The results that are being discussed in this overview paper apply to the three primary forms of sampling under the assumption that the signal s is a realization of a continuous-time (Gaussian or sparse) stochastic process that is ruled by a stochastic differential equation with known parameters (the operator L and the Lévy exponent f of the excitation). For simplicity of presentation, we shall concentrate on one-dimensional sampling, keeping in mind that most of the results that are being discussed here have multidimensional extensions.

1) Uniform Sampling: This is the process of converting a function into a sequence of equally spaced samples. For simplicity of notation, we take the samples on the Cartesian grid (cardinal setting) [1], [2], [3]

$$s(x), x \in \mathbb{R} \longrightarrow \{s[k] = s(k)\}_{k\in\mathbb{Z}}$$

2) Nonuniform Sampling: Here, the samples are taken at a series of known locations $\cdots < x_{k-1} < x_k < x_{k+1} < \cdots$ [4]

$$s(x) \longrightarrow \{s_k = s(x_k)\}_{k\in\mathbb{Z}}$$

978-1-4673-7353-1/15/$31.00 ©2015 IEEE

3) Generalized or Compressed Sampling: This is the richest form of sampling. It returns a series of linear measurements

$$s(x) \longrightarrow \{m_k = \langle s, \phi_k \rangle\}_{k\in\mathbb{Z}}$$

where the ϕ_k are appropriate analysis functions. Mathematically, each measurement corresponds to a linear functional of s. Classically, this extended form is called generalized sampling [3], [5]. Clearly, the two first configurations are particular cases of the third with $\phi_k = \delta(\cdot - k)$ and $\phi_k = \delta(\cdot - x_k)$, respectively.

The general problem of sampling is to obtain the most faithful reconstruction $\tilde{s}(x)$ of $s(x)$ for all $x \in \mathbb{R}$ from its discrete measurements. In the classical setting, this reconstruction is linear [3], [6], [4], [7].

In the case of compressed sensing, where the ϕ_k need to carefully chosen, a reconstruction from a reduced set of measurements is possible under the assumption that the signal s has a sparse representation in some privileged basis [8], [9], [10]. The reconstruction algorithm, however, is nonlinear: It is typically based on the minimization of a cost functional that favors sparse solutions [11].

The purpose of this paper and of the special session on "Sampling and Stochastic Processes" is to provide statistical arguments that support both types of reconstruction algorithms. The main conclusions, in a nutshell, are:

- A linear reconstruction (with splines that are tailored to the spectral properties of the process) is optimal under the Gaussian assumption.
- The non-Gaussian processes that are ruled by the same kind of stochastic differential equations are inherently sparse (in a matched wavelet basis), which justifies the deployment of nonlinear reconstruction methods that favor sparsity. However, more research is required to obtain estimators with good statistical properties—that is, minimum-mean-square-error (MMSE) rather than the more conventional maximum a posteriori (MAP) whose performance can be deceptive [12].

A. Mathematical Context

$\mathcal{S}'(\mathbb{R})$ is Schwartz' space of tempered distributions. All subsequent equalities involving Dirac impulses $\delta(\cdot - x_k)$ or the innovation w are in the weak sense of distributions. For

instance, the statement $\mathrm{L}s = w$ (in the sense of distributions) is equivalent to

$$\langle \varphi, \mathrm{L}s \rangle = \langle \varphi, w \rangle$$

for all test functions $\varphi \in \mathcal{S}(\mathbb{R})$ (Schwartz' space of smooth and rapidly decreasing function).

II. Brief Overview of Splines

The leading thread of our exposition is the intimate connection between splines and differential operators; namely, the property that an admissible operator L specifies a particular brand of splines [13], [14], [15, Chapter 6].

Definition 1: A operator L is called *spline-admissible* if

1) it is linear shift-invariant (LSI); that is, if L is linear and $\mathrm{L}\{s(\cdot - x_0)\} = \mathrm{L}\{s\}(\cdot - x_0)$ for any signal s in its range;
2) there exists a function $\rho_{\mathrm{L}}(x)$ of slow growth (the Green's function of L) such that $\mathrm{L}\{\rho_{\mathrm{L}}\} = \delta$ where δ is the Dirac distribution;
3) the null space of the operator

$$\mathcal{N}_{\mathrm{L}} = \{p_0(x) \in \mathcal{S}'(\mathbb{R}) : \mathrm{L}\{p_0\} = 0\}$$

is either empty—$\mathcal{N}_{\mathrm{L}} = \{0\}$—or finite-dimensional.

The frequency response of L is denoted by $\hat{L}(\omega)$ with $\hat{L}(\omega) = \int_{\mathbb{R}} \mathrm{L}\{\delta\}(x)\ \mathrm{e}^{-\mathrm{j}\omega x}\mathrm{d}x$ when the impulse response is absolutely integrable. The composition of the null space of L is determined by the zeros of $\hat{L}(\omega)$. Specifically, a zero of multiplicity N at $\omega = \omega_0$ corresponds to components of the form $p_0(x) = \mathrm{e}^{\mathrm{j}\omega_0 x} \sum_{n=0}^{N-1} b_n x^n$ (modulated polynomials).

The generic example of an admissible operator from the theory of linear systems is

$$\mathrm{D}^n + a_{n-1}\mathrm{D}^{n-1} + \cdots + a_0\mathrm{I} \tag{1}$$

where $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}x}$ and a_n are constant coefficients. Another interesting case is the fractional derivative D^γ of order $\gamma \in \mathbb{R}_+$ which corresponds to a multiplication by $(\mathrm{j}\omega)^\gamma$ in the frequency domain. Its Green's function is

$$\rho_{\mathrm{D}^\gamma}(x) = \frac{x_+^{\gamma-1}}{\Gamma(\gamma)}$$

where $x_+ = \max(0, x)$ and Γ is Euler's gamma function.

Definition 2: The function $s(x)$ is an L-spline with knots $(x_k)_{k\in\mathbb{Z}}$ if $\mathrm{L}s(x) = \sum_{k\in\mathbb{Z}} a_k \delta(x - x_k)$ where (a_k) is a sequence of (possibly slowly increasing) real weights.

Hence, we can view a spline as the solution of a differential equation driven by a sequence of weighted Dirac impulses. By invoking the properties in Definition 1, we can solve this equation to obtain the generic form of a spline

$$s(x) = p_0(x) + \sum_{k\in\mathbb{Z}} a_k \rho_{\mathrm{L}}(x - x_k) \tag{2}$$

where the specification of the null-space component p_0 requires some additional boundary conditions.

The main use of splines is for the reconstruction of a function from a set of nonuniform samples $\{s(x_k) = s_k\}_{k\in\mathbb{Z}}$. Specifically, when $\mathrm{L} = \mathrm{K}^*\mathrm{K}$ is self-adjoint and $s_k \in \ell_2(\mathbb{Z})$, it can be shown that there is a unique solution of the form (2) that solves the interpolation problem (see [16])

$$s_{\mathrm{rec}} = \arg\min_{s\in\mathcal{V}} \|\mathrm{K}s\|_{L_2(\mathbb{R})} = \langle \mathrm{L}f, f \rangle \text{ s.t. } s(x_k) = s_k$$

with $\mathcal{V} = \{s : \|\mathrm{K}s\|_{L_2(\mathbb{R})} < \infty\}$ where $\mathrm{K} = \mathrm{L}^{1/2}$. This shows that splines are optimal in the sense that they minimize some corresponding L_2 energy functional [17], [18].

When the sampling is uniform, the spline-interpolation problem admits an efficient filter-based solution [19], [20]. Specifically, the reconstructed signal is expressed as

$$s_{\mathrm{rec}}(x) = \sum_{k\in\mathbb{Z}} c[k]\beta_{\mathrm{L}}(x - k)$$

where β_{L} is the B-spline associated with the operator L. The coefficients of the expansion are then given by $c[k] = (h_{\mathrm{int}} * s)[k]$ where $s[k] = s(x)|_{x=k}$ are the samples of the signal and h_{int} is the digital filter whose frequency response is

$$H_{\mathrm{int}}(\mathrm{e}^{\mathrm{j}\omega}) = \frac{1}{\sum_{k\in\mathbb{Z}} \beta_{\mathrm{L}}(k)\mathrm{e}^{-\mathrm{j}\omega k}}.$$

III. Stochastic Models of Signals

The parallel with splines is that stochastic models can also be tied to a differential operator L: the so-called whitening operator that decouples the process and uncovers its "innovation", which is the unpredictable part [21], [22]. This is equivalent to specifying a stochastic process as the solution of the stochastic differential equation (SDE)

$$\mathrm{L}s = w \tag{3}$$

where w is a continuous-domain white Lévy noise—or innovation [23]. The term Lévy noise refers to the broadest possible family of generalized stochastic processes that are stationary and independent at every point [15]. While the family includes the white Gaussian noises of the traditional theory of stochastic processes, it is considerably richer, the great majority of its members being sparse [24].

An important point is that (3) only holds in the sense of distributions since the innovation $w \in \mathcal{S}'(\mathbb{R})$ is too rough to have a classical pointwise interpretation [25]. If L is spline-admissible, then it is generally possible to invert this equation, which yields the formal solution

$$s = \mathrm{L}^{-1}w \tag{4}$$

or, more explicitly,

$$s(x) = \int_{\mathbb{R}} h_{\mathrm{L}}(x, y)w(y)\mathrm{d}y \tag{5}$$

where $h_{\mathrm{L}}(\cdot, y) = \mathrm{L}^{-1}\{\delta(\cdot - y)\}$ is the kernel (or generalized impulse response) of L^{-1} [15]. The connection with the results in Section II is that $\mathrm{L}\{h_{\mathrm{L}}(\cdot, y_0)\} = \delta(\cdot - y_0)$, which shows that $h_{\mathrm{L}}(x, y_0)$ with y_0 fixed is an L-spline.

In the simplest scenario where $\mathcal{N}_{\mathrm{L}} = \{0\}$, the inverse operator L^{-1} is LSI with $h(x, y) = \rho_{\mathrm{L}}(x - y)$, so that the stochastic process $s = \rho_{\mathrm{L}} * w$ is stationary. Otherwise, $s = \mathrm{L}^{-1}w$ will

generally be non-stationary, the better known example being the Lévy processes with $\mathrm{L} = \mathrm{D}$ and $h_{\mathrm{D}}(x, y) = \mathbb{1}_{(0,x]}(y)$, which is a piecewise-constant spline.

The innovation model (4) generates a whole variety of signals whose correlation structure is imposed by the mixing operator L^{-1} (shaping filter), while their level of sparsity is determined by the innovation w. The latter is uniquely characterized by its Lévy exponent

$$f(\omega) = \log \mathbb{E}\{\mathrm{e}^{\omega\langle \mathrm{rect}, w\rangle}\} = \log \hat{p}_{X_{\mathrm{rect}}}(\omega)$$

where $\hat{p}_{X_{\mathrm{rect}}}(\omega)$ is the characteristic function of the random variable $X_{\mathrm{rect}} = \langle \mathrm{rect}, w\rangle$ (canonical observation of the innovation through a rectangular window). When the input w is a white Gaussian noise (i.e., $f(\omega) = -|\omega|^2$), the model is able to generate the complete gamut of Gaussian stochastic processes, which are the only non-sparse members of the family. Another fundamental type of excitation is the impulsive noise $w_{\mathrm{Poisson}} = \sum_k a_k \delta(\cdot - x_k)$, where the impulse locations (x_k) follow a Poisson distribution with rate λ and the amplitudes (a_k) are i.i.d. with pdf p_A. The corresponding output signal (generalized Poisson process) is a random spline—the direct stochastic counterpart of (2) [26]. Other interesting instances of the model are the symmetric-α-stable (SαS) processes with $f(\omega) = -|\omega|^\alpha$, $\alpha \in (0, 2)$ [27].

IV. Splines and MMSE Reconstruction

Estimation theory tells us that the optimal reconstruction of the stochastic process s from its nonuniform samples $\{s(x_k)\}$ is given by the conditional mean

$$\tilde{s}(x) = \mathbb{E}\left\{s(x) | \{s(x_k), k \in \mathbb{Z}\}\right\}.$$

Moreover, when the process is Gaussian, the optimal reconstruction at x is known to be a linear combination of the measurement values: $\tilde{s}(x) = \sum_{k\in\mathbb{Z}} c_k(x) s(x_k)$ where the regression coefficients c_k are functions of the location x. These coefficients can be found by solving the so-called *normal equations* that involve the covariance function $C_s(x, y) = \mathbb{E}\{s(x)s(y)\}$ of the process. In the present scenario where the signal satisfies the innovation model (3), the covariance function is given by

$$C_s(x, y) = \sigma_w^2 (\mathrm{L}^{-1}\mathrm{L}^{-1*})\{\delta(\cdot - x)\}(\cdot - y)$$

where $\sigma_w^2 = -\frac{\mathrm{d}^2 f(0)}{\mathrm{d}\omega^2}$ is the variance of the noise. The crucial observation here is that $C_s(x, y)$ actually corresponds to the kernel $h_{\mathrm{L}^*\mathrm{L}}(x, y)$ associated with the inverse of the self-adjoint operator $(\mathrm{L}^*\mathrm{L})$. Based on the property that the latter is an $\mathrm{L}^*\mathrm{L}$-spline of the variable x with a single knot at y, it can be shown that the optimal reconstruction is a nonuniform $\mathrm{L}^*\mathrm{L}$-spline with knots at the sampling locations x_k. It follows that the optimal reconstruction is of the same form as (2) with the underlying Green's function ρ_{L} being substituted by $\rho_{\mathrm{L}^*\mathrm{L}}$. This leads to conclusion that the MMSE signal reconstruction is given by an $\mathrm{L}^*\mathrm{L}$-spline interpolant. This statistical optimality of splines is a classical result that has been used to justify the interpolation method known as kriging in geostatistics, and the use of reproducing kernels (radial-basis functions) for the interpolation of scattered data [28], [29], [30]. It is important to mention that this spline interpolant also yields the linear minimum-mean-square-error (LMMSE) estimator when the underlying process is non-Gaussian (with finite variance).

In the case where the data is uniformly sampled—and possibly corrupted by noise—the MMSE estimator under the assumption of stationarity amounts to a hybrid Wiener filter which has a convenient representation in terms of B-spline basis functions [31]. The approach can also be extended to the class of fractional Brownian motions, which are self-similar at the expense of some lack of stationarity [32]. Another related—and truly remarkable—result is that the piecewise-linear interpolator ($\mathrm{D}^*\mathrm{D}$-spline) is MMSE optimal not only for Brownian motion [33], but also for the complete family of (non-Gaussian) Lévy processes [34, Theorem 2].

For particular configurations of analysis functions, it is possible as well to obtain multi-spline extensions of such solutions for the generalized sampling problem; in particular, for the Hermite interpolation problem where the reconstruction is based on the samples of the function and its derivatives [35].

V. Sparse Processes and Finite-Rate of Innovation

The nonuniform L-spline described by (2) is the perfect example of a signal with a *finite rate of innovation*, which is non-bandlimitted, but can still be recovered from uniform samples provided that the signal is pre-filtered and sampled at a sufficient rate [36]. Alternatively, we can view such a signal as a realization of a generalized Poisson process which is the solution of the SDE (3) driven by impulsive noise [26]. The rate of innovation is then given by the Poisson parameter λ that represents the average number of Dirac impulses per unit length.

While such an explicit description of the solutions of (3) is not available for non-impulsive innovations, it is still possible to view $s = \mathrm{L}^{-1}w$ as a limit of a sequence of random L-splines with increasing rates of innovation (i.e., $\lambda \to \infty$) and some corrected amplitude distribution given by

$$p_{A,\lambda}(x) = \int_{\mathbb{R}} \mathrm{e}^{\frac{1}{\lambda} f(\omega)} \mathrm{e}^{-\mathrm{j}\omega x} \frac{\mathrm{d}x}{2\pi}$$

where $f(\omega)$ is the Lévy exponent of the innovation w. The relevant theory is developed in [37] for the class of CARn processes associated with the generic operator (1). In particular, we note that $\lambda\left(\mathrm{e}^{\frac{1}{\lambda}f(\omega)} - 1\right) = f(\omega) + O\left(\frac{f^2(\omega)}{\lambda}\right)$, which shows that the result is compatible with the compound-Poisson model for which $f_{\mathrm{Poisson}}(\omega) = \lambda\,(\hat{p}_A(\omega) - 1)$.

While this makes for an elegant link with splines, we should keep in mind that the rate of innovation alone is not necessarily a good predictor of the sparsity or compressibility of a signal. A striking example is provided by the family of SαS processes with $\alpha \in (0, 2]$ whose rate of innovation is infinite, but whose level of sparsity varies as $1/\alpha$, as discussed next.

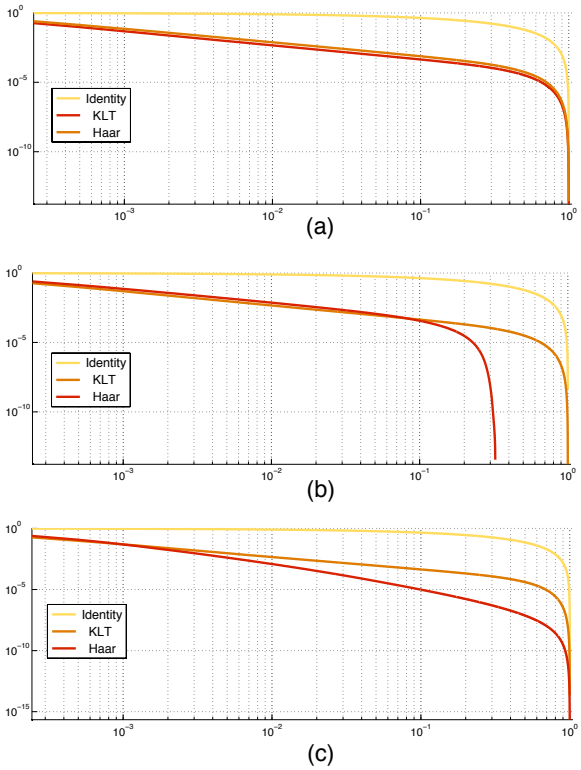

Fig. 1. Haar wavelets vs. KLT=DCT: M-term approximation errors for different brands of Lévy processes. The vertical axis represents the relative quadratic error and the horizontal one the relative number of transform coefficients. (a) Gaussian (Brownian motion). (b) Compound Poisson with Gaussian jump distribution and $\mathrm{e}^{-\lambda} = 0.9$. (c) Alpha-stable (symmetric Cauchy). The results are averages over 1000 realizations.

VI. Compressibility of Sparse Processes

The fundamental assumption that makes compressed sensing feasible is that the underlying signal admits a sparse representation in an appropriate basis. Remarkably, it is possible again to use L-splines to construct operator-like wavelet bases that provide a parsimonious representation of the sparse stochastic processes described in Section III [38], [39], [40]. For the Lévy processes, it holds that $\mathrm{L} = \mathrm{D}$, which corresponds to the Haar wavelet: the shortest wavelet with a derivative-like behavior. The graphs in Fig.1b-c illustrate the property that the wavelet decomposition yields a better M-term approximation than the DCT for the sparse varieties of Lévy processes. This is in contrast with the results in Figure 1a where the optimality of the Karhunen-Loève basis for the representation of a Gaussian process is confirmed. The latter is undistinguishable from the DCT. The signal in Figure 1b is the compound-Poisson process (random, piecewise-constant spline). It is a finite-rate-of-innovation signal that admits a perfect M-term wavelet approximation past some critical threshold. In the case of the third signal, which is SαS with $\alpha = 1$, the Haar wavelet transform always performs better than the DCT, so that it is arguably the sparsest of the lot.

By considering a variation of the model (3) where the excitation noise w is 2π-periodic, one can explain the above empirical observations by characterizing the Besov smoothness properties of sparse stochastic processes. Specifically, when L is an nth-order ordinary differential operator of the form (1) and $w = w_\alpha$ is an SαS innovation, then $s_\alpha = \mathrm{L}^{-1} w_\alpha$ can be shown to be included in the periodic Besov space $B^{n-1+1/\alpha}_{\alpha,\infty}([-\pi,\pi])$ with probability one [41]. By invoking the approximation properties of Besov spaces [42], this implies that

$$\|s_\alpha - s_{\alpha,M}\|_{L_2} = O(M^{-\tau_0}) \quad \text{with} \quad \tau_0 = n + \frac{1}{\alpha} - 1 - \epsilon$$

for any $\epsilon > 0$, where $s_{\alpha,M}$ denotes the M-term approximation of s_α in a suitable (e.g., wavelet-like) basis.

Acknowledgment

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 267439 and the Swiss National Science Foundation under Grant 200020-144355.

References

[1] A. Jerri, "Shannon sampling theorem—Its various extensions and applications—Tutorial review," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.
[2] I. J. Schoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA: Society of Industrial and Applied Mathematics, 1973.
[3] M. Unser, "Sampling—50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.
[4] A. Aldroubi and K. Gröchenig, "Nonuniform sampling and reconstruction in shift-invariant spaces," *SIAM Review*, vol. 43, pp. 585–620, 2001.
[5] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Transactions on Circuits and Systems—II: Analog and Digital Signal Processing*, vol. 45, no. 8, pp. 959–969, August 1998.
[6] J. Kybic, T. Blu, and M. Unser, "Generalized sampling: A variational approach—Part II: Applications," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1977–1985, August 2002.
[7] Y. C. Eldar, "Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors," *Journal of Fourier Analysis and Applications*, vol. 9, no. 1, pp. 77–96, 2003.
[8] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[9] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.
[10] B. Adcock and A. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," *Foundations of Computional Mathematics*, in press.
[11] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.
[12] U. S. Kamilov, P. Pad, A. Amini, and M. Unser, "MMSE estimation of sparse Lévy processes," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 137–147, January 1, 2013.
[13] C. Micchelli, "Cardinal L-splines," in *Studies in Spline Functions and Approximation Theory*, S. Karlin, C. Micchelli, A. Pinkus, and I. Schoenberg, Eds. Academic Press, 1976, pp. 203–250.
[14] M. H. Schultz and R. S. Varga, "L-splines," *Numerische Mathematik*, vol. 10, no. 4, pp. 345–369, 1967.
[15] M. Unser and P. D. Tafti, *An Introduction to Sparse Stochastic Processes*. Cambridge University Press, 2014.
[16] J. Kybic, T. Blu, and M. Unser, "Generalized sampling: A variational approach—Part I: Theory," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1965–1976, August 2002.
[17] P. Prenter, *Splines and Variational Methods*. New York: Wiley, 1975.

[18] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev spaces," in *Constructive Theory of Functions of Several Variables*, W. Schempp and K. Zeller, Eds. Berlin: Springer-Verlag, 1977, pp. 85–100.

[19] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, November 1999.

[20] M. Unser and T. Blu, "Cardinal exponential splines: Part I—Theory and filtering algorithms," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1425–1449, April 2005.

[21] H. W. Bode and C. E. Shannon, "A simplified derivation of linear least square smoothing and prediction theory," *Proceedings of the IRE*, vol. 38, no. 4, pp. 417–425, 1950.

[22] T. Kailath, "The innovations approach to detection and estimation theory," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 680–695, May 1970.

[23] M. Unser, P. Tafti, and Q. Sun, "A unified formulation of Gaussian versus sparse stochastic processes—Part I: Continuous-domain theory," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1945–1962, March 2014.

[24] A. Amini and M. Unser, "Sparsity and infinite divisibility," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2346–2358, April 2014.

[25] J. Fageot, A. Amini, and M. Unser, "On the continuity of characteristic functionals and sparse stochastic modeling," *The Journal of Fourier Analysis and Applications*, vol. 20, no. 6, pp. 1179–1211, December 2014.

[26] M. Unser and P. D. Tafti, "Stochastic models for sparse and piecewise-smooth signals," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 989–1005, March 2011.

[27] G. Samorodnitsky and M. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, 1994.

[28] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, no. 8, pp. 1246–1266, 1963.

[29] G. Kimeldorf and G. Wahba, "A correspondence between Bayesian estimation on stochastic processes and smoothing by splines," *The Annals of Mathematical Statistics*, vol. 41, no. 2, pp. 495–502, 1970.

[30] D. E. Myers, "Kriging, cokriging, radial basis functions and the role of positive definiteness," *Computers and Mathematics with Applications*, vol. 24, no. 12, pp. 139–148, 1992.

[31] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 2146–2159, June 2005.

[32] T. Blu and M. Unser, "Self-similarity: Part II—Optimal estimation of fractal processes," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1364–1378, April 2007.

[33] P. Lévy, *Le Mouvement Brownien*. Paris, France: Gauthier-Villars, 1954.

[34] A. Amini, P. Thévenaz, J. Ward, and M. Unser, "On the linearity of Bayesian interpolators for non-Gaussian continuous-time AR(1) processes," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 5063–5074, August 2013.

[35] V. Uhlmann, J. Fageot, H. Gupta, and M. Unser, "Statistical optimality of Hermite splines," in *Sampling Theory and Applications*, 2015, p. this issue.

[36] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.

[37] J. Fageot, J.-P. Ward, and M. Unser, "Interpretation of continuous-time autoregressive processes as random exponential splines," in *Sampling Theory and Applications*, 2015, p. this issue.

[38] I. Khalidov and M. Unser, "From differential equations to the construction of new wavelet-like bases," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1256–1267, April 2006.

[39] I. Khalidov, M. Unser, and J. Ward, "Operator-like wavelet bases of $L_2(\mathbb{R}^d)$," *The Journal of Fourier Analysis and Applications*, vol. 19, no. 6, pp. 1294–1322, December 2013.

[40] P. Pad and M. Unser, "On the optimality of operator-like wavelets for sparse AR(1) processes," in *Proceedings of the Thirty-Eighth IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'13)*, Vancouver BC, Canada, May 26-31, 2013, pp. 5598–5602.

[41] J.-P. Ward, J. Fageot, and M. Unser, "Compressibility of symmetric-α-stable processes," in *Sampling Theory and Applications*, 2015, p. this issue.

[42] R. A. Devore, "Nonlinear approximation," *Acta Numerica*, vol. 7, pp. 51–150, 1998.