Compressibility of symmetric-α-stable processes

John Paul Ward, Julien Fageot, and Michael Unser
Biomedical Imaging Group
Ecole polytechnique fédérale de Lausanne
CH-1015 Lausanne VD, Switzerland
Email: john.ward@epfl.ch

***Abstract*—Within a deterministic framework, it is well known that n-term wavelet approximation rates of functions can be deduced from their Besov regularity. We use this principle to determine approximation rates for symmetric-α-stable (SαS) stochastic processes. First, we characterize the Besov regularity of SαS processes. Then the n-term approximation rates follow. To capture the local smoothness behavior, we consider sparse processes defined on the circle that are solutions of stochastic differential equations.**

I. Introduction

Stochastic models are commonly used in engineering and financial applications [1], [2], [3], [4]. The most widely considered stochastic models are based on Gaussian probabilities. However, this assumption is too restrictive for many applications, and more general models are needed to accurately represent real data. Symmetric-α-stable (SαS) models directly generalize Gaussian models, and they maintain many desirable properties, such as satisfying a generalized central limit theorem [5]. The non-Gaussian stable laws have heavy-tailed distributions, so they are particularly relevant for applications [6]. Moreover, as we shall see, the family of SαS processes is defined by a parameter $\alpha \in (0, 2]$, which determines its sparsity with respect to wavelet representation. Here, we characterize sparsity by n-term wavelet approximation rates; i.e., one process is said to be sparser than another if it satisfies a faster rate of approximation.

Our interest in a sparse wavelet representation is compressibility; we can more efficiently store and transmit sparse processes by first transforming them into a wavelet domain. We shall show that Gaussian processes (corresponding to $\alpha = 2$) are the least sparse, and as α decreases, the sparsity increases. Sparsity is also a useful assumption in inverse imaging problems. If an image is known to be sparse a priori, then this information can be incorporated to produce a better reconstruction algorithm [7], [8], [9].

Our model is the stochastic differential equation

$$\mathrm{L}s = w, \tag{1}$$

where L is a differential operator, s is a stochastic process, and w is a SαS white noise process. The two defining components are the operator and the white noise. The white noise w is a random generalized function on the circle $\mathbb{T} = [-1/2, 1/2) \subset \mathbb{R}$. We consider standard differential-type operators of the following type: L is an order $\gamma > 0$ operator that reduces the regularity of a function by γ. More precisely, L maps the Besov space (of functions with mean zero) $B^{\tau}_{p,q}(\mathbb{T})$ isomorphically to $B^{\tau-\gamma}_{p,q}(\mathbb{T})$.

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

In the remainder of this section, we collect our notation. In Section 2, we recall some standard results from the theory of wavelet approximation. In Section 3, we specify SαS processes and review some of their properties. Section 4 is devoted to our main result: the n-term approximation of SαS processes, and we conclude in Section 5.

A. Notation and Background

The Lebesgue space $L_p(\mathbb{T})$ is the collection of functions for which

$$\|f\|_{L_p(\mathbb{T})} := \left(\int_{\mathbb{T}} |f(t)|^p \, \mathrm{d}t \right)^{1/p} \tag{2}$$

is finite. For $p \geq 1$, (2) is a norm, and for $0 < p < 1$, it is a quasi-norm.

The space of continuous functions on the circle is denoted as $C(\mathbb{T})$. The notation for the collection of functions with $k \in \mathbb{N}$ continuous derivatives on the circle is $C^k(\mathbb{T})$. The L_2 Sobolev space of order $k \in \mathbb{Z}$ is $H^k_2(\mathbb{T})$. For $k > 0$, $H^k_2(\mathbb{T})$ is the collection of functions with $k \in \mathbb{Z}$ distributional derivatives in $L_2(\mathbb{T})$.

The Schwartz space of infinitely differentiable test functions on the circle $\mathbb{T}$ is denoted as $\mathcal{S}(\mathbb{T})$. The corresponding space of generalized functions is $\mathcal{S}'(\mathbb{T})$. These spaces are nuclear spaces, and

$$\mathcal{S}(\mathbb{T}) = \bigcap_{k \in \mathbb{Z}} H^k_2(\mathbb{T}) \tag{3}$$

$$\mathcal{S}'(\mathbb{T}) = \bigcup_{k \in \mathbb{Z}} H^k_2(\mathbb{T}). \tag{4}$$

For our purpose, we consider the related spaces with mean zero. Such generalized functions are well suited to wavelet approximation rates, since wavelets also have mean zero. In addition, this assumption simplifies the definition of stochastic processes. For example, the distributional derivative becomes a bijective mapping of Sobolev and Besov spaces of functions with mean zero.

We define the nuclear space of infinitely differentiable test functions with mean zero to be $\mathcal{N}(\mathbb{T})$. The corresponding space of generalized functions (continuous linear functionals on $\mathcal{N}(\mathbb{T})$) is $\mathcal{N}'(\mathbb{T})$.

We have introduced $\mathcal{N}'(\mathbb{T})$ because nuclear spaces are the ideal infinite dimensional extension of finite dimensional vector spaces. Many of the properties associated with finite dimensional vector spaces, for example Bochner's theorem, have a generalization to this setting. The Bochner-Minlos theorem implies that a continuous, positive definite functional C on $\mathcal{N}(\mathbb{T})$ is the Fourier transform of a finite Radon measure μ on $\mathcal{N}'(\mathbb{T})$; i.e., there is a measure μ such that

$$C(\varphi) = \int_{\mathcal{N}'(\mathbb{T})} \exp\{i \langle T, \varphi \rangle\} \, \mathrm{d}\mu(T) \tag{5}$$

for $\varphi \in \mathcal{N}(\mathbb{T})$. The functionals $\exp\left\{-\|\varphi\|^{\alpha}_{L_\alpha(\mathbb{T})}\right\}$ are positive definite and continuous for $0 < \alpha \leq 2$, and they are the characteristic functionals for SαS white noises, cf. Section III.

II. Wavelets, Besov spaces, and n-term approximation

A. Daubechies Wavelets on the Circle

There are several different equivalent definitions of Besov spaces. We choose to use the wavelet-based definition since it is suited to our presentation. Essentially, it says that a generalized function on $\mathbb{T}$ is in a given Besov space if its appropriately normalized wavelet coefficients are in a Besov sequence space.

Partly due to their compact support, we consider Daubechies wavelets [10]. By periodizing the standard Daubechies wavelets, we obtain an orthonormal basis of $L_2(\mathbb{T})$

$$\left\{ \Psi^{j,k}_{G,m} = 2^{j/2} \Psi^{0,k}_{G,0}(2^j \cdot - m) \;\middle|\; j \in \mathbb{N}_{\geq 0}, G \in G^j, m \in \mathbb{P}_j \right\}. \tag{6}$$

The index $j \in \mathbb{N}_{\geq 0}$ corresponds to a scaling parameter, and G is used to denote gender (M for mother or F for father). The coarsest scale is $j = 0$, which includes the scaling functions, so $G^0 = \{M, F\}$ has 2 elements and $G^j = \{M\}$ has one element for $j > 0$. The parameter k denotes the smoothness of the wavelet, and it determines the support of the wavelet. For $k > 0$, the classical Daubechies mother wavelet on the real line has support greater than one. Here, we require the support of the wavelets to be a proper subset of the circle. Consequently, the coarsest scale is scaled by 2^L, where the parameter $L \in \mathbb{N}$ ensures that this condition is satisfied. For the remainder of this paper, we set L (as a function of k) to be the smallest integer that guarantees this support condition. The wavelet translates are indexed by m, and the set of translations at scale j is

$$\mathbb{P}_j = \left\{ m \in \mathbb{Z} \;\middle|\; 0 \leq m < 2^{j+L} \right\}. \tag{7}$$

More details on the adaptation of wavelet bases for periodic domains can be found in [11, Section 1.3]. In fact, we follow the notation of that book. Our $\Psi^{j,k}_{G,m}$ corresponds to $\Psi^{j,\text{per}}_{G,m}$ of [11, Proposition 1.34].

Definition 1. *The notation $\Psi^{0,k}_{M,0}$ denotes a Daubechies mother wavelet in $C^k(\mathbb{T})$. Furthermore, the Lebesgue measure of the support of $\Psi^{0,k}_{M,0}$ is less than one.*

The wavelet decomposition of a sufficiently smooth generalized function f is

$$f = \sum_{j,G,m} \lambda^{j,G}_m 2^{-(j+L)/2} \Psi^{j,k}_{G,m} \tag{8}$$

$$\lambda^{j,G}_m = \left\langle f, 2^{(j+L)/2} \Psi^{j,k}_{G,m} \right\rangle. \tag{9}$$

B. Besov Spaces of Zero Mean

We now come to the definition of Besov spaces $B^\tau_{p,q}(\mathbb{T})$, which are defined by two primary parameters, p and τ, and one secondary (minor adjustment) parameter q. The parameter $p \in (0, \infty]$ is similar to the index defining the Lebesgue spaces $L_p(\mathbb{T})$, and $\tau \in \mathbb{R}$ indicates smoothness. Therefore, roughly speaking, for $\tau \in \mathbb{N}$, a function in $B^\tau_{p,q}(\mathbb{T})$ has τ derivatives in $L_p(\mathbb{T})$. In Figure 1, we provide a structural diagram, representing the collection of Besov spaces.

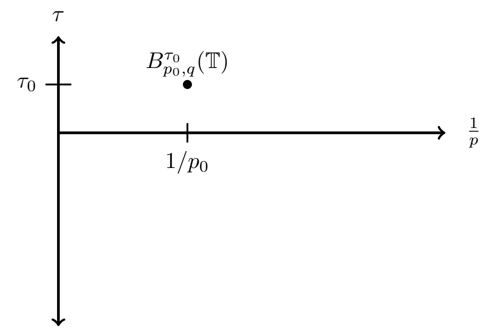

Fig. 1. Function space diagram. A point $(1/p_0, \tau_0)$ represents all of the Besov spaces $B^{\tau_0}_{p_0,q}(\mathbb{T})$ for $0 < q \leq \infty$.

Definition 2 (Definition 1.3.2, [11])**.** *Let $\tau \in \mathbb{R}$ and $0 < p, q < \infty$. The Besov sequence space $b^\tau_{p,q}$ is the collection of sequences*

$$\lambda = \left\{ \lambda^{j,G}_m \;\middle|\; j \in \mathbb{N}_{\geq 0}, G \in G^j, m \in \mathbb{P}_j \right\}, \tag{10}$$

indexed as described above, with finite (quasi-)norm

$$\|\lambda\|_{b^\tau_{p,q}} := \left(\sum_{j=0}^{\infty} 2^{j(\tau - 1/p)q} \sum_{G \in G^j} \left(\sum_{m \in \mathbb{P}_j} \left|\lambda^{j,G}_m\right|^p \right)^{q/p} \right)^{1/q}. \tag{11}$$

If $p = \infty$ or $q = \infty$, there is an analogous definition.

In the special case $p = q$, we have

$$\|\lambda\|_{b^\tau_{p,p}} := \left(\sum_{j,G,m} \left| 2^{j(\tau - 1/p)} \lambda^{j,G}_m \right|^p \right)^{1/p}, \tag{12}$$

which is a weighted ℓ_p space. We shall work with these spaces, which have simplified norms, and derive more general results using embedding properties between Besov spaces.

Definition 3. *Let $k \in \mathbb{N}$, $0 < p, q \leq \infty$, and $\tau \in \mathbb{R}$ such that*

$$k > \max\{\tau, (1/p - 1)_+ - \tau\}. \tag{13}$$

A generalized function $f \in \mathcal{N}'(\mathbb{T})$ is in the Besov space $B^{\tau}_{p,q}(\mathbb{T})$ if it can be represented as

$$f = \sum_{j,G,m} \lambda_m^{j,G} 2^{-(j+L)/2} \Psi^{j,k}_{G,m} \tag{14}$$

for some sequence $\lambda \in b^{\tau}_{p,q}$.

Criteria for unconditional convergence and many additional properties can be derived. We refer the interested reader to [11, Section 1.3].

The above definition is somewhat unsatisfactory, since the sequence λ is not specified. However, if a preliminary bound on the smoothness of a generalized function f is known, then we can choose sufficiently smooth wavelets (by specifying a sufficiently large parameter k) to obtain the following result.

Proposition 4. *Suppose $f \in B^{-k}_{2,2}(\mathbb{T})$ (which is equivalent to the Sobolev space $H_2^{-k}(\mathbb{T})$) for some $k \in \mathbb{N}$, and let $0 < p, q \leq \infty$ and $\tau \in \mathbb{R}$ such that $k > \max\{\tau, (1/p-1)_+ - \tau\}$. Then $f \in B^{\tau}_{p,q}(\mathbb{T})$ if and only if the wavelet coefficients*

$$\left\langle f, 2^{(j+L)/2} \Psi^{j,k+1}_{G,m} \right\rangle \tag{15}$$

are in the Besov sequence space $b^{\tau}_{p,q}$.

C. n-term Approximation

Besov spaces also have a classical definition. One of the chief advantages of the definition above is to show that they are isomorphic to sequence spaces. This reduces the question of n-term approximation to the study of sequence spaces.

In [12], the authors cover several important n-term approximation results. In particular, their Section 8 describes how to obtain approximation rates in sequence spaces.

Definition 5. *Let $f \in L_2(\mathbb{T})$, and suppose we have wavelets that are sufficiently smooth. Then $\Sigma_n(f)$ denotes the best n-term wavelet approximation to f, where the error is measured in the $L_2(\mathbb{T})$ norm.*

Basically, $\Sigma_n(f)$ is obtained as follows. The function f is expanded in a wavelet basis of smoothness k

$$f = \sum_{j,G,m} \lambda_m^{j,G} 2^{-(j+L)/2} \Psi^{j,k}_{G,m}. \tag{16}$$

Then, the collection of the n largest wavelet coefficients are identified. Let Λ_n denote the index triples (j, G, m) associated with these coefficients. Then the n-term approximation is

$$\Sigma_n(f) = \sum_{(j,G,m)\in\Lambda_n} \lambda_m^{j,G} 2^{-(j+L)/2} \Psi^{j,k}_{G,m}. \tag{17}$$

The following result can be derived using the techniques described in [12].

Lemma 6. *If $f \in B^{1/p-1/2}_{p,p}(\mathbb{T})$ for some $p < 2$, then there is a constant $C > 0$ such that*

$$\|f - \Sigma_n(f)\|_{L_2(\mathbb{T})} \leq C n^{-(1/p-1/2)} \|f\|_{B^{1/p-1/2}_{p,p}(\mathbb{T})}. \tag{18}$$

III. White noise and sparse processes

A. Definition of SαS White Noise

Our definition of continuous SαS white noises is analogous to the finite dimensional case, which we recall here. A probability density function (pdf) $\mathscr{P}_X$ for a scalar-valued random variable X is SαS with scale $c > 0$ if its characteristic function $\widehat{\mathscr{P}}_X$ is of the form $\exp\{-c^\alpha |\cdot|^\alpha\}$; i.e.,

$$\widehat{\mathscr{P}}_X(\omega) = \mathbb{E}\left\{e^{i\omega X}\right\} = e^{-c^\alpha |\omega|^\alpha}. \tag{19}$$

A vector of independent, identically distributed (i.i.d.) SαS random variables $\boldsymbol{X} = (X_1, \dots, X_N)$ is called a discrete white noise process. In this case, the characteristic function for $\boldsymbol{X}$ is

$$\widehat{\mathscr{P}}_{\boldsymbol{X}}(\boldsymbol{\omega}) = e^{-c^\alpha \sum_{n=1}^N |\omega_n|^\alpha} \tag{20}$$

$$= e^{-c^\alpha \|\boldsymbol{\omega}\|^\alpha_{\ell_\alpha}}. \tag{21}$$

Note the simple form for the characteristic function, which fully identifies the pdf of the random variable.

We define continuous white noise processes on the circle $\mathbb{T}$ in an analogous way: via characteristic functionals. Specifically, a continuous SαS white noise w_α on $\mathbb{T}$ is a random generalized function. For every test function $\varphi \in \mathcal{N}(\mathbb{T})$, the random variable $\langle w_\alpha, \varphi\rangle$ follows a SαS probability law. Also, if two test functions have disjoint support, then the corresponding random variables are independent.

Definition 7. *A SαS white noise w_α is a random process with characteristic functional*

$$\widehat{\mathscr{P}}_{w_\alpha}(\varphi) = \exp\left\{-c^\alpha \|\varphi\|^\alpha_{L_\alpha(\mathbb{T})}\right\},$$

for some $c > 0$.

The existence of w_α is simplified by the Bochner-Minlos theorem, which implies that $\widehat{\mathscr{P}}_{w_\alpha}$ specifies a probability measure on $\mathcal{N}'(\mathbb{T})$. Further information on SαS white noises and their properties can be found in [9].

B. Regularity of SαS White Noise

There are many benefits in defining a white noise through its characteristic functional; however, regularity estimates do not immediately follow. Such results are important, and we provide a bound below that is derived using wavelet techniques. Our result uses the fact that Besov spaces are measurable with respect to the inverse Fourier transform of $\widehat{\mathscr{P}}_{w_\alpha}$, which is a Radon probability measure on $\mathcal{N}'(\mathbb{T})$.

Theorem 8. *The sample paths of a SαS white noise, with $0 < \alpha < 2$, are in the Besov space $B^{1/\alpha-1-\epsilon}_{\alpha,\alpha}(\mathbb{T})$ with probability one, for every $\epsilon > 0$.*

For the Gaussian case $\alpha = 2$, there is a slight modification; the sample paths are in the Besov spaces $B^{-1/2-\epsilon}_{p,p}(\mathbb{T})$ with $p = \infty$ rather than $p = 2$.

Theorem 9. *The sample paths of a SαS Gaussian white noise ($\alpha = 2$) are in the Besov space $B^{-1/2-\epsilon}_{\infty,\infty}(\mathbb{T})$ with probability one, for every $\epsilon > 0$.*

Proofs for these regularity results will be presented in [13].

We apply embedding theorems to determine a collection of Besov spaces that also contains the sample paths with probability one. These spaces are depicted in Figure 2 and Figure 3. For Figure 2, we use two embedding theorems: (1) $B^{1/\alpha-1-\epsilon}_{\alpha,\alpha}(\mathbb{T}) \subset B^{1/\alpha-1-\epsilon'}_{p,q}(\mathbb{T})$ for $p \le \alpha$ and $\epsilon' > \epsilon$, and (2) $B^{1/\alpha-1-\epsilon}_{\alpha,\alpha}(\mathbb{T}) \subset B^{\tau}_{p,q}(\mathbb{T})$ for $p > \alpha$ and $(1/p, \tau)$ below the Sobolev embedding line (slope one) passing through $(1/\alpha, 1/\alpha - 1 - \epsilon)$ [14, Section 3.5]. Only the former embedding theorem is required for the Gaussian case.

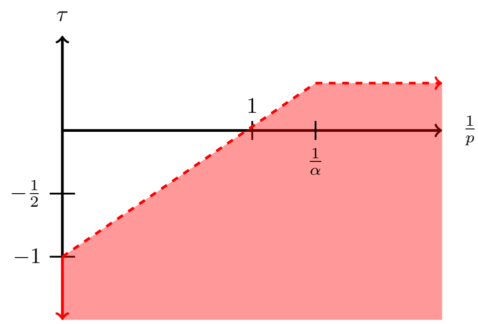

Fig. 2. The shaded region comprises the Besov spaces having probability one with respect to the SαS $(0 < \alpha < 2)$ Radon probability measure.

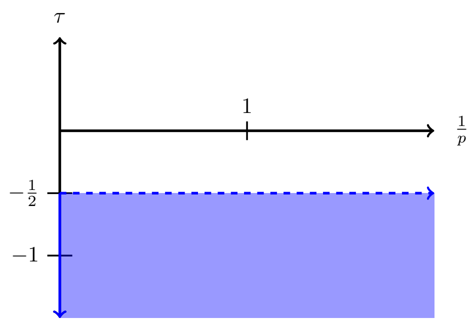

Fig. 3. The shaded region comprises the Besov spaces having probability one with respect to the SαS (Gaussian, $\alpha = 2$) Radon probability measure.

C. SαS Processes

A SαS process s, satisfies (1), where w is a SαS white noise.

Definition 10. *We say that a Fourier multiplier operator* $\mathrm{L} : \mathcal{N}'(\mathbb{T}) \to \mathcal{N}'(\mathbb{T})$ *is admissible of order* $\gamma > 0$ *if it isomorphically maps* $B^{\tau}_{p,q}(\mathbb{T})$ *to* $B^{\tau-\gamma}_{p,q}(\mathbb{T})$ *for all choices of the parameters* p, q, τ.

This is a very natural class for defining stochastic processes: it is very general and contains all of the standard families of differential operators. It comprises all operators that reduce the Besov or Sobolev regularity of a function by some increment $\gamma > 0$. Examples of admissible operators include the distributional derivative D (order $\gamma = 1$); polynomials in D

$$\mathrm{D}^N + a_{N-1}\mathrm{D}^{N-1} + \cdots + a_1\mathrm{D} + \mathrm{I} \tag{22}$$

of order $\gamma = N$; the fractional Laplacian $(-\Delta)^{\gamma/2}$; and the Matérn operators $(1-\Delta)^{\gamma/2}$.

The fact that we consider spaces of functions with zero mean is important here. Otherwise, the derivative and Laplacian operators would not be isomorphisms, and these cases would involve additional technicalities.

For certain operators L, the resulting processes are classical. When $\mathrm{L} = \mathrm{D}$, s is a Lévy process. In general, when L is of the form (22), s is an AR(N) process (autoregressive process of order N) [15], [16].

Theorem 11. *The sample paths of a SαS stochastic process* s *defined by a SαS white noise* w_α $(0 < \alpha < 2)$ *and an order* $\gamma > 0$ *admissible operator* L *are in the Besov space* $B^{\gamma+1/\alpha-1-\epsilon}_{\alpha,\alpha}(\mathbb{T})$ *with probability one, for every* $\epsilon > 0$.

The sample paths of an analogous Gaussian $(\alpha = 2)$ *process* s *are in* $B^{\gamma-1/2-\epsilon}_{\infty,\infty}(\mathbb{T})$ *with probability one, for every* $\epsilon > 0$.

IV. n-TERM APPROXIMATION OF SαS PROCESSES

Using the Besov regularity of SαS processes from the last section, we immediately deduce their n-term wavelet approximation rates from Lemma 6.

Theorem 12. *Let* $0 < \alpha \le 2$ *and* $\gamma > 1/2$. *The* n*-term wavelet approximation for a sample path of an order* γ *SαS process satisfies the following rate of decay. For any* $0 < \epsilon < \gamma - 1/2$, $\tau_0 := \gamma + 1/\alpha - 1 - \epsilon$, *and* $1/p := \tau_0 + 1/2$, *we have*

$$\|s - \Sigma_n(s)\|_{L_2(\mathbb{T})} \le C n^{-\tau_0} \|s\|_{B^{\tau_0}_{p,p}(\mathbb{T})}$$

with probability one.

The remainder of this section is devoted to a discussion of this result. In Theorem 11, we observe that the Gaussian result is not what is expected by taking the limit $(\alpha \to 2)$ of the SαS case. However, this does not impact the results of Theorem 12, where the Gaussian result matches the limiting behavior.

As expected, the rate of decay of the n-term bound is increasing with γ, when α is fixed. This correlates with the principle that smoother functions satisfy higher approximation rates.

For fixed γ, the rate of decay of the n-term bound increases monotonically as α goes to zero. We interpret this to mean that the Gaussian processes are the least *sparse* among the collection of SαS processes. Finally, note that $\tau_0 \to +\infty$ when $\alpha \to 0$, so for small α, SαS processes are highly compressible.

V. CONCLUSION

We have presented n-term wavelet approximation rates for SαS stochastic processes on the circle $\mathbb{T}$, which are based on Besov regularity properties of the underlying white noise. Our analysis indicates that Gaussian processes have the least sparse representation within the specified family.

We considered processes on the circle; however, there is a local nature to our results which should carry over to the real line and more general classes of domains.

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.

REFERENCES

[1] R. A. Davis, K. Knight, and J. Liu, "M-estimation for autoregressions with infinite variance," *Stochastic Processes and their Applications*, vol. 40, no. 1, pp. 145–180, 1992.

[2] P. Kidmose, "Alpha-stable distributions in signal processing of audio signals," in *41st Conference on Simulation and Modelling*, 2000, pp. 87–94.

[3] C. L. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*. Wiley-Interscience, 1995.

[4] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[5] G. Samoradnitsky and M. S. Taqqu, *Stable non-Gaussian random processes: stochastic models with infinite variance*. CRC Press, 1994, vol. 1.

[6] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: stable processes and their applications," *Proceedings of the IEEE*, vol. 81, no. 7, pp. 986–1010, 1993.

[7] E. Bostan, U. S. Kamilov, M. Nilchian, and M. Unser, "Sparse stochastic processes and discretization of linear inverse problems," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2699–2710, 2013.

[8] M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*. Springer, 2010.

[9] M. Unser and P. D. Tafti, *An Introduction to Sparse Stochastic Processes*. Cambridge University Press, 2014.

[10] I. Daubechies, *Ten lectures on wavelets*. SIAM, 1992, vol. 61.

[11] H. Triebel, *Function spaces and wavelets on domains*, ser. Tracts in Mathematics. European Mathematical Society, 2008, vol. 7.

[12] A. Cohen, R. A. DeVore, and R. Hochmuth, "Restricted nonlinear approximation," *Constructive Approximation*, vol. 16, no. 1, pp. 85–113, 2000.

[13] J. Fageot, M. Unser, and J. P. Ward, "On the Besov regularity of periodic Lévy noises," under review.

[14] H.-J. Schmeisser and H. Triebel, *Topics in Fourier analysis and function spaces*. Wiley Chichester, 1987.

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

[16] H. Kirshner, M. Unser, and J. P. Ward, "On the unique identification of continuous-time autoregressive models from sampled data," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1361–1376, 2014.