On the Linearity of Bayesian Interpolators for Non-Gaussian Continuous-Time AR(1) Processes

Arash Amini, Philippe Thévenaz, John Paul Ward, and Michael Unser, *Fellow, IEEE*

***Abstract*—Bayesian estimation problems involving Gaussian distributions often result in linear estimation techniques. Nevertheless, there are no general statements as to whether the linearity of the Bayesian estimator is restricted to the Gaussian case. The two common strategies for non-Gaussian models are either finding the best linear estimator or numerically evaluating the Bayesian estimator by Monte Carlo methods. In this paper, we focus on Bayesian interpolation of non-Gaussian first-order autoregressive (AR) processes where the driving innovation can admit any symmetric infinitely divisible distribution characterized by the Lévy-Khintchine representation theorem. We redefine the Bayesian estimation problem in the Fourier domain with the help of characteristic forms. By providing analytic expressions, we show that the optimal interpolator is linear for all symmetric α-stable distributions. The Bayesian interpolator can be expressed in a convolutive form where the kernel is described in terms of exponential splines. We also show that the limiting case of Lévy-type AR(1) processes, the system of which has a pole at the origin, always corresponds to a linear Bayesian interpolator made of a piecewise linear spline, irrespective of the innovation distribution. Finally, we show the two mentioned cases to be the only ones within the family for which the Bayesian interpolator is linear.**

***Index Terms*—Alpha-Stable Innovation, Autoregressive, Bayesian Estimator, Interpolation, Ornstein-Uhlenbeck Process.**

I. Introduction

Auto-Regressive (AR) processes are well-studied models in statistics and signal processing. They are used to model real-world signals such as speech [1], among others. These processes can be generated by applying an all-pole filter on white noise (innovation). The problem which is often studied in signal processing and spectral estimation is to recover the parameters of the model, for instance the poles of the shaping filter, based on a finite number of measurements [2]. The estimated parameters can be used to further predict the process [3].

Conventional AR models are founded on Gaussian innovations, and most of the results are obtained with this assumption. However, there are cases in real-world applications, such as stock-market data, where the asymptotic behavior is non-Gaussian [4], [5]. Early investigations of the non-Gaussian case can be found in [6]; the research work in this field is still ongoing [7], [8].

Manuscript received June 14, 2012 and revised January 21, 2013.

The authors are with the Biomedical Imaging Group (BIG), École polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland.

Emails: {arash.amini, philippe.thevenaz, john.ward and michael.unser}@epfl.ch.

This work is supported by the European Research Center (ERC) under grant ERC-2010-AdG 267439-FUN-SP and Swiss National Science Foundation under Grant 200020-121763.

In many applications there is a need to estimate some unknown values based on observations that contain related information. In a Bayesian framework where we know the statistics of the unknowns and observations, one may think of the posterior mean (alternatively known as the regression of the unknown on the observations) as the optimal estimator. In fact, the posterior mean estimator, which is also referred to as the Bayesian filter, minimizes the least-square error whenever it is finite.

Bayesian filtering of AR processes is traditionally studied in finite-dimensional state-space. In this approach, the main goal is to estimate the state vector, which automatically yields the desired information. The best-known example of this technique is the Kalman filter, which is the Bayesian filter when the innovation process is Gaussian. The Kalman filter also works under noisy measurements, where the noise is additive and Gaussian. Some extensions to non-Gaussian and heavy-tailed noises are studied in [9], [10]. In this paper, however, we exclusively focus on the noiseless scenario and the derivation of closed form solutions for AR(1) interpolators.

The main difficulty for obtaining the Bayesian filter for non-Gaussian innovations is that there are very few cases where an explicit form for the posterior distribution exists. Indeed, early works in the non-Gaussian case often started by approximating the posterior distribution [11]. Among the notable methods one can name the Gaussian sum filter [12] and the extended Kalman filter [13]. Further extensions of the Bayesian filter in cases where the process follows a dynamic generalized linear model are investigated in [14].

Instead of evaluating the Bayesian filter for an approximate posterior prior, which is usually a linear estimator, one can look for the optimal linear estimator in the sense of the least-square error. The corresponding methods are called linear least-square estimators. The linear estimator is frequently expressed in terms of the covariance function; the Wiener filter, used in denoising applications, is a typical example. For the interpolation problem, the link between interpolating splines and optimal linear estimators for stochastic processes has been established in [15], [16].

More recent approaches towards achieving the Bayesian filter rely on numerical techniques [17]. The Monte Carlo methods are among the most successful candidates [18], [19].

A. Scope

In this paper, we focus on real-valued continuous-time AR processes of order 1 with various innovation distributions and investigate the problem of Bayesian interpolation between the

measurements. More precisely, we are interested in estimating $s(x_0)$, where s is an AR(1) process from which we have observed the samples $s(\bar{x}_0), \ldots, s(\bar{x}_m)$ with $\bar{x}_0 < \cdots < \bar{x}_m$. We assume $x_0 \in [\bar{x}_0, \bar{x}_m]$ to avoid extrapolation, and we use the posterior mean estimator given by

$$\hat{s}(x_0) = \mathbb{E}\big\{s(x_0) \,\big|\, \{s(\bar{x}_k)\}_{k=0}^{m}\big\}.$$

The distinguishing property of this stochastic family is that the AR(1) processes (not necessarily Gaussian), together with the Lévy processes, form the set of all Markovian processes. These processes are referred to in statistical finance as the Ornstein-Uhlenbeck processes when the distribution is Gaussian [20], while there are non-Gaussian generalizations as well [21]. These processes are widely adopted models for interest rates and currency exchange rates [20]. Among other applications, one can name derivative securities [22], electricity pricing [23], and pairs trading [24]. The interpolation problem is motivated by the fact that, in some financial applications, such as hedging, one needs to represent the available discrete-space data in the form of smooth curves [25]. Besides, interpolation is also a crucial operation in image processing, and its performance can be improved substantially by tuning the basis functions to some underlying AR model [26].

Due to the continuous-time definition of the process, we can no longer apply the common finite-dimensional state-space approaches. We adopt instead the generalized-function approach to stochastic processes presented in [27] which demonstrates a link between spline theory and Gaussian/sparse stochastic processes.

B. Contribution

The problem studied in this paper is to determine, for various innovation statistics, whether the Bayesian interpolator is a linear function of the measurements. Our surprising conclusion is that the linearity of the Bayesian interpolator is not limited to Gaussian innovations. More precisely, we show that the estimator is linear for all stable innovations. Furthermore, at each point it depends only on the two neighboring measurements.

We explicitly derive the expression of the optimal interpolator. For non-stable innovations, however, we show that the Bayesian interpolator cannot be a linear function of the measurements, with the exception of Lévy processes. In fact, the optimal interpolator for Lévy processes is always the linear B-spline, irrespective of the innovation. Although it is natural to think of extending the results to AR(n), the tools used in this paper only address the case of AR(1).

C. Outline

The rest of the paper is organized as follows: To provide the stochastic framework upon which our results are based, we revisit the continuous-time AR(1) processes in Section II, from the perspective of generalized functions as in [27]. In Section III, we summarize our contributions in the form of three theorems. Next, we present fundamental properties of AR(1) processes and their benefit in deriving the Bayesian interpolator in Sections IV and V, respectively. The proofs of the theorems stated in Section III are provided in Section VI. We test the performance of the interpolators for stable innovations in Section VII, where we apply them to a few realizations. Finally, Section VIII concludes the paper.

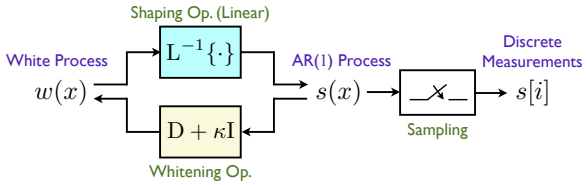

Fig. 1. Generation of the stochastic AR(1) process s based on the excitation white noise w. The inverse linear operator L^{-1} includes the possible boundary condition.

II. AR(1) Model

The model in this paper is a special case of [27] adapted for AR(1) processes. The schematic diagram of the continuous-time model is given in Figure 1. The process of interest s is formed by the innovation w as

$$\frac{\mathrm{d}}{\mathrm{d}x}s(x) + \kappa s(x) = w(x), \tag{1}$$

where w is a stationary white process, $\kappa \in \mathbb{R}$ is a constant, and the derivative operator is interpreted in the weak sense of generalized functions. Equation (1) suggests the filter $\mathrm{D} + \kappa\mathrm{I}$ as the whitening operator L, where D and I stand for the derivative and identity operators, respectively. This whitening operator has a one-dimensional null space spanned by the function $\mathrm{e}^{-\kappa x}$. Therefore, the shaping operator denoted by L^{-1} (inverse of the whitening operator), which transforms the innovations into the main process, is not uniquely defined.

A. Shaping Operator

The system $\mathrm{D} + \kappa\mathrm{I}$ has a unique stable right inverse for $\kappa \neq 0$. The stable inverse, which we use as the shaping operator, is shift-invariant and corresponds to the impulse response $\mathrm{e}^{-\kappa x}\chi_{\mathbb{R}_0^+}(x)$ for $\kappa > 0$, and $\mathrm{e}^{-\kappa x}\chi_{\mathbb{R}_0^+}(-x)$ for $\kappa < 0$, where $\chi_{\mathbb{R}_0^+}(\cdot)$ denotes the characteristic function of the nonnegative real numbers (step function). In either case, the impulse response is represented by $\frac{1}{\mathrm{j}\omega+\kappa}$ in the Fourier domain. The use of a linear shift-invariant operator results in a stationary process s.

The main difficulty for $\kappa = 0$ is that the corresponding whitening operator is D, whose shift-invariant inverse (respectively, adjoint inverse operator) is not stable. Therefore, the shaping operator L^{-1} cannot be a convolutive operator, which implies that the resultant process s will not be stationary [27]. Lévy processes are special cases associated with the shaping operator $\mathrm{L}^{-1} = \int_0^x$. This choice imposes $s(0) = 0$, which acts as a boundary condition. More generally, we can consider the form

$$\mathrm{L}^{-1}\{w\}(x) = \int_0^x w(\tau)\mathrm{d}\tau + \langle \phi, w \rangle, \tag{2}$$

where ϕ is an anti-causal function that decreases rapidly and $\langle \phi, w \rangle = \int w(\tau)\phi(\tau)\mathrm{d}\tau$ in the sense of generalized functions.

The existence of ϕ in (2) allows for arbitrary but linear boundary conditions. The anti-causal choice of ϕ shows that, for all $x > 0$, the random variable $\mathrm{L}^{-1}\{w\}(x)$ is statistically independent of $w(\tau)$ for $\tau > x$. This will later help us in simplifying the estimation procedure.

Since the innovation process is white and the impulse response of the shaping operator for $(-\kappa)$ is the time-reversal of the one for $\kappa \neq 0$, we expect to obtain the interpolation results of $(-\kappa)$ by time-reversing the results for κ. This is, indeed, confirmed by the structure of the interpolation kernels in Theorem 1. Thus, without loss of generality, we shall assume $\kappa \geq 0$ in the rest of the paper.

B. Innovation Process

To describe the family of white processes (innovations), we need to use Gelfand's theory of generalized random processes. In this approach, the process is characterized by the statistics of its inner products with a set of test functions. The set of acceptable test functions is a function space $\mathcal{E}$ which usually includes a set of functions with finite support. Let $\varphi \in \mathcal{E}$ be an arbitrary test function and w denote the continuous-time white process. The inner product $\langle \varphi, w \rangle = \int_{\mathbb{R}} \varphi(\tau) w(\tau) \mathrm{d}\tau$ defines a random variable for which we can compute the characteristic form. It turns out that the mapping between test functions and characteristic forms is key in defining the process in Gelfand's theory. The characteristic form of a process w is defined as

$$\forall \varphi \in \mathcal{E}: \ \widehat{\mathscr{P}}_w(\varphi) = \mathbb{E}\left\{ \mathrm{e}^{\mathrm{j}\langle \varphi, w \rangle} \right\}, \tag{3}$$

where $\mathbb{E}\{\cdot\}$ stands for the expected-value operator. It is not hard to check that $\widehat{\mathscr{P}}_w(0) = 1$ and $\widehat{\mathscr{P}}_w$ is a positive-definite functional over $\mathcal{E}$. According to the Bochner-Minlos theorem, the converse is also true: Under suitable conditions[1] on $\mathcal{E}$, any positive-definite functional $\widehat{\mathscr{P}}_w$ over $\mathcal{E}$ such that $\widehat{\mathscr{P}}_w(0) = 1$ defines a unique probability measure on the dual space of $\mathcal{E}$ (existence and uniqueness of the process). Thus, the characteristic form is an alternative way of uniquely specifying the statistics of a stochastic process.

For the particular case of white processes, Gelfand and Vilenkin consider the generic form

$$\widehat{\mathscr{P}}_w(\varphi) = \mathrm{e}^{\int_{\mathbb{R}} f(\varphi(\tau)) \mathrm{d}\tau}, \tag{4}$$

where f is a scalar function. This function, which is usually referred to as the Lévy exponent, is not arbitrary and should satisfy some constraints. The Lévy-Khintchine representation theorem states the necessary and sufficient conditions for the function f to define a valid characteristic form as [28]

$$\begin{aligned} f(\omega) &= \mathrm{j} b_1 \omega - b_2^2 \omega^2 \\ &+ \int_{\mathbb{R}\setminus\{0\}} \left(\mathrm{e}^{\mathrm{j}a\omega} - 1 - \mathrm{j}a\omega \mathbb{1}_{|a|<1}(a) \right) V(\mathrm{d}a), \end{aligned} \tag{5}$$

where $\mathbb{1}_{|a|<1}(a) = 1$ for $|a| < 1$ and 0 otherwise, b_1, b_2 are constants, and V (the Lévy measure) is a positive measure that satisfies

$$\int_{\mathbb{R}\setminus\{0\}} \min(1, a^2) V(\mathrm{d}a) < \infty. \tag{6}$$

[1]Precisely, the space $\mathcal{E}$ needs to be nuclear. For instance, this condition is satisfied by the Schwartz space of smooth and rapidly decreasing functions.

Formulation (5) shows that $f(0) = 0$. Moreover, the white process is uniquely characterized by the triplet (b_1, b_2, V). In this paper, we only consider innovation processes with a symmetric distribution. This implies that $b_1 = 0$ and V is a symmetric measure in the sense that, for all measurable sets I, $V(I) = V(-I)$. With this assumption, (5) simplifies to

$$f(\omega) = -b_2^2 \omega^2 - \int_{\mathbb{R}\setminus\{0\}} (1 - \cos(a\omega)) \, V(\mathrm{d}a). \tag{7}$$

Stable innovations are of particular interest to us. For these processes, we have that $f(\omega) = -|\sigma\omega|^\alpha$, where $0 < \alpha \leq 2$ is the stability index and σ is the shape factor (which plays almost the same role as the standard deviation in the Gaussian case). The characterizing triplet associated with the α-stable innovation with $0 < \alpha < 2$ is given by $(b_1, b_2, V) = (0, 0, V_\alpha)$, where $V_\alpha(\mathrm{d}a) = \frac{|\sigma|^\alpha}{|a|^{\alpha+1}} \mathrm{d}a$. Although the function f is continuous with respect to α, the characterizing triplet for $\alpha = 2$ (Gaussian innovation) does not follow the same structure as $0 < \alpha < 2$ and is given by $(0, \sigma, 0)$.

C. Discrete Samples

The last thing to mention about the model is the discretization procedure. We assume that the AR(1) process $s(x)$ is sampled at integers $x = 0, 1, \ldots, m$, which corresponds to a finite number of uniform samples with a unit sampling period. For the sake of simplicity, we use $s[k]$ to denote the sample $s(x)|_{x=k}$ for $k = 0, 1, \ldots, m$. The purpose of interpolation is to estimate $s(x)$ for a given $x \in [0, m]$ based on the samples $\{s[k]\}_{k=0}^m$.

III. Main Results

Assume that $s(x)$ for $x \in \mathbb{R}$ is a realization of a random process from which we have only the samples $s[k] = s(x)|_{x=k}$ for $k \in \mathbb{Z}$. Now, the goal of interpolation is to estimate the value $s(x_0)$ for every $x_0 \in \mathbb{R}$ given the samples $\{s[k]\}_{k\in\mathbb{Z}}$. As an example, for a bandlimited stationary process s, the optimal Bayesian interpolator is known to take the form [29]

$$\hat{s}(x_0) = \mathbb{E}\{ s(x_0) \,|\, \{s[k]\}_{k\in\mathbb{Z}} \} = \sum_{k\in\mathbb{Z}} a_k s[k], \tag{8}$$

where $\{a_k\}_{k\in\mathbb{Z}}$ are constants expressed in terms of the sinc function and the point x_0. An important point is that, though these constants depend on x_0 and the statistics of the process, they are independent of the samples. This is the reason why interpolators of the form (8) are referred to as linear interpolators, stressing their linearity with respect to the samples.

When s is not bandlimited, however, the conditional expectation in (8) does not necessarily result in a linear estimator. The Gaussian processes are well-studied examples for which the optimal estimator is often linear. One of the distinguishing properties of the Gaussian distribution is that it is closed with respect to linear combinations, so that the weighted sum of two Gaussian random variables is again a Gaussian random variable. Nevertheless, this property is not limited to Gaussian distributions. It is generalizable to a family of distributions known as α-stables [30], where α is an index with $0 < \alpha \leq 2$. The Gaussian distribution corresponds to

the extreme case $\alpha = 2$. For the particular case of AR(1) processes, we show in Theorem 1 that the linearity of the Bayesian interpolator is not limited to the Gaussian statistics; for all $0 < \alpha \leq 2$, innovations with an α-stable distribution result in linear interpolators.

Theorem 1. *For the AR(1) process s associated with the whitening operator $\mathrm{D}+\kappa\mathrm{I}$ for $\kappa \neq 0$ and a symmetric α-stable innovation, the Bayesian interpolator at the point $x_0 > 0$ is given by*

$$\hat{s}(x_0) = \sum_{k\in\mathbb{Z}} s[k]\ \beta_\kappa^{(\alpha)}(x_0 - k), \tag{9}$$

where the interpolation kernel $\beta_\kappa^{(\alpha)}$ is given by

$$\beta_\kappa^{(\alpha)}(x) = \begin{cases} \mathrm{e}^{-\kappa(1-\frac{\alpha}{2})x}\ \frac{\sinh\left(\frac{\alpha\kappa}{2}(1-|x|)\right)}{\sinh\left(\frac{\alpha\kappa}{2}\right)}, & x \in [-1,\ 1], \\ 0, & x \notin [-1,\ 1]. \end{cases} \tag{10}$$

The condition $x_0 > 0$ in Theorem 1 (which will also appear in Theorems 2 and 3) is due to the possible boundary condition that will be explained in Section II-A. When boundary conditions are disregarded, this restriction can be removed.

Theorem 1 suggests a convolutive form for the optimal interpolator—see Appendix A for a discussion regarding the optimality of the Bayesian interpolator for fat-tailed distributions. It depends both on the parameter κ in the whitening operator and the stability index α. Observe that, the interpolation kernel in (9) is non-zero only for one data point on either side of x_0.

For $\alpha \neq 2$, the interpolation kernel $\beta_\kappa^{(\alpha)}$ is asymmetric, because of the term $\exp(-\kappa(1-\frac{\alpha}{2})x)$. The exclusion of the choice $\kappa = 0$ in Theorem 1 is to ensure that the kernels are well-defined. However, the kernels are convergent for either case of $\kappa \to 0$ and $\alpha \to 0$, since

$$\begin{cases} \lim_{\kappa\to 0} \beta_\kappa^{(\alpha)}(x) &= 1 - |x|, \\ \lim_{\alpha\to 0} \beta_\kappa^{(\alpha)}(x) &= \mathrm{e}^{-\kappa x}(1 - |x|). \end{cases} \tag{11}$$

Another property of the kernels is that

$$\beta_{-\kappa}^{(\alpha)}(x) = \beta_\kappa^{(\alpha)}(-x), \tag{12}$$

which again confirms that the limiting function for $\kappa \to 0$ should be symmetric. We illustrate in Figure 2 some of the interpolation kernels for $\kappa = 2$; the curves for $\kappa = -2$ can be achieved by flipping the horizontal axis.

Remark 1. *It is already known that, if $X_0, X_1, \ldots, X_m$ follow a joint symmetric α-stable law, then $\mathbb{E}\{X_0|X_1\}$ is linear with respect to X_1, while $\mathbb{E}\{X_0|X_1, \ldots, X_n\}$ is not necessarily linear with respect to any of $X_1, \ldots, X_n$ [30]. What Theorem 1 reveals is that the joint distribution of the samples of AR(1) processes are special cases of joint α-stable laws that preserve the linearity of posterior means for arbitrary dimensions.*

Our next result shows that, for $\kappa = 0$, the interpolation kernel is the same as the limiting function $\lim_{\kappa\to 0} \beta_\kappa^{(\alpha)}(x)$. However, the surprising result is that the statement is true for all innovation statistics and not just α-stables.

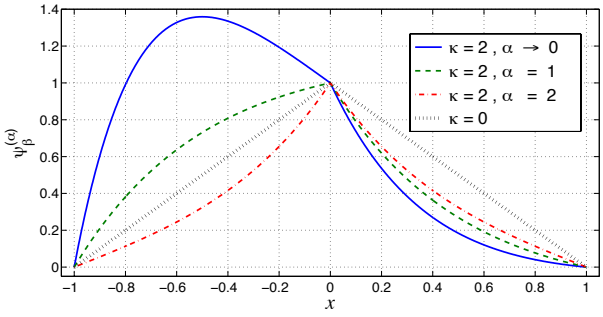

Fig. 2. Interpolating kernel $\beta_\kappa^{(\alpha)}$ for some values of κ and α. For the special case $\kappa = 0$, the kernel is independent of α. Except when $\alpha = 2$ or $\kappa = 0$, the kernel is asymmetric.

Theorem 2. *The optimal interpolator for a Lévy-type process s associated with the whitening operator D and a finite first-order moment is linear. Moreover, it is a B-spline of degree 1, meaning*

$$\hat{s}(x_0) = \sum_{k\in\mathbb{Z}} s[k]\beta_0(x_0 - k), \tag{13}$$

where

$$\beta_0(x) = \begin{cases} 1 - |x|, & x \in [-1,\ 1], \\ 0, & x \notin [-1,\ 1]. \end{cases} \tag{14}$$

Theorem 2 reveals that, when the whitening operator is $\mathrm{D}+\kappa\mathrm{I}$ with $\kappa = 0$, the optimal interpolator is linear and does not depend on the statistics. In other words, the linearity of the interpolator is not a property that is completely determined by the statistics. As is reflected in Theorem 2, some of the mathematical tools exploited in this paper require that the distributions have a finite first-order moment ($\mathbb{E}\{|x|\} \leq \infty$).

It is possible to extend the results of Theorem 1 and 2 to the case of nonuniform samples, as given in Proposition 1. The proof is similar to that of Theorems 1 and 2 and therefore, omitted.

Proposition 1. *Assume that an AR(1) process associated with the whitening operator $\mathrm{D} + \kappa\mathrm{I}$ is sampled at $\bar{x}_0 < \cdots < \bar{x}_m$, and that we are interested in Bayesian interpolation of $s(x_0)$ with $\bar{x}_i < x_0 < \bar{x}_{i+1}$.*

(i) If $\kappa \neq 0$ and the process follows a symmetric α-stable law, then

$$\begin{aligned} \hat{s}(x_0) = \ & s(\bar{x}_i)\beta_\kappa^{(\alpha)}\left(\frac{x_0 - \bar{x}_i}{\bar{x}_{i+1} - \bar{x}_i}\right) \\ & + s(\bar{x}_{i+1})\beta_\kappa^{(\alpha)}\left(\frac{x_0 - \bar{x}_{i+1}}{\bar{x}_{i+1} - \bar{x}_i}\right). \end{aligned} \tag{15}$$

(ii) If $\kappa = 0$, then

$$\begin{aligned} \hat{s}(x_0) = \ & s(\bar{x}_i)\beta_0\left(\frac{x_0 - \bar{x}_i}{\bar{x}_{i+1} - \bar{x}_i}\right) \\ & + s(\bar{x}_{i+1})\beta_0\left(\frac{x_0 - \bar{x}_{i+1}}{\bar{x}_{i+1} - \bar{x}_i}\right). \end{aligned} \tag{16}$$

Our last result in Theorem 3 completes the classification of linear regimes by demonstrating that, except for those cases

already covered in the previous theorems, we cannot expect the linearity of the interpolator.

Theorem 3. *For the AR(1) process s associated with the whitening operator $\mathrm{D}+\kappa\mathrm{I}$ for $\kappa \neq 0$ and a non-α-stable innovation that has finite first-order moment, the Bayesian interpolator cannot be a linear function of the samples at all points.*

IV. Preliminaries

We prefer to use the same framework for $\kappa \neq 0$ and $\kappa = 0$. However, due to the existence of the boundary condition in $\kappa = 0$, the process s is non-stationary, which complicates our analysis. Hence, we work with the generalized-increment process defined as

$$u_T(x) = s(x) - \mathrm{e}^{-\kappa T} s(x-T), \tag{17}$$

where T is a positive real. To further simplify the notations, let

$$\varphi_T(x) = \mathrm{e}^{-\kappa x}\Big(\chi_{\mathbb{R}_0^+}(x) - \chi_{\mathbb{R}_0^+}(x-T)\Big). \tag{18}$$

Also, we denote by $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ the Fourier and inverse Fourier transforms, respectively, defined as

$$\left\{\begin{array}{lcl} \mathcal{F}\{q\}(\omega) & = & \int_{\mathbb{R}} q(x)\mathrm{e}^{-\mathrm{j}\omega x}\mathrm{d}x, \\ \mathcal{F}^{-1}\{\hat{q}\}(x) & = & \frac{1}{2\pi}\int_{\mathbb{R}} \hat{q}(\omega)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}\omega. \end{array}\right. \tag{19}$$

Since most of the proofs in this paper involve properties of the generalized-increment process, we have summarized them in Lemma 1.

Lemma 1. *Let x_1, x_2 be nonnegative and T_1, T_2 be positive real numbers. For the generalized increment processes u_{T_1} and u_{T_2} we have that*

(i) *u_{T_1} and u_{T_2} are stationary processes;*
(ii) *if $(x_1 - x_2) \notin]-T_2, T_1[$, then the random variables $u_{T_1}(x_1)$ and $u_{T_2}(x_2)$ are independent;*
(iii) *if $x_1 \geq \max\{T_1, x_2 + T_1\}$, then $u_{T_1}(x_1)$ and $s(x_2)$ are independent;*
(iv) *if $p_{u_{T_1}}$ represents the probability density function of $u_{T_1}(x_1)$, then we have that*

$$\mathcal{F}_x\{p_{u_{T_1}}\}(-\omega) = \exp\left(\int_0^{T_1} f(\omega \mathrm{e}^{-\kappa\tau})\mathrm{d}\tau\right), \tag{20}$$

where f is the Lévy exponent of the innovations.

Proof. (i) First, we express the generalized increments in terms of the innovation process. To this end, note that

$$\begin{aligned} u_T(x) &= s(x) - s(x-T) \\ &= \mathrm{L}^{-1}\{w\}(x) - \mathrm{e}^{-\kappa T}\mathrm{L}^{-1}\{w\}(x-T). \end{aligned} \tag{21}$$

To proceed further, we study the cases $\kappa = 0$ and $\kappa \neq 0$ separately.

- For $\kappa = 0$, we rewrite (21) as

$$\begin{aligned} u_T(x) &= \int_0^x w(\tau)\mathrm{d}\tau + \langle \phi, w\rangle \\ &\quad -\Big(\int_0^{x-T} w(\tau)\mathrm{d}\tau + \langle \phi, w\rangle\Big) \\ &= \int_{x-T}^{x} w(\tau)\mathrm{d}\tau = \langle \varphi_T(x-\cdot), w\rangle, \end{aligned} \tag{22}$$

  where $\varphi_T(x) = \chi_{\mathbb{R}_0^+}(x) - \chi_{\mathbb{R}_0^+}(x-T)$ and the dot notation in $\varphi_T(x-\cdot)$ stands for the variable over which the inner product is defined (here, τ).
- For $\kappa \neq 0$, the application of L^{-1} on w can be seen as the convolution (in the sense of generalized functions) of the innovation with the impulse response $\mathrm{e}^{-\kappa x}\chi_{\mathbb{R}_0^+}(x)$, or, alternatively, $\langle \mathrm{e}^{-\kappa(x-\cdot)}\chi_{\mathbb{R}_0^+}(x-\cdot), w\rangle$. This shows that

$$\begin{aligned} u_T(x) &= \langle \mathrm{e}^{-\kappa(x-\cdot)}\chi_{\mathbb{R}_0^+}(x-\cdot), w\rangle \\ &\quad -\mathrm{e}^{-\kappa T}\langle \mathrm{e}^{-\kappa(x-T-\cdot)}\chi_{\mathbb{R}_0^+}(x-T-\cdot), w\rangle \\ &= \langle \varphi_T(x-\cdot), w\rangle. \end{aligned} \tag{23}$$

Both (22) and (23) indicate the same result, which is that

$$u_T(x) = \langle \varphi_T(x-\cdot), w\rangle \tag{24}$$

for all κ. One can verify that (24) is valid even when there exists a boundary condition for $\kappa \neq 0$. Equation (24) reveals that the generalized-increment process is generated by applying a linear shift-invariant filter on the innovations. Thus, the process is stationary.

(ii) According to (18), the support of φ_T is limited to the interval $[0, T[$. Thus, the support intervals of $\varphi_{T_1}(x_1 - \cdot)$ and $\varphi_{T_2}(x_2 - \cdot)$ are $]x_1 - T_1,\, x_1]$ and $]x_2 - T_2\,, x_2]$, respectively. The condition $(x_1 - x_2) \notin]-T_2, T_1[$ guarantees that the supports of the functions $\varphi_{T_1}(x_1 - \cdot)$ and $\varphi_{T_2}(x_2 - \cdot)$ are disjoint. Therefore, their inner products with a white process are independent by definition.

(iii) According to the definition of L^{-1} in Section II-A, $s(x_2)$ consists of up to two constituents.

1) The integral of the innovation process from $(-\infty)$ to x_2 for $\kappa \neq 0$, or from 0 to x_2 for $\kappa = 0$. This term is statistically independent of $w(\tau)$ for $\tau > x_2$.
2) A boundary-condition term ($\kappa = 0$) formed by observing the innovations through an anti-causal window ϕ. This term is statistically independent of $w(\tau)$ for $\tau > 0$.

In summary, $s(x_2)$ is independent of $w(\tau)$ for $\tau > \max\{0,\, x_2\}$. On the other hand, recalling (24), we know that $u_{T_1}(x_1)$ depends on $w(\tau)$ for $\tau \in]x_1 - T_1,\, x_1]$. Hence, the condition $x_1 \geq \max\{T_1,\, x_2 + T_1\}$ guarantees that $u_{T_1}(x_1)$ and $s(x_2)$ depend on disjoint intervals of the white process and are consequently independent.

(iv) The Fourier transform of the pdf $p_{u_{T_1}}$, known as the characteristic function and represented as $\hat{p}_{u_{T_1}}$, is given by

$$\begin{aligned} \hat{p}_{u_{T_1}}(-\omega) &= \int_{\mathbb{R}} p_{u_{T_1}}(x)\mathrm{e}^{\mathrm{j}x\omega}\mathrm{d}x = \mathbb{E}\{\mathrm{e}^{\mathrm{j}\omega\langle \varphi_{T_1}(x-\cdot), w\rangle}\} \\ &= \widehat{\mathscr{P}}_w\left(\varphi_{T_1}(x-\cdot)\right) \\ &= \exp\left(\int_0^{T_1} f(\omega\mathrm{e}^{-\kappa\tau})\mathrm{d}\tau\right). \end{aligned} \tag{25}$$

■

Corollary 1. *When the innovation process follows a symmetric α-stable distribution with $f(\omega) = -|\sigma\omega|^\alpha$, by using Lemma 1-(iv), we have that*

$$\mathcal{F}\{p_{u_T}\}(\omega) = \begin{cases} \mathrm{e}^{-|\sigma\omega|^\alpha \frac{1-\mathrm{e}^{-\alpha\kappa T}}{\alpha\kappa}}, & \kappa \neq 0, \\ \mathrm{e}^{-T|\sigma\omega|^\alpha}, & \kappa = 0. \end{cases} \quad (26)$$

Corollary 2. *Since the characteristic functions are Fourier transforms of probability density functions which are absolutely integrable, they are uniformly continuous. Similarly, if the innovation is such that u_T has a finite first-order moment, the derivative of the characteristic form is uniformly continuous, too.*

Corollary 3. *For $\kappa \neq 0$, by change of variables we rewrite the equation for the characteristic function of u_T in (25) as*

$$\hat{p}_{u_T}(-\omega) = \exp\left(\frac{1}{\kappa}\int_{\omega\mathrm{e}^{-T\kappa}}^{\omega} \frac{f(\tau)}{\tau}\mathrm{d}\tau\right). \quad (27)$$

Thus,

$$\hat{p}'_{u_T}(\omega) = \frac{\partial}{\partial\omega}\hat{p}_{u_T}(\omega) = \frac{f(-\omega) - f(-\omega\mathrm{e}^{-T\kappa})}{\kappa\omega}\hat{p}_{u_T}(\omega). \quad (28)$$

This shows that, for $\kappa \neq 0$, the characteristic function is always differentiable at $\omega \neq 0$.

V. Bayesian Interpolation

The Bayesian filter is usually considered as the optimal estimator because it minimizes the mean-square error whenever it is finite. However, there are cases where the posterior mean exists while the mean-square error is unbounded. Therefore, it is no longer possible to speak about the optimality of the Bayesian estimator with respect to the mean-square criterion. Estimation problems involving heavy-tail distributions such as α-stables with $1 < \alpha < 2$ are usually among these cases. A brief discussion about the optimality of the Bayesian interpolator is provided in Appendix A. In particular, we show the optimality for given realizations.

In this section, we show how to benefit from the notion of generalized increments in the interpolation problem. In general, the Bayesian interpolator (or the posterior mean estimator at the desired point) depends on all the samples $\{s[k]\}_{k=0}^m$, which suggests the use of $(m+1)$-dimensional joint distributions. Nevertheless, Lemma 2 shows that we can efficiently reduce the size of the sufficient statistics by using generalized increments.

Lemma 2. *Let $x_0 = n+r$, where $0 \leq n < m$ is an integer and $0 < r < 1$ is a real number. Then, the Bayesian interpolator of the AR(1) process at the point x_0, given the samples $\{s[k]\}_{k=0}^m$, is given by*

$$\hat{s}(x_0) = \mathrm{e}^{-\kappa r}s[n] + \mathbb{E}\left\{u_r(n+r) \;\middle|\; u_1(n+1) = s[n+1] - \mathrm{e}^{-\kappa}s[n]\right\}. \quad (29)$$

Proof. We start by the definition of the Bayesian interpolator (posterior mean)

$$\begin{aligned} \hat{s}(n+r) &= \mathbb{E}\left\{s(n+r) \;\middle|\; \{s[k]\}_{k=0}^m\right\} \\ &= \mathrm{e}^{-\kappa r}s[n] + \mathbb{E}\left\{u_r(n+r)\middle|\{s[k]\}_{k=0}^m\right\}. \end{aligned} \quad (30)$$

Since there is a bijection between the sets $\{s[k]\}_{k=0}^m$ and $s[0] \cup \{u_1(k)\}_{k=1}^m$, the condition in the expectation of (30) can be replaced according to

$$\begin{aligned} &\hat{s}(n+r) - \mathrm{e}^{-\kappa r}s[n] \\ &= \mathbb{E}\left\{u_r(n+r) \;\middle|\; s[0] \cup \{u_1(k)\}_{k=1}^m\right\} \\ &= \mathbb{E}\left\{u_r(n+r) \;\middle|\; u_1(n+1) = s[n+1] - \mathrm{e}^{-\kappa}s[n]\right\}, \end{aligned} \quad (31)$$

where the validity of the second equality comes from the fact that $u_r(n+r)$ is statistically independent of $s[0]$ and $u_1(k)$ for $k \neq n+1$ (Lemma 1). ■

Up to this point, we have simplified the general form of the Bayesian interpolator. However, the main challenge is that our statistical information of the model is given in the form of characteristic functions. Since the probability density functions are related to the characteristic functions by means of Fourier transforms, we need to reformulate the Bayesian interpolator in the Fourier domain.

Lemma 3. *Let f be the Lévy exponent of the innovation process and X be a random variable with $\mathrm{e}^{f(\omega)}$ as its characteristic function. If f is such that either*

(i) *the white innovation has a finite first-order moment $\mathbb{E}\{|X|\} < \infty$, or*

(ii) *$\kappa \neq 0$ and the pdf p_X of X is continuous, with $|x|^{1+\epsilon}p_X(x)$ being bounded for some $\epsilon > 0$,*

then, we have that

$$\begin{aligned} &\mathbb{E}\left\{u_r(n+r) \;\middle|\; u_1(n+1) = \theta\right\} = \\ &-\mathrm{j}\frac{\int_{\mathbb{R}} \hat{p}'_{u_r}\left(\omega\mathrm{e}^{-\kappa(1-r)}\right)\hat{p}_{u_{1-r}}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega}{\int_{\mathbb{R}}\hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega}, \end{aligned} \quad (32)$$

where $\hat{p}_{u_T}$ stands for the characteristic function of the random variable u_T (Fourier transform of its pdf) and $\hat{p}'_{u_T}$ denotes its derivative in the Fourier domain.

The proof of Lemma 3 is provided in Appendix B. The significance of this lemma is the establishment of a link between the conditional expectation involved in the Bayesian interpolator and the characteristic functions of the generalized increments. Lemma 1-(iv) allows us to relate the characteristic functions to the Lévy exponent of the innovation process.

In Lemma 4, we summarize the results of this section. Again, for fluidity of the paper, we have postponed the proof to Appendix C.

Lemma 4. *Let $0 \leq n < m$ be an integer and $0 < r < 1$ be a real number. If the white innovation satisfies at least one of the conditions in Lemma 3, then*

$$\begin{aligned} \hat{s}(n+r) &= \mathrm{e}^{-\kappa r}s[n] \\ &+ \theta\,\mathrm{e}^{\kappa(1-r)}\frac{\int_{\mathbb{R}}\mu_\kappa^{(f)}(\omega;1-r)\,\hat{p}_{u_1}(\omega)\,\mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}{\int_{\mathbb{R}}\mu_\kappa^{(f)}(\omega;0)\,\hat{p}_{u_1}(\omega)\,\mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}, \end{aligned} \quad (33)$$

where $\theta = s[n+1] - \mathrm{e}^{-\kappa}s[n]$ and

$$\mu_\kappa^{(f)}(\omega;x) = \begin{cases} \frac{f(-\omega\mathrm{e}^{-\kappa x}) - f(-\omega\mathrm{e}^{-\kappa})}{\kappa\omega}, & \kappa \neq 0, \\ (x-1)f'(-\omega), & \kappa = 0, \end{cases} \quad (34)$$

in which f' is the derivative of the Lévy exponent of the innovation process.

VI. Proofs

A. Proof of Theorem 1

A symmetric α-stable innovation corresponds to the Lévy exponent $f(\omega) = -|\sigma\omega|^\alpha$. The density functions are known to be continuous and asymptotically decaying proportionally to $|x|^{-(1+\alpha)}$ [30]. This implies that $|x|^{1+\epsilon}p_x(x)$ is bounded for $0 < \epsilon \leq \alpha$, where p_x is an α-stable pdf. In order to apply Lemma 4 for stable laws, we evaluate $\mu_\kappa^{(f)}(\omega; x)$ from (34) as

$$\mu_\kappa^{(f)}(\omega; x) = \frac{|\sigma\omega|^\alpha}{\kappa\omega}\left(\mathrm{e}^{-\alpha\kappa} - \mathrm{e}^{-\alpha\kappa x}\right). \tag{35}$$

Hence, the expression in Lemma 4 for the Bayesian interpolator at the point $x_0 = n + r$, where $n = \lfloor x_0 \rfloor$ and $r = \{x_0\}$ (integral and fractional parts of x_0, respectively), simplifies to

$$\hat{s}(n+r) = \mathrm{e}^{-\kappa r}s[n] + \theta\,\mathrm{e}^{\kappa(1-r)}\frac{\mathrm{e}^{-\alpha\kappa(1-r)} - \mathrm{e}^{-\alpha\kappa}}{1-\mathrm{e}^{-\alpha\kappa}}, \tag{36}$$

which, by substituting θ with $s[n+1] - \mathrm{e}^{-\kappa}s[n]$, yields

$$\begin{aligned}\hat{s}(n+r) &= \underbrace{\mathrm{e}^{-\kappa r}\frac{1-\mathrm{e}^{-\alpha\kappa(1-r)}}{1-\mathrm{e}^{-\alpha\kappa}}}_{\beta_\kappa^{(\alpha)}(r)} s[n] \\ &\quad + \underbrace{\mathrm{e}^{\kappa(1-r-\alpha)}\frac{\mathrm{e}^{\alpha\kappa r}-1}{1-\mathrm{e}^{-\alpha\kappa}}}_{\beta_\kappa^{(\alpha)}\left(-(1-r)\right)} s[n+1]. \end{aligned}\tag{37}$$

Furthermore, the interpolator at the point $x_0 \notin \mathbb{Z}$ depends on the two neighboring samples, which is consistent with the support interval of $\beta_\kappa^{(\alpha)}$. Finally, when $x_0 \notin \mathbb{Z}$, the Bayesian interpolator trivially reproduces the sample. This is also confirmed by the interpolation formula as $\beta_\kappa^{(\alpha)}(0) = 1$ and $\beta_\kappa^{(\alpha)}(k) = 0$ for $k \in \mathbb{Z} \setminus \{0\}$. □

B. Proof of Theorem 2

The choise $\kappa = 0$ reduces the whitening operator to the differential operator D. This corresponds to the second case of Definition (34) and leads to

$$\mu_\kappa^{(f)}(\omega; x) = (x-1)f'(-\omega) \tag{38}$$

which, by using Lemma 4, results in

$$\begin{aligned}\hat{s}(n+r) &= s[n] + (s[n+1] - s[n])(1-r) \\ &= r\,s[n] + (1-r)s[n+1]. \end{aligned}\tag{39}$$

This, in fact, completes the proof. □

C. Proof of Theorem 3

We prove here the contrapositive of the statement of Theorem 3. In other words, we assume the linearity of the interpolator and conclude the stability of innovations. Because of technical details involved in the proof, we divide it into a sequence of claims.

Claim 1. *The linearity of the Bayesian interpolator happens only if there exists a continuous function $g_\kappa(r)$ such that, for all ω and $0 < r < 1$,*

$$f(\omega\mathrm{e}^{r\kappa}) = g_\kappa(r)f(\omega\mathrm{e}^{\kappa}) + (1 - g_\kappa(r))\,f(\omega). \tag{40}$$

Proof. Suppose the Bayesian interpolator is a linear function of the samples. Then, by multiplying the samples by a scalar constant, we expect the interpolated values to get scaled by the same constant. Equivalently, the terms multiplied by the samples in the formulation of Lemma 4 should be independent of the sample values. Consequently,

$$\frac{\int_{\mathbb{R}} \mu_\kappa^{(f)}(\omega; 1-r)\,\hat{p}_{u_1}(\omega)\,\mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}{\int_{\mathbb{R}} \mu_\kappa^{(f)}(\omega; 0)\,\hat{p}_{u_1}(\omega)\,\mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}$$

can only depend on r and not on θ. This implies that

$$\frac{\mathcal{F}_\omega^{-1}\{\mu_\kappa^{(f)}(\omega; 1-r)\,\hat{p}_{u_1}(\omega)\}(\theta)}{\mathcal{F}_\omega^{-1}\{\mu_\kappa^{(f)}(\omega; 0)\,\hat{p}_{u_1}(\omega)\}(\theta)} = g_\kappa(r). \tag{41}$$

By moving the denominator to the right-hand side and taking the Fourier transform of both sides with respect to θ, we further simplify the condition to

$$\mu_\kappa^{(f)}(\omega; 1-r) = g_\kappa(r)\mu_\kappa^{(f)}(\omega; 0), \tag{42}$$

or

$$\begin{aligned} &f(-\omega\mathrm{e}^{(r-1)\kappa}) - f(-\omega\mathrm{e}^{-\kappa}) = \\ &g_\kappa(r)\left(f(-\omega) - f(-\omega\mathrm{e}^{-\kappa})\right), \end{aligned}\tag{43}$$

since $\kappa \neq 0$. In addition, because of the continuity of f, the function g_κ is also continuous. By replacing ω with $-\omega\mathrm{e}^{\kappa}$, (43) transforms into (40). ■

Claim 2. *Using previous assumptions and terminology, for all ω and $0 < r < 1$, we have that*

$$\frac{f(\omega\mathrm{e}^{\kappa}) - f(\omega)}{f(\omega) - f(\omega\mathrm{e}^{-\kappa})} = \frac{(1-g_\kappa(r))\,(1-g_\kappa(1-r))}{g_\kappa(r)\,g_\kappa(1-r)} = \xi_\kappa, \tag{44}$$

where ξ_κ is a positive constant.

Proof. We apply Claim 1 on two instances of ω and r to evaluate $f(\omega\mathrm{e}^{r\kappa})$ and $f(\omega\mathrm{e}^{(r-1)\kappa})$ as

$$\begin{aligned} f(\omega\mathrm{e}^{r\kappa}) &= g_\kappa(r)f(\omega\mathrm{e}^{\kappa}) + (1-g_\kappa(r))\,f(\omega), \\ f(\omega\mathrm{e}^{(r-1)\kappa}) &= g_\kappa(r)f(\omega) + (1-g_\kappa(r))\,f(\omega\mathrm{e}^{-\kappa}). \end{aligned}\tag{45}$$

Again applying Claim 1, we can relate the three values $f(\omega)$, $f(\omega\mathrm{e}^{r\kappa})$, and $f(\omega\mathrm{e}^{(r-1)\kappa})$ as

$$\begin{aligned} f(\omega) &= g_\kappa(1-r)f(\omega\mathrm{e}^{r\kappa}) \\ &\quad + (1-g_\kappa(1-r))\,f(\omega\mathrm{e}^{(r-1)\kappa}), \end{aligned}\tag{46}$$

which yields

$$\frac{f(\omega\mathrm{e}^{\kappa}) - f(\omega)}{f(\omega) - f(\omega\mathrm{e}^{-\kappa})} = \frac{(1-g_\kappa(r))\,(1-g_\kappa(1-r))}{g_\kappa(r)\,g_\kappa(1-r)} \tag{47}$$

by replacing the values of $f(\omega\mathrm{e}^{r\kappa})$ and $f(\omega\mathrm{e}^{(r-1)\kappa})$ from (45). The left-hand side of (47) can only be a function of ω while the right-hand side can only vary with r. Thus, the fractions are equal to a constant which we denote by ξ_κ. The fraction in (47) becomes a square for $r = \frac{1}{2}$, which shows that ξ_κ is positive. Note that, $\xi_\kappa = 0$ results in $f(\omega)$ being a constant function for $\omega \neq 0$ which is not admissible. ■

Claim 3. *If f is such that the estimator is linear, then the following is valid for all $\omega \in \mathbb{R}$ and $m \in \mathbb{Z}$:*

$$f(\omega\mathrm{e}^{m\kappa}) = (\xi_\kappa)^m f(\omega). \tag{48}$$

Proof. A direct consequence of Claim 2 is that

$$(\xi_\kappa)^j = \prod_{i=1}^{j} \frac{f(\omega \mathrm{e}^{i\kappa}) - f(\omega \mathrm{e}^{(i-1)\kappa})}{f(\omega \mathrm{e}^{(i-1)\kappa}) - f(\omega \mathrm{e}^{(i-2)\kappa})} = \frac{f(\omega \mathrm{e}^{j\kappa}) - f(\omega \mathrm{e}^{(j-1)\kappa})}{f(\omega) - f(\omega \mathrm{e}^{-\kappa})}, \quad (49)$$

where j is a positive integer. With a similar argument, one can show that (49) is also valid for nonpositive values of j. We consider two cases.

1) $\kappa < 0$: From (49) and for $m \geq 1$, we know that

$$f(\omega \mathrm{e}^{m\kappa}) - f(\omega) = \sum_{j=1}^{m} \big(f(\omega \mathrm{e}^{j\kappa}) - f(\omega \mathrm{e}^{(j-1)\kappa})\big) = \big(f(\omega) - f(\omega \mathrm{e}^{-\kappa})\big) \sum_{j=1}^{m} (\xi_\kappa)^j. \quad (50)$$

When $m \to \infty$, the left-hand side of (50) remains bounded, which suggests that $0 < \xi_\kappa < 1$. Moreover, since f is continuous, for the limiting case of $m \to \infty$ we can write that

$$\underbrace{f(0)}_{0} - f(\omega) = \frac{\xi_\kappa}{1 - \xi_\kappa} \big(f(\omega) - f(\omega \mathrm{e}^{-\kappa})\big). \quad (51)$$

2) $\kappa > 0$: From (49) and for $m \leq -1$, we know that

$$f(\omega) - f(\omega \mathrm{e}^{m\kappa}) = \sum_{j=m}^{-1} \big(f(\omega \mathrm{e}^{(j+1)\kappa}) - f(\omega \mathrm{e}^{j\kappa})\big) = \big(f(\omega) - f(\omega \mathrm{e}^{-\kappa})\big) \sum_{j=m}^{-1} (\xi_\kappa)^{j+1}. \quad (52)$$

When $m \to -\infty$, the left-hand side of (52) remains bounded, which suggests that $\xi_\kappa > 1$. Moreover, since f is continuous, for the limiting case of $m \to -\infty$ we can write that

$$f(\omega) - \underbrace{f(0)}_{0} = \frac{\xi_\kappa}{\xi_\kappa - 1} \big(f(\omega) - f(\omega \mathrm{e}^{-\kappa})\big). \quad (53)$$

We excluded the possibility $f(\omega) = f(\omega \mathrm{e}^{-\kappa})$ in our argument because it would force f to be a constant function (Claim 1) which, in conjunction with the condition $f(0) = 0$, would result in the trivial solution $f \equiv 0$.

We see from (51) and (53) that the final result turns out to be

$$f(\omega) = \xi_\kappa f(\omega \mathrm{e}^{-\kappa}). \quad (54)$$

Since (54) is valid for all ω, it is easy to conclude the claim by induction. ■

Claim 4. *If the innovation model results in a linear estimator, then for all $0 < r < 1$ we have that*

$$f(\omega \mathrm{e}^{r\kappa}) = (\xi_\kappa)^r f(\omega). \quad (55)$$

Proof. We can rewrite the statement in Claim 1 by employing Claim 3 as

$$f(\omega \mathrm{e}^{r\kappa}) = \big(\underbrace{1 + (\xi_\kappa - 1) g_\kappa(r)}_{\tilde{g}_\kappa(r)}\big) f(\omega). \quad (56)$$

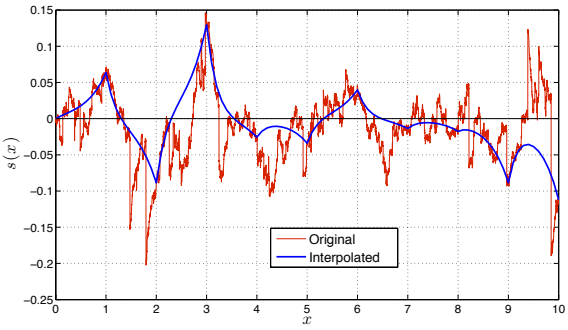

Fig. 3. Realization of an AR(1) process with $\kappa = 5$ and $\alpha = 1.5$, and the function interpolated from samples at the integer points.

When $0 < r_1, r_2$ are such that $r_1 + r_2 < 1$, we can write

$$\tilde{g}_\kappa(r_1 + r_2) f(\omega) = f(\omega \mathrm{e}^{(r_1+r_2)\kappa}) = \tilde{g}_\kappa(r_1) f(\omega \mathrm{e}^{r_2\kappa}) = \tilde{g}_\kappa(r_1)\, \tilde{g}_\kappa(r_2)\, f(\omega). \quad (57)$$

Hence, for all $0 < r_1, r_2$ such that $r_1 + r_2 < 1$, we have that

$$\tilde{g}_\kappa(r_1 + r_2) = \tilde{g}_\kappa(r_1)\, \tilde{g}_\kappa(r_2). \quad (58)$$

Due to the continuity of $\tilde{g}_\kappa$ inherited from g_κ, we conclude that

$$\tilde{g}_\kappa(r) = c^r, \quad (59)$$

where $0 < r < 1$ and c is a positive constant. This can be interpreted as $g_\kappa(r) = \frac{c^r - 1}{\xi_\kappa - 1}$. We can check that $c = \xi_\kappa$ is the only possibility that satisfies (44). Thus, $\tilde{g}_\kappa(r) = (\xi_\kappa)^r$ in (56), which proves the claim. ■

We are now equipped to resume the proof of Theorem 3. Let x be an arbitrary real number and assume $x = n + r$, where n and r are the integral and fractional parts of x, respectively. For any ω and by applying Claims 3 and 4, we readily obtain

$$f(\omega \mathrm{e}^{x\kappa}) = f(\omega \mathrm{e}^{r\kappa} \mathrm{e}^{n\kappa}) = (\xi_\kappa)^n f(\omega \mathrm{e}^{r\kappa}) = (\xi_\kappa)^n (\xi_\kappa)^r f(\omega) = (\xi_\kappa)^x f(\omega). \quad (60)$$

Now, for $\omega \neq 0$ we have that

$$f(\omega) = f\big(\mathrm{sgn}(\omega) \mathrm{e}^{\frac{\log|\omega|}{\kappa}\kappa}\big) = (\xi_\kappa)^{\frac{\log|\omega|}{\kappa}} f\big(\mathrm{sgn}(\omega)\big) = -|\sigma\omega|^\alpha, \quad (61)$$

where

$$\begin{cases} \alpha = \frac{\log \xi_\kappa}{\kappa}, \\ \sigma = \big(-f(\pm 1)\big)^{\frac{1}{\alpha}}. \end{cases} \quad (62)$$

Note that, due to the symmetry of f, we have that $f(1) = f(-1)$. This value is also nonpositive because of the general fact that, for all ω, we have that $\Re\big(f(\omega)\big) \leq f(0)$ [28]. Since f admits only real values in our case, we conclude that $f(\omega) \leq f(0) = 0$. □

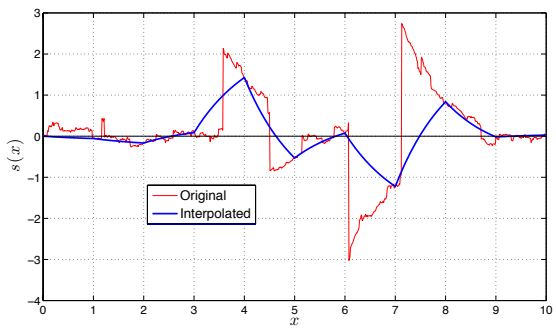

Fig. 4. Realization of an AR(1) process with $\kappa = 1$ and $\alpha = 1.1$, and the function interpolated from samples at the integer points.

VII. Simulations

To test the performance of our interpolator for stable innovations, we have applied it to MATLAB simulated data. Since we are limited to discrete signals in computer simulations, we have used a fine grid to represent the continuous-time stochastic process. We present in Figures 3 and 4 the realizations of two α-stable AR(1) processes with $\alpha = 1.5$ and $\alpha = 1.1$, respectively, and their interpolated versions using samples at integer points.

As is most evident in Figure 3, the curves connecting the points deviate from straight lines and are not even piecewise monotonic (*e.g.*, the part corresponding to the interval [9,10]). In fact, the statistics of the model show that, for each pair of adjacent samples, the distribution of the values between them is biased in favor of zero. It can be verified that the α and κ values used in Figure 3 define an interpolation kernel $\beta_\kappa^{(\alpha)}$ which is dominated by the linear interpolator β_0. Thus, the interpolated curve connecting two adjacent samples deviates from the straight line towards the horizontal axis. Remark 2 proved in Appendix D, explains the tendency towards zero in more generality.

Remark 2. *For $\kappa \neq 0$, let us define the average error between the interpolation kernel $\beta_\kappa^{(\alpha)}$ and the linear B-spline β_0 as*

$$\Delta_\kappa^{(\alpha)} = \int_{-1}^{1} \Big(\beta_0(x) - \beta_\kappa^{(\alpha)}(x)\Big)\mathrm{d}x. \tag{63}$$

Then, $\Delta_\kappa^{(\alpha)}$ has the same sign as $(\alpha - 1)$.

To clarify the message of Remark 2, let us consider that all the obtained samples are equal. Thus, by interpolating the samples using the linear B-spline, we obtain a horizontal line. For $\alpha > 1$ (respectively, $\alpha < 1$), Remark 2 implies that the interpolated curve using the Bayesian kernel, compared to the horizontal line of the linear B-spline, is biased towards (respectively, away from) the horizontal axis. Although the value of the bias depends on the parameter κ, its sign is fully determined by α.

From Figure 3, it is understood that the optimal interpolator takes advantage of knowing the system parameters and better follows the process than the uninformed outcome that would be provided by a first-degree (*i.e.*, linear) B-spline.

It is shown in [31] that α-stable priors become more compressible as α decreases. This means that, in the realization of an α-stable process, the intervals with large amplitudes are few and narrow. As α decreases, these intervals become even smaller while the range of amplitudes increases. Therefore, at small α, it is likely that the samples miss these large-amplitude intervals. This explains why a decrease from $\alpha = 1.5$ in Figure 3 to $\alpha = 1.1$ in Figure 4 results in a degradation of the quality of the interpolated signal.

VIII. Conclusion

In this paper, we studied the interpolation of the first-order autoregressive processes generated from stable innovations, including non-Gaussian ones. We introduced the minimum conditional mean-square error (MCMSE) criterion for stable laws, which relaxes the hypothesis of a finite variance. We applied this criterion to derive the optimal interpolator in a general settings and found that it is linear with respect to the samples. Moreover, it depends on the stability index that characterizes stable innovations. In particular, for a suitable value of this index, our analysis encompasses the classical Gaussian case. Finally, we have extended to general stable innovations the link between exponential-spline interpolators and the Gaussian case.

Acknowledgment

The authors would like to thank M. Nilchian and P. Pad for fruitful discussions.

Appendix A
Optimality of Bayesian Interpolator

Assume that $X, Y_1, \ldots, Y_k$ are dependent random variables. We want to estimate the unobserved value of X based on the measurements $Y_1 = y_1, \ldots, Y_k = y_k$. The minimum mean-square error (MMSE) estimator is the function $\hat{X} = \zeta(y_1, \ldots, y_k)$ that minimizes

$$\mathbb{E}_{X,\{Y_i\}}\{(X - \hat{X})^2\} = \mathbb{E}_{\{Y_i\}}\Big\{ \underbrace{\mathbb{E}_{X|\{Y_i\}}\big\{(X - \hat{X})^2 \;\big|\; \{Y_i\}\big\}}_{\text{conditional mean}-\text{square error}} \Big\}. \tag{64}$$

It is well-known that $\hat{X} = \mathbb{E}\big\{X \;\big|\; \{Y_i = y_i\}\big\}$ (*i.e.*, the posterior mean estimator) is the desired minimizer when the mean-square error is finite. In the following, we discuss cases where the mean-square error is infinite while the posterior distribution has finite variance.

The posterior-mean estimator also minimizes the conditional mean-square error, for which the error is equal to the posterior variance. Hence, in cases where the posterior variance is finite, the Bayesian estimator (posterior mean) is optimal for each set of measurements (*i.e.*, $\{y_i\}$), while the average error over all possible sets of measurements may be unbounded.

For the interpolation problem, as explained in Section V, the posterior distribution of $s(n + r)$ (n integer and $0 < r < 1$) is given by

$$p_s\big(s(n+r) \,\big|\, \{s[k]\}\big) = p_u\big(u_r(n+r) \,\big|\, u_1(n+1) = \theta\big), \tag{65}$$

where $\theta = s[n+1] - \mathrm{e}^{-\kappa} s[n]$. The second-order moment of the posterior distribution is in turn given by

$$\frac{\int_{\mathbb{R}} x^2 \, p_{u_r}(x) \, p_{u_{1-r}}(\theta - x\,\mathrm{e}^{-\kappa(1-r)})\mathrm{d}x}{p_{u_1}(\theta)}. \tag{66}$$

For heavy-tail innovations, the algebraic-decay orders of p_{u_r} and $p_{u_{1-r}}$ are equal. Thus, if $\lim_{|x|\to\infty} |x|^{1.5+\epsilon} p_{u_r}(x) = 0$ for some $\epsilon > 0$, the expression in (66) and, consequently, the posterior variance, is bounded. The α-stable innovations for $\alpha > 0.5$ are examples of these cases.

In summary, the Bayesian estimator results in an unbounded mean-square error for some heavy-tail innovation distributions, but the estimator might still be optimal for given realizations according to the minimum conditional mean-square error criterion.

APPENDIX B
PROOF OF LEMMA 3

First, note that the boundedness of the first-order moment in the case $\kappa = 0$ is to guarantee the existence of $\hat{p}'_{u_T}(\omega)$ (see Corollary 2). Then, recalling the definition of generalized increments in (17), we have that

$$u_1(n+1) = \mathrm{e}^{-\kappa(1-r)} u_r(n+r) + u_{1-r}(n+1). \tag{67}$$

Moreover, Lemma 1 implies that $u_r(n+r)$ and $u_{1-r}(n+1)$ are independent. Thus,

$$\begin{aligned} &p\left(u_r(n+r) = x \,\middle|\, u_1(n+1) = \theta\right) \\ &\quad = \frac{p_{u_r}(x) p_{u_{1-r}}(\theta - x\mathrm{e}^{-\kappa(1-r)})}{p_{u_1}(\theta)}, \end{aligned} \tag{68}$$

where we use p to denote probability density functions. By employing (67), the desired conditional expectation is rewritten as

$$\begin{aligned} &\mathbb{E}\left\{u_r(n+r) \,\middle|\, u_1(n+1) = \theta\right\} \\ &= \frac{\int_{\mathbb{R}} x\, p_{u_r}(x)\, p_{u_{1-r}}(\theta - x\mathrm{e}^{-\kappa(1-r)})\,\mathrm{d}x}{p_{u_1}(\theta)} \\ &= \frac{\int_{\mathbb{R}} \overline{\mathcal{F}_x\{x p_{u_r}(x)\}(\omega)}\, \mathcal{F}_x\{p_{u_{1-r}}(\theta - x\mathrm{e}^{-\kappa(1-r)})\}(\omega)\mathrm{d}\omega}{2\pi \mathcal{F}^{-1}\{\hat{p}_{u_1}\}(\theta)} \\ &= \mathrm{j}\mathrm{e}^{\kappa(1-r)} \frac{\int_{\mathbb{R}} \hat{p}'_{u_r}(\omega) \hat{p}_{u_{1-r}}(\omega \mathrm{e}^{\kappa(1-r)}) \mathrm{e}^{\mathrm{j}\omega\theta \mathrm{e}^{\kappa(1-r)}}\,\mathrm{d}\omega}{\int_{\mathbb{R}} \hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega} \\ &= \mathrm{j}\frac{\int_{\mathbb{R}} \hat{p}'_{u_r}(\omega \mathrm{e}^{-\kappa(1-r)}) \hat{p}_{u_{1-r}}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}{\int_{\mathbb{R}} \hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega}, \end{aligned} \tag{69}$$

where we used Parseval's theorem to obtain the second equality. To prove the validity of Parseval's theorem in our case, note that $p_{u_r}(x)$ and $p_{u_{1-r}}(\theta - x\mathrm{e}^{-\kappa(1-r)})$ are absolutely summable functions. For the case of Condition (i) in Lemma 3, the boundedness of the first-order moment guarantees that $x p_{u_r}(x)$ is also absolutely integrable, and therefore, both Fourier transforms are uniformly continuous functions and the proof is complete. Although Condition (ii) is weaker in the sense that $\hat{p}_{u_r}$ is not continuous and has a singularity at $\omega = 0$, it still implies that $\int_{\mathbb{R}} x\, p_{u_r}(x)\, p_{u_{1-r}}(\theta - x\mathrm{e}^{-\kappa(1-r)})\,\mathrm{d}x$ is finite, which is again a sufficient condition [32].

APPENDIX C
PROOF OF LEMMA 4

By using Lemma 2 and Lemma 3, we only need to prove

$$\begin{aligned} &\mathrm{j}\frac{\int_{\mathbb{R}} \hat{p}'_{u_r}\left(\omega \mathrm{e}^{-\kappa(1-r)}\right) \hat{p}_{u_{1-r}}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega}{\int_{\mathbb{R}} \hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega} = \\ &\theta\, \mathrm{e}^{\kappa(1-r)} \frac{\int_{\mathbb{R}} \mu_\kappa^{(f)}(\omega; 1-r)\, \hat{p}_{u_1}(\omega)\, \mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}{\int_{\mathbb{R}} \mu_\kappa^{(f)}(\omega; 0)\, \hat{p}_{u_1}(\omega)\, \mathrm{e}^{\mathrm{j}\omega\theta}\,\mathrm{d}\omega}. \end{aligned} \tag{70}$$

To this end, we focus on the derivative of the characteristic function $\hat{p}_{u_T}(\omega)$. For $\kappa \neq 0$, Corollary 3 implies

$$\begin{aligned} \hat{p}'_{u_T}(\omega) &= \frac{f(-\omega) - f(-\omega \mathrm{e}^{-T\kappa})}{\kappa\omega} \hat{p}_{u_T}(\omega) \\ &= \mathrm{e}^{(1-T)\kappa} \mu_\kappa^{(f)}(\omega \mathrm{e}^{(1-T)\kappa}; 1-T) \hat{p}_{u_T}(\omega). \end{aligned} \tag{71}$$

When $\kappa = 0$, Lemma 1 suggests that $\hat{p}_{u_T}(-\omega) = \exp(T f(\omega))$. Hence,

$$\hat{p}'_{u_T}(\omega) = -T f'(-\omega) \hat{p}_{u_T}(\omega) = \mu_0^{(f)}(\omega; 1-T) \hat{p}_{u_T}(\omega), \tag{72}$$

which is the same as (71) for $\kappa = 0$. In fact, the function $\mu_\kappa^{(f)}$ is defined in such a way that it satisfies (71) for all κ.

From (71), it is easy to verify

$$\begin{cases} \frac{\hat{p}'_{u_r}(\omega \mathrm{e}^{-\kappa(1-r)})}{\hat{p}_{u_r}(\omega \mathrm{e}^{-\kappa(1-r)})} &= \mathrm{e}^{\kappa(1-r)}\, \mu_\kappa^{(f)}(\omega; 1-r), \\ \frac{\hat{p}'_{u_1}(\omega)}{\hat{p}_{u_1}(\omega)} &= \mu_\kappa^{(f)}(\omega; 0). \end{cases} \tag{73}$$

Therefore,

$$\begin{aligned} &\hat{p}'_{u_r}\left(\omega \mathrm{e}^{-\kappa(1-r)}\right) \hat{p}_{u_{1-r}}(\omega) \\ &\quad = \mathrm{e}^{\kappa(1-r)} \mu_\kappa^{(f)}(\omega; 1-r) \hat{p}_{u_r}(\omega \mathrm{e}^{-\kappa(1-r)}) \hat{p}_{u_{1-r}}(\omega) \\ &\quad = \mathrm{e}^{\kappa(1-r)} \mu_\kappa^{(f)}(\omega; 1-r) \hat{p}_{u_1}(\omega). \end{aligned} \tag{74}$$

Our next step is to rewrite the denominator of the fraction in the left-hand side of (70) as

$$\begin{aligned} &\int_{\mathbb{R}} \hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega \\ &\quad = \underbrace{\frac{\hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}}{\mathrm{j}\theta}\Big|_{\omega=-\infty}^{\infty}}_{=0} - \frac{\int_{\mathbb{R}} \hat{p}'_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega}{\mathrm{j}\theta} \\ &\quad = -\frac{\int_{\mathbb{R}} \mu_\kappa^{(f)}(\omega; 0)\hat{p}_{u_1}(\omega)\mathrm{e}^{\mathrm{j}\omega\theta}\mathrm{d}\omega}{\mathrm{j}\theta}. \end{aligned} \tag{75}$$

By combining (74) and (75) into (70), one can easily deduce the claim of the Lemma.

APPENDIX D
PROOF OF REMARK 2

The interpolation kernel $\beta_\kappa^{(\alpha)}$ can be expressed as weighted sum of exponential functions. This helps us in finding the closed form expression for $\Delta_\kappa^{(\alpha)}$ as

$$\Delta_\kappa^{(\alpha)} = 1 - \frac{\frac{\sinh(\kappa/2)}{\kappa/2} \cdot \frac{\sinh\left((\alpha-1)\kappa/2\right)}{(\alpha-1)\kappa/2}}{\frac{\sinh(\alpha\kappa/2)}{\alpha\kappa/2}}. \tag{76}$$

Note that $\frac{\sinh(x)}{x} = \frac{\sinh(|x|)}{|x|}$. Let θ_1 and θ_2 be the minimum and maximum of the set $\left\{\left|\frac{\kappa}{2}\right|, \left|\frac{\kappa}{2}(\alpha-1)\right|\right\}$, respectively. Then,

we have that $|\frac{\alpha\kappa}{2}| = \theta_1 \pm \theta_2$, where $\pm$ reflects the sign of $(\alpha - 1)$. This enables us to rewrite (76) as

$$\begin{aligned}\Delta_\kappa^{(\alpha)} &= 1 - \frac{\frac{\sinh(\theta_1)}{\theta_1} + \frac{\sinh(\theta_2)}{\theta_2}}{\frac{\sinh(\theta_1 \pm \theta_2)}{\theta_1 \pm \theta_2}} \\ &= 1 - \frac{\frac{1}{\theta_2} \pm \frac{1}{\theta_1}}{\frac{1}{\tanh(\theta_2)} \pm \frac{1}{\tanh(\theta_1)}}. \end{aligned} \quad (77)$$

Consider the function $h_1(x) = x - \tanh(x)$. Since $h_1(0) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}x} h_1(x) = 1 - \frac{1}{\cosh^2(x)} \geq 0$, the function $h_1(x)$ is nonnegative for $x \geq 0$. Thus,

$$\left\{ \begin{array}{rcl} \tanh(\theta_1) & \leq & \theta_1 \\ \tanh(\theta_2) & \leq & \theta_2 \end{array} \right. \Rightarrow \frac{\frac{1}{\theta_2} + \frac{1}{\theta_1}}{\frac{1}{\tanh(\theta_2)} + \frac{1}{\tanh(\theta_1)}} \leq 1. \quad (78)$$

This confirms that $\Delta_\kappa^{(\alpha)}$ is positive when $\alpha > 1$. For $\alpha < 1$, let us consider the function $h_2(x) \frac{1}{\tanh(x)} - \frac{1}{x}$. Since $\frac{\mathrm{d}}{\mathrm{d}x} h_2(x) = \frac{1}{x^2} - \frac{1}{\sinh^2(x)}$ and $\sinh^2(x) \leq x^2$ (Taylor series of sinh), h_2 is an increasing function. Therefore,

$$\begin{aligned} & \frac{1}{\tanh(\theta_2)} - \frac{1}{\theta_2} \leq \frac{1}{\tanh(\theta_1)} - \frac{1}{\theta_1} \\ \Rightarrow \quad & \frac{\frac{1}{\theta_2} - \frac{1}{\theta_1}}{\frac{1}{\tanh(\theta_2)} - \frac{1}{\tanh(\theta_1)}} \geq 1. \end{aligned} \quad (79)$$

Similarly, this confirms that $\Delta_\kappa^{(\alpha)}$ is negative for $\alpha < 1$. It is also easy to verify from (76) that $\Delta_\kappa^{(1)} = 0$.

REFERENCES

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

[2] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Prentice Hall, 1997.

[3] H. Akaike, "Fitting autoregressive models for prediction," *Ann. Inst. Stat. Math.*, vol. 21, no. 1, pp. 243–247, 1969.

[4] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modeling Extremal Events: For Insurance and Finance (Stochastic Modeling and Applied Probability)*, 2nd ed. Springer, 2008.

[5] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*. Wiley-Interscience, 2001.

[6] R. A. Davis, K. Knight, and J. Liu, "M-estimation for autoregressions with infinite variance," *Stoch. Proc. App.*, vol. 40, pp. 145–180, Feb. 1992.

[7] S. Ling, "Self-weighted least absolute deviation estimation for infinite variance autoregressive models," *J. R. Stat. Soc. B*, vol. 67, no. 3, pp. 381–393, Jun. 2005.

[8] Y. Cho and L. K. Saul, "Learning dictionaries of stable autoregressive models for audio scene analysis," in *Int. Conf. on Machine Learning (ICML), Montreal, Canada*, Jun. 14–18, 2009.

[9] H. Ahn and R. E. Feldman, "Optimal filtering of a Gaussian signal in the presence of Lévy noise," *SIAM J. App. Math.*, vol. 60, no. 1, pp. 359–369, Nov.-Dec. 1999.

[10] D. Sornette and K. Ide, "The Kalman-Lévy filter," *Physica D*, vol. 151, no. 2-4, pp. 142–174, May 2001.

[11] A. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[12] D. Alspach and H. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Auto. Control*, vol. 17, no. 4, pp. 439–448, Aug. 1972.

[13] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs: Prentice Hall, 1979.

[14] M. West, P. J. Harrison, and H. S. Migon, "Dynamic generalized linear models and Bayesian forecasting," *J. Amer. Stat. Assoc.*, vol. 80, no. 389, pp. 73–83, Mar. 1985.

[15] H. Weinert and G. Sidhu, "A stochastic framework for recursive computation of spline functions—Part I: Interpolating splines," *IEEE Trans. Inform. Theo.*, vol. 24, no. 1, pp. 45–50, Jan. 1978.

[16] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Trans. Sig. Proc.*, vol. 53, no. 6, pp. 2146–2159, Jun. 2005.

[17] G. Kitagawa, "Non-Gaussian state-space modeling of nonstationary time series," *J. Amer. Stat. Assoc.*, vol. 82, no. 400, pp. 1032–1041, Dec. 1987.

[18] ——, "Monte Carlo filter and smoother for non-Gaussian nonlinear state-space models," *J. Comput. and Graph. Stat.*, vol. 5, no. 1, pp. 1–25, Mar. 1996.

[19] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. and Comput.*, vol. 10, no. 3, pp. 197–208, Jul. 2000.

[20] T. A. Marsh and E. R. Rosenfeld, "Stochastic processes for interest rates and equilibrium bond prices," *J. Finance*, vol. 38, no. 2, pp. 635–646, May 1983.

[21] O. E. Barndorff-Nielsen, "Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics," *J. R. Stat. Soc. B*, vol. 63, no. 2, pp. 167–241, 2001.

[22] J. E. Griffin and M. F. J. Steel, "Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility," *J. Econometrics*, vol. 134, no. 2, pp. 605–644, Oct. 2006.

[23] F. E. Benth, J. Kallsen, and T. MeyerBrandis, "A nonGaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing," *App. Math. Finance*, vol. 14, no. 2, pp. 153–169, May 2007.

[24] S. Mudchanatongsuk, J. A. Primbs, and W. Wong, "Optimal pairs trading: A stochastic control approach," in *Amer. Control Conf.*, Jun. 11-13, 2008, pp. 1035–1039.

[25] P. S. Hagan and G. West, "Interpolation methods for curve construction," *App. Math. Finance*, vol. 13, no. 2, pp. 89–129, Jun. 2006.

[26] H. Kirshner and M. Porat, "On the role of exponential splines in image interpolation," *IEEE Trans. Image Proc.*, vol. 18, no. 10, pp. 2198–2208, Oct. 2009.

[27] M. Unser, P. Tafti, and Q. Sun, "A unified formulation of Gaussian *vs.* sparse stochastic processes: Part I—Continuous-domain theory," *arXiv:1108.6150v1*.

[28] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*. Chapman & Hall, 1994.

[29] S. P. Lloyd, "A sampling theorem for stationary (wide sense) stochastic processes," *Trans. Amer. Math. Soc.*, vol. 92, no. 1, pp. 1–12, Jul. 1952.

[30] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes*. Chapman & Hall/CRC, 1994.

[31] A. Amini, M. Unser, and F. Marvasti, "Compressibility of deterministic and random infinite sequences," *IEEE Trans. Sig. Proc.*, vol. 59, no. 11, pp. 5193–5201, Nov. 2011.

[32] K. B. Howell, *Principles of Fourier Analysis*. Chapman & Hall/CRC, 2001.

Arash Amini received B.Sc., M.Sc. and Ph.D. degrees in electrical engineering (Communications and Signal Processing) in 2005, 2007, and 2011, respectively, and the B.Sc. degree in petroleum engineering (Reservoir) in 2005, all from Sharif University of Technology (SUT), Tehran, Iran. Since April 2011, he is a post-doctoral researcher in Biomedical Imaging Group (BIG), École polytechnique fédérale de Lausanne, Lausanne, Switzerland. His research interests include different aspects of sampling, specially compressed sensing.

Philippe Thévenaz received the Diploma degree in microengineering from École polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland, and the Ph.D. degree from the Institute of Microtechnology (IMT), University of Neuchâtel, Neuchâtel, Switzerland, in 1986 and 1993, respectively.

He was a Visiting Fellow with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, from 1993 to 1998, where he developed splines and multiresolution signal representations, geometric image transformations, and biomedical image registration. Since 1998, he has been at EPFL as a Senior Researcher.

John Paul Ward received a B.S. degree in mathematics from the University of Georgia and a Ph.D. in mathematics from Texas A&M University in 2005 and 2010, respectively. He is currently a post-doctoral researcher in the Biomedical Imaging Group at the Swiss Federal Institute of Technology (EPFL).

Michael Unser (M'89SM'94F'99) received the M.S. (summa cum laude) and Ph.D. degrees in Electrical Engineering in 1981 and 1984, respectively, from the Ecole Polytechnique Fdrale de Lausanne (EPFL), Switzerland. From 1985 to 1997, he worked as a scientist with the National Institutes of Health, Bethesda USA. He is now full professor and Director of the Biomedical Imaging Group at the EPFL.

His main 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. He has published 200 journal papers on those topics, and is one of ISIs Highly Cited authors in Engineering (http://isihighlycited.com).

Dr. Unser has held the position of associate Editor-in-Chief (2003-2005) for the IEEE Transactions on Medical Imaging and has served as Associate Editor for the same journal (1999-2002; 2006-2007), the IEEE Transactions on Image Processing (1992-1995), and the IEEE Signal Processing Letters (1994-1998). He is currently member of the editorial boards of Foundations and Trends in Signal Processing, and Sampling Theory in Signal and Image Processing. He co-organized the first IEEE International Symposium on Biomedical Imaging (ISBI2002) and was the founding chair of the technical committee of the IEEE-SP Society on Bio Imaging and Signal Processing (BISP).

Dr. Unser received the 1995 and 2003 Best Paper Awards, the 2000 Magazine Award, and two IEEE Technical Achievement Awards (2008 SPS and 2010 EMBS). He is an EURASIP Fellow and a member of the Swiss Academy of Engineering Sciences.