Digital Signal Processing 57 (2016) 79–92

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Parameters estimation for continuous-time heavy-tailed signals modeled by α-stable autoregressive processes

Zeinab Hashemifard [a], Hamidreza Amindavar [a,*], Arash Amini [b]

[a] Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran
[b] Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Available online 30 June 2016

Keywords:
Infinite variance α-stable process
Continuous-time autoregressive model
M-estimator
Parameter estimation
Heavy-tailed signal

ABSTRACT

In this paper, we focus on the heavy-tailed stochastic signals generated through continuous-time autoregressive (CAR) models excited by infinite-variance α-stable processes. Our goal is to estimate the parameters of the continuous-time model, such as the autoregressive coefficients and the distribution parameters related to the excitation process for the α-stable CAR process with $0 < \alpha < 2$ based on the state-space representation. Likewise, we investigate the closed form expressions for the parameters of equivalent model in the discrete-time setting via regular samples of the process. We analyze the estimator based on the Monte Carlo simulations and illustrate the estimator consistency to the desired values when sampling frequency and sample size tend to infinity. We also apply the proposed method to the two types of real-world data, financial and ground magnetometer data, to evaluate its performance in real environments.

1. Introduction

There are two main categories of statistical signal models, namely, discrete and continuous. Because of the simplicity of the model and the involved mathematical tools, the discrete-domain framework is by far the widely adopted choice in engineering applications, while discrete-time modeling is an approximation for real-time phenomena. However, in different fields of studies, continuous-time models have been frequently used, such as automatic control and signal processing [1,2]. Typical benefits associated to continuous modeling over the discrete models are in the signal processing applications. In this respect, the problem of missing data in discrete domain can be alleviated through modeling the signal in the continuous domain. Similarly, the problem of treating with discrete samples of an irregular sampled process can be simplified when the signal is modeled in the continuous domain from the beginning. Besides, interpolation of intermediate data points within the range of discrete samples can be properly resolved by defining the signal model in the continuous-time domain. Another substantial advantage of continuous-time modeling rather than commonly used discrete models is the flexibility of continuous-time models to varying the sampling frequency, whereas discrete-time modeling is conditioned on the sampling frequency. As a result, a discrete model may switch to another one when the sampling frequency is changed. Consequently, the continuous-domain models generally provide more accurate and detailed insight regarding the signals of interest. In this respect, there are already advanced stochastic models in the continuous-domain such as the continuous-time autoregressive (CAR) processes [3–5] that could simplify the modeling task. The CAR stochastic processes play remarkable role in the signal processing applications, such as speech analysis and synthesis, adaptive filtering, identification of continuous-time systems and financial modeling [6–9].

In the past decade, the impulsiveness nature of real-world signals has been the focus of a number of researches [10]. A rich class of probability distributions that allows heavy tail is the stable distribution family. Lévy-stable processes have been proposed as a model that can accurately represent many kinds of real-world data and physical systems such as hitting times for a Brownian motion, the gravitational field of stars, financial time series, ground magnetometer data, sea clutter, noise processes in the impulsive environment, co-channel interference in ad-hoc and cellular network environments, impulsive bioacoustic signals and teletraffic data [11–15].

The Lévy-stable CAR processes provide the advantages of continuous-time modeling for the statistically dependent and heavy-tailed processes. Estimating the model parameters of CAR process driven by α-stable process is a challenging task due to continuous-time modeling and infinite variance assumption. Unlike the Gaussian case, the optimal estimators in the case of

* Corresponding author.
E-mail addresses: hashemifard@aut.ac.ir (Z. Hashemifard), hamidami@aut.ac.ir (H. Amindavar), aamini@sharif.edu (A. Amini).

http://dx.doi.org/10.1016/j.dsp.2016.06.013
1051-2004/

heavy-tailed stochastic processes are nonlinear, and thus, computationally more difficult. The estimation problems such as denoising and interpolation have been studied in [10,16], where the model has been assumed to be known. Furthermore, in [17] a general family of the Lévy processes have been considered for the excitation signals of multivariate continuous-time autoregressive moving average (MCARMA) processes. The autoregressive and moving average coefficients of the MCARMA model have been all assumed to be known. Essentially, the authors have only studied the parameters estimation of the excitation distribution using the generalized method of moments on the approximated increments of the excitation process. In [2], the challenge of estimating the model parameters in the case of finite variance CAR processes has been partially studied. In particular, it has been shown that the uniform samples of the pth order CAR process can lead to exact identification of the differential equation, under finite second-order moment assumption of the innovation process. The latter assumption allows for the evaluation of the autocorrelation function and the spectrum of the process. As the second-order moment (variance) is infinite in the non-Gaussian stable distributions, the autocorrelation function (and consequently, spectrum) cannot be defined for the α-stable processes with $\alpha < 2$. This fact rules out the standard method of solving Yule–Walker equations. The alternative used in [18] for estimation of the parameters in a univariate symmetric stable discrete-domain autoregressive process is based on the so-called *covariation*. Feng et al. [19] have used a similar estimation technique in a radar application where the correlated non-Gaussian clutter background has been modeled by symmetric α-stable (SαS) fractional autoregressive system. In [20,21], using the M-estimator which is based on the least L_ν-norm ($\nu < \alpha$), the parameters of discrete autoregressive signals with the α-stable innovation have been investigated. The authors have shown that in the case of finite first moment assumption of the heavy-tailed noise, the M-estimator performs well. The CAR processes of order $p = 1$ are widely known as Ornstein–Uhlenbeck (OU) processes which involve the smallest number of model parameters among the CAR family. The statistical estimation of the model parameters of OU processes with infinite variance assumption has been considered by the least squares estimator in [22,23], while the least-square (LS) estimator cannot properly handle the outliers in the heavy-tailed scenarios.

In this study, we investigate the model parameters estimation of CAR(p) process generated by linear stochastic differential equation excited by infinite-variance Lévy-stable process. In the discrete-time models, for the heavy-tailed processes, the least absolute deviation estimator is more efficient than the LS estimator [24], hence, we extend this issue to the use of the M-estimators in the continuous-time setting for the α-stable linear models employed in this research. We employ the covariation technique along with M-estimators to learn the autoregressive coefficients of the model. We assume to have access to uniform samples of the CAR(p) process on a fine grid. The sampled CAR(p) processes in regular time grids yield discrete ARMA($p, p-1$) processes with generally dependent excitation processes [3]. In this regard, we show that under α-stable assumption, we have independent excitation process in the discrete-time setting. After that, we investigate renewed equations according to sampled CAR(p) process and closed form expressions relevant to model coefficients in the discrete-domain. The statistical analyses of the estimator show that the accuracy and deviation of the estimated parameters of the continuous-time model depend on the sampling frequency and sample size. The proposed learning approach is also successfully applied to the two types of real-world data; ground magnetometer data and financial time series.

The rest of the paper is organized as follows. In Section 2, we explain the employed model by briefly reviewing the Lévy-stable processes and the α-stable CAR processes. In Section 3, we describe the parameter learning procedure. The learning consists of estimating both the differential equation coefficients and the parameters identifying the excitation distribution. Using the Monte Carlo simulation, we assess the consistency and statistical properties of the estimator in Section 4. Performance of the proposed method is numerically evaluated in Section 5. We also apply the estimator on the real-world data sets in Section 6 and after that, in the last section, the concluding remarks are discussed.

2. Signal model

This section introduces our continuous-time stochastic model that includes two main parts: continuous-time innovation process and linear differential equation. To induce heavy-tailedness for the signal model, it is assumed that the excitation process is generated from α-stable distribution with $0 < \alpha < 2$ [10]. Furthermore, the dependency structure of signal model depends on the linear differential equation. In this work, we deal with CAR(p) process as a stationary solution of a p-order differential equation.

2.1. Lévy-stable process

A Lévy process $L(t)$, $t > 0$ is a continuous-domain stochastic process with stationary and independent increments. The widely-used concept of white noise in the engineering literature corresponds to the generalized derivative of Lévy processes; in particular, the white Gaussian noise is the derivative of Brownian motion. A Lévy process is completely characterized by its characteristic function (CF), $\Phi_L(\omega)$,

$$\Phi_{L(t)}(\omega) = E\{e^{j\omega L(t)}\} = \exp\{t\Psi_L(\omega)\}, \ \omega \in \mathbb{R} \tag{1}$$

An α-stable Lévy process with the distribution parameters $(\alpha, \rho, \mu, \beta)$, is a real-valued process which satisfies

$$\Psi(\omega) = \begin{cases} j\mu\omega - \rho^\alpha |\omega|^\alpha \{1 + j\beta \operatorname{sgn}(\omega)\tan(\alpha\pi/2)\} & \text{if } \alpha \neq 1 \\ j\mu\omega - \rho|\omega| \left\{1 + j\frac{2}{\pi}\beta \operatorname{sgn}(\omega)\log|\omega|\right\} & \text{if } \alpha = 1 \end{cases} \tag{2}$$

Particularly, in a symmetric α-stable distribution, SαS(α, ρ), by setting $\mu = 0$ and $\beta = 0$ the characteristic function exponent has the form

$$\Psi(\omega) = -\rho^\alpha |\omega|^\alpha. \tag{3}$$

In (2), μ is the shift parameter that determines the mean of the distribution of $L(1)$ when $1 < \alpha \leq 2$ and its median when $0 < \alpha \leq 1$. The spread of the PDF around the mean/median is specified by the scale parameter $\rho > 0$, and the symmetry of the distribution around its location parameter is indicated by β (for a symmetric case $\beta = 0$). The characteristic exponent α is a fundamental parameter in an α-stable distribution that determines the shape, the tail and the decay rate of the distribution, while the smaller α represents the more heavy tail. The extreme case of $\alpha = 2$ corresponds to the Gaussian distribution, nevertheless, any value of $\alpha < 2$ results in a distribution with infinite variance. The usual representation of an α-stable distribution is via its characteristic function, as there is no known closed-form expression for its PDF except for the Gaussian ($\alpha = 2$, $\beta = 0$), Cauchy ($\alpha = 1$, $\beta = 0$) and Lévy distribution ($\alpha = 0.5$, $\beta = 1$).

2.2. α-stable CAR(p) process

Since we focus on the heavy-tailed stochastic process modeled by the α-stable CAR(p) process, we review some pertinent results. A CAR(p) process can be modeled in terms of a continuous-time innovation process and a stochastic integral. Naturally, a CAR(p) process $x(t)$ is defined as a stationary solution of a linear differential equation. The process of interest $x(t)$ is defined as,

$$\mathrm{D}^p x(t) + \sum_{i=1}^{p} a_i \mathrm{D}^{p-i} x(t) = \mathrm{D}L(t), \tag{4}$$

where D denotes the differentiation with respect to t (i.e. $\mathrm{D} \triangleq d/dt$), $L(t)$ represents a Lévy process, and $a_1, \ldots, a_p$ are constant coefficients of stochastic differential equation. In the signal processing literature, $\mathrm{D}L(t)$ shows the white noise process in (4), because Lévy processes have stationary and independent increments. In (4), our model includes two main parts; continuous-time excitation process and whitening operator. The whitening operator can be considered as linear operator, where the shaping operator (inverse of linear operator) converts the white noise to the signal of interest $x(t)$. Thus, the inverse of whitening operator determines the dependency structure of signal. The whitening operator that converts our interested signal $x(t)$ into the white noise is defined as

$$a(\mathrm{D}) = \mathrm{D}^p + a_1 \mathrm{D}^{p-1} + \cdots + a_p I \tag{5}$$

where I denotes the identity factor. Since Lévy processes are not differentiable in the usual sense, the expression in (4) should be equivalently written in the state-space form [3]. Let $\underline{Y}(t)$ be a vector that its ith component shows the $(i-1)$th derivative of $x(t)$ as,

$$\underline{Y}(t) \triangleq \begin{bmatrix} x(t) & \mathrm{D}x(t) & \cdots & \mathrm{D}^{p-1}x(t) \end{bmatrix}^T. \tag{6}$$

Therefore, (4) can be rewritten as the form of

$$\mathrm{d}\underline{Y}(t) = \mathbf{A}\underline{Y}(t)dt + \underline{c}\,\mathrm{d}L(t) \tag{7}$$

$$x(t) = \underline{b}^T \underline{Y}(t), \tag{8}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{bmatrix}, \tag{9}$$

$$\underline{b} = [\,1 \quad 0 \quad \cdots \quad 0 \quad 0\,]^T, \underline{c} = [\,0 \quad 0 \quad \cdots \quad 0 \quad 1\,]^T. \tag{10}$$

The usual differential operator in calculus is shown by d (for instance $\int f(x)\,\mathrm{d}x$ or $\frac{\mathrm{d}}{\mathrm{d}x} f(x)$). Additionally, in our notation, vectors are denoted as underlined characters while matrices are upper case characters in bold face. Last investigations have shown that the equidistant sampled CAR(1) process produces the discrete AR(1) process with generally i.i.d. noise, although this equivalency is not held for higher order CAR(p) processes [3]. Considering this fact helps solve the Itō differential equation (7), because the vector $\underline{Y}(t)$ is similar to a multivariate CAR(1) process, and consequently its sampled version is analogous to multivariate AR(1) process with a multivariate excitation process [3]. Hence, this equation clearly has the solution of

$$\underline{Y}(t) = \mathrm{e}^{\mathbf{A}t}\underline{Y}(0) + \int_0^t \mathrm{e}^{\mathbf{A}(t-u)}\underline{c}\,\mathrm{d}L(u), \ \forall t \geq 0. \tag{11}$$

If $p = 1$, $x(t)$ is known as a Lévy OU process. The strictly causal stationary solution of (11) exists, if $\underline{Y}(0)$ is independent of $\underline{Y}(t)$ for $t > 0$ and it has the similar distribution to $\int_{-\infty}^{0} \mathrm{e}^{-\mathbf{A}u}\underline{c}\,\mathrm{d}L(u)$, and also if the polynomial $a(\mathrm{D})$ has the roots with negative real parts [3]. As a result, for access the strictly causal stationary CAR process $x(t)$ from $\underline{Y}(t)$, these assumptions are essential. In a similar manner, using the signal processing literature, an autoregressive process $x(t)$ is stationary if its rational transfer function in the Laplace domain has the poles with the negative real part. For this reason, it similarly requires to negative real parts of the zeroes of polynomial $a(\varsigma) = \varsigma^p + a_1\varsigma^{p-1} + \cdots + a_p$, where ς is the variable of Laplace domain.

We can rewrite the Lévy process $\underline{Y}(t)$ stated in (11) as,

$$\underline{Y}(t) = \mathrm{e}^{\mathbf{A}(t-s)}\underline{Y}(s) + \int_s^t \mathrm{e}^{\mathbf{A}(t-u)}\underline{c}\,\mathrm{d}L(u), \ \forall t > s \geq 0. \tag{12}$$

The sampled $\underline{Y}(t)$ and also the sampled $x(t)$ are exactly achieved by imposing $s = (n-1)T$ and $t = nT$ in (12),

$$\underline{Y}[n] = \mathrm{e}^{\mathbf{A}T}\,\underline{Y}[n-1] + \underline{Z}[n], \ n = 1, 2, \ldots \tag{13}$$

and

$$x[n] = \underline{b}^T \underline{Y}[n], \tag{14}$$

where

$$\underline{Z}[n] \triangleq \int_{(n-1)T}^{nT} \mathrm{e}^{\mathbf{A}(nT-u)}\underline{c}\,\mathrm{d}L(u). \tag{15}$$

The time sampling is T. In our notation, parenthesis is used for the argument of the continuous-time signal (e.g. $\underline{Y}(t)$) and bracket is employed for discrete-time one (e.g. $\underline{Y}[n]$), so $\underline{Y}[n] \triangleq \underline{Y}(nT)$.

Due to independent increments of the Lévy process, $\underline{Z}[n]$ statistically represents the multivariate independent sequence for $n = 1, 2, \cdots$ and therefore it yields the multivariate first order Markov process of $\underline{Y}[n]$. To analytically find the CF of $\underline{Z}[n]$, we need a property of the Lévy process integrators that indicates if $q(t) \triangleq \int_{t_1}^{t_2} g(u)\mathrm{d}L(u)$, where $g(.)$ is a bounded continuous function in $\mathbb{R}$, then the CF exponent of $q(t)$ is stated as [25]

$$\ln E\{\exp(\mathrm{j}\omega q(t))\} = \int_{t_1}^{t_2} \Psi_L(\omega g(u))\mathrm{d}u. \tag{16}$$

Therefore, using (15) and (16), the multivariate CF of $\underline{Z}[n]$ is deduced as

$$\Phi_{\underline{Z}}(\underline{\omega}) = E\{\exp(j\underline{\omega}^T \underline{Z})\} = \exp(\Psi_{\underline{Z}}(\underline{\omega})), \tag{17}$$

where

$$\Psi_{\underline{Z}}(\underline{\omega}) = \int_{(n-1)T}^{nT} \Psi_L(\underline{\omega}^T \mathrm{e}^{\mathbf{A}(nT-u)}\underline{c})\mathrm{d}u. \tag{18}$$

Briefly, we discussed the statistical model of the heavy-tailed dependent signal $x(t)$ represented by the α-stable CAR(p) process. Using the state-space representation, we achieved to multivariate CAR(1) process $\underline{Y}(t)$ in (7), where the relation between $\underline{Y}(t)$ and $x(t)$ is presented in (6) and (8). Under some assumptions, the multivariate CAR(1) process $\underline{Y}(t)$ has a strictly stationary solution as stated in (11). Therefore, the sampled version of $\underline{Y}(t)$ yields to a multivariate discrete AR(1) $\underline{Y}[n]$ in (13), with excitation $\underline{Z}[n]$ that its CF exponent is demonstrated in (18).

We present in Fig. 1, the realizations of the first and third order α-stable driven CAR processes with $\alpha = 1.1$ and $\alpha = 1.9$. In addition, we depict the realizations of the increments of their innovation processes.

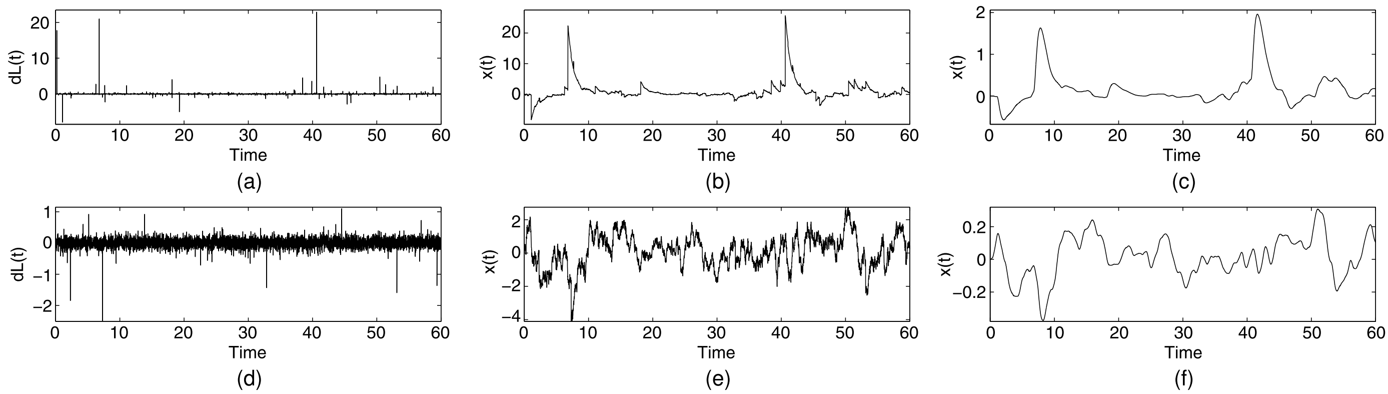


Fig. 1. In the first row, three realizations of (a) the increments of the Lévy-stable process with $\alpha = 1.1$ and the generated (b) CAR(1), (c) CAR(3) processes. In the second row, three realizations of (d) the increments of the Lévy-stable process with $\alpha = 1.9$ and the generated (e) CAR(1), (f) CAR(3) processes.

3. Model parameters estimation

In this work, we consider the CAR(p) signals driven by the symmetric Lévy-stable processes. We propose an estimation technique for learning the autoregressive coefficients, $\underline{a} = \{a_1, \ldots, a_p\}$, expressed in (4) and likewise employ a method for the unknown parameters related to SαS(α, ρ) distribution of the excitation process in (3). The proposed learning method in the continuous-time setting encompasses three main parts that are explained below and summarized in Table 1. In Subsection 3.1, the characteristic exponent α of the excitation distribution and also ρ_x are initially estimated. In Subsection 3.2, in order to estimate the autoregressive coefficients of the continuous-time model, an estimation procedure based on the covariation issue and M-estimation algorithm is proposed. After that, in Subsection 3.3, the scale parameter of the excitation distribution using ρ_x and the estimated autoregressive coefficients is extracted. These described steps are summarized in the first, second and third parts of Table 1, respectively. Furthermore, in the last subsection, the signal model and its parameters in the discrete-time setting are characterized.

3.1. Characteristic exponent estimation of Lévy-stable CAR(p) process

To study the distribution of $x(t)$ as a CAR(p) process driven by SαS(α, ρ) process, the causal strictly stationary process $\underline{Y}(t)$ in (11) can be written in a moving average representation of the form [3]

$$\underline{Y}(t) = \int_{-\infty}^{+\infty} K(t-u)\mathrm{d}L(u), \tag{19}$$

when

$$K(t) \triangleq \mathrm{e}^{\mathbf{A}t}\underline{c}u(t), \tag{20}$$

and $u(t)$ is a step function; $u(t) = 1$ for $t \geq 0$, $u(t) = 0$ for $t < 0$. In this case, using (8) we have

$$x(t) = \int_{-\infty}^{t} \underline{b}^T \mathrm{e}^{\mathbf{A}(t-u)}\underline{c}\mathrm{d}L(u). \tag{21}$$

The independent increments of Lévy process $L(t)$ have SαS distribution with characteristic exponent α. Equation (21) indicates a moving average representation of the independent Lévy increments that yields to α-stable distribution with the same α, because a linear combination of independent α-stable random processes with the same α is an α-stable process.

In order to analytically drive the scale parameter of $x(t)$ and also find relation between the scale parameters of $x(t)$ and excitation process that are denoted as ρ_x and ρ respectively, we can use the univariate CF of $x(t)$ with the expression $\Phi_x(\omega) = \exp(\Psi_x(\omega))$ where

$$\begin{aligned}\Psi_x(\omega) &= \int_{-\infty}^{t} \Psi_L(\omega\underline{b}^T \mathrm{e}^{\mathbf{A}(t-u)}\underline{c})\mathrm{d}u \\ &= -\rho^\alpha|\omega|^\alpha \underbrace{\int_0^\infty |\underline{b}^T\mathrm{e}^{\mathbf{A}u}\underline{c}|^\alpha \mathrm{d}u}_{\triangleq k} = -\rho_x^\alpha|\omega|^\alpha,\end{aligned} \tag{22}$$

when

$$\rho_x \triangleq \rho k^{1/\alpha}, \tag{23}$$

and

$$k \triangleq \int_0^\infty |\underline{b}^T\mathrm{e}^{\mathbf{A}u}\underline{c}|^\alpha \mathrm{d}u. \tag{24}$$

Equation (22) clearly indicates that $x(t)$ has the characteristic component α that is equal to characteristic component of Lévy-stable excitation process, but different scale parameter ρ_x. The expression in (23) shows that the scale parameter ρ_x depends on the scale parameter ρ of excitation process and the autoregressive coefficients $\underline{a}$.

The density parameters Estimation of α-stable process have been frequently considered in the literature [26,27]. Due to the lack of closed form expression for the most of α-stable distributions, the maximum likelihood procedure for the parameters estimation results in considerable computation cost. There are several well-known estimation approaches for the density parameters of α-stable such as empirical characteristic function (ECF), quantile and logarithmic moment. By comparison of these methods in [26], the authors have shown that the ECF approach has the highest accuracy and the best stability for the density parameters estimation, without any restricted assumption on the parameters. As a result, in order to estimate the characteristic exponent α and the scale parameter ρ_x, we employ the ECF of $x(nT)$ as

$$\widehat{\Phi}_x(\omega) = \frac{1}{N}\sum_{i=1}^{N} \exp(\mathrm{j}\omega x(iT)). \tag{25}$$

Using the CF expression of $x(t)$ in (22), we achieve to

$$\log(-\log(|\widehat{\Phi}_x(\omega)|^2)) = \log(2\rho_x{}^\alpha) + \alpha\log(|\omega|). \tag{26}$$

It is possible to estimate α and ρ_x through a linear regression on (26) and employing the method explained in [26]. The first part of Table 1 mentions the estimation algorithm of α and ρ_x.

3.2. Autoregressive parameters estimation

Here, we develop an estimation algorithm for the autoregressive coefficients expressed in (4) related to an infinite variance CAR(p) process $x(t)$ driven by SαS(α, ρ) process. The main challenges of the estimation procedure for the model parameters of $x(t)$ are three points: Firstly, due to the lack of finite second moment, we are prevented from employing the methods based on autocorrelation issue. Secondly, no closed form solution for the probability distribution of the process causes that the inference procedures according to distribution function undertake significant computational costs. Thirdly, the samples of CAR(p) process $x(t)$ in the discrete-time setting can not be modeled by a p-order autoregressive process as a p-order Markov model. In our learning algorithm for the autoregressive coefficients of $x(t)$, we employ the sampled data and estimated parameter $\hat{\alpha}$, and also utilize the covariation concept and the M-estimation method. To this end, the covariation technique is initially employed, when $\alpha > 1$. Next, the estimated parameters are used as initial values for the M-estimation of autoregressive coefficients. Details of the learning algorithm for autoregressive coefficients are explained in the following and summarized in the second part of Table 1.

Traditionally, the autoregressive parameters of discrete-time processes with finite variance property are estimated by Yule–Walker equations with an appropriate accuracy. Besides, in the last decade, under infinite variance assumption, the model parameter of the first order CAR process has been estimated by the LS estimator [22,23]. Nevertheless, in the discrete-time setting the method based on the sample covariance has been proposed as a measure for the parameters estimation of univariate SαS autoregressive processes. It has been shown that the confidence interval in the covariation method performs better than those based on the autocorrelation function [18]. We extend this topic to estimate the parameters relevant to autoregressive coefficients of the continuous-time setting. To this end, we use the discretized state-space equation (13) that converts the system model to a multivariate Lévy OU process with an exponential matrix, $e^{\mathbf{A}T}$, corresponding to the autoregressive parameters. We multiply the both sides of (13) by a vector $\xi(\underline{Y}_{n-1})^T$ as a p-variate function of $\underline{Y}_{n-1}$. After that by applying an expectation operator, it achieves to

$$E\{\underline{Y}_n\xi(\underline{Y}_{n-1})^T\} = \mathbf{e}^{\mathbf{A}T}E\{\underline{Y}_{n-1}\xi(\underline{Y}_{n-1})^T\} + E\{\underline{Z}_n\xi(\underline{Y}_{n-1})^T\}. \quad (27)$$

Since $\underline{Z}_n$ and $\underline{Z}_{n-1}$ are independent, it yields to $E\{\underline{Z}_n\xi(\underline{Y}_{n-1})^T\} = 0$. On the other hand, by defining

$$C_{n,n-k} \triangleq E\{\underline{Y}_n\xi(\underline{Y}_{n-k})^T\}, \quad k=0,1, \quad (28)$$

and employing stationary property of $\underline{Y}_n$ which yields to $C_{n-1,n-1} = C_{n,n}$, we access to

$$C_{n,n-1} = e^{\mathbf{A}T}C_{n,n}. \quad (29)$$

Since $\underline{Y}[n]$ is a multivariate α-stable variable, the fractional moments of lower order than α for every entry of $\underline{Y}[n]$ exist. Considering this fact and the covariation concept help us choose an appropriate function for $\xi(\cdot)$ that causes limited values for the entries of matrix $C_{n,n-k}$.

$$\xi(\underline{Y}[n]) \triangleq \begin{bmatrix} s_1[n] & \cdots & s_p[n] \end{bmatrix}^T, \quad (30)$$

where

$$s_i[n] \triangleq sign(y_i[n]) \mid y_i[n] \mid^{\nu-1}, \; \nu < \alpha \quad \text{if} \quad y_i[n] \neq 0.$$
$$s_i[n] \triangleq 0 \quad if \quad y_i[n] = 0. \quad (31)$$

The sample version of the empirical covariation matrix is stated as

$$\hat{C}_{n,n-k} = \frac{1}{N}\sum_{n=1}^{N}\underline{Y}_n\xi(\underline{Y}_{n-k})^T, \quad k=0,1. \quad (32)$$

The covariation matrix $C_{n,n-k}$ for $\nu < 2$ may not be a positive definite matrix and the heavy-tailed process $\underline{Y}[n]$ for $\alpha < 2$ makes the non-symmetric covariation matrix $C_{n,n-k}$ non-invertible. Hence, to derive an exponential matrix estimator $\widehat{e^{\mathbf{A}T}}$, (32) is substituted in (29) and then a LS estimator is used as

$$\widehat{e^{\mathbf{A}T}} = \arg\min_{e^{\mathbf{A}T}}(\|\hat{C}_{n,n-1} - e^{\mathbf{A}T}\hat{C}_{n,n}\|_2). \quad (33)$$

Further, the known structure of matrix $\mathbf{A}$ from (9) is inserted in (33) to arrive at more accurate estimation by a nonlinear optimization method of

$$[a_p, \cdots, a_1] = \arg\min_{\underline{a}}(\|\hat{C}_{n,n-1} - f(\underline{a})\hat{C}_{n,n}\|_2), \quad (34)$$

where

$$f(\underline{a}) \triangleq \exp\left(\begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \\ -a_p & \cdots & -a_2 & -a_1 \end{bmatrix} T\right). \quad (35)$$

The non-symmetric and non-invertible properties of $C_{n,n-k}$ reduce the accuracy of the estimator. In this regard, lower values of α (e.g. $\alpha \in (0,1)$) cause very impulsive signal $L(t)$ and then the estimator of (34) does not converge to a proper value. Consequently, for $\alpha \geq 1$, the estimated parameters of (34) as initial values in another optimization task associated to the M-estimation approach are used. But, for $\alpha < 1$, the covariation-based estimator of (34) is not utilized, whereas the M-estimation approach without initialization is used.

In the M-estimator for the discrete autoregressive process, the loss function is $\Im(x) = |x|^\nu$ which especially yields to LS and least absolute deviation (LAD) when $\nu = 2$ and $\nu = 1$, respectively. Unlike the LS estimator, the M-estimator criterion provides a suitable measure for the heavy-tailed process with infinite variance [28]. The strong consistency of the M-estimator for discrete-time autoregressive processes with the finite first moment has been proven in the literature [28]. Also, several simulation studies for the discrete α-stable AR(p) process have depicted that for $\alpha \in (0,1)$, the loss function $\Im(x) = |x|^\nu$ is optimal when $\nu = \alpha$. Therefore, we use these results in our case of the multivariate first order Markov process $\underline{Y}[n]$. Since the discrete multivariate excitation process $\underline{Z}[n]$ in (13) has statistically independent multivariate distribution for $n = 0, 1, \ldots$, we apply the M-estimator with loss function $\Im(x) = |x|^\nu$ to (13) as

$$\widehat{e^{\mathbf{A}T}} = \arg\min_{e^{\mathbf{A}T}}\left(\sum_{n=2}^{N}\|\underline{Y}[n] - e^{\mathbf{A}T}\underline{Y}[n-1]\|_\nu\right). \quad (36)$$

It is also beneficial to consider again the structure of matrix $\mathbf{A}$ in the optimization method for a more precise estimation of $\{a_1, \ldots, a_p\}$.

$$[a_p \; \cdots \; a_1] = \arg\min_{\underline{a}}\left(\sum_{n=2}^{N}\|\underline{Y}[n] - f(\underline{a})\underline{Y}[n-1]\|_\nu\right), \quad (37)$$

where $f(\underline{a})$ is defined in (35). Also, ν is selected according to estimated parameter of α. It is substantial that besides the higher accuracy, the proposed method encompasses more advantages than a simple M-estimator, because the number of parameters that should

Table 1
Learning method.

An outline of the proposed procedure for the model parameters estimation of the α-stable CAR(p) processes.
Inputs: $x(t)$ and p; Outputs: $\hat{\underline{a}}$ and $(\hat{\alpha}, \hat{\rho})$
I. Estimate of the characteristic exponent α and scale parameter ρ_x,
• Sampling $x(t)$.
• Compute ECF of $x(t)$ using (25).
• Estimate α and ρ_x by ECF and linear regression on (26).
II. Estimate of the differential equation coefficients $\underline{a}$,
• For $n = 1, ..., N$
• Approximate the derivatives of $x(t)\|_{t=nT}$ using (38), then approximate the sampled version of (6) (i.e. $\underline{Y}[n]$).
• If $\alpha > 1$,
• Estimate the covariation matrix by (32).
• Compute the initial values $\underline{a}_0$ by a LS in (34).
• Estimate autoregressive coefficients $\underline{a}$ by minimizing the loss function of M-estimator in (37), with the initial values $\underline{a}_0$.
• Else,
• Estimate autoregressive coefficients $\underline{a}$ by minimizing the objective function of M-estimator in (37) without initialization.
III. Estimate of the scale parameter ρ,
• Compute $\mathbf{e}^{\mathbf{A}T}$.
• Compute k from (24).
• Extract ρ from (39) using k and ρ_x.

be estimated decreases to p from $p \times p$ and the valid structure of $\mathbf{A}$ is preserved. Furthermore, satisfactory initial values for $\{a_1, ..., a_p\}$ from (34), improve the speed of the convergence of the optimization method.

In summary, our estimation procedure for autoregressive coefficients depends on the estimated value of α. If $\alpha \geq 1$, the objective function of M-estimator is optimized as stated in (37) with initial values that are firstly computed from the covariation approach through the minimization in (35). But, for $\alpha < 1$ the optimization of the M-estimator in (37) is considered without initialization. Furthermore, the value of ν in M-estimation is conditioned on α. In the literature [28], simulation results have shown that for the M-estimation with loss function $\Im(x) = |x|^{\nu}$, the optimal choice of ν, when $\alpha \in [1, 2)$ is $\nu = 1$ and the best selection of ν for $\alpha \in (0, 1)$ is $\nu = \alpha$.

The explained optimization procedure needs the $(p-1)$ derivatives of $x(t)$ in the sampling time instants to construct $\underline{Y}_n$. The approximation of derivatives for the Lévy continuous-time process needs the observed samples of the CAR process with an appropriate sampling frequency. There are two different assumptions for sampling frequency, that both yield to proper approximation of required signal derivatives. In the first case, an efficient sampling time T for proper approximation of derivatives is assumed and then all the signal observations are on the uniformly-spaced grids with the fixed size T. In the second case, our main observations on the uniformly-spaced grids are assumed with a low and arbitrary frequency sampling $1/T$, while there is access to $p-1$ high frequency samples with a mesh size h ($h < T$) in each time grid [29]. Throughout the second assumption, we do not need a full time high frequency sampling. The local high frequency data besides low frequency samples offer the required data for an appropriate approximation. The most commonly used approximations of derivatives are the forward, backward and the central difference approximations. The consistent backward estimation of the derivatives is defined as [29]

$$\hat{\mathrm{D}}^k x(t) \triangleq \frac{1}{h^k} \sum_{i=0}^{k} (-1)^i \binom{k}{i} x(t - ih), \text{ for } k = 0, ..., p-1, \tag{38}$$

where the error is $O(h)$ in the Taylor series expansions. Hence, it is possible to construct $\hat{\underline{Y}}_n$ from (38) and employ in the optimization task of (34) and (37) to estimate the model parameters. Nevertheless, to reduce the approximation error of the signal derivatives to $O(h^2)$, we should employ the central difference approximation. Further, to achieve more precise approximations, we can employ a greater number of samples using closed-form expressions represented in [30].

3.3. Scale parameter estimation of excitation distribution

In the learning procedure, first, we estimate the characteristic exponent α related to excitation process. After that, using the proposed technique, the coefficients $\underline{a}$ of the stochastic differential equation in (4) are estimated. Here, inference of the model parameters of the SαS(α, ρ) excitation distribution is completed by determining its scale parameter ρ, that is mentioned in the third part of Table 1. From (23), the scale parameter of excitation distribution depends on the scale parameter of $x(t)$ and the autoregressive coefficients $\underline{a}$. Using ρ_x which is derived by ECF and linear regression on (26), and also the autoregressive coefficients that are estimated from (37), it is explicitly possible to derive k from (24) and then the scale parameter ρ corresponding to the excitation process by (23)

$$\rho = \frac{\rho_x}{k^{1/\alpha}}. \tag{39}$$

Briefly, in our learning procedure, we investigated the parameter estimation of the continuous-time signals modeled by the SαS driven differential equation under infinite variance assumption. The proposed learning method in the continuous-time setting encompasses three main parts that are summarized in Table 1.

3.4. Characterizing the signal model in the discrete-time setting

Since we usually access the discrete-time samples of received signals in the signal processing applications, it is essential to understand the characteristics of the discrete-time model obtained by sampling the continuous-time process. For the stationary and causal process $x(t)$ described by (4), the sampled process $x[n]$ represents an ARMA$(p, p-1)$ model with generally dependent excitation process. It satisfies the following equation [3]

$$\Theta(B)x[n] = s[n], \tag{40}$$

where

$$\Theta(z) = \prod_{i=1}^{p} (1 - \mathrm{e}^{\lambda_i T} z^{-1}), \tag{41}$$

B is a backshift operator; $B^k x[n] = x[n-k]$, and $\{\lambda_1, \cdots, \lambda_p\}$ are the roots of polynomial $a(D)$. The polynomial function $\Theta(z)$ in the discrete domain is associated to polynomial $a(\varsigma)$ in the continuous setting, where ς is the Laplace symbol. On the other words, $\Theta(z)$ is the discrete counterpart of $a(\varsigma)$. The polynomial function $\Theta(z)$ in the time domain is stated as

$$\theta[n] = Z^{-1}\{\Theta(z)\} = \sum_{k=0}^{p} q[k]\delta[n-k]. \tag{42}$$

In (40), generally $s[n]$ is a dependent sequence with the expression of [3]

$$s[n] = \underline{b}^T \sum_{m=0}^{p-1}\sum_{k=0}^{m} q[k]\mathrm{e}^{\mathbf{A}(m-k)T}\underline{M}_{n-m}, \tag{43}$$

where

$$\underline{M}_n = \int_{(n-1)T}^{nT} \mathrm{e}^{\mathbf{A}(nT-u)}\underline{c}\mathrm{d}L(u). \tag{44}$$

Considering (42), it is explicit that $\theta[n]$ is an FIR filter of length $(p+1)$ expressed by $\{q[k]\}_{k=0,\ldots,p}$ which converts $x[n]$ to $s[n]$ with expression (43). During the last decades, under finite variance assumption for the Lévy process [3], $s[n]$ has been modeled by a moving average of white noise sequence with uncorrelated and not necessarily independent components as $s[n] = \Phi(B)u[n]$, where the white noise process $u[n]$ facilitates the computational complexities. Unlike, in our scenario, due to infinite variance of the Lévy-stable process (when $\alpha < 2$), the correlation concept is not appealing, therefore we need to derive the dependency structure of the $s[n]$ sequence. We initially derive a closed form expression for the coefficients of $\theta[n]$ in (42) using (41) as

$$\begin{aligned}
&\text{for } k=0: \quad q[k] = 1\\
&\forall k \in (0:p): \quad q[k] = (-1)^k \times \underbrace{\sum_{i_1=1}^{p-k+1}\sum_{i_2=i_1+1}^{p-k+2}\cdots\sum_{i_k=i_{k-1}+1}^{p}}_{k\ times}\\
&\qquad\qquad \times \exp\Bigg\{\underbrace{(\lambda_{i_1}+\cdots+\lambda_{i_k})}_{k\ terms}T\Bigg\}.
\end{aligned} \tag{45}$$

After that, by substituting (44) in the expression of (43), we explicitly show that in each sampling instant, $t = nT$, $s[n]$ is generated from a sum of p independent terms relating to $(p-1)$ former, that satisfies the $(p-1)$th order Markov model, as

$$s[n] = \sum_{m=0}^{p-1}\int_{(n-m-1)T}^{(n-m)T}\left(\sum_{k=0}^{m} q[k]\underline{b}^T\mathrm{e}^{\mathbf{A}((n-k)T-u)}\underline{c}\right)\mathrm{d}L(u). \tag{46}$$

Based on the celebrated property of the Lévy processes that satisfies the independent nature of the non-overlapping increments, we conclude the independency among the p terms in the external summation from (46). As mentioned above, the expression of $s[n]$ demonstrates the $(p-1)$th order Markov model. We investigate the statistical dependency of $s[n]$ for the α-stable process as a substantial class of Lévy processes. We propose the following theorem to present $s[n]$ is a linear moving average of the i.i.d. α-stable process.

Theorem 1. *For the samples of a CAR(p) signal $x(t)$ expressed in (4) and driven by $S\alpha S(\alpha,\rho)$ process $L(t)$, the filtered process $s[n]$ in (40), demonstrates the $(p-1)$th order of linear Moving Average of i.i.d. $S\alpha S$ signal $u[n]$ with time-invariant coefficients γ_i as follows*

$$s[n] = \sum_{i=1}^{p}\gamma_i u[n-(i-1)], \tag{47}$$

where

$$\begin{aligned}
&u[n] \overset{i.i.d.}{\sim} S\alpha S(\alpha,\rho),\\
&\gamma_i = \left(\int_0^T \left|\underline{b}^T\left(\sum_{j=0}^{i-1} q[j]\mathrm{e}^{\mathbf{A}(m+(i-1-j)T)}\right)\underline{c}\right|^{\alpha}\mathrm{d}m\right)^{1/\alpha}.
\end{aligned} \tag{48}$$

The proof of Theorem 1 is provided in Appendix A. The main significance of this theorem is that it explicitly identifies the dependency structure of $s[n]$ and provides closed form expression for the linear coefficients γ_i, $i = 1, \ldots, p$.

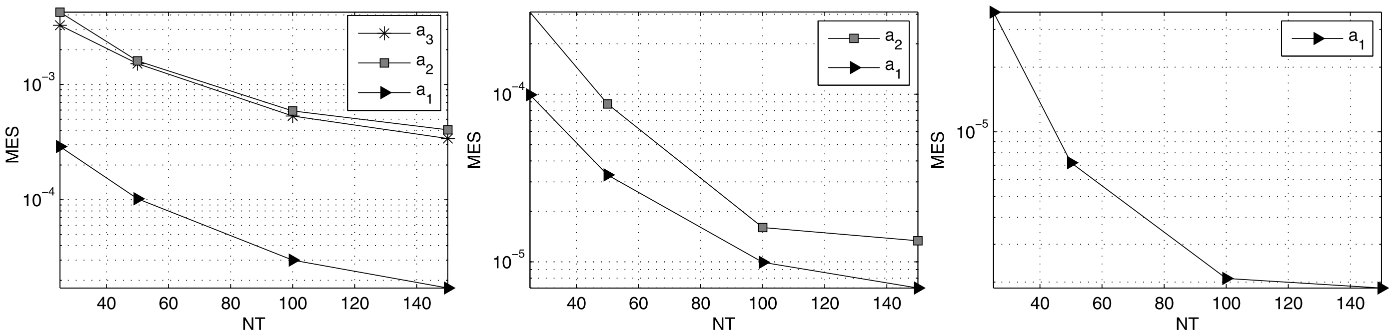
4. Estimator analysis

We discuss the convergence of exponential matrix estimation related to the autoregressive coefficients of α-stable CAR(p) process from discretized equation (36). Also, the statistical properties of the estimator are shown through numerically computing the estimator PDF for each coefficient a_i and evaluating the effects of the sampling frequency and sample size on the limiting distribution of the estimator.

4.1. Convergence

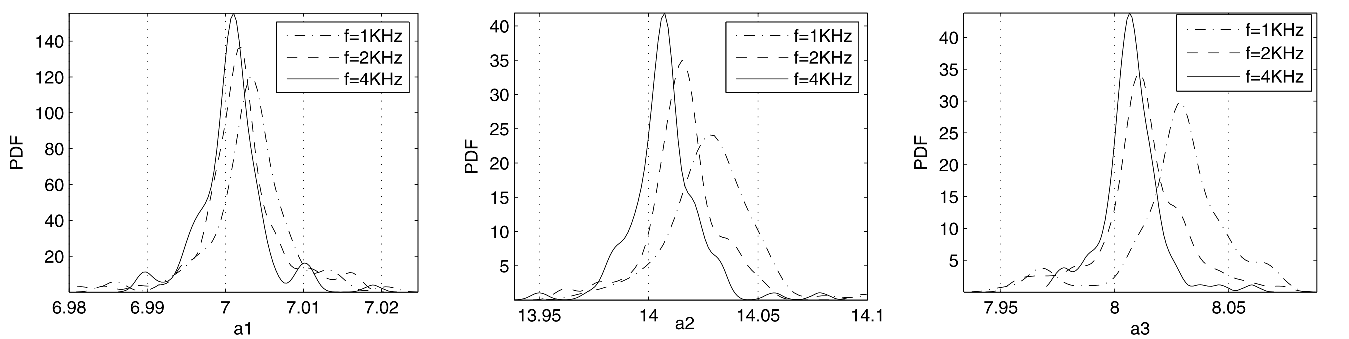
Here, same as the literature for continuous-time model parameters estimation [23,31], asymptotic behavior of the estimator is considered when $T \to 0$ (i.e. $f \to \infty$) and $NT \to \infty$. In the proposed estimation procedure, the state-space representation of α-stable CAR(p) process produces a multivariate AR(1) model as shown in (13), where, $\underline{Z}[n]$ is an independent heavy-tailed signal for $n = 1, \cdots$. Using this representation, the model parameters estimation leads to the convex objective function derived in (36) that can be considered as a Least L_ν multivariate linear autoregression with infinite variance error. For heavy-tailed process, the M-estimation procedure is more appropriate compared to LS estimator, since less weights are assigned to outliers [28]. In [32,33], the authors established that M-estimation of multivariate linear regression under α-stable assumption is consistent, when $\alpha > 1$ and $\nu < \alpha$. Consequently, we deduce the consistent estimator for the exponential matrix in (36) for the case of $\alpha > 1$. Also, in [28], the strong consistency of the LAD estimator for $\alpha > 1$ has been demonstrated under assumption that Z_1 has a unique median at zero. On the other hand, the authors proved that for model parameters estimates of the alpha-stable AR(p) process when $\alpha < 1$, the least L_ν estimator is consistent if $E|Z[n]|^\nu < \infty$ and the function of $m(x) = E|Z_n - x|^\nu$ has a unique minimum. But, it is difficult to show the unique minimum for $m(x)$. However, the simulation results in [34,35] depicted that in the M-estimation approach with α-stable noise the optimal choice for ν when $\alpha < 1$, is $\nu = \alpha$. In summary, in the estimation algorithm, we have consistency for the estimator of exponential matrix in (36) when $\alpha > 1$.

Further, in our algorithm, we then use the known structure of $\mathbf{A}$ in (37) to decrease the computational complexity in the optimization algorithm and increase the rate of the convergence. This causes to estimate the unknown vector of $\underline{a}$ including p coefficients, instead of the exponential matrix with $p \times p$ entries. Hence, a nonlinear objective function with much less unknown coefficients is appeared in (37). In this regards, computing the initial value for $\underline{a}$ via the covariation method in (34) accelerates the convergence rate of the optimization approach. To evaluate the asymptotic behavior of our estimator for $\underline{a}$, we utilize the mean square error (MSE) sense. To this end, the MSE of every coefficient

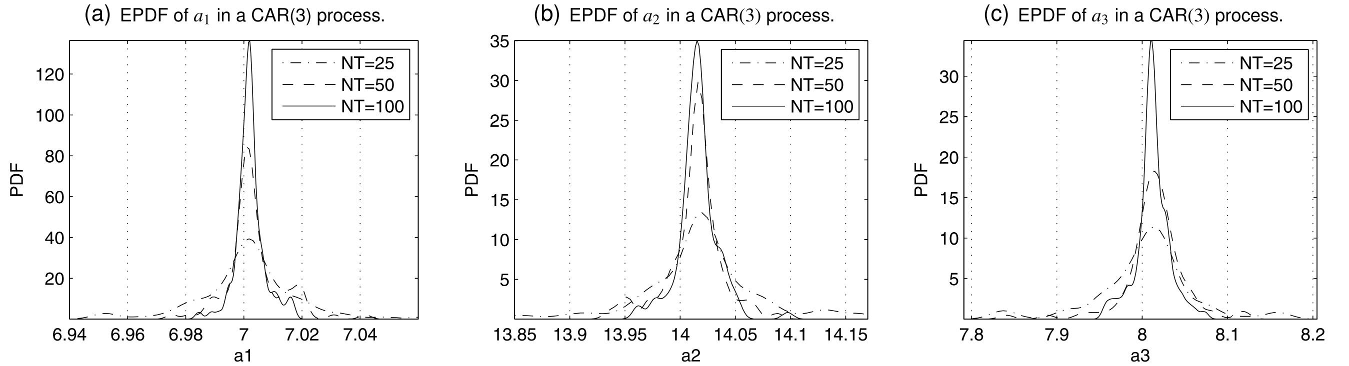


(a) CAR(3). (b) CAR(2). (c) CAR(1).

Fig. 2. MSE of the estimate of every coefficient a_i as a function of NT when frequency sampling is $f = 2$ KHz, for three cases of CAR(p) process: $p = 3$, $p = 2$ and $p = 1$.



(a) EPDF of a_1 in a CAR(3) process. (b) EPDF of a_2 in a CAR(3) process. (c) EPDF of a_3 in a CAR(3) process.



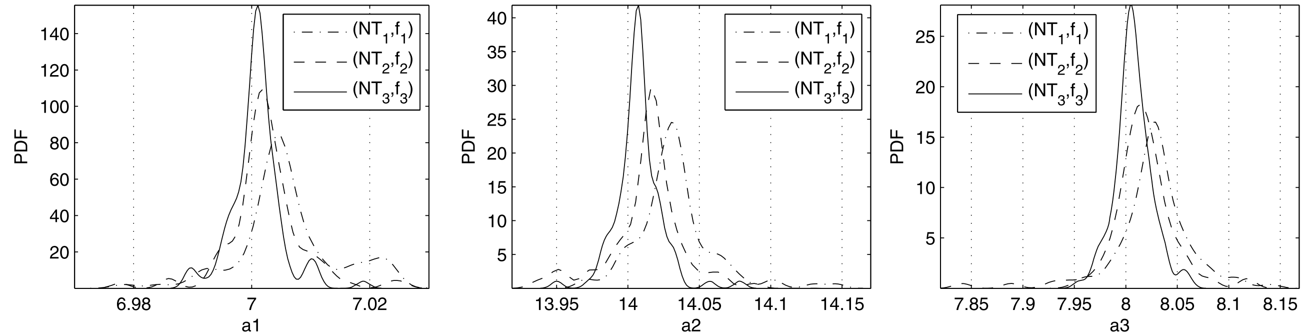
(d) EPDF of a_1 in a CAR(3) process. (e) EPDF of a_2 in a CAR(3) process. (f) EPDF of a_3 in a CAR(3) process.

Fig. 3. Effects of f and NT on the autoregressive coefficients estimation of the α-stable CAR(3) process. (a)–(c) Shift the PDF of the estimator for every a_i to the true value by increasing the sampling frequency in a fixed NT. (d)–(f) Decreasing the dispersion of the estimator for every a_i around the estimated parameter by increasing NT in a fixed sampling frequency.

a_i in three cases of CAR(p) process is numerically computed, when $p = 3$, $p = 2$ and $p = 1$. As shown in Fig. 2, when NT increases, the MSE of the estimates decreases and the estimates converge to their true values, in a sampling frequency of $f = 2$ KHz. This reveals that the estimator asymptotically tends to accurate value in MSE sense. Furthermore, since the theoretical computation of limiting distribution for the estimate of every a_i is complex, we numerically provide a brief discussion for the characteristics of the estimator PDF in the next section.

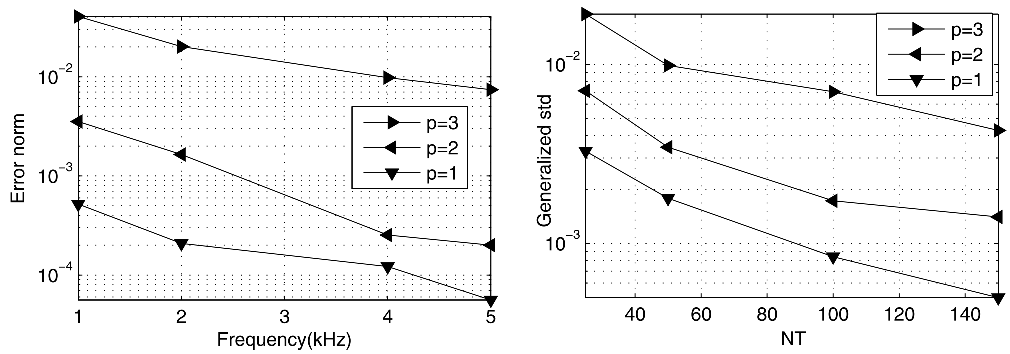
4.2. Statistical properties

Due to the complexity of the estimator function employed in (37) for estimating the coefficients of the differential equation $\{a_i\}_{i=1,\dots p}$, deriving the analytical expression for the asymptotic distribution function of the estimator is not feasible. Hence, we investigate the statistical behavior of the estimator via Monte Carlo simulation by computing the empirical probability density function (EPDF) for every estimated parameter $\hat{a}_i$ which is provided by multiple running of the learning algorithm. We numerically present the significant role of increasing two effective factors, f and NT, on the statistical properties of the estimator. First, for a fixed $NT = 50$, we increase the sampling frequency of a CAR(3) signal driven by SαS process and compute the EPDF of the estimator $\hat{a}_i$ through 100 independent runs, as shown in Figs. 3a–3c. By increasing the sampling frequency f, the mode of EPDF shifts to true value. After that, in a fixed sampling frequency of $f = 2$ KHz, by increasing NT, the reduction in the dispersion of the EPDF of the estimator $\hat{a}_i$ is clearly depicted in Figs. 3d–3f. Finally, we increase the sampling frequency and NT altogether and illustrate their effects on the consistency and dispersion of the estimator in Fig. 4. In Fig. 5a the error norm of the estimated vector $\underline{\hat{a}}$, versus frequency for $p = 1, 2, 3$ in a fixed and large $NT = 50$ is depicted. It shows that the accuracy of the estimated vector increases when the sampling frequency enlarges. On the other hand, since the EPDF of the estimator has heavier tail than normal distribution, we compute the average of the dispersion parameter of $\underline{\hat{a}}$ versus NT, as generalized standard deviation, in a fixed sampling frequency $f = 2$ KHz as shown in Fig. 5b. It reveals the effects of increasing NT on decreasing the dispersion of the estimates. As a result, by



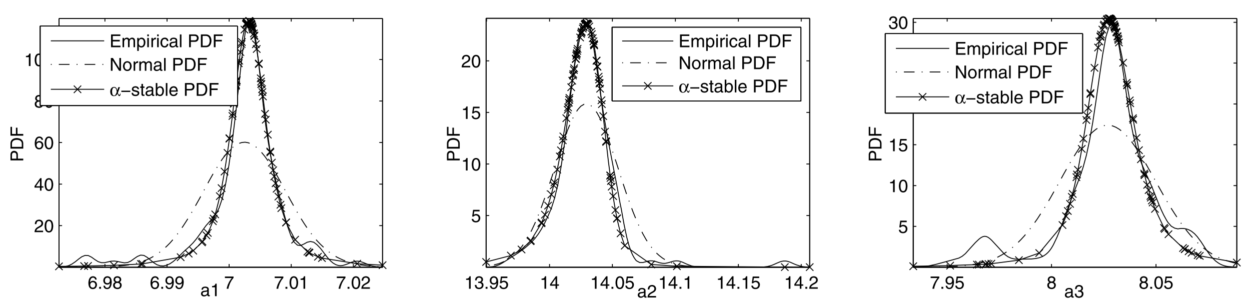
(a) EPDF of a_1 in a CAR(3) process. (b) EPDF of a_2 in a CAR(3) process. (c) EPDF of a_3 in a CAR(3) process.

Fig. 4. The influence of increasing NT and f together on the estimator for every autoregressive coefficient a_i of an α-stable CAR(3) process, where $(NT_3, f_3) > (NT_2, f_2) > (NT_1, f_1)$.

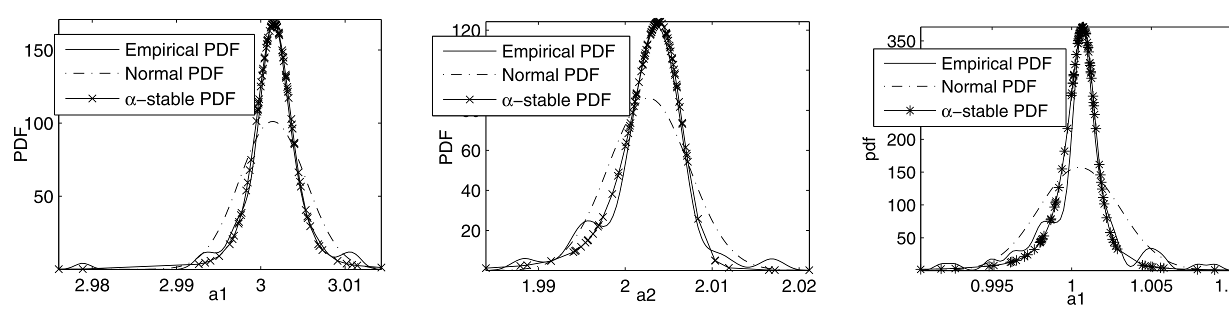


(a) α-stable CAR(p) with $\alpha = 1.1$ (b) α-stable CAR(p) with $\alpha = 1.1$

Fig. 5. (a) The error norm related to the estimation of the vector of the autoregressive coefficients, $\underline{a}$, for a CAR(p) process versus f. (b) The generalized standard deviation of the estimator $\underline{a}$ for a CAR(p) process versus NT.



(a) PDF of a_1 in a CAR(3) process. (b) PDF of a_2 in a CAR(3) process. (c) PDF of a_3 in a CAR(3) process.



(d) PDF of a_1 in a CAR(2) process. (e) PDF of a_2 in a CAR(2) process. (f) PDF of a_1 in a CAR(1) process.

Fig. 6. Nonparametric PDF approximation corresponding to the estimator of the every autoregressive coefficient a_i in the α-stable CAR(p) model and also the best fitted α-stable PDF and the best fitted Gaussian PDF, when (a)–(c) $p = 3$, (d)–(e) $p = 2$ and (f) $p = 1$.

increasing both factors of NT and f, the estimates tend to accurate values.

We compare the limiting distribution of our estimates with normal and α-stable distributions for the assumption of frequency sampling $f = 1$ KHz and sample size $N = 10^5$ that is used in the next section for computing the results given in Table 3. The EPDF of every $\hat{a}_i$ for three cases of α-stable CAR(p) process, $p = 1$, $p = 2$ and $p = 3$, with $\alpha = 1.1$, through 100 independent runs is depicted in Fig. 6. The figure also shows the best fitted normal PDF and the best fitted α-stable PDF in addition to the computed

Table 2
Anderson–Darling statistic.

	$p=3$			$p=2$		$p=1$
	a_1	a_2	a_3	a_1	a_2	a_1
Gaussian fit	3.7631	4.0808	3.4598	4.3482	3.7937	1.2803
α-stable fit	2.6502	0.6472	0.7902	0.7336	0.6790	0.6320

EPDF. As shown, the EPDF of the estimated values for every $\hat{a}_i$, is very close to the fitted α-stable distribution which presents a heavy-tailed behavior that is in contrast to normal assumption. Testing the goodness of fit of a distribution is based on the difference between the empirical distribution function, $F(a)$, and the fitted distribution function, $\hat{F}(a)$, according to the most well-known goodness-of-fit statistics such as Kolmogorov–Smirnov and Anderson–Darling statistics [36]. In the Anderson–Darling statistic, discrepancy measure is given by the Cramer–von Mises family $Q = n\int_{-\infty}^{+\infty}\{\hat{F(a)} - F(a)\}^2\chi(a)\mathrm{d}F(a)$, when the weight function $\chi(a)$ is $\{F(a)\{1 - F(a)\}\}^{-1}$. The Anderson–Darling distance makes more weight on the samples in the tails of the distribution. For the critical values of hypothesis testing in the Anderson–Darling test, no asymptotic results are explicitly characterized for α-stable law in [37]. Hence, we do not use hypothesis testing and just compare the test values. Certainly, the lower values show the better fitted distribution. Therefore, not only visually but also based on the goodness-of-fit measures in Table 2, stable distributions provide a better fit compared to the Gaussian, especially for the smaller values of α, because a Gaussian process is a special case of α-stable processes and does not include the heavy-tailed behavior. Hence, for the results of Table 3, instead of the mean we generally derive the median of the estimated parameters and also in place of standard deviation, the scale parameter is used to illustrate their dispersion.

5. Simulation study

We have employed the Euler scheme and also Yuima package [38] to generate several realizations of Lévy-stable driven CAR(p) processes for use in MATLAB. To assess the performance of our proposed procedure, Table 3 shows the results of learning method corresponding to differential equation coefficients, $\underline{a}$, and in addition the parameters of excitation distribution (α, ρ) for different orders of CAR process, $p = 1, 2, 3$, and also for variant values of α with the sampling frequency of $f = 1$ KHz and the sample size of $N = 10^5$ in the 100 independent runs. To estimate the unknown parameters, we use the learning procedure given in Table 1. In the literature [34], the authors numerically demonstrated that the best choices for the M-estimator with Loss function of L_ν norm are $\nu = 1$ when $\alpha > 1$ and $\nu = \alpha$ when $\alpha <= 1$. Hence, we use this point in our learning algorithm for the estimates given in Table 3. Based on the estimator analyses in the previous section which show that the estimator presents a heavy-tailed behavior when $f = 1$ KHz and $N = 10^5$, the scale parameter of the estimator is used as the generalized standard deviation (Gstd) to present the dispersion of the estimates. In addition, the median of the estimator (μ_{est}) is employed instead of the mean corresponding the estimated parameters. The initialization of the optimization method (μ_0) for $\{\hat{a}_i\}_{i=1,\dots,p}$ using the covariation method, and the error of the estimated parameters (Error) are given in Table 3. As shown, the proposed learning procedure explicitly provides more accurate estimator for the lower values of α.

To the best of our knowledge, statistical inference for the α-stable CAR(p) process has not been explored previously apart from the first order univariate or multivariate CAR process as the stable OU process. Since in the past methods of the multivariate OU process or in the finite variance Lévy CAR process, the unknown matrix has been directly estimated, we compare our method with their approaches (i.e. Yule–Walker and initialized LS estimator) [23] through the normalized error of the estimated exponential matrix $\mathbf{e^A}$ versus the multiple values of α in Fig. 7. The LS estimator of a matrix without an appropriate initial value takes a long time to even converge, hence, we initialize the LS estimator by the covariation-based method to accelerate the convergence and enhance the accuracy of the results. Due to the low accuracy of the covariation-based method for $\alpha < 1$, the LS method for $\alpha = 0.8$ is initialized by the Yule–Walker estimates. However, if we do not properly initialize LS method in $\alpha = 0.8$, similar to our proposed learning procedure which is not initialized for $\alpha < 1$, the accuracy of the LS estimates by considering a limitation on the number of iterations due to very long runtime, decreases as depicted in Fig. 7. Our method provides more accuracy for the estimates and significantly decreases the required time for the convergence to the true values, because it uses a proper initialization for $\alpha > 1$ and also reduces the unknown parameters from $p \times p$ to p.

Besides, we investigated the sampled α-stable CAR(p) process in the discrete-time setting. Theorem 1 proves that its samples present a discrete ARMA($p, p-1$) model with absolutely i.i.d. excitation process and provides a closed form expression for the moving average coefficients. To represent the consequence of the theorem, we have generated a realization of a CAR(3) process $x(t)$, driven by an α-stable process $L(t)$. The sampled process $x[n]$ should be a discrete ARMA(3, 2) signal with the autoregressive

Table 3
The estimated parameters of the α-stable driven CAR(P) processes.

		$p=1$, $\underline{a}=[1]$			$p=2$, $\underline{a}=[3,2]$				$p=3$, $\underline{a}=[7,14,8]$				
		α	ρ	a_1	α	ρ	a_1	a_2	α	ρ	a_1	a_2	a_3
$\alpha=0.8$, $\rho=1$	μ_{est}	0.7991	1.0201	1.0005	0.8021	1.0341	3.0015	2.0030	0.8025	1.015	7.0035	14.0290	8.0281
	Error	9.00e−4	2.01e−2	1.50e−3	2.10e−3	3.41e−2	1.52e−3	3.01e−3	2.52e−3	1.50e−2	3.53e−3	2.90e−2	2.81e−2
	Gstd	1.21e−3	4.40e−3	1.15e−5	7.12e−3	1.80e−2	2.49e−5	4.37e−5	3.60e−3	6.62e−2	5.85e−5	1.61e−4	2.73e−4
$\alpha=1.1$, $\rho=1$	μ_0	–	–	1.1078	–	–	2.8271	2.1350	–	–	6.6668	14.3915	9.0059
	μ_{est}	1.1003	1.0120	1.0005	1.0999	1.030	3.0018	2.0035	1.1049	0.9420	7.0040	14.0297	8.0279
	Error	3.41e−4	1.20e−2	5.00e−4	1.46e−4	3.01e−2	1.83e−3	3.57e−3	4.92e−3	5.80e−2	4.02e−3	2.97e−2	2.79e−2
	Gstd	4.61e−3	6.40e−3	9.67e−4	5.31e−3	3.59e−2	1.62e−3	2.82e−3	5.11e−3	7.41e−3	2.8e−3	1.30e−2	1.67e−2
$\alpha=1.5$, $\rho=1$	μ_0	–	–	1.0672	–	–	2.9322	2.1222	–	–	7.1021	14.1151	8.2003
	μ_{est}	1.4998	1.0071	1.0008	1.5002	1.0502	3.0056	1.9970	1.4951	0.9475	7.0123	14.0406	8.0368
	Error	2.00e−4	7.11e−3	8.22e−3	1.79e−4	5.02e−2	5.61e−3	3.05e−3	4.93e−3	5.25e−2	1.23e−2	4.06e−2	3.68e−2
	Gstd	5.3e−3	3.62e−3	1.94e−2	5.52e−3	1.4e−2	3.20e−2	3.76e−2	6.30e−3	6.11e−2	5.36e−2	1.65e−1	1.60e−1
$\alpha=1.9$, $\rho=1$	μ_0	–	–	1.05	–	–	3.1204	2.1104	–	–	7.0901	14.3164	8.3423
	μ_{est}	1.8976	1.0108	0.9806	1.9013	0.9840	3.0324	2.0333	1.8983	1.013	7.0705	14.1943	7.9562
	Error	2.44e−3	1.08e−2	1.94e−2	1.30e−3	1.60e−2	3.24e−2	3.33e−2	1.74e−3	1.30e−2	7.05e−2	1.9e−1	4.38e−2
	Gstd	4.91e−3	1.30e−2	9.73e−2	4.72e−3	2.70e−2	1.60e−1	2.06e−1	3.61e−3	3.53e−2	4.60e−1	8.02e−1	9.12e−1

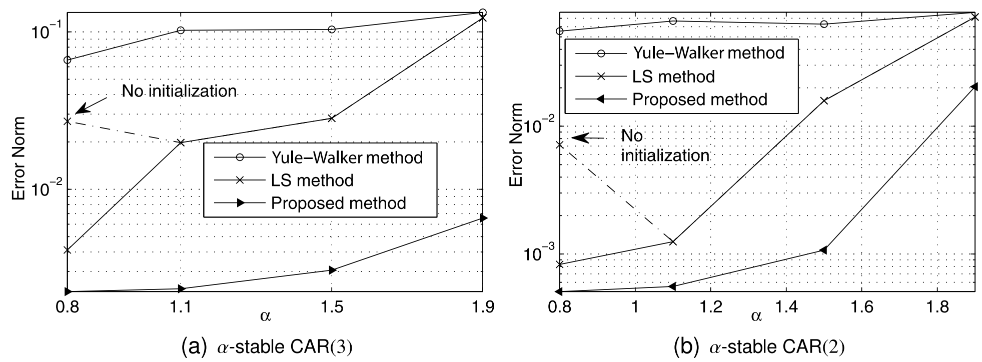


Fig. 7. A comparison between the proposed method, Yule–Walker and LS method versus the different values of α for α-stable CAR(p) process when (a) $p = 3$ and (b) $p = 2$.

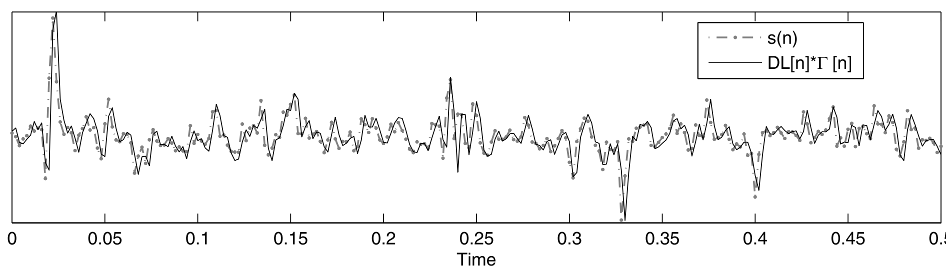


Fig. 8. The consequence of Theorem 1. $s[n]$ follows $\mathrm{D}L[n] * \Upsilon[n]$, where $\mathrm{D}L[n]$ represents the increments of α-stable process and $\Upsilon[n]$ is an FIR filter with coefficients derived from (48).

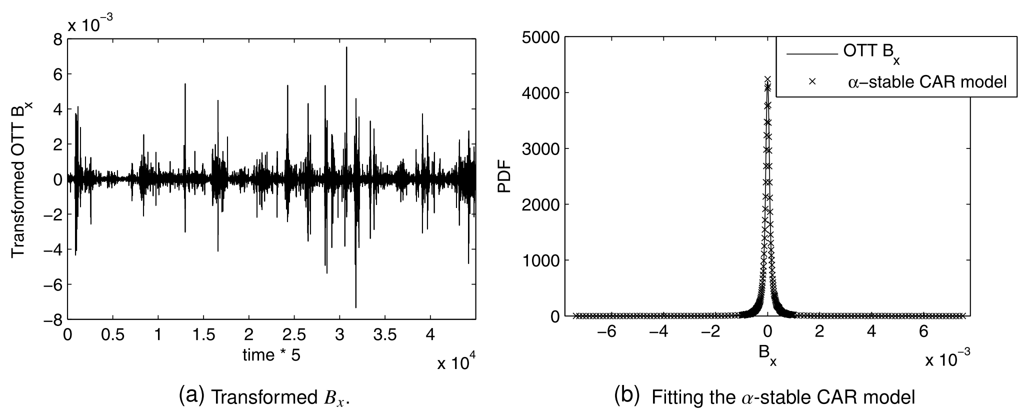


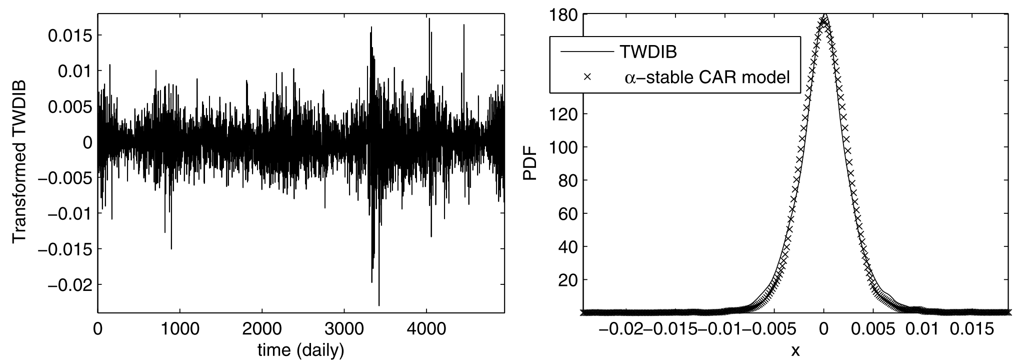
Fig. 9. (a) Transformed horizontal component B_x of the magnetic field with 5-minute resolution for the year 1987 at OTT station. (b) The α-stable CAR fitting for the PDF of the transformed component B_x at OTT station.

and moving average coefficients expressed in (45) and (48) respectively. Moreover, the filtered process $s[n] \triangleq x[n] * \theta[n]$ acquired by the FIR filter $\theta[n]$ from (42), should be similar to $\mathrm{D}L * \Upsilon[n]$ when $\Upsilon[n]$ is an FIR filter with coefficients computed in (48) and $\mathrm{D}L[n]$ is generated through stationary and independent increments of the excitation process $L(t)$. It is shown in Fig. 8 that $s[n]$ follows $\mathrm{D}L * \Upsilon[n]$ which verifies the validity of the derived expressions in (47) and (48).

6. Applications to real-world data

In this section, we demonstrate the utility of the proposed method by applying to some data sets whose distributions have heavier tails than the exponential distribution. We fit the α-stable CAR model to the twelve real time series which are drawn from two types of real-world data, namely, ground magnetometer data and financial time series. Since our data sets are generated from non-stationary processes, we transform them to stationary forms by computing $x_i = \ln(r_{i+1}/r_i)$, where x_i is the stationary transformation of the real-time data r_i. Through this conversion, no information is lost and the original time series can be achievable from the transformed data [11]. Figs. 9a and 10a show the transformed time series for the two kinds of real-world data sets.

To evaluate the proposed method on the real data, we fit an α-stable distribution to the EPDF of every time series, x, and estimate its distribution parameters $(\alpha_x, \rho_x, \mu_x, \beta_x)$. Since it is assumed that our signal model has a symmetric distribution, we only utilize the data sets with the small values for β and μ as shown in Tables 4 and 5. According to the described method for the estimation of the distribution parameters of the excitation process (α, ρ), we have $\alpha = \alpha_x$ and $\rho = \rho_x / k^{(1/\alpha)}$ from (23). Hence, it is possible to estimate the differential equation coefficients, $\underline{a}$, and the scale parameter ρ of the excitation process using the proposed procedure in Table 1 and the assumed order model. To show that the data sets are indeed consistent with the estimated α-stable CAR models, we compare the EPDF of the real time series with the EPDF of the generated data from the learned CAR model. In this regard, first, the EPDF of the real time series, $f(x)$, is computed. After that, we generate a sample path of the model by the estimated model parameters to calculate its EPDF, $\widehat{f}(x)$. The results of



(a) Transformed TWDIB (b) Fitting the α-stable CAR model

Fig. 10. (a) Transformed daily financial time series of TWDIB in the period of 1995-01-4 to 2015-06-26. (b) The α-stable CAR fitting for the PDF of TWDIB.

Table 4
Estimated model parameters for the ground magnetometer data.

Station	α	ρ_x	μ_x	β_x	ρ	$\underline{a}$	Error	MSE	RSE
ME	0.9443	2.08e−04	1.57e−07	0.0135	2.68e−4	[1.3445]	0.0108	3.92e−07	0.0208
RPC	1.2550	7.54e−05	6.60e−07	0.0213	9.95e−05	[1.1287]	0.0237	1.88e−06	0.0793
RIT	1.1855	0.0027	5.22e−06	0.0268	0.0046	[1.5974]	0.0361	4.36e−06	0.1128
OTT	1.2486	8.05e−05	−7.96e−07	0.0176	5.75e−04	[4.8081, 6.4680]	0.0244	2.11e−06	0.0773
CB	1.2728	0.0020	5.40e−05	−0.0011	0.0143	[4.9218, 6.3626]	0.0319	3.71e−06	0.1140
NEW	1.1326	7.78e−05	2.59e−07	0.0217	5.19e−04	[4.1125, 6.5369]	0.0207	1.81e−06	0.0601

Table 5
Estimated model parameters for the financial data.

Financial time series	α	ρ_x	μ_x	β_x	ρ	$\underline{a}$	Error	MSE	RSE
TWDIB	1.8216	0.0017	−7.96e−06	0.0406	0.0054	[4.5589]	0.0122	5.03e−07	0.1080
CCFSI/CM	1.5769	0.0051	4.47e−04	−0.0256	0.0065	[0.9279]	0.0170	9.75e−07	0.1234
CCFSI/FEM	0.8293	0.0121	3.66e−04	−0.0525	0.0071	[0.7648]	0.0275	2.57e−06	0.0934
CUSFER	1.5943	0.0019	−1.45e−06	−0.0034	0.0116	[2.8250, 6.7719]	0.0236	2.81e−06	0.1167
7-YTCMR	1.4973	0.0067	−1.83e−04	−0.0281	0.0408	[2.9283, 6.6527]	0.0226	1.72e−06	0.1328
GFPLBM	1.5596	0.0057	1.01e−04	0.0545	0.0322	[2.4274, 6.9757]	0.0283	4.05e−06	0.1358

three error functions with the following statements, are reported in Tables 4 and 5 which are based on average of 100 independent runs.

$$Error = \sqrt{\sum_{i=1}^{N_p}\left(f(x_i)-\widehat{f}(x_i)\right)^2};\ MSE = \frac{1}{N_p}\sum_{i=1}^{N_p}\left(f(x_i)-\widehat{f}(x_i)\right)^2;$$
$$RSE = \frac{\sqrt{\sum_{i=1}^{N_p}\left(f(x_i)-\widehat{f}(x_i)\right)^2}}{\sqrt{\sum_{i=1}^{N_p}\left(f(x_i)-f_{av}\right)^2}}. \tag{49}$$

N_p is the number of EPDF samples and f_{av} is the average value for the EPDF of the real data. The goodness of fit that is determined by RSE varies between 0 and 1, where the small value for RSE indicates that the employed model performs well in simulating the real data [39]. In the following, we briefly explain about the real-world data sets and the results that are achieved by employing the α-stable CAR model.

6.1. Ground magnetometer data

Lévy processes are an excellent tool for modeling the heavy-tailed behavior in the interplanetary medium and the magnetosphere [12,39]. Here, we model a component of the ground magnetometer measurements by the α-stable CAR process. The magnetometer arrays are placed in the different locations of the world and measure the Earth's magnetic fluctuations. The parameters of the measurements encompass some components such as the north component of the horizontal intensity B_x [39]. We use ground magnetometer measurements of the horizontal component B_x of the magnetic field at 6 stations, namely, Meanook (ME), Rapid-City, SD (RPC), Rankin-Inlet (RIT), Ottawa (OTT), Cambridge-Bay (CB), Newport (NEW) [40]. We examine the ground station data averaged over 5-minute intervals covering the first 6 months of the year 1978.

The α-stable CAR processes are employed to model these ground magnetometer data. The estimated model parameters are reported in Table 4. As shown, the signals of the first half stations are modeled by the first order CAR processes and the signals of the second half stations are represented by the second order CAR models. The proposed learning approach is applied to these real data sets. The results indicate that our estimation procedure works on the real data and the α-stable CAR provides a good approach to model the horizontal component B_x of the magnetic field. As an example, the transformed time series from the OTT station is depicted in Fig. 9a. Fig. 9b shows that the EPDF of this time series is consistent with the EPDF of the simulated data from the estimated model.

6.2. Financial data

Stable distributions, a rich class of probability distributions, are an appropriate model for representing many types of economic data. Due to heavy-tailed behavior of financial data, α-stable processes accurately model multiple kinds of financial time series [11]. We fit the α-stable CAR model to some financial data sets and report their results in Table 5. As shown, the first half data sets are modeled by the first order CAR processes and the remained data sets with the second order. Our learning approach is examined on these financial data sets. Table 5 depicts that our algorithm works on the real data. Furthermore, the results of the error functions illustrate a reasonably good agreement between the model and the experimental data. The transformed financial time series of the TWDIB is shown in Fig. 10a. In Fig. 10b, the EPDF of this data and the EPDF of the generated data from the learned model reveal the compatibility of the model with the data set.

The daily financial data that are used in Table 5 are respectively, Trade Weighted U.S. Dollar Index: Broad (TWDIB) in the period of 1995-01-4 to 2015-06-26, Contributions to the Cleveland Financial Stress Index: Credit Markets (CCFSI/CM) in the period of 1991-09-25 to 2015-06-30, Contributions to the Cleveland Financial Stress Index: Foreign Exchange Markets (CCFSI/FEM) in the period of 1991-09-25 to 2015-06-25, Canada/U.S. Foreign Exchange Rate (CUSFER) in the period of 1971-01-04 to 2015-06-26, 7-Year Treasury Constant Maturity Rate (7-YTCMR) in the period of 1969-07-01 to 2015-06-30, Gold Fixing Price 10:30 A.M. (London time) in London Bullion Market based in British Pounds (GFPLBM) in the period of 1968-04-01 to 2015-06-29 [41].

7. Conclusion

We studied the model parameters estimation of the heavy-tailed stochastic processes described by p-order CAR models excited by α-stable processes under infinite-variance assumption. The proposed learning procedure is divided into two parts: estimating the parameters of the excitation distribution and finding the CAR coefficients. The former is accomplished by estimating the characteristic exponent of the SαS process. In contrast to solving conventional Yule–Walker equations which rely on finite second-order moments of the distributions, we made use of the covariation equations and M-estimation optimization methods to estimate the CAR coefficients. To perform the mentioned tasks, given the samples of a stable CAR(p) process, we used the equivalent state-space representation. Furthermore, since the sampled CAR(p) processes yield ARMA($p, p-1$) processes in discrete-time setting, we derived closed-form expressions that describe the model parameters of ARMA($p, p-1$) process. Our Monte Carlo simulations indicate that the method works well when the sampling frequency (f) and the whole interval of signal (NT) are large, especially for the signals with the smaller values of α. In addition, our proposed method was applied on the ground magnetometer data and the financial time series to evaluate the performance of our learning algorithm in the realistic world. The experimental results show a good agreement between the model and these real-world data.

Appendix A

Proof of Theorem 1. We show that $s[n]$ follows a moving average of i.i.d. α-stable process $u[n]$, under α-stable assumption for the Lévy excitation process. Using (46), we have

$$s[n-l] = \int_{(n-l-1)T}^{(n-l)T} q_0 \underline{b}^T e^{\mathbf{A}((n-l)T-u)} \underline{c} dL(u)$$
$$+ \int_{(n-l-2)T}^{(n-l-1)T} \underline{b}^T \left(q_0 e^{\mathbf{A}((n-l)T-u)} + q_1 e^{\mathbf{A}((n-l-1)T-u)} \right) \underline{c} dL(u) + \cdots$$
$$+ \int_{(n-l-p)T}^{(n-l-p+1)T} \underline{b}^T \left(q_0 e^{\mathbf{A}((n-l)T-u)} + \cdots + q_{p-1} e^{\mathbf{A}((n-l-p+1)T-u)} \right)$$
$$\times \underline{c} dL(u) \tag{A.1}$$

Every term is generated from the Lévy process $L(t)$ in a distinct time duration, thus $s[n-l]$ is constituted from a linear combination of p independent α-stable processes with the identical parameter α and different scales which are inferred by CF corresponding to $s[n-l]$, as following

$$\Phi_{s[n-l]}\omega = \exp\{j\omega s[n-l]\}$$
$$= \exp\left\{ -\left(\int_0^T \left| q_0 \underline{b}^T e^{\mathbf{A}m} \underline{c} \right|^\alpha dm \right) \rho^\alpha |\omega|^\alpha \right\}$$
$$. \exp\left\{ -\left(\int_T^{2T} \left| \underline{b}^T \left(q_0 e^{\mathbf{A}m} + q_1 e^{\mathbf{A}(m-T)} \right) \underline{c} \right|^\alpha dm \right) \rho^\alpha |\omega|^\alpha \right\} \cdots$$
$$. \exp\left\{ -\left(\int_{(p-1)T}^{pT} \left| \underline{b}^T \left(q_0 e^{\mathbf{A}m} + \cdots + q_{p-1} e^{\mathbf{A}(m-(p-1)T)} \right) \underline{c} \right|^\alpha dm \right) \right.$$
$$\left. . \rho^\alpha |\omega|^\alpha \right\}$$
$$= \exp\left\{ -\gamma_1^\alpha \rho^\alpha |\omega|^\alpha \right\} \exp\left\{ -\gamma_2^\alpha \rho^\alpha |\omega|^\alpha \right\} \cdots \exp\left\{ -\gamma_p^\alpha \rho^\alpha |\omega|^\alpha \right\}. \tag{A.2}$$

Therefore $s[n-l]$ is a linear combination of p independent but not identical α-stable random variables. From (A.2), these independent α-stable random variables have the same characteristic exponent α but different scale parameters γ_i. On the other hand, for example, a random variable r with SαS(α, ρ) distribution has the identical distribution with the random variable $\rho.v$ when ρ is a constant parameter and v is a random variable with SαS$(\alpha, 1)$ distribution. In other words, we can similarly deduce that $s[n-l]$ is a linear combination of p independent and identical α-stable random variables $u[n] \sim S\alpha S(\alpha, \rho)$ with the coefficients extracted from γ_i, where

$$\gamma_k^\alpha = \int_{(k-1)T}^{kT} \left| \underline{b}^T \left(\sum_{i=0}^{k-1} q_i e^{\mathbf{A}(m-iT)} \right) \underline{c} \right|^\alpha dm, \; k = 1, \ldots, p. \tag{A.3}$$

So (48) can be exactly derived that represents the time invariant coefficients. Hence it yields (47), a linear moving average of i.i.d. sequence $u[n]$ with symmetric α-stable distribution and equal statistical characteristics corresponding to the distribution of excitation increments. □

References

[1] G.C. Goodwin, D.Q. Mayne, Continuous-time stochastic model reference adaptive control, IEEE Trans. Autom. Control 36 (11) (1991) 1254–1263.
[2] H. Kirshner, M. Unser, J. Ward, On the unique identification of continuous-time autoregressive models from sampled data, IEEE Trans. Signal Process. 62 (6) (2014) 1361–1376.
[3] P. Brockwell, Recent results in the theory and applications of CARMA processes, Ann. Inst. Stat. Math. 66 (4) (2014) 647–685.
[4] E. Schlemm, R. Stelzer, et al., Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes, Bernoulli 18 (1) (2012) 46–63.
[5] M. Unser, P.D. Tafti, An Introduction to Sparse Stochastic Processes, Cambridge University Press, 2014.
[6] I.-T. Lim, B.G. Lee, Lossless pole-zero modeling of speech signals, IEEE Trans. Speech Audio Process. 1 (3) (1993) 269–276.
[7] S.S. Haykin, Adaptive Filter Theory, Pearson Education India, 2008.
[8] L. Ljung, System Identification, Springer, 1998.
[9] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events: For Insurance and Finance, vol. 33, Springer Science & Business, Media, 2013.
[10] A. Amini, U.S. Kamilov, E. Bostan, M. Unser, Bayesian estimation for continuous-time sparse stochastic processes, IEEE Trans. Signal Process. 61 (4) (2013) 907–920.

[11] J.P. Nolan, Modeling financial data with stable distributions, in: Handbook of Heavy Tailed Distributions in Finance, vol. 1, 2003, pp. 105–130.
[12] L. Burlaga, C. Wang, N. Ness, A model and observations of the multifractal spectrum of the heliospheric magnetic field strength fluctuations near 40 au, Geophys. Res. Lett. 30 (10) (2003) 1543.
[13] G. Jubelin, A. Khenchaf, A statistical model of sea clutter in panchromatic high resolution images, in: International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2012, pp. 424–427.
[14] L.J. Hadjileontiadis, S.M. Panas, On modeling impulsive bioacoustic signals with symmetric α-stable distributions: application in discontinuous adventitious lung sounds and explosive bowel sounds, in: Proceedings of the 20th Annual International Conference of Engineering in Medicine and Biology Society, IEEE, 1998, pp. 13–16.
[15] S.I. Resnick, et al., Heavy tail modeling and teletraffic data, Ann. Stat. 25 (5) (1997) 1805–1869.
[16] A. Amini, P. Thévenaz, J.P. Ward, M. Unser, On the linearity of Bayesian interpolators for non-Gaussian continuous-time AR(1) processes, IEEE Trans. Inf. Theory 59 (8) (2013) 5063–5074.
[17] P.J. Brockwell, E. Schlemm, Parametric estimation of the driving Lévy process of multivariate CARMA processes from discrete observations, J. Multivar. Anal. 115 (2013) 217–251.
[18] C.M. Gallagher, A method for fitting stable autoregressive models using the autocovariation function, Stat. Probab. Lett. 53 (4) (2001) 381–390.
[19] X. Feng, S. Wang, J. Yang, X. Zhu, Radar clutter suppression based on SαS fractional autoregressive model, Chin. J. Electron. 21 (4) (2012) 745–750.
[20] M. Calder, R. Davis, Inference for linear processes with stable noise, in: A Practical Guide to Heavy Tails, Birkhauser Boston Inc., 1998, pp. 159–176.
[21] E.E. Kuruoglu, P.J. Rayner, W.J. Fitzgerald, Least L_p-norm estimation of autoregressive model coefficients of symmetric α-stable processes, IEEE Signal Process. Lett. 4 (7) (1997) 201–203.
[22] Y. Hu, H. Long, Least squares estimator for Ornstein–Uhlenbeck processes driven by α-stable motions, Stoch. Process. Appl. 119 (8) (2009) 2465–2480.
[23] V. Fasen, Statistical estimation of multivariate Ornstein–Uhlenbeck processes and applications to co-integration, J. Econom. 172 (2) (2013) 325–337.
[24] H.-Z. An, Z.-g. Chen, On convergence of LAD estimates in autoregression with infinite variance, J. Multivar. Anal. 12 (3) (1982) 335–345.
[25] K. Bichteler, Stochastic Integration with Jumps, vol. 89, Cambridge University Press, 2002.
[26] X. Wang, K. Li, P. Gao, S. Meng, Research on parameter estimation methods for α-stable noise in a laser gyroscope's random error, Sensors 15 (8) (2015) 18550–18564.
[27] I.A. Koutrouvelis, Regression-type estimation of the parameters of stable laws, J. Am. Stat. Assoc. 75 (372) (1980) 918–928.
[28] R.A. Davis, K. Knight, J. Liu, M-estimation for autoregressions with infinite variance, Stoch. Process. Appl. 40 (1) (1992) 145–180.
[29] P.J. Brockwell, J.-P. Kreiss, T. Niebuhr, Bootstrapping continuous-time autoregressive processes, Ann. Inst. Stat. Math. 66 (1) (2014) 75–92.
[30] I.R. Khan, R. Ohba, Closed-form expressions for the finite difference approximations of first and higher derivatives based on Taylor series, J. Comput. Appl. Math. 107 (2) (1999) 179–193.
[31] Y. Hu, H. Long, Parameter estimation for Ornstein–Uhlenbeck processes driven by α-stable Lévy motions, Commun. Stoch. Anal. 1 (2) (2007) 175–192.
[32] Z. Bai, C.R. Rao, Y. Wu, M-estimation of multivariate linear regression parameters under a convex discrepancy function, Stat. Sin. 2 (1) (1992) 237–254.
[33] R.A. Davis, W. Wu, M-estimation for linear regression with infinite variance, Probab. Math. Stat. Wroclaw Univ. 17 (1997) 1–20.
[34] J. Liu, Infinite variance and non-linear time series models, Ph.D. thesis, Colorado State University, Department of Statistics, 1987.
[35] K. Knight, Rate of convergence of centred estimates of autoregressive parameters for infinite variance autoregressions, J. Time Ser. Anal. 8 (1) (1987) 51–60.
[36] T.B. Arnold, J.W. Emerson, Nonparametric goodness-of-fit tests for discrete null distributions, R J. 3 (2) (2011) 34–39.
[37] R. Weron, Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach, vol. 403, John Wiley & Sons, 2006.
[38] A. Brouste, M. Fukasawa, H. Hino, S.M. Iacus, K. Kamatani, Y. Koike, R. Nomura, Y. Shimizu, M. Uchida, N. Yoshida, The YUIMA project: a computational framework for simulation and inference of stochastic differential equations, J. Stat. Softw. 57 (4) (2014) 1–51.
[39] Z.-G. Yu, V. Anh, Y. Wang, D. Mao, J. Wanliss, Modeling and simulation of the horizontal component of the geomagnetic field by fractional stochastic differential equations in conjunction with empirical mode decomposition, J. Geophys. Res. Space Phys. (1978–2012) 115 (A10) (2010).
[40] Ground magnetometer data, available online at http://www-ssc.igpp.ucla.edu/SPDS/geomag/, accessed: 2015-05-20.
[41] FRED economic data, Available online at https://research.stlouisfed.org/fred2/, accessed: 2015-05-18.

Zeinab Hashemifard received the B.Sc. degree from Shariaty University, Tehran, Iran, and the M.Sc. degree from Amirkabir University of Technology (Polytechnique), Tehran, both in Electrical Engineering in 2008, and 2011, respectively. She is currently pursuing the Ph.D. degree in electrical engineering at the Amirkabir University of Technology, Tehran. Her research interests include statistical signal processing, sparse signal modeling, estimation and detection theory.

Hamidreza Amindavar received the B.S.E.E. degree in 1985, the M.S.E.E. degree in 1987, the M.Sc. degree in applied mathematics in 1991, and the Ph.D. degree in electrical engineering in 1991, all from the University of Washington, Seattle, WA, USA. He has been a faculty member in the Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran, since 1993. His research interests include remote sensing detection, statistical signal and image processing, human live detection, RADAR and SONAR signal processing, multiresolution signal analysis, and multiuser detection.

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. He was a researcher at École Polytechnique Fédérale (EPFL), Lausanne, Switzerland, from 2011 to 2013, working on statistical approaches towards modeling sparsity in continuous-domain. Since September 2013, he is an Assistant Professor at Sharif University of Technology, Tehran, Iran. He is also an associate editor of IEEE Signal Processing Letters, since July 2014. His research interests include various topics in statistical signal processing, specially compressed sensing.