Fractional-Order Correntropy Filters for Tracking Dynamic Systems in α-Stable Environments

Vinay Chakravarthi Gogineni, Sayed Pouria Talebi, Stefan Werner, *Senior Member, IEEE*, and Danilo P. Mandic, *Fellow, IEEE*

***Abstract*—In an increasing number of modern filtering applications, the encountered signals consist of frequent sharp spikes, that cannot be accurately modeled using Gaussian random processes. Modeling the behavior of such signals requires the more general framework of α-stable random processes. In order to present an inclusive filtering solution, this work derives a new class of fractional-order correntropy adaptive filters that are robust to the jittery α-stable signals. In contrast to conventional correntropy filters, the proposed objective function is compatible with the characteristic function of α-stable processes and captures fractional moments; therefore, the resulting algorithms do not depend on non-existing second-order moments. The work also includes performance and convergence analysis of the derived algorithms. Finally, simulations are conducted to illustrate the effectiveness of the proposed filtering techniques, which indicate that the proposed filters can outperform their counterparts and show less sensitivity to changes in the α parameter.**

***Index Terms*—α-stable signals, correntropy criterion, fractional-order filters, dynamic system tracking.**

I. INTRODUCTION

Traditionally, signal processing and machine learning techniques have been derived based on the assumption that the signal and noise are Gaussian [1], [2]. This assumption has played a crucial role in mathematical tractability and computationally efficiency of filtering solutions. However, the Gaussian model for signal/noise is being questioned in an increasing number of applications such as underwater acoustics [3], wideband communications [4], financial data modeling [5] and audio signal processing [6], [7] in which the encountered signals exhibit sharp spikes. The class of symmetric α-stable (SαS) random processes has proven to be a very flexible tool for modeling the behavior of such signals [3]–[9]. Considering that, with the exception of the Gaussian case, SαS random processes do not possess finite second-order moments, classical Wiener filtering techniques based on minimizing the second-order moment of an error measure suffer considerable performance degradation when applied to lower-order SαS signals [6], [9]. To tackle this problem, a class of filtering techniques has been proposed that is based on minimizing the fractional-norm and mixed p-norm of an error measure [9]–[13]. Though these algorithms offered improved performance over the traditional methods, they are computationally expensive and lack comprehensive performance and convergence analysis frameworks.

More recently, a class of adaptive filtering techniques for tracking dynamic systems, based on minimizing the fractional-order norm of an error measure using the framework of fractional-order calculus [14]–[16], has been proposed in [17]–[19]. Although these fractional-order adaptive filters achieve improved performance over the conventional Wiener filtering techniques, their performance is sensitive with respect to the value of characteristic exponent, α. Furthermore, residual jitters are still present in their steady-state estimates.

On the other hand, correntropy criterion based adaptive filtering techniques have been successfully applied for processing signals corrupted by impulsive noise. [20]–[28]. However, as the second-order, or, higher-order moment of the error is involved in the evaluation of maximum correntropy criterion (MCC) [22], or, generalized maximum correntropy criterion (GMCC) [24] similarity measure, the correntropy criterion based adaptive filters can not be extended to a situation where both signal and noise processes are modeled as SαS signals.

In order to provide a comprehensive adaptive filtering solution for the problem of tracking the state of a dynamic system, where the dynamic system itself is only observable through SαS input/output signals, is considered. Since both signal and noise processes are modeled as SαS signals, neither the conventional fractional-order adaptive filters nor the correntropy based adaptive filters alone will be enough here. For this, a class of adaptive filters based on maximizing a new *fractional-order correntropy* criterion in a gradient ascent manner is derived. The resulting fractional-order correntropy adaptive filters are computationally efficient to implement and are useful in a wide range of applications such as estimation and tracking in adverse conditions. Our main contributions here are as follows:

- By intrinsically combining the principles of correntropy-type local similarity measure and fractional-order calculus, we propose a class of fractional-order correntropy adaptive filters that effectively regulate the presence of strong jittery behavior of SαS signals.
- Stability of the proposed class of adaptive filters is analyzed and the conditions for convergence are derived.
- Detailed simulations are conducted to demonstrate the effectiveness of the proposed class of adaptive filters.

Mathematical Notations: We denote scalars, column vectors and matrices with lower case, bold lower case and bold uppercase letters, while $\mathbf{I}$ represents the identity matrix of

This work was supported, in part, by the ERCIM Alain Bensoussan Fellowship Programme, the Academy of Finland under Grant 296849, and the Research Council of Norway.

Vinay C. Gogineni and Stefan Werner are with the Norwegian University of Science and Technology, Norway (e-mail: vc.gogineni@ntnu.no and stefan.werner@ntnu.no).

Sayed Pouria Talebi is with the Aalto University, Finland (e-mail: pouria.talebi@aalto.fi).

Danilo P. Mandic is with the Imperial College London, London (e-mail: d.mandic@imperial.ac.uk).

appropriate size. Matrix transpose is denoted by $(\cdot)^{\mathrm{T}}$ and operator vec$\{\cdot\}$ transforms a matrix into a vector by stacking its successive columns. The symbol $\otimes$ is the right Kronecker product operator. Finally, $(\cdot)^{\langle\tau\rangle}$ denotes the elementwise implementation of the function $g(z) = |z|^{\tau}\mathrm{sign}(z)$, where sign$(\cdot)$ and $|\cdot|$ return the sign and the absolute values of their input, respectively.

II. Preliminaries

In many real-world environments the signal and noise processes are likely to have strong variations. The simple Gaussian assumption on these signals is not reasonable as their densities are heavier tailed than that of the Gaussian density. These signals can be well approximated by the α-stable distributions. In literature, researchers used α-stable distributions to model various real-world phenomena such as the random fluctuations of gravitational fields, network traffic, data file sizes on the web and economic market indexes [29].

In general, α-stable distributions do not have an inclusive closed form expression for their probability distribution functions. For example, the Lévy distribution ($\alpha = 0.5$), Cauchy distribution ($\alpha = 1$) and Gaussian distribution ($\alpha = 2$) have different probability density functions. However, the class of real-valued α-stable random processes with elliptically symmetric distributions, which often referred as symmetric α-stable (SαS), have the characteristic function of the form [30], [31]

$$\Phi_{\mathbf{z}}(\mathbf{s}) = \mathrm{E}[\exp(i\mathbf{s}^{\mathrm{T}}\mathbf{z})] = \exp(i\mathbf{s}^{\mathrm{T}}\boldsymbol{\xi})\exp\big(-(\frac{1}{2}\mathbf{s}^{\mathrm{T}}\boldsymbol{\Gamma}_{\mathbf{z}}\mathbf{s})^{\frac{\alpha}{2}}\big), \tag{1}$$

where $\Phi_{\mathbf{z}}(\cdot)$ is the characteristic function of $\mathbf{z}$, $i^2 = -1$, with E$[\cdot]$ denoting the statistical expectation operator. The positive definite covariance matrix $\boldsymbol{\Gamma}_{\mathbf{z}}$ determines the elliptical shape of the distribution of $\mathbf{z}$ that is centered at mean vector $\boldsymbol{\xi}$. The characteristic exponent $\alpha \in (0, 2]$ in (1) governs the tail heaviness of the density function [30], [31]. Small values of α correspond to strong impulsiveness, resulting in heavier tails.

Excluding the Gaussian case, SαS random processes have only finite statistical moments of orders strictly less than α [9], [31], [32]. When it comes to filtering solutions, it is implicitly assumed that $\alpha \in (1, 2]$, so that conditional expectations can be established. Therefore, without loss of generality, this work is limited to real-valued SαS random processes with $\alpha \in (1, 2]$.

III. Proposed Solution

We consider here the problem of tracking the state of a given time-varying system using the input and output signals. At each time instant n, the system state is represented by a parameter matrix $\mathbf{H}_n$ whose internal dynamics are modeled as

$$\mathbf{H}_n = \mathbf{A}\mathbf{H}_{n-1} + \mathbf{V}_n, \tag{2}$$

where matrix $\mathbf{A}$ describes the deterministic system evolution, and $\mathbf{V}_n$ is an SαS matrix sequence that models random variations in the system. The input and output signals of the system are assumed to be related via the linear model

$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{w}_n, \tag{3}$$

where $\mathbf{x}_n$ is the input regression vector used to identify the system, $\mathbf{y}_n$ is the observable output signal vector, and $\mathbf{w}_n$ denotes the background noise vector. Both $\mathbf{x}_n$ and $\mathbf{w}_n$ are SαS random processes.

The *a posteriori* estimate at the previous time index $n-1$ will be mapped onto the current time instant n to obtain the *a priori* estimate of $\mathbf{y}_n$ as stated below:

$$\hat{\mathbf{y}}_n = \hat{\mathbf{H}}_{n|n-1}\mathbf{x}_n \ \text{ with } \ \hat{\mathbf{H}}_{n|n-1} = \mathbf{A}\hat{\mathbf{H}}_{n-1|n-1}, \tag{4}$$

where $\hat{\mathbf{y}}_n$ is the *a priori* estimate of $\mathbf{y}_n$, $\hat{\mathbf{H}}_{n-1|n-1}$ denotes the *a posteriori* estimate at time index $n-1$ and its projection onto time index n is denoted by $\hat{\mathbf{H}}_{n|n-1}$. Here, the objective is to obtain the *a posteriori* estimate at time index n, i.e., $\hat{\mathbf{H}}_{n|n}$ given the observed system response $\mathbf{y}_n$. For this purpose, we need to optimize a suitable cost function so that the *a priori* estimate $\hat{\mathbf{y}}_n$ matches $\mathbf{y}_n$ as closely as possible.

Correntropy provides a similarity measure of two random variables in the vicinity of the kernel bandwidth [21]. By varying the kernel width, the observation window in which the similarity measure is assessed, can be controlled. This adjustable window mechanism helps to regulate the effects of sharps spikes presented in signal/noise [21]. Recently in [22]–[24], correntropy has proven to be a very useful tool in the domain of non-Gaussian signal processing where the noise is modeled as impulsive noise. Inspired from these works, here we adopt the principles of correntropy in designing a suitable objective function. However, the second-order moment of the error measure presented in the maximum correntropy criterion function becomes the main hurdle in obtaining the convergence criterion of the derived algorithm. Furthermore, the usage of ordinary calculus makes the algorithm unstable. In order to overcome these issues, we propose the *fractional-order correntropy* criterion, which is a function of the fractional error-norm. Therefore, we aim to select the state estimates $\{\hat{\mathbf{H}}_{n|n}, n = 1, 2, \cdots\}$ so that they maximize the following criterion:

$$\mathcal{J}_n = \mathrm{E}\Big[\exp\big(-\frac{\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}}{2\beta^2}\big)\Big], \tag{5}$$

where $\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'} = \boldsymbol{\epsilon}_n^T\boldsymbol{\epsilon}_n^{\langle\alpha'-1\rangle}$ is the fractional error-norm, and $\boldsymbol{\epsilon}_n = \mathbf{y}_n - \hat{\mathbf{y}}_n$ is the *a priori* estimation error between the predicted system response and the observed system response. The parameter $\alpha' \in (1, \alpha)$ is a real-valued positive constant that guarantees a concave shape to the cost function (5) and $\beta > 0$ specifies the bandwidth of the kernel.

Similarly as in [17]–[19], the state estimate can be approximated through gradient ascent iterations as follows:

$$\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} + \eta\,\nabla^{\alpha'-1}\mathcal{J}_n, \tag{6}$$

where $\nabla^{\alpha'-1}$ denotes the $(\alpha'-1)$-order gradient operator and η is a positive real-valued gain.

Using the concepts of fractional differentials [14]–[16], the $(\alpha'-1)$-order gradient $\nabla^{\alpha'-1}\mathcal{J}_n$ can be evaluated. We first note that, the function $\exp\big(-\frac{\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}}{2\beta^2}\big)$ is differentiable with respect to $\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}$ and $\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}$ is $(\alpha'-1)$-th differentialble with respect to $\mathbf{H}_n$. Hence, from [14, Lemma 13], we obtain

$$\begin{aligned}\nabla^{\alpha'-1}\mathcal{J}_n &= \nabla^{\alpha'-1}\exp\big(-\frac{\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}}{2\beta^2}\big)\\ &= -\exp\big(-\frac{\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}}{2\beta^2}\big)\frac{(\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'})^{\alpha'-2}}{(2\beta^2)^{\alpha'-1}}\nabla^{\alpha'-1}\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}.\end{aligned} \tag{7}$$

Now, using [14, Lemma 12], we finally have

$$\nabla^{\alpha'-1}\mathcal{J}_n = \exp\big(-\frac{\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}}{2\beta^2}\big)\,\tau_n\,\boldsymbol{\epsilon}_n\,(\mathbf{x}_n^{\langle\alpha'-1\rangle})^{\mathrm{T}}, \tag{8}$$

where $\tau_n = \big((\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'})^{\alpha'-2}\big)/\big((2\beta^2)^{\alpha'-1}\big)$. Absorbing the scalar term $\eta\tau_n$ into the adaptation gain μ, and substituting (8) in (6), at every time index n, the updation rule of the proposed fractional-order correntropy adaptive filter is given by

$$\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} + \mu f(\boldsymbol{\epsilon}_n)\boldsymbol{\epsilon}_n(\mathbf{x}_n^{\langle\alpha'-1\rangle})^{\mathrm{T}}, \tag{9}$$

where $f(\boldsymbol{\epsilon}_n) = \exp\big(-\frac{\|\boldsymbol{\epsilon}_n\|_{\alpha'}^{\alpha'}}{2\beta^2}\big)$ is a function of the fractional-norm of the estimation error $\boldsymbol{\epsilon}_n$.

Remark 1. Since $f(\boldsymbol{\epsilon}_n) \in (0,1]$ and $(\alpha'-1) \in (0,1)$, the proposed algorithm is able to regulate the sharp spikes present both in signal and noise processes. Hence, the proposed algorithm always performs at par or better than the conventional fractional-order adaptive filter.

Remark 2. When $\alpha' = 2$, the proposed algorithm reduces to the conventional maximum correntropy criterion adaptive filter. On the other hand, for a large β value, the conventional fractional-order adaptive filter can be a special case of the proposed algorithm.

Remark 3. The proposed fractional-order correntropy adaptive filter incurs a small amount of additional computational overhead compared to the conventional fractional-order adaptive filter (i.e., $L+2$ extra multiplications, $L-1$ extra additions and an extra $\exp(\cdot)$ function, where L is the length of the *a priori* estimation error vector). This slight increase in overhead is quite acceptable in view of the improvement in performance achieved as shown in simulation results.

Remark 4. By introducing the normalized updates with respect to input regressor $\mathbf{x}_n$ (i.e., normalizing with the fractional-order norm of input regressor), we obtain the following fractional-order correntropy adaptive filter with normalized adaptation gain:

$$\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} + \frac{\mu}{\|\mathbf{x}_n\|_{\alpha'}^{\alpha'}} f(\boldsymbol{\epsilon}_n)\boldsymbol{\epsilon}_n(\mathbf{x}_n^{\langle\alpha'-1\rangle})^{\mathrm{T}}. \tag{10}$$

IV. Convergence Analysis

In this section we study the behavior of the proposed algorithm and obtain a sufficient condition for its convergence. Denoting the state estimation error at time index n as $\boldsymbol{\Upsilon}_n = \mathbf{H}_n - \hat{\mathbf{H}}_{n|n}$ and using (2)-(4), the estimation error $\boldsymbol{\epsilon}_n$ can be expressed as

$$\begin{aligned}\boldsymbol{\epsilon}_n &= \mathbf{H}_n\mathbf{x}_n + \mathbf{w}_n - \mathbf{A}\hat{\mathbf{H}}_{n-1|n-1}\mathbf{x}_n\\ &= \mathbf{A}\mathbf{H}_{n-1}\mathbf{x}_n + \mathbf{V}_n\mathbf{x}_n + \mathbf{w}_n - \mathbf{A}\hat{\mathbf{H}}_{n-1|n-1}\mathbf{x}_n\\ &= \mathbf{A}\boldsymbol{\Upsilon}_{n-1}\mathbf{x}_n + \mathbf{V}_n\mathbf{x}_n + \mathbf{w}_n.\end{aligned} \tag{11}$$

Substituting (11) in (9), the state update equation can be reformulated as

$$\begin{aligned}\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} &+ \mu f(\boldsymbol{\epsilon}_n)\mathbf{A}\boldsymbol{\Upsilon}_{n-1}\boldsymbol{\mathcal{X}}_n + \mu f(\boldsymbol{\epsilon}_n)\mathbf{V}_n\boldsymbol{\mathcal{X}}_n\\ &+ \mu f(\boldsymbol{\epsilon}_n)\mathbf{Q}_n,\end{aligned} \tag{12}$$

where $\boldsymbol{\mathcal{X}}_n = \mathbf{x}_n(\mathbf{x}_n^{\langle\alpha'-1\rangle})^{\mathrm{T}}$ and $\mathbf{Q}_n = \mathbf{w}_n(\mathbf{x}_n^{\langle\alpha'-1\rangle})^{\mathrm{T}}$. Recalling that $\hat{\mathbf{H}}_{n|n-1} = \mathbf{A}\hat{\mathbf{H}}_{n-1|n-1}$ and subtracting both sides of (12) from $\mathbf{H}_n$, the recursion for the state estimation error evolution is obtained as

$$\boldsymbol{\Upsilon}_n = \mathbf{A}\boldsymbol{\Upsilon}_{n-1}\boldsymbol{\mathcal{B}}_n + \mathbf{V}_n\boldsymbol{\mathcal{B}}_n - \mu f(\boldsymbol{\epsilon}_n)\mathbf{Q}_n, \tag{13}$$

where $\boldsymbol{\mathcal{B}}_n = (\mathbf{I} - \mu f(\boldsymbol{\epsilon}_n)\boldsymbol{\mathcal{X}}_n)$. Using the vec$\{\cdot\}$ operator [33] and denoting $\boldsymbol{\gamma}_n = \mathrm{vec}\{\boldsymbol{\Upsilon}_n\}$, (13) can be rewritten as

$$\begin{aligned}\boldsymbol{\gamma}_n = (\boldsymbol{\mathcal{B}}_n \otimes \mathbf{A})\,\boldsymbol{\gamma}_{n-1} &+ (\boldsymbol{\mathcal{B}}_n \otimes \mathbf{I})\,\mathrm{vec}\{\mathbf{V}_n\}\\ &- \mu f(\boldsymbol{\epsilon}_n)\mathrm{vec}\{\mathbf{Q}_n\}.\end{aligned} \tag{14}$$

To establish the convergence conditions, we make the following assumptions:

A1: The random processes $\mathbf{x}_n, \mathbf{w}_n, \mathbf{V}_n$ are assumed to be mutually and temporally independent with mean zero.

A2: The quantity $f(\boldsymbol{\epsilon}(n))$ is assumed to be independent of other quantities.

Remark 5. At each time instant n, we always have $0 < f(\boldsymbol{\epsilon}(n)) \leq 1$. Furthermore, in worst case scenario, i.e., when $f(\boldsymbol{\epsilon}(n)) = 1$, proposed algorithm reduces to conventional fractional-order filter. So **A2** is a reasonable assumption to make. It does not alter the convergence behavior of the proposed algorithm.

Theorem 1. Assume the data model (3), state transition model (2) and the assumptions **A1**-**2** to hold. Then a sufficient condition for the fractional-order correntropy adaptive filter to converge in mean is

$$\max\big\{0, (|\lambda_{\max}(\mathbf{A})| - 1)/\theta\big\} < \mu < (|\lambda_{\max}(\mathbf{A})| + 1)/\theta, \tag{15}$$

where $\theta = \mathrm{E}[f(\boldsymbol{\epsilon}_n)]\lambda_{\max}(\mathrm{E}[\boldsymbol{\mathcal{X}}_n])|\lambda_{\max}(\mathbf{A})|$ with $\lambda_{\max}(\cdot)$ denoting the maximum eigenvalue of its argument matrix.

Proof. Taking the statistical expectation $\mathrm{E}[\cdot]$ on both sides of (14) and using the assumptions **A1**-**2**, we obtain

$$\mathrm{E}[\boldsymbol{\gamma}_n] = (\mathrm{E}[\boldsymbol{\mathcal{B}}_n] \otimes \mathbf{A})\,\mathrm{E}[\boldsymbol{\gamma}_{n-1}], \tag{16}$$

where $\mathrm{E}[\boldsymbol{\mathcal{B}}_n] = \mathbf{I} - \mu\mathrm{E}[f(\boldsymbol{\epsilon}_n)]\mathrm{E}[\boldsymbol{\mathcal{X}}_n)]$. Iterating the recursion (16), backwards down to $n = 0$, we have

$$\mathrm{E}[\boldsymbol{\gamma}_n] = \Big(\prod_{i=1}^{n}\big(\mathrm{E}[\boldsymbol{\mathcal{B}}_i] \otimes \mathbf{A}\big)\Big)\mathrm{E}[\boldsymbol{\gamma}_0]. \tag{17}$$

A sufficient condition for $\lim_{n\to\infty}\mathrm{E}[\boldsymbol{\gamma}_n]$ to attain a finite value is that $|\lambda_{\max}(\mathrm{E}[\boldsymbol{\mathcal{B}}_n])| < 1$ for all n. Using the properties of Kronecker product [33], the above convergence condition can be equivalently stated as $|1 - \mu\mathrm{E}[f(\boldsymbol{\epsilon}_n)]\lambda_{\max}(\mathrm{E}[\boldsymbol{\mathcal{X}}_n])|\ |\lambda_{\max}(\mathbf{A})| < 1$. By solving the above convergence condition, we arrive at (15). □

Since the condition on α' (i.e., $\alpha' \in (1, \alpha)$) ensures the existence of $\mathrm{E}[\boldsymbol{\mathcal{X}}_n]$, the bounds on μ can be evaluated.

Recalling that $0 < f(\boldsymbol{\epsilon}_n) \le 1$, implying $0 < \mathrm{E}[f(\boldsymbol{\epsilon}_n)] \le 1$, it is seen that the convergence conditions of the conventional fractional-order filter obtained in [17] are also sufficient for the convergence of the proposed algorithm.

Following the similar steps (11)-(17), it can be shown that the convergence of the normalized version (10) is guaranteed for $0 < \mu < 1$.

V. Numerical Simulations

In this section, we demonstrate the performance of the proposed fractional-order correntropy adaptive filters via a series of numerical simulations.

A. Tracking a Dynamic System

We consider the filtering problem introduced in Section III, where

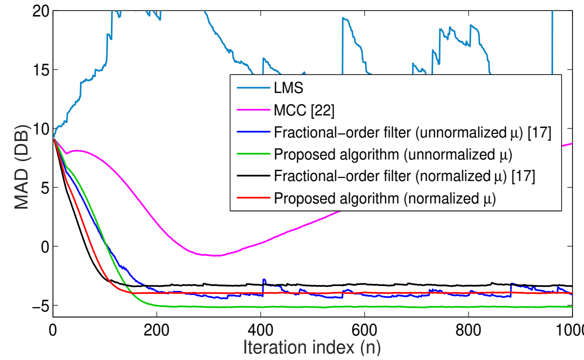
$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0.04 & 0 \\ 0 & 1 & 0 & 0.04 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{V}_n = \begin{bmatrix} 0.0008 & 0 \\ 0 & 0.0008 \\ 0.04 & 0 \\ 0 & 0.04 \end{bmatrix} \boldsymbol{\nu}_n,$$

with $\boldsymbol{\nu}_n$ being a 2×4 zero-mean SαS random matrix sequence having $\boldsymbol{\Gamma}_{\boldsymbol{\nu}_n} = 10^{-4} \times \mathbf{I}_4$. For tracking the system, a 4×1 regression vector $\mathbf{x}_n$ having $\boldsymbol{\Gamma}_{\mathbf{x}_n} = 0.1 \times \mathbf{I}_4$, was used. The background noise vector $\mathbf{w}_n$ has

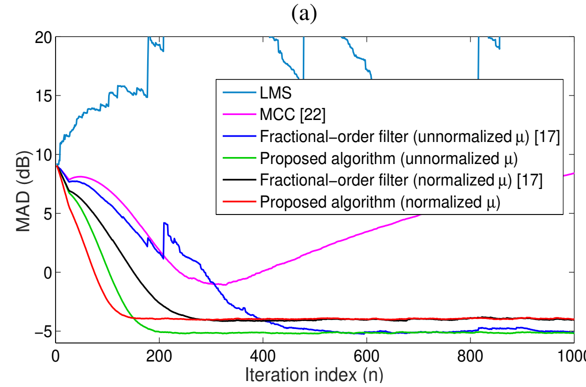
$$\boldsymbol{\Gamma}_{\mathbf{w}_n} = \begin{bmatrix} 10^{-2} & 0 \\ 0 & 10^{-2} \end{bmatrix} \otimes \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix}.$$

In the simulations, the parameter α' was set to $\alpha - 0.5$. The mean absolute deviation (MAD) (defined as $\mathrm{E}[\|\boldsymbol{\gamma}_n\|_1]$) was considered as performance metric.

The characteristic exponent α was fixed at 1.8. The adaptation gain μ and the kernel width parameter β of the proposed filtering approaches were set to 0.2 and 1.7, respectively. The proposed filtering techniques were simulated for tracking the system and the corresponding results are displayed by plotting MAD (in dB) vs iteration index n, obtained by averaging over 5000 independent experiments. The resulting MAD curves are plotted in Fig. 2. For comparison, Fig. 2 also includes MAD performance curves of the conventional fractional-order filters [17], maximum correntropy criterion filter (MCC) [22] and the least mean square filter (LMS). In Fig. 1(a), the adaptation gain of all filters was set to $\mu = 0.2$, to provide a fair comparison of their steady-state performance. However, in Fig. 1(b), the adaptation gain of the conventional fractional-order filters was selected so that they achieved a similar steadstate MAD performances as that of the proposed filters, to provide a fair comparison for their convergence speed. From Fig. 2, we see that the proposed filtering techniques converged properly and are able to track the system state efficiently when compared to conventional fractional-order filters. In contrast, the conventional approaches that depend on the second-order moment of the error measure, i.e, LMS and MCC, failed to track the system state. Furthermore, the unnormalized conventional fractional-order filter and LMS exhibited jitters (i.e., sharp spikes) in their MAD performance curves. This jittery behavior is due to the heavier tailed distributions of input regressor $\mathbf{x}_n$ and noise $\{\mathbf{V}_n, \mathbf{w}_n\}$ processes. On the other hand, since the fractional-order correntropy is insensitive to the jittery behavior of the SαS signals, the proposed approaches have not exhibited any jitters in the MAD performance curves.



(a)



(b)

Fig. 1: MAD performance curves of the proposed adaptive filtering techniques. Performance curves of the conventional fractional-order approaches [17], MCC [22] and LMS are included for comparison. (a) equal convergence rate (b) equal steady-state MAD.

B. Robustness against α Value

In order to examine the robustness of the proposed fractional-order correntropy adaptive filters against the characteristic exponent value α, we carried out the same simulation exercise for various values of α and adaptation gain μ. The steady-state MAD value of the proposed approaches against the α and μ is plotted in Fig. 2. For comparative assessment, we also plotted the steady-state MAD values of conventional fractional-order filters. From Fig. 2, we observe that the MAD performance of the conventional fractional-order filters degrades rapidly as α value decreases. The jitters in input $\mathbf{x}_n$ and noise $\{\mathbf{V}_n, \mathbf{w}_n\}$ will result in a similar behavior in state estimation error $\boldsymbol{\Upsilon}_n$ that manifest itself as a a degraded MAD performance. Furthermore, the unnormalized version of the conventional fractional-order filter exhibits strong jitters in its steady-state MAD performance. These results confirm that the conventional fractional-order filters are not suitable to the scenario where the input regression and noise sequences are heavier tailed (correspond to small α values). On the other hand, as the value of α decreases the amount of degradation in estimation performance is insignificant in the case of the proposed class of filters. From simulations, it is evident that the fractional-order correntropy adaptive filters can achieve better performance by regulating the effect of jittery behavior of input and noise processes. Note that this improvement is achieved at negligible computational overhead as compared to the conventional fractional-order filters.

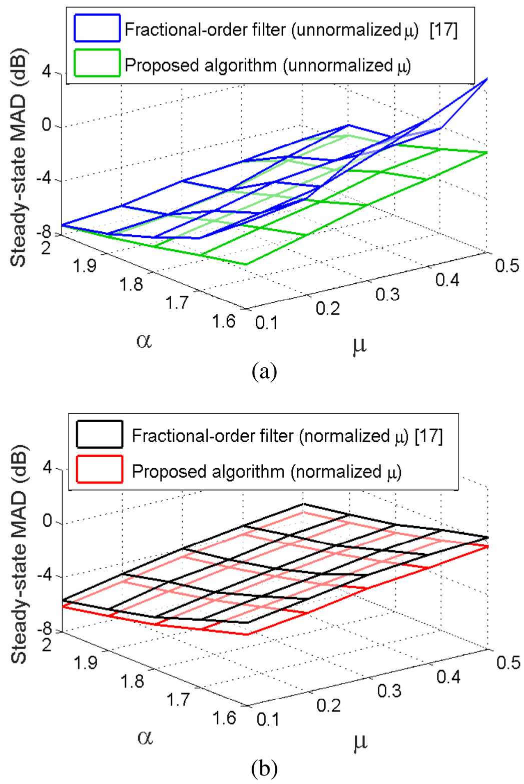


Fig. 2: Steady-state MAD vs α and μ. (a) unnormalized algorithms (b) normalized algorithms.

VI. Conclusions

We considered the problem of estimation and tracking of dynamic systems where the encountered signals exhibit sharp spikes and are modeled as symmetric α-stable signals. To this end, we proposed a class of fractional-order correntropy adaptive filters that effectively overcome the jittery behavior of the signals. The conditions for their stability were established. The performance of the proposed class of algorithms was demonstrated through simulations. Simulations confirmed the superiority of the proposed approaches over the state-of-the-art algorithms.

References

[1] T. Kailath, A. H. Sayed and B. Hassib, *Linear Estimation*, Prentice-Hall, 2000.

[2] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, John Wiley & Sons, 2001.

[3] M. Bouvet and S. C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 37, no. 5, pp. 621-626, May 1989.

[4] N. Azzaoui and L. Clavier, "Statistical channel model based on α-stable random processes and application to the 60 GHz ultra wide band channel," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1457-1467, May 2010.

[5] J. P. Nolan, "Modeling financial data with stable distributions," *Handbook of heavy tailed distributions in finance*, S. T. Rachev, Ed., vol. 1 of *Handbooks in Finance*, pp. 105-130, 2003.

[6] P. Kidmose, "Alpha-stable distributions in signal processing of audio signals," in *Proc. Conf. Simulat. Model.*, Sep. 2000, pp. 87-94

[7] S. Leglaive, U. Simşekli, A. Liutkus, R. Badeau and G. Richard, "Alpha-stable multichannel audio source separation," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Mar. 2017, pp. 576-580.

[8] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, Wiley-Interscience, 1995.

[9] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *Proc. IEEE*, vol. 81, no. 7, pp. 986-1010, Jul. 1993.

[10] E. Masry, "Alpha-stable signals and adaptive filtering," *IEEE Trans. Signal Process.*, vol. 48, no. 11, pp. 3011-3016, Nov. 2000.

[11] A. Navia-Vazquez and J. Arenas-Garcia, "Combination of recursive least p-norm algorithms for robust adaptive filtering in alpha-stable noise," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1478-1482, Mar. 2012.

[12] H. Zayyani, "Continuous mixed p-norm adaptive algorithm for system identification," *IEEE Signal Process. Lett.*, vol. 21, no. 9, pp. 1108-1110, Sep. 2014.

[13] G. Aydin, O. Arikan and A. E. Cetin, "Robust adaptive filtering algorithms for α-stable random processes," *IEEE Transa. Circuits and Syst. II: Analog and Digital Signal Process.*, vol. 46, no. 2, pp. 198-202, Feb. 1999.

[14] G. Jumarie, "On the derivative chain-rules in fractional calculus via fractional difference and their application to systems modelling," *Central European Journal of Physics*, vol. 11, no. 6, pp. 617-633, Jun. 2013.

[15] G. Jumarie,, "Modified Riemann-Liouville derivative and fractional Taylor series of nondifferentiable functions further results," *Computers & Mathematics with Applications*, vol. 51, no. 9, pp. 1367-1376, 2006.

[16] G. Cottone, M. D. Paola and R. Metzler, "Fractional calculus approach to the statistical characterization of random variables and vectors," *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 5, pp. 909-920, 2010.

[17] S. P. Talebi, S. Werner, S. Li and D. P. Mandic, "Tracking dynamic systems in α-stable environments," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process.*, May 2019, pp. 4853-4857.

[18] S. P. Talebi, S. Werner and D. P. Mandic, "Complex-valued nonlinear adaptive filters with applications in α-stable environments," *IEEE Signal Process. Lett.*, vol. 26, no. 9, pp. 1315-1319, Sep. 2019.

[19] S. P. Talebi, S. Werner and D. P. Mandic, "Distributed adaptive filtering of α-stable signals," *IEEE Signal Process. Lett.*, vol. 25, no. 10, pp. 1450-1454, Oct. 2018.

[20] W. Liu, P. P. Pokharel and J. C. Principe, "Correntropy: A localized similarity measure," in *Proc. IEEE Int. Joint Conf. on Neural Networks*, Jul. 2006, pp. 4919-4924.

[21] W. Liu, P. P. Pokharel and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286-5298, Nov. 2007.

[22] A. Singh and J. C. Principe, "Using correntropy as a cost function in linear adaptive filters," in *Proc. Int. Joint Conf. on Neural Networks*, Jun. 2009, pp. 2950-2955.

[23] B. Chen, L. Xing, J. Liang, N. Zheng and J. C. Principe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 880-884, Jul. 2014.

[24] B. Chen, L. Xing, H. Zhao, N. Zheng and J. C. Principe, "Generalized correntropy for robust adaptive filtering," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3376-3387, Jul. 2016.

[25] B. S. Dees, Y. Xia, S. C. Douglas and D. P. Mandic, "Correntropy-based adaptive filtering of noncircular complex data," *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Apr. 2018, pp. 4339-4343.

[26] V. C. Gogineni and S. Mula, "Improved proportionate-type sparse adaptive filtering under maximum correntropy criterion in impulsive noise environments," *Digital Signal Process.*, vol. 79, pp. 190-198, Aug. 2018.

[27] Y. Li, Z. Jiang, W. Shi, X. Han and B. Chen, "Blocked maximum correntropy criterion algorithm for cluster-sparse system identifications," *IEEE Trans. on Circuits and Syst. II: Express Briefs*, vol. 66, no. 11, pp. 1915-1919, Nov. 2019.

[28] K. Xiong, W. Shi and S. Wang, "Robust multikernel maximum correntropy filters," *IEEE Trans.Circuits and Syst. II: Express Briefs.*

[29] R. Adler, R. E. Feldman, and M. S. Taqqu, Eds., *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhauser, Boston, 1998.

[30] J. P. Nolan, "Multivariate elliptically contoured stable distributions: Theory and estimation," *Computational Statistics*, vol. 28, no. 5, pp. 2067-2089, Oct. 2013.

[31] G. Samoradnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman & Hall, 1994.

[32] V. Uchaikin and V. Zolotarev, *Chance and Stability: Stable Distributions and their Applications*, VSP, 1999.

[33] G. Alexander, *Kronecker Products and Matrix Calculus: With Applications*, New York: Halsted, 1981.