TRACKING DYNAMIC SYSTEMS IN α-STABLE ENVIRONMENTS

Sayed Pouria Talebi†, Stefan Werner‡, Shengxi Li, and Danilo P. Mandic**

†Signal Processing and Acoustics Department, Aalto University, FI-02150 Finland
‡Electronic Systems Department, Norwegian University of Science and Technology, NO-7491 Norway
*Electrical and Electronic Engineering Department, Imperial College London, SW7 2AZ, UK
E-mails: pouria.talebi@aalto.fi, stefan.werner@ntnu.no, & {shengxi.li17,d.mandic}@imperial.ac.uk

ABSTRACT

In order to accommodate for modern adaptive filtering applications, the classic adaptive filtering paradigm is considered from a more general perspective. The new formulation allows for time dependent variations in the state of the system and more importantly it relaxes the Gaussian assumption to the generalized setting of α-stable distributions. In this work, based on the principles of gradient descent and fractional-order calculus, a cost-effective technique for tracking the state of such a system is derived. For rigour, performance of the derived filtering technique is analyzed and convergence conditions are established.

Index Terms— α-stable signals, fractional-order calculus, fractional-order filtering, adaptive filtering/tracking.

1. INTRODUCTION

In order to accommodate for mathematical tractability, the overwhelming majority of stochastic machine learning, signal processing, and control techniques assume a Gaussian model for the signal and noise [1–3]. However, in a growing number of modern applications, the encountered signal/noise exhibits sharp spikes which results in distributions that decay slower than the Gaussian case [4–11]. Owing to their signature stable property [4] and the generalized central limit theorem [12], accurate modeling of such signals has been shown to be possible through the framework of α-stable random processes [4–8].

The class of real-valued α-stable random processes with elliptically symmetric distributions, hereafter referred to as symmetric α stable (SαS), have characteristic function of the form [13,14]

$$\mathbf{\Phi}_{\mathbf{z}}(\mathbf{s}) = \mathsf{E}\left\{e^{i\mathbf{s}^{\mathsf{T}}\mathbf{z}}\right\} = e^{i\mathbf{s}^{\mathsf{T}}\boldsymbol{\xi}} e^{-\left(\frac{1}{2}\mathbf{s}^{\mathsf{T}}\mathbf{\Gamma}_{\mathbf{z}}\mathbf{s}\right)^{\frac{\alpha}{2}}} \tag{1}$$

where $\mathbf{\Phi}_{\mathbf{z}}(\cdot)$ is the characteristic function of $\mathbf{z}$, $i^2 = -1$, $(\cdot)^{\mathsf{T}}$ denotes the transpose operator, and $\mathsf{E}\{\cdot\}$ denotes the statistical expectation operator, while the positive definite matrix $\mathbf{\Gamma}_{\mathbf{z}}$ determines the elliptical shape of the distribution of $\mathbf{z}$ that is centered at $\boldsymbol{\xi}$. The characteristic exponent $\alpha \in (0, 2]$ in (1) governs the tail heaviness of the density function [8,13]. Small values of α correspond to severe impulsiveness, resulting in heavier tails, while values close to 2 correspond to more Gaussian type behavior. For the special case when $\alpha = 2$, random vector $\mathbf{z}$ has a Gaussian distribution with the covariance matrix $\mathbf{\Gamma}_{\mathbf{z}}$ and mean vector $\boldsymbol{\xi}$.

The work of Sayed Pouria Talebi and Stefan Werner was supported in part by Academy of Finland under Grant 296849 and in part by Research Council of Norway.

From the entire class of α-stable random processes, only the Gaussian case has well-defined second and higher-order statistical moments [12,13]. Indeed, excluding the Gaussian case, SαS random processes have only finite statistical moments of orders strictly less than α [8,12,13]. Therefore, when it comes to filtering solutions, it is implicitly implied that $\alpha \in (1, 2]$, so that conditional expectations can be established. Without loss of generality, this work is limited to real-valued SαS random processes with $\alpha \in (1, 2]$, where $\boldsymbol{\xi}$ in (1) corresponds to the mean vector.

The restriction to finite statistical moments of order less than α hinders the effectiveness of classical filtering techniques when applied to general SαS signals [15–19], as these techniques are usually based on minimizing the second-order moment of an error measure. To this end, a number of filtering techniques have been developed for processing α-stable signals [5,16,18,20,21]. These filtering techniques typically aim to minimize a fractional norm of the error measure. However, their use of ordinary calculus results in filtering techniques that are not cost-effective and/or do not have analytically tractable behavior, prohibiting the establishment of closed-form convergence criteria. Notable recent developments in this area include the particle filtering technique in [15], which is based upon tracking parameters of the characteristic function in (1) and presents a solution for tracking SαS state vector sequences. Furthermore, the authors of this work have formulated a gradient-descent adaptive filtering technique in [19], based on fractional-order calculus and statistics [22–24]. More importantly, using the characteristic function of SαS processes an optimal filtering solution was also established in [19].

In order to present an all-inclusive adaptive filtering solution, 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. Then, an adaptive solution based on minimizing fractional-order norms of an error measure in a gradient descent manner is derived. The introduced adaptive filter is cost-effective to implement and is applicable to a wide range of estimation/tracking scenarios. Moreover, in incidences limited to the Gaussian case and given a number of simplifying assumptions, the proposed adaptive filtering technique can be simplified to classical approaches. Furthermore, performance of the derived adaptive filtering algorithm is analyzed and requirements for its convergence are established. Finally, the introduced concepts are verified through simulations.

***Mathematical Notations*:** Scalars, column vectors, and matrices are denoted by lowercase, bold lowercase, and bold uppercase letters, with $\mathbf{I}$ representing the identity matrix of appropriate size. The transpose and statistical expectation operators are denoted by $(\cdot)^{\mathsf{T}}$ and $\mathsf{E}\{\cdot\}$, while $\otimes$ denotes the Kronecker product. The operator $\mathsf{vec}\{\cdot\}$ transforms a matrix into a vector by stacking its columns.

Finally, $(\cdot)^{\langle\tau\rangle}$ denotes the element wise implementation of the function $f(z) = |z|^{\tau}\text{sign}(z)$, with sign$(\cdot)$ denoting the sign function.

2. PROBLEM FORMULATION

The goal is to track the state of a given system through observation of its input/output signals. The state of the system at time instant n is represented via parameter matrix $\mathbf{H}_n$, while the input and output signals are interrelated as

$$\mathbf{y}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{w}_n \tag{2}$$

where at time instant n, $\mathbf{y}_n$ and $\mathbf{w}_n$ denote the output and background noise vectors, with $\mathbf{x}_n$ representing the regression vector used to identify the system. The system in question is considered to be time variant with internal dynamics modeled as

$$\mathbf{H}_n = \mathbf{A}\mathbf{H}_{n-1} + \mathbf{V}_n \tag{3}$$

where $\mathbf{A}$ is a matrix representing the deterministic system evolution and $\{\mathbf{V}_n : n = 1, 2, \ldots\}$ is an SαS matrix sequence representing random system mutations. The random processes $\{\mathbf{x}_n, \mathbf{w}_n, \mathbf{V}_n\}$ are assumed to be mutually and temporally independent with zero-mean.

Remark 1: Although the models in (2)-(3) are considered to be linear and the system evolution matrix $\mathbf{A}$ is taken to be time invariant for brevity and simplicity in presentation, the introduced concepts can be readily generalized.

Remark 2: If all random sequences are considered to be Gaussian and the system is assumed to be time invariant, that is, $\mathbf{H}_n \to \mathbf{H}$; then, the problem in this work simplifies to that considered in [25, 26], which leads to the standard least mean square (LMS) algorithm.

Remark 3: Given the simplifying assumptions that all random sequences are Gaussian and the system is directly accessible via a linear observer; then, the problem in this work simplifies to that considered by Kalman in [27] which leads to the standard Kalman filter.

3. THE PROPOSED FRACTIONAL-ORDER FILTER

Let $\hat{\mathbf{y}}_n$ denote the estimate of $\mathbf{y}_n$ that is obtained through the strictly linear model

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

where $\hat{\mathbf{H}}_{n-1|n-1}$ denotes the system state estimate at time instant $(n-1)$ and $\hat{\mathbf{H}}_{n|n-1}$ denotes the projection of $\hat{\mathbf{H}}_{n-1|n-1}$ onto time instant n. The aim here becomes that of updating the available estimate of the system state given the accessible information at time instant n, that is, obtaining $\hat{\mathbf{H}}_{n|n}$ given the observed system response $\mathbf{y}_n$. This is achieved through selecting the system state estimates $\{\hat{\mathbf{H}}_{n|n} : n = 1, 2, \ldots\}$ so that they minimize the cost function

$$\mathcal{J}_n = \boldsymbol{\epsilon}_n^{\mathsf{T}}\boldsymbol{\epsilon}_n^{\langle\alpha'-1\rangle} \quad \text{with} \quad \boldsymbol{\epsilon}_n = \mathbf{y}_n - \hat{\mathbf{y}}_n \tag{5}$$

where $1 < \alpha' < \alpha$. The selected error measure, $\boldsymbol{\epsilon}_n$, indicates the discrepancy between the predicted system response to input $\mathbf{x}_n$, that is, $\hat{\mathbf{y}}_n$ as formulated in (4), and the observed system response $\mathbf{y}_n$. Furthermore, it is straightforward to show the cost function approaches its unique minimum point as $\hat{\mathbf{H}}_{n|n} \to \mathbf{H}_n$ (see Section 4).

Remark 4: The constraint on parameter $\alpha' \in (1, \alpha)$ is to guarantee a convex shape for the cost function. This constraint also becomes crucial in Section 4, when establishing convergence criteria.

Akin to the approach proposed by the authors in their previous work [19] and based on the same principle, the system state estimate is updated at each time instant in a gradient descent manner so that

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

where $\nabla^{\alpha'-1}$ represents the $(\alpha'-1)$-order gradient operator, while μ denotes a positive real-valued adaptation gain. After some mathematical manipulations, the update law in (6) yields

$$\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} + \mu\boldsymbol{\epsilon}_n\left(\mathbf{x}_n^{\langle\alpha'-1\rangle}\right)^{\mathsf{T}} \tag{7}$$

where the framework introduced in [23,24] was used for calculating fractional differentials and constant multiplicative terms are absorbed into the adaptation gain μ.

A block diagram of the proposed filtering operations is shown in Fig. 1, where $\mathbf{H}_0$ represents the initial state of the dynamic system, that is, the state of the system at $n = 0$. Moreover, the linear model in (4) uses the filter output at time instant n, $\hat{\mathbf{H}}_{n|n}$, to estimate the system output at time instant $(n+1)$. Notice that for the case where $\mathbf{V}_n$ vanishes and $\mathbf{A} = \mathbf{I}$, the proposed filter simplifies to that of our previous work in [19]. Furthermore, if it is also assumed that $\alpha = 2$; then, as $\alpha' \to 2$ the proposed algorithm simplifies to the LMS.

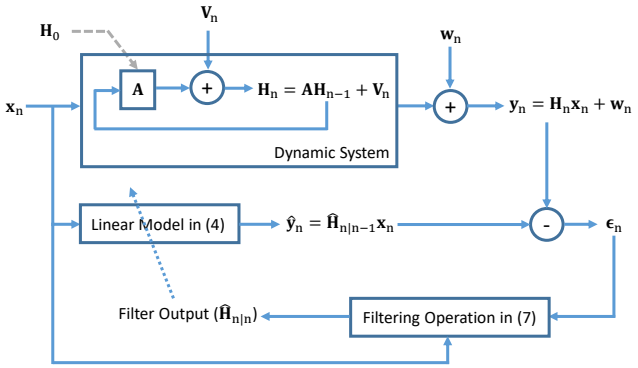

Fig. 1: Operations of the proposed filtering approach. The initial value of the dynamic system, $\mathbf{H}_0$, is used at the initial time instant only. The output of the filter, $\hat{\mathbf{H}}_{n|n}$, is used to update the linear model for the next time instant.

4. PERFORMANCE AND STABILITY ANALYSIS

In order to analyze the behavior of the proposed filtering technique and establish its convergence criteria, the error measure in (5) is first formulated in terms of the state estimation error. To this end, upon replacing (2) and (4) into (5) we have

$$\boldsymbol{\epsilon}_n = \mathbf{y}_n - \hat{\mathbf{y}}_n = \mathbf{H}_n\mathbf{x}_n + \mathbf{w}_n - \mathbf{A}\hat{\mathbf{H}}_{n-1|n-1}\mathbf{x}_n. \tag{8}$$

Now, substitution of (3) into (8) yields

$$\begin{aligned}\boldsymbol{\epsilon}_n =& \mathbf{A}\mathbf{H}_{n-1}\mathbf{x}_n + \mathbf{V}_n\mathbf{x}_n + \mathbf{w}_n - \mathbf{A}\hat{\mathbf{H}}_{n-1|n-1}\mathbf{x}_n \\ =& \mathbf{A}\boldsymbol{\Upsilon}_{n-1}\mathbf{x}_n + \mathbf{V}_n\mathbf{x}_n + \mathbf{w}_n\end{aligned} \tag{9}$$

where $\boldsymbol{\Upsilon}_{n-1} = \mathbf{H}_{n-1} - \hat{\mathbf{H}}_{n-1|n-1}$ denotes the state estimation error at time instant $(n-1)$.

Using the expression in (9), the update law in (7) can be reformulated to give

$$\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} + \mu\mathbf{A}\boldsymbol{\Upsilon}_{n-1}\boldsymbol{\mathcal{X}}_n + \mu\mathbf{V}_n\boldsymbol{\mathcal{X}}_n + \boldsymbol{\mathcal{Q}}_n \tag{10}$$

where $\boldsymbol{\mathcal{X}}_n = \mathbf{x}_n \left(\mathbf{x}_n^{\langle \alpha'-1\rangle}\right)^{\mathsf{T}}$ and $\boldsymbol{\mathcal{Q}}_n = \mu \mathbf{w}_n \left(\mathbf{x}_n^{\langle \alpha'-1\rangle}\right)^{\mathsf{T}}$.

Substituting (4) into (10) allows the state estimation error to be expressed in a regressive fashion as

$$\boldsymbol{\Upsilon}_n = \mathbf{A}\boldsymbol{\Upsilon}_{n-1} - \mu\mathbf{A}\boldsymbol{\Upsilon}_{n-1}\boldsymbol{\mathcal{X}}_n + \mathbf{V}_n - \mu\mathbf{V}_n\boldsymbol{\mathcal{X}}_n - \boldsymbol{\mathcal{Q}}_n$$

which can be rearranged as

$$\boldsymbol{\Upsilon}_n = \mathbf{A}\boldsymbol{\Upsilon}_{n-1}\left(\mathbf{I} - \mu\boldsymbol{\mathcal{X}}_n\right) + \mathbf{V}_n\left(\mathbf{I} - \mu\boldsymbol{\mathcal{X}}_n\right) - \boldsymbol{\mathcal{Q}}_n. \tag{11}$$

Alternatively, the expression in (11) can be formulated in a vector format to give

$$\begin{aligned}\mathsf{vec}\left\{\boldsymbol{\Upsilon}_n\right\} = &\left(\left(\mathbf{I} - \mu\boldsymbol{\mathcal{X}}_n^{\mathsf{T}}\right) \otimes \mathbf{A}\right) \mathsf{vec}\left\{\boldsymbol{\Upsilon}_{n-1}\right\} \\ &+ \left(\left(\mathbf{I} - \mu\boldsymbol{\mathcal{X}}_n^{\mathsf{T}}\right) \otimes \mathbf{I}\right) \mathsf{vec}\left\{\mathbf{V}_n\right\} - \mathsf{vec}\left\{\boldsymbol{\mathcal{Q}}_n\right\}\end{aligned} \tag{12}$$

which corresponds to a closed-form expression for the evolution of state estimation error from one time instant to the next.

Taking the statistical expectation of (12) and considering that $\{\mathbf{x}_n, \mathbf{w}_n, \mathbf{V}_n\}$ are assumed zero-mean and independent results in

$$\mathsf{E}\left\{\mathsf{vec}\left\{\boldsymbol{\Upsilon}_n\right\}\right\} = \left(\left(\mathbf{I} - \mu\mathsf{E}\left\{\boldsymbol{\mathcal{X}}_n^{\mathsf{T}}\right\}\right) \otimes \mathbf{A}\right)\mathsf{E}\left\{\mathsf{vec}\left\{\boldsymbol{\Upsilon}_{n-1}\right\}\right\}. \tag{13}$$

From (13) it becomes clear that any misadjustment in initialization will decay exponentially fast given

$$\rho\left(\left(\mathbf{I} - \mu\mathsf{E}\left\{\boldsymbol{\mathcal{X}}_n^{\mathsf{T}}\right\}\right) \otimes \mathbf{A}\right) < 1 \tag{14}$$

where $\rho(\cdot)$ returns the spectral radius. The condition in (14) results in a bound on the allowed adaptation gain, given by

$$\max\left\{0, \frac{\rho(\mathbf{A}) - 1}{\rho(\mathbf{A})\,\rho\left(\mathsf{E}\left\{\boldsymbol{\mathcal{X}}_n\right\}\right)}\right\} < \mu < \frac{\rho(\mathbf{A}) + 1}{\rho(\mathbf{A})\,\rho\left(\mathsf{E}\left\{\boldsymbol{\mathcal{X}}_n\right\}\right)} \tag{15}$$

which will ensure convergence.

Remark 5: Recall that $\boldsymbol{\mathcal{X}}_n = \mathbf{x}_n\left(\mathbf{x}_n^{\langle \alpha'-1\rangle}\right)^{\mathsf{T}}$ with $\alpha' \in (1, \alpha)$. The bound on α' ensures that $\mathsf{E}\left\{\boldsymbol{\mathcal{X}}_n\right\}$ does exist and that a bound for the adaptation gain can be determined.

Once again, consider the regression in (11). If the convergence condition in (15) is met; then, for $1 < p < \alpha$, the p^{th} order statistical moment of $\boldsymbol{\Upsilon}_n$, that is, $\mathsf{E}\left\{\|\mathsf{vec}\mathsf{E}\left\{\boldsymbol{\Upsilon}_n\right\}\|_p^p\right\}$, converges to a stabilizing solution, where the criterion $1 < p < \alpha$ is set to guarantee the statistical moment in question exists and is finite. In cases where statistics of the regression vector, $\mathbf{x}_n$, are not available, a pragmatic method for ensuring convergence is to normalize the adaptation step-size at each time instant. In this setting, the update equation in (7) becomes

$$\hat{\mathbf{H}}_{n|n} = \hat{\mathbf{H}}_{n|n-1} + \frac{\mu}{\|\mathbf{x}_n\|_{\alpha'}^{\alpha'}}\boldsymbol{\epsilon}_n\left(\mathbf{x}_n^{\langle \alpha'-1\rangle}\right)^{\mathsf{T}}. \tag{16}$$

Taking into account that $\|\mathbf{x}_n\|_{\alpha'}^{\alpha'} = \mathbf{x}_n^{\mathsf{T}}\mathbf{x}_n^{\langle \alpha'-1\rangle}$ is equal to the trace of $\boldsymbol{\mathcal{X}}_n$, upon repeating the analysis in (8)-(15) it becomes clear that for the case of the normalized adaptation step-size in (16), convergence will be guaranteed for $0 < \mu < 1$. Finally, attention of the reader is drawn to the similarity between the expression in (12) and the equation governing performance of the Kalman filter (see Chapter 9 in [3] and Chapter 7 in [1]). Indeed, when the system state, $\mathbf{H}_n$, is accessible through a linear observer, the expression in (12) transforms into a quasi-Lyapunov equation.[1] However, due to the space limitations of this venue, this matter is not explored further.

[1]For the case that is also restricted to Gaussian signal/noise processes, (12) transforms into to a Lyaponov equation.

5. PERFORMANCE VERIFICATION

In this section, performance of the proposed adaptive filtering technique is illustrated. To this end, the filtering problem formulated in Section 2 was considered, where

$$\mathbf{A} = \begin{bmatrix}1 & 0 & 0.04 & 0\\ 0 & 1 & 0 & 0.04\\ 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{bmatrix} \text{ and } \mathbf{V}_n = \begin{bmatrix}0.08 & 0.008\\ 0 & 0.08\\ 4 & 0.4\\ 0 & 4\end{bmatrix}\boldsymbol{\nu}_n$$

with $\boldsymbol{\nu}_n$ denoting a 2×4 zero-mean SαS random matrix sequence generated so that $\boldsymbol{\Gamma}_{\mathsf{vec}\{\boldsymbol{\nu}_n\}} = 10^{-4} \times \mathbf{I}$. A regression vector of length four with $\boldsymbol{\Gamma}_{\mathbf{x}_n} = \mathbf{0}.1 \times \mathbf{I}$ was used to track the system, while

$$\boldsymbol{\Gamma}_{\mathbf{w}_n} = \begin{bmatrix}10^{-2} & 0\\ 0 & 10^{-2}\end{bmatrix} \otimes \begin{bmatrix}1 & 0.2\\ 0.2 & 1\end{bmatrix}.$$

Furthermore, in all the conducted simulations $\alpha' = \alpha - 0.5$. The mean absolute error (MAE), given by $\mathsf{E}\left\{|\boldsymbol{\epsilon}_{l,n}|\right\}$, and mean absolute deviation (MAD), given by $\mathsf{E}\left\{|\boldsymbol{\Upsilon}_n|\right\}$, were considered as performance metrics.

In the first set of simulations, tracking performance of the proposed adaptive filtering technique was considered. In this case, the characteristic exponent was $\alpha = 1.8$ and the adaptation gain was $\mu = 0.2$. The MAE and MAD performance metrics are shown in Fig. 2 and the transient behavior of $\mathsf{E}\left\{\mathsf{vec}\left\{\boldsymbol{\Upsilon}_n\right\}\right\}$ is shown in Fig. 3. In addition, to demonstrate importance of the fractional-order approach, Fig. 2 also includes filtering results for the conventional approach using the first-order gradient operator to minimize cost function $J_n = \boldsymbol{\epsilon}_n^{\mathsf{T}}\boldsymbol{\epsilon}_n$, which corresponds to the case of $\alpha' \to 2$. From Fig. 2 and Fig. 3, observe that the proposed adaptive filter, both with the normalized update term in (16) and with the update term in (7), converged and was able to track the system state. Moreover, from Fig. 2 it becomes clear that the conventional approach, that is, for $\alpha' \to 2$, was unable to track the system state. Furthermore, note that the conventional approach exhibited large jitters (sharp spikes) in its MAE and MAD performance. These jitters are due to the regression and background noise sequences being heavy tailed. On the other hand, the proposed adaptive filtering technique had a smother MAE and MAD performance especially in the case where the step-size was normalized.

In the second set of simulations, the MAD performance profile of the proposed adaptive filtering technique was considered. The steady-state MAD performance of the proposed adaptive filtering technique for varying adaptation gain and characteristic exponent values are shown in Fig. 4, while Fig. 5 shows performance of the proposed adaptive filtering technique with normalized step-size, as formulated in (16). From Fig. 4 and Fig. 5, observe that the steady-state MAD performance of the filter degrades as the characteristic exponent decreases. This is due to the heavier tailed distributions of the noise, $\{\mathbf{V}_n, \mathbf{w}_n\}$, and regression, $\mathbf{x}_n$, processes. This follows from (11) and (12) from which it can be inferred that encountering jitters (sharp jumps in value or spikes) in $\mathbf{x}_n$, $\mathbf{w}_n$, or $\mathbf{V}_n$ will result in a similar behavior in $\boldsymbol{\Upsilon}_n$ that manifests it self as a degraded MAD performance. It should also be noted that Fig. 4 and Fig. 5 indicate that using the normalized update term formulated in (16) reduces MAD performance dependency on the characteristic exponent, as a normalized step-size counters the effect caused by jitters in $\mathbf{x}_n$.

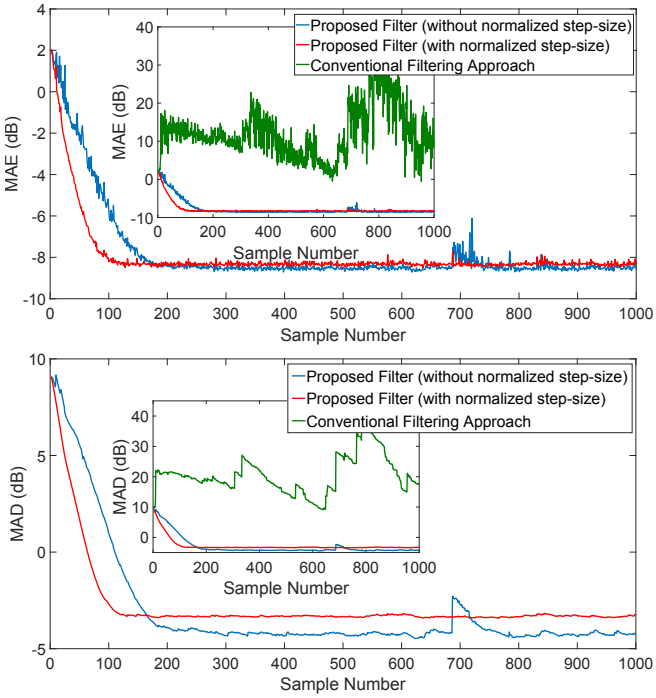

Fig. 2: The MAE and MAD performances of the proposed adaptive filtering technique are shown in the top and bottom graphs respectively. The approach without the normalized step-size formulated in (7) is shown in blue, while the approach with the normalized step-size formulated in (16) is shown in red. The conventional filtering performance for which $\alpha' \rightarrow 2$ is shown in green.

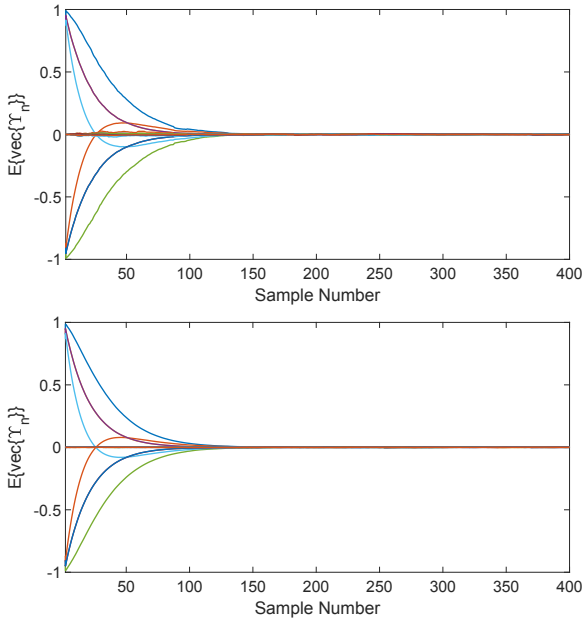

Fig. 3: Transient behavior of all components of $\mathsf{E}\{\mathsf{vec}\{\boldsymbol{\Upsilon}_n\}\}$. The top graph shows results obtained by the approach without the normalized step-size formulated in (7), while the bottom graph shows results obtained by the approach with the normalized step-size formulated in (16).

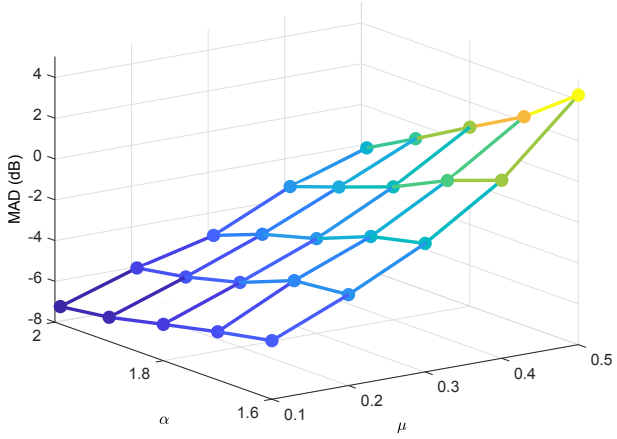

Fig. 4: The MAD performance for the proposed filtering technique for varying adaptation gain and characteristic exponent values. Note that the step-size was not normalized.

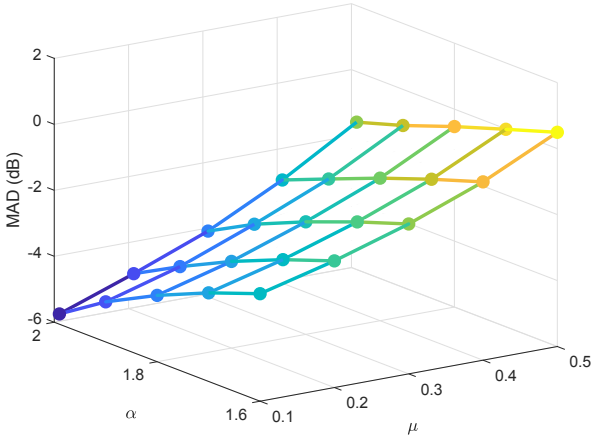

Fig. 5: The MAD performance for the proposed filtering technique for varying adaptation gain and characteristic exponent values. Note that the step-size was normalized.

6. CONCLUSION

In order to suit estimation and tracking needs arising in modern adaptive filtering applications, the classical adaptive filtering paradigm has been revised. The formulated problem is general in the sense that it can be transformed into the LMS (*cf.* Kalman) filtering approaches as a special case and can accommodate SαS signal models. Then, based on minimizing fractional powers of an error measure using differentials of orders less than one, a gradient-descent type adaptive solution to the formulated problem has been developed. The performance of the derived algorithm has been analyzed, establishing the required convergence criteria. The effectiveness of the derived adaptive filtering technique has been validated over illustrative simulation examples.

7. REFERENCES

[1] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*, Prentice Hall, 1986.

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

[3] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 2000.

[4] J. P. Nolan, "Modeling financial data with stable distributions," *Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance, S. T. Rachev, Ed. Elsevier-North Holland, Amsterdam*, 2005.

[5] J. P. Nolan and D. Ojeda-Revah, "Linear and nonlinear regression with stable errors," *Journal of Econometrics*, vol. 172, no. 2, pp. 186–194, February 2013.

[6] 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 Transactions on Communications*, vol. 58, no. 5, pp. 1457–1467, May 2010.

[7] S. Leglaive, U. Simsekli, A. Liutkus, R. Badeau, and G. Richard, "Alpha-stable multichannel audio source separation," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 576–580, March 2017.

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

[9] J. Wang, E. E. Kuruoglu, and T. Zhou, "Alpha-stable channel capacity," *IEEE Communications Letters*, vol. 15, no. 10, pp. 1107–1109, October 2011.

[10] S. Niranjayan, N. C. Beaulieu, S. Niranjayan, and N. C. Beaulieu, "BER optimal linear combiner for signal detection in symmetric alpha-stable noise: Small values of alpha," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 886–890, March 2010.

[11] Y. Gao, J. Ni, J. Chen, and X. Chen, "Steady-state and stability analyses of diffusion sign-error LMS algorithm," *Signal Processing*, vol. 149, pp. 62–67, 2018.

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

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

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

[15] S. P. Talebi and D. P. Mandic, "Distributed particle filtering of α-stable signals," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1862–1866, December 2017.

[16] E. Masry, "Adaptive filtering algorithms and alpha stable signals," *In Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 9–12, 2000.

[17] L. Mihaylova, P. Brasnett, A. Achim, D. Bull, and N. Canagarajah, "Particle filtering with alpha-stable distributions," *In Proceedings of IEEE/SP* 13^{th} *Workshop on Statistical Signal Processing*, pp. 381–386, July 2005.

[18] E. Masry, "Alpha-stable signals and adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 48, no. 11, pp. 3011–3016, November 2000.

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

[20] A. Navia-Vazquez and J. Arenas-Garcia, "Combination of recursive least p-norm algorithms for robust adaptive filtering in alpha-stable noise," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1478–1482, March 2012.

[21] G. Aydin, O. Arikan, and A. E. Cetin, "Robust adaptive filtering algorithms for α-stable random processes," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 2, pp. 198–202, February 1999.

[22] 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.

[23] 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, 2013.

[24] 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.

[25] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *IRE WESCON Convention Record, Part 4*, pp. 96–104, August 1960.

[26] B. Widrow and S. D. Stearns, *Adaptive signal processing*, Prentice Hall, 1985.

[27] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82, Series D, pp. 35–45, 1960.