Talebi, Sayed Pouria; Werner, Stefan; Mandic, Danilo P.

**Complex-Valued Nonlinear Adaptive Filters with Applications in α-Stable Environments**

*Published in:*
IEEE Signal Processing Letters

*DOI:*
10.1109/LSP.2019.2929874

Published: 01/09/2019

*Document Version*
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

*Please cite the original version:*
Talebi, S. P., Werner, S., & Mandic, D. P. (2019). Complex-Valued Nonlinear Adaptive Filters with Applications in α-Stable Environments. *IEEE Signal Processing Letters*, *26*(9), 1315-1319. Article 8768059.
https://doi.org/10.1109/LSP.2019.2929874

# Complex-Valued Nonlinear Adaptive Filters with Applications in $\alpha$-Stable Environments

Sayed Pouria Talebi, Stefan Werner, and Danilo P. Mandic

***Abstract*—A nonlinear adaptive filtering framework for processing complex-valued signals is derived. The introduced adaptive filter extends the fractional-order framework of the authors for dealing with real-valued signals to the complex domain via the augmented statistical approach to complex-valued signal processing. This results in a versatile class of adaptive filtering techniques which allows the classical Gaussian assumption to be extended to the generalized context of $\alpha$-stables. For rigour, the performance of the introduced adaptive filtering framework is analyzed, its convergence criteria is established, and its application in tracking signals of chaotic systems is demonstrated using simulations.**

***Index Terms*—Nonlinear adaptive filtering, complex-valued $\alpha$-stable signals, fractional-order adaptive filtering.**

## I. Introduction

Classical adaptive filtering, learning, and control techniques have been derived and implemented in the real domain [1]–[4]. However, in a great number of engineering applications, complex-valued modeling of physical systems provides a straightforward mathematical framework for solving the differential equations that govern their behaviour; thus, allowing for fast and convenient analysis of their performance [5]. In addition to communication engineering, perhaps the best known example of such systems are electrical circuits, such as power transmission systems [6]–[9]. Moreover, when equipped with the power of augmented statistics, complex-valued modeling and signal processing have been shown to be advantageous in an increasing number of applications, such as wind profile forecasting [10,11], frequency domain signal analysis [12], and information processing [13]–[15].

Traditionally, complex-valued filtering techniques have been developed with the implicit assumption that signals of interest are Gaussian and second-order circular (proper), i.e., Gaussian with rotation invariant distributions. However, recent developments have by and large put this assumption under scrutiny [5,16]. In particular, adaptive filtering techniques developed with impropriety in mind have been shown to possess advantageous performance in an increasing number of applications [5,12,16]–[18]. Moreover, modern filtering and estimation applications often deal with signals that exhibit sharp spikes, the distributions of which decay slower than the Gaussian case [19]–[24]. Modeling these signals has been shown to be possible through the generalized framework of $\alpha$-stable random processes [22,25]. Although a number of approaches for filtering such signals have been recently introduced [26]–[30], these approaches are restricted to real-valued signals. Thus, a comprehensive framework for dealing with complex-valued $\alpha$-stable signals is still lacking.

The focus of this work is on a flexible class of $\alpha$-stable random processes, termed complex-valued elliptically symmetric $\alpha$-stable, hereafter referred to as C-S$\alpha$S for short. This is motivated by the versatility of C-S$\alpha$S processes that allow for the derivation of mathematically tractable filtering solutions. In addition, as a large class of $\alpha$-stable random processes can be approximated via a combination of C-S$\alpha$S random processes, the applicability of adaptive filtering techniques derived on this basis is maintained.

In this work, a class of nonlinear adaptive filtering techniques for real-time processing of C-S$\alpha$S signals is derived. This is achieved using the augmented statistical approach for dealing with complex-valued signals in order to extend the framework first proposed by the authors in [31,32] for dealing with real-valued signals to the complex-valued setting. In addition, performance of the derived filtering technique is analyzed and its convergence criteria are established. Finally, the introduced filtering framework is generalized so that a mixture of the derived fractional-order filtering operators can be used in unison for improved performance.

***Mathematical Notations***: The real-valued sign operator is denoted by $\text{sign}(\cdot)$. The real and complex domains are denoted by $\mathbb{R}$ and $\mathbb{C}$. Scalars, column vectors, and matrices are denoted respectively by lowercase, bold lowercase, and bold uppercase letters, while $\mathbf{I}$ denotes an identity matrix of appropriate size. The transpose, Hermitian transpose, and complex conjugate operators are respectively denoted by $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $(\cdot)^{*}$. The statistical expectation operator is denoted by $\mathsf{E}\{\cdot\}$, while the operator $\mathsf{vec}\{\cdot\}$ stacks the columns of a matrix transforming it into a vector. Finally, $\Re\{x\}$ and $\Im\{x\}$ return the real and imaginary components of $x$, while $j^2 = -1$.

## II. Preliminaries and Problem Formulation

In the context of the $\mathbb{CR}$-calculus [5,33], a complex-valued function $f(x) : \mathbb{C} \rightarrow \mathbb{C}$ is considered in terms of its real-valued components, $f(x) = \Re\{f(x)\} + j\Im\{f(x)\}$, and the conjugate coordinate basis $\mathbf{x}^a = [x, x^*]$. Then, using the mapping

$$\begin{bmatrix} x \\ x^* \end{bmatrix} = \begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix} \begin{bmatrix} \Re\{x\} \\ \Im\{x\} \end{bmatrix} \tag{1}$$

Sayed Pouria Talebi is with The Department of Signal Processing and Acoustics, Aalto University, Espoo FI-02150 Finland, E-mail: pouria.talebi@aalto.fi.

Stefan Werner is with The Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim NO-7491 Norway, E-mail: stefan.werner@ntnu.no.

Danilo P. Mandic is with The Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. E-mail: d.mandic@imperial.ac.uk.

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

a relation is established between the derivatives taken in $\mathbb{R}^2$ and those taken directly in $\mathbb{C}$. This relation is subsequently adopted to introduce a framework for calculating derivatives and establishing gradients of complex-valued functions.

The approach of augmenting a complex-valued variable, $\mathbf{x}$, with its complex conjugate in order to form the so-called augmented variable, $\mathbf{x}^a = [\mathbf{x}^\mathsf{T}, \mathbf{x}^\mathsf{H}]^\mathsf{T}$, has also been instrumental in complex-valued statistics, where the relation between $\mathbf{x}^a$ and $\{\Re\{\mathbf{x}\}, \Im\{\mathbf{x}\}\}$ is used to provide a framework for the full statistical description of complex-valued random processes [5]. For instance, the full second-order statistical information of a complex-valued random vector, $\mathbf{x}$, is only describable through its augmented covariance matrix given by

$$\mathbf{C}_{\mathbf{x}^a} = \mathsf{E}\left\{\mathbf{x}^a\mathbf{x}^{a\mathsf{H}}\right\} = \begin{bmatrix}\mathbf{C}_\mathbf{x} & \mathbf{P}_\mathbf{x} \\ \mathbf{P}_\mathbf{x}^\mathsf{H} & \mathbf{C}_\mathbf{x}^\mathsf{H}\end{bmatrix}$$

where $\mathbf{C}_\mathbf{x} = \mathsf{E}\left\{\mathbf{x}\mathbf{x}^\mathsf{H}\right\}$ is the standard covariance, and $\mathbf{P}_\mathbf{x} = \mathsf{E}\left\{\mathbf{x}\mathbf{x}^\mathsf{T}\right\}$ is referred to as the pseudo-covariance [5].

The most fundamental problem in complex-valued filtering and learning paradigms is the adaptive estimation of weighting matrices $\{\mathbf{H}, \mathbf{G}\}$ in the so-called widely-linear model

$$\mathbf{y}_n = \mathbf{H}\mathbf{x}_n + \mathbf{G}\mathbf{x}_n^* \qquad (2)$$

based on the input sequence $\{\mathbf{x}_n : n = 1, 2, \ldots\}$ and observation sequence $\{\mathbf{y}_n : n = 1, 2, \ldots\}$ [5,34]. A number of solutions to this problem have been developed in recent years, among which the augmented complex least mean square (ACLMS) [11,15,35] has found widespread acceptance. However, these solutions are based on the assumption that $\{\mathbf{y}_n, \mathbf{x}_n\}$ are jointly Gaussian signals.

The class of elliptically symmetric $\alpha$-stable signals have attracted a great deal of attention for modeling signals encountered in modern filtering applications [21,26,31,36]. A complex-valued signal is referred to as C-S$\alpha$S if its real and imaginary components are jointly elliptically symmetric $\alpha$-stable [25]. The class of C-S$\alpha$S signals admit characteristic function of the form[1] [25]

$$\mathbf{\Phi}_{\mathbf{x}^a}(\mathbf{s}^a) = \mathsf{E}\left\{e^{\left(j\frac{\mathbf{s}^{a\mathsf{H}}\mathbf{x}^a}{2}\right)}\right\} = e^{\frac{1}{2}j\mathbf{s}^{a\mathsf{H}}\boldsymbol{\zeta}^a}e^{-\left(\frac{1}{8}\mathbf{s}^{a\mathsf{H}}\mathbf{\Sigma}\mathbf{s}^a\right)^{\alpha/2}} \qquad (3)$$

where the positive definite matrix, $\mathbf{\Sigma}$, determines the elliptical shape of the distribution of $\mathbf{x}$ which is centered at $\boldsymbol{\zeta}$. The tail heaviness of the density is determined by the characteristic exponent $\alpha \in (0, 2]$. A small value of $\alpha$ indicates severe impulsiveness, resulting in heavier tails, whereas when $\alpha \to 2$ the distribution exhibits more Gaussian type behaviour. Indeed, for the case of $\alpha = 2$, the distribution is Gaussian. Given that C-S$\alpha$S signals only posses finite statistical moments of orders strictly less than $\alpha$, classical signal processing techniques, based on minimizing second-order moments of an error measure, do not perform well when applied to C-S$\alpha$S signals [19,26,28,31]. Moreover, another consequence of this property is that for filtering purposes, in order to ensure finite conditional expectations exist, hereafter it is assumed that $\alpha \in (1, 2]$.

[1]For more information on characteristic function of complex-valued signals the keen reader is referred to [5,37].

In this work, the adaptive filtering problem formulated in (2) is considered, where the Gaussian assumption on the input/output signal is extended to the generality of C-S$\alpha$S signals. This makes it possible to accommodate for modeling signals encountered in modern filtering applications, e.g. [19]–[24,26,31]. The proposed solution draws upon ideas from our previous work dealing with real-valued signals in [31,32], which was developed using concepts from fractional-order calculus and statistics [38]–[41].

## III. Proposed Filtering Approach

In order to simplify the presentation, the widely-linear model in (2) is rearranged into the augmented formulation

$$\mathbf{y}_n^a = \mathbf{W}\mathbf{x}_n^a \quad \text{with} \quad \mathbf{W} = \begin{bmatrix}\mathbf{H} & \mathbf{G} \\ \mathbf{G}^* & \mathbf{H}^*\end{bmatrix} \qquad (4)$$

where the task at hand becomes that of estimating the weight matrix $\mathbf{W}$. To this end, consider the adaptive filter

$$\hat{\mathbf{y}}_n^a = \hat{\mathbf{W}}_n\mathbf{x}_n^a \qquad (5)$$

where $\hat{\mathbf{W}}_n$ is the estimate of desired weight matrix $\mathbf{W}$ at time instant $n$. The estimates, $\{\hat{\mathbf{W}}_n : n = 1, 2, \ldots\}$ are selected to iteratively minimize the cost function

$$\mathcal{J}_n = \boldsymbol{\epsilon}_n^{a\mathsf{H}} g_\rho(\boldsymbol{\epsilon}_n^a) \quad \text{with} \quad \begin{cases}\boldsymbol{\epsilon}_n^a = \mathbf{y}_n^a - \hat{\mathbf{y}}_n^a \\ \rho \in (0, \alpha - 1)\end{cases} \qquad (6)$$

where $\boldsymbol{\epsilon}_n^a$ denotes the discrepancy between the output signal, $\mathbf{y}_n^a$, and its predicted value, $\hat{\mathbf{y}}_n^a$, while $g_\rho(\cdot)$ is a function, hereafter referred to as the fractional-order operator, that performs the transform

$$\forall z \in \mathbb{C} : z \leftarrow \text{sign}(\Re\{z\})|\Re\{z\}|^\rho + j\text{sign}(\Im\{z\})|\Im\{z\}|^\rho \qquad (7)$$

on each element of its augmented input vector.

*Remark* 1. After some mathematical manipulation, the cost function in (6) can be reformulated in terms of the real and imaginary components of $\boldsymbol{\epsilon}_n^a$ to give

$$\mathcal{J}_n = \|\Re\{\boldsymbol{\epsilon}_n^a\}\|_{\rho+1}^{\rho+1} + \|\Im\{\boldsymbol{\epsilon}_n^a\}\|_{\rho+1}^{\rho+1}.$$

This makes it clear that, the condition $\rho \in (0, \alpha-1)$ guarantees $\mathcal{J}_n$ is convex and $\mathsf{E}\{\mathcal{J}_n\}$ exists.

The weight matrix estimate, $\hat{\mathbf{W}}_n$, is updated at each time instant using the gradient-descent principle so that

$$\hat{\mathbf{W}}_{n+1} = \hat{\mathbf{W}}_n - \mu\nabla^\rho\mathcal{J}_n \qquad (8)$$

where $\mu$ is a positive real-valued adaptation gain and $\nabla^\rho\mathcal{J}_n$ denotes the $\rho$-order gradient of $\mathcal{J}_n$, which is used to form the direction of descent. Akin to the $\mathbb{CR}$-calculus approach, the cost function is considered in terms of the real-valued components of its input. This allows for the use of conventional real-valued techniques for calculating the gradient, which is then mapped onto the complex domain. Thus, considering the framework set in [38,39] and our previous work [31,32], from (8) we have

$$\hat{\mathbf{W}}_{n+1} = \hat{\mathbf{W}}_n + \mu\boldsymbol{\epsilon}_n^a\left(g_\rho(\mathbf{x}_n^a)\right)^\mathsf{H} \qquad (9)$$

where all constant terms have been incorporated into the adaptation gain $\mu$.

## IV. Performance Analysis

From the error term defined in (6) and the adaptive filter formulated in (5) we have

$$\boldsymbol{\epsilon}_n^a = \mathbf{y}_n^a - \hat{\mathbf{y}}_n^a = \mathbf{W}\mathbf{x}_n^a - \hat{\mathbf{W}}_n\mathbf{x}_n^a = \boldsymbol{\Upsilon}_n\mathbf{x}_n^a \tag{10}$$

where

$$\boldsymbol{\Upsilon}_n = \mathbf{W} - \hat{\mathbf{W}}_n \tag{11}$$

denotes the weight matrix estimation error. Subsequently, replacing $\boldsymbol{\epsilon}_n^a$ from (10) into (9) gives

$$\hat{\mathbf{W}}_{n+1} = \hat{\mathbf{W}}_n + \mu\boldsymbol{\Upsilon}_n\mathbf{x}_n^a\left(g_\rho(\mathbf{x}_n^a)\right)^{\mathsf{H}} \tag{12}$$

which using (11) yields

$$\begin{aligned}\boldsymbol{\Upsilon}_{n+1} =& \boldsymbol{\Upsilon}_n - \mu\boldsymbol{\Upsilon}_n\mathbf{x}_n^a\left(g_\rho(\mathbf{x}_n^a)\right)^{\mathsf{H}} \\ =& \boldsymbol{\Upsilon}_n\left(\mathbf{I} - \mu\mathbf{x}_n^a\left(g_\rho(\mathbf{x}_n^a)\right)^{\mathsf{H}}\right) = \boldsymbol{\Upsilon}_n\left(\mathbf{I} - \mu\boldsymbol{\Psi}_n\right)\end{aligned} \tag{13}$$

where $\boldsymbol{\Psi}_n = \mathbf{x}_n^a\left(g_\rho(\mathbf{x}_n^a)\right)^{\mathsf{H}}$.

The expression in (13) demonstrates the transformation of weight matrix estimation error from one time instant to the next and can be reformulated to give

$$\mathsf{vec}\left\{\boldsymbol{\Upsilon}_{n+1}^{\mathsf{T}}\right\} = \mathbf{T}_n\mathsf{vec}\left\{\boldsymbol{\Upsilon}_n^{\mathsf{T}}\right\} \tag{14}$$

where $\mathbf{T}_n$ is block diagonal matrix of appropriate size with identical block diagonal elements, so that

$$\mathbf{T}_n = \text{block-diag}\{\mathbf{I} - \mu\boldsymbol{\Psi}_n^{\mathsf{T}}\}.$$

From (14), it becomes clear that for $\beta \in (1, \alpha)$, $\|\mathsf{vec}\left\{\boldsymbol{\Upsilon}_n\right\}\|_\beta^\beta$ is convergent when

$$\forall n:\ 0 < \mu < \frac{1}{\lambda_{max}(\boldsymbol{\Psi}_n)} \tag{15}$$

where $\lambda_{max}(\boldsymbol{\Psi}_n)$ is the largest eigenvalue of $\boldsymbol{\Psi}_n$.

*Remark* 2. Following the process in (10)-(15) and assuming that $\{\mathbf{x}_n^a : n = 1, 2, \ldots\}$ is a zero-mean, stationary, and temporally independent sequence, it follows that $\mathsf{E}\left\{\|\mathsf{vec}\left\{\boldsymbol{\Upsilon}_n\right\}\|_\beta^\beta\right\}$ converges if $0 < \mu < 1/\lambda_{max}(\mathsf{E}\left\{\boldsymbol{\Psi}_n\right\})$. After some mathematical manipulation it can be shown that

$$\begin{aligned}\boldsymbol{\Psi}_n^{\{l,k\}} =& \mathbf{x}_n^{a\{l\}}\text{sign}(\Re\{\mathbf{x}_n^{a\{k\}}\})|\Re\{\mathbf{x}_n^{a\{k\}}\}|_\rho^\rho \\ &- j\mathbf{x}^{a\{l\}}\text{sign}(\Im\{\mathbf{x}_n^{a\{k\}}\})|\Im\{\mathbf{x}_n^{a\{k\}}\}|_\rho^\rho\end{aligned} \tag{16}$$

where $\boldsymbol{\Psi}_n^{\{l,k\}}$ denotes the element on the $l^{\text{th}}$ row and $k^{\text{th}}$ column of $\boldsymbol{\Psi}_n$, with $\mathbf{x}_n^{a\{k\}}$ representing the $k^{\text{th}}$ element of $\mathbf{x}_n^a$. From (16) it is clear that the condition $\rho \in (0, \alpha - 1)$ set in (6) guarantees that $\mathsf{E}\left\{\boldsymbol{\Psi}_n\right\}$ exists.

*Remark* 3. For the Gaussian case, where $\alpha = 2$, as $\rho \to 1$, the proposed filtering technique simplifies into the conventional ACLMS in [11,35]. The same statement follows for the derived convergence criteria.

## V. Mixed Fractional-Order Operator Approach

Assume, without loss of generality, that $\hat{\mathbf{W}}_1 = \mathbf{0}$. Then, from the expression in (9) we have

$$\hat{\mathbf{W}}_{n+1} = \mu\sum_{l=1}^{n}\boldsymbol{\epsilon}_l^a\left(g_\rho(\mathbf{x}_l^a)\right)^{\mathsf{H}}. \tag{17}$$

The expression in (17) allows for the proposed approach to be generalized to the case that admits $m$ different fractional-order operators to be used. In this setting, following the approach introduced in Section III, we have

$$\hat{\mathbf{W}}_{n+1} = \mu\sum_{l=1}^{n}\sum_{k=1}^{m}\boldsymbol{\epsilon}_l^a\left(g_{\rho_k}(\mathbf{x}_l^a)\right)^{\mathsf{H}} \tag{18}$$

where $\forall k:\ \rho_k \in (0, \alpha - 1)$.

From (18), the evolution of the weight matrix estimates can be formulated as

$$\begin{aligned}\hat{\mathbf{W}}_{n+1} - \hat{\mathbf{W}}_n =& \mu\boldsymbol{\epsilon}_n^a\sum_{k=1}^{m}\left(g_{\rho_k}(\mathbf{x}_n^a)\right)^{\mathsf{H}} \\ =& \mu\boldsymbol{\Upsilon}_n\mathbf{x}_n^a\sum_{k=1}^{m}\left(g_{\rho_k}(\mathbf{x}_n^a)\right)^{\mathsf{H}}.\end{aligned} \tag{19}$$

Subsequently, a substitution of (11) into (19) yields

$$\hat{\mathbf{W}}_{n+1} - \hat{\mathbf{W}}_n = \mu\left(\mathbf{W} - \hat{\mathbf{W}}_n\right)\mathbf{x}_n^a\sum_{k=1}^{m}\left(g_{\rho_k}(\mathbf{x}_n^a)\right)^{\mathsf{H}}$$

which can be rearranged to give

$$\begin{aligned}\hat{\mathbf{W}}_{n+1} =& \mu\mathbf{W}\mathbf{x}_n^a\sum_{k=1}^{m}\left(g_{\rho_k}(\mathbf{x}_n^a)\right)^{\mathsf{H}} \\ &+ \hat{\mathbf{W}}_n\left(\mathbf{I} - \mu\mathbf{x}_n^a\sum_{k=1}^{m}\left(g_{\rho_k}(\mathbf{x}_n^a)\right)^{\mathsf{H}}\right).\end{aligned} \tag{20}$$

Applying the statistical expectation of (20) and assuming that $\{\mathbf{x}_n^a : n = 1, 2, \ldots\}$ is a zero-mean, stationary, and temporally independent sequence; then, for $n > 1$, we have

$$\mathsf{E}\left\{\hat{\mathbf{W}}_{n+1}\right\} = \mu\mathbf{W}\boldsymbol{\Gamma}\sum_{l=1}^{n+1}\left(\mathbf{I} - \mu\boldsymbol{\Gamma}\right)^{l-1} \tag{21}$$

where $\boldsymbol{\Gamma} = \mathsf{E}\left\{\mathbf{x}_n^a\sum_{k=1}^{m}\left(g_{\rho_k}(\mathbf{x}_n^a)\right)^{\mathsf{H}}\right\}$. From (21) it is clear now that $\mathsf{E}\left\{\hat{\mathbf{W}}_{n+1}\right\} \to \mathbf{W}$ as $n \to \infty$ conditional on

$$0 < \mu < \frac{1}{\lambda_{max}(\boldsymbol{\Gamma})}.$$

*Remark* 4. In so-called *mixed norm approaches*, different norms of an error measure are simultaneously minimized; this, is typically achieved based on their first-order gradient. However, this still does not describe the introduced mixed fractional-order approach. The weight matrix estimate in (18) represents the summation of $m$ different fractional gradient operators, which minimize their corresponding fractional-order cost functions through the framework derived in Section III.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LSP.2019.2929874



## VI. Performance Evaluation

In the first set of simulations, the adaptive filtering problem defined in (4)-(5) was considered, where

$$\mathbf{H} = \begin{bmatrix}1.4 + j2.1 & -0.52 - j0.52\end{bmatrix}$$
$$\mathbf{G} = \begin{bmatrix}-1.5 + j1.21 & 2 + j2.3\end{bmatrix}$$

while the input sequence $\{\mathbf{x}_n^a : n = 1, 2, \ldots\}$ was a zero-mean temporally independent C-S$\alpha$S sequence with $\alpha = 1.6$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 20 & 0 & j16 & 0 \\ 0 & 20 & 0 & j8 \\ -j16 & 0 & 20 & 0 \\ 0 & -j8 & 0 & 20 \end{bmatrix}.$$

The adaptation gain was $\mu = 4 \times 10^{-3}$ for this simulation.

Two metrics were used to measure performance of the proposed filtering techniques, the mean absolute error (MAE) given by $\mathsf{E}\{|\boldsymbol{\epsilon}_n^a|\}$, and the mean absolute deviation (MAD) given by $\mathsf{E}\{|\boldsymbol{\Upsilon}_n|\}$. The mean-values were calculated empirically via averaging of results obtained from $10^3$ independent realizations of the experiment. Performances of the derived adaptive filtering techniques are compared to that of the traditional ACLMS in Figure 1. Observe that, in contrast to the ACLMS, the class of derived filters converged. Moreover, a decrease in the parameter $\rho$ reduced jitters (sharp spikes) in the MAD and MAE behaviour, while at the same time decreasing convergence rates.[2] Figure 1 also includes performance of the mixed fractional-order operator with operators $\rho_1 = 1/2$, $\rho_2 = 1/8$, and $\rho_3 = 1/16$.

In the second set of simulations, a nonlinear complex-valued Ikeda map signal was considered. The signal was generated through the nonlinear recursive relation

$$\Re\{x_{n+1}\} = 1 + 0.65\Big(\Im\{x_n\}\cos(u_n) - \Im\{x_n\}\sin(u_n)\Big)$$
$$\Im\{x_{n+1}\} = 0.65\Big(\Re\{x_n\}\sin(u_n) + \Im\{x_n\}\cos(u_n)\Big)$$

where $u_n = 0.4 - 6(1 + |x_n|^2)^{-1}$.

The past four samples of the Ikeda signal were used as input to the widely-linear model in (2) in order to predict the upcoming output. The ACLMS and the proposed filtering technique, in the mixed fractional-order operator formulation, were implemented to estimate the weighting matrix. The fractional-order operators used for this simulation were $\rho_1 = 1$, $\rho_2 = 1/2$, and $\rho_3 = 1/4$. Thus, in essence, the gradient of the ACLMS (that is for $\rho_1 = 1$) is combined with two of its fractional-order counterparts (that is $\rho_2 = 1/2$ and $\rho_3 = 1/4$). The adaptation gain, $\mu$, was chosen so that both methods achieved a similar steady-state MAE. The MAE performance of the ACLMS and the proposed filtering framework, calculated empirically via averaging of results obtained from $10^3$ independent realizations of the experiment, are shown in Figure 2. Although both methods were able to track the signal, the proposed filter outperformed the ACLMS in terms of convergence rate.

[2]This indicates a trade-off in selection of $\rho$. However, a precise formulation of the effect of $\rho$ on the filtering performance remains an open problem.

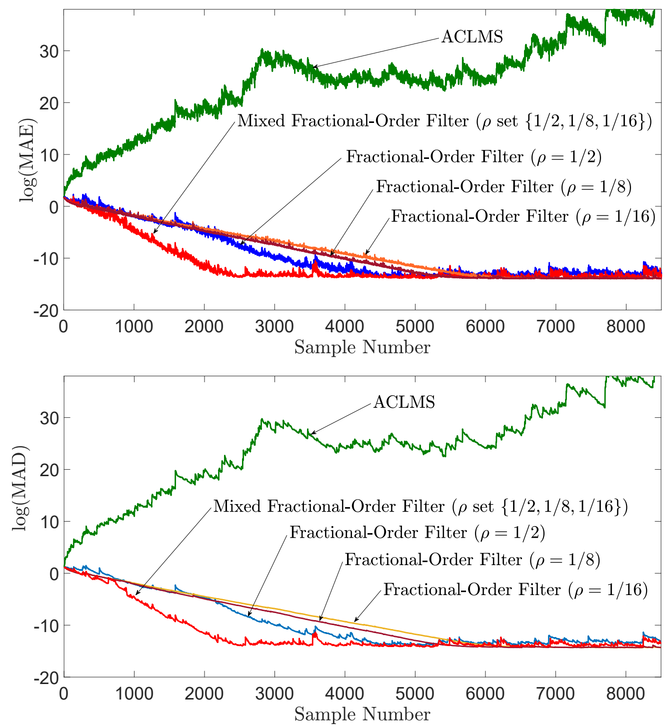


Figure 1. The MAE (top) and MAD (bottom) performance of the proposed filtering technique. Performance of the proposed filtering technique is denoted as *Fractional-Order Filter*, while the mixed fractional-order approach is designated as *Mixed Fractional-Order Filter*. Performance of the standard ACLMS is provided for comparison.

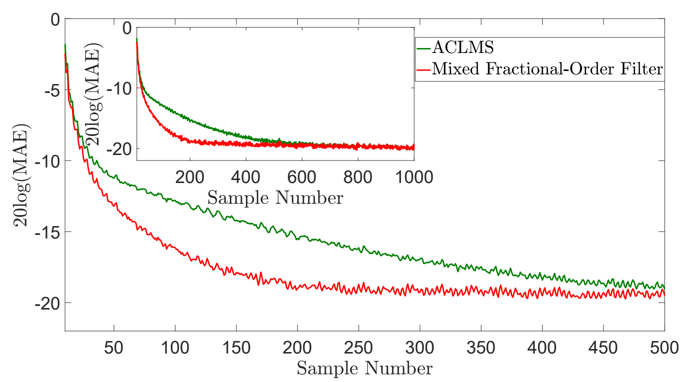


Figure 2. Learning curves for the Ikeda map. The MAE performance of the ACLMS and proposed filtering technique, designated as *Mixed Fractional-Order Filter*, are shown.

## VII. Conclusion

A class of nonlinear adaptive filtering techniques for processing C-S$\alpha$S signals has been derived. The proposed adaptive filter is based on the principle of gradient-descent and augmented statistical approach to processing complex-valued signals. In essence, the derived approach represents an extension of the ACLMS, where the first-order gradient is generalized to fractional-order gradients in order to accommodate C-S$\alpha$S signals with $\alpha \in (1, 2]$. Moreover, convergence criteria for the derived filtering techniques have been established and the introduced concepts have been verified using simulations. The obtained results also indicate that when dealing with systems exhibiting chaotic behaviour, the performance of the ACLMS adaptive filter can benefit if the adaptation step is mixed with the derived fractional-order filtering operators.

## REFERENCES


[1] R. F. Stengel, *Optimal Estimation and Control*. Dover Publications, 1994.

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

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

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

[5] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear, and Neural Models*. New York: Wiley, 2009.

[6] E. Clarke, *Circuit Analysis of A.C. Power Systems*. New York: Wiley, 1943.

[7] S. P. Talebi, S. Kanna, and D. P. Mandic, "A non-linear state space frequency estimator for three-phase power systems," *In Proceedings of International Joint Conference on Neural Networks*, pp. 1–7, July 2015.

[8] R. Arablouei, S. Werner, and K. Doğançay, "Estimating frequency of three-phase power systems via widely-linear modeling and total least-squares," *In Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pp. 464–467, December 2013.

[9] J. Sun, E. Aboutanios, and D. B. Smith, "Iterative weighted least squares frequency estimation for harmonic sinusoidal signal in power systems," *In Proceedings of European Signal Processing Conference*, pp. 176–180, September 2018.

[10] S. L. Goh, M. Chen, D. H. Popovic, K. Aihara, D. Obradovic, and D. P. Mandic, "Complex-valued forecasting of wind profile," *Renewable Energy*, vol. 31, no. 11, pp. 1733–1750, 2006.

[11] D. Mandic, S. Javidi, S. Goh, A. Kuh, and K. Aihara, "Complex-valued prediction of wind profile using augmented complex statistics," *Renewable Energy*, vol. 34, no. 1, pp. 196–201, 2009.

[12] P. Clark, I. Kirsteins, and L. Atlas, "Multiband analysis for colored amplitude-modulated ship noise," *In Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2010.

[13] A. Hirose, "Recent progress in applications of complex-valued neural networks," *In Proceedings of International Conference on Artificial Intelligence and Soft Computing*, pp. 42–46, 2010.

[14] J. Hu and J. Wang, "Global stability of complex-valued recurrent neural networks with time-delays," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 6, pp. 853–865, June 2012.

[15] S. L. Goh and D. P. Mandic, "An augmented CRTRL for complex-valued recurrent neural networks," *Neural Networks*, vol. 20, no. 10, pp. 1061–1066, 2007.

[16] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101–5125, November 2011.

[17] P. Chevalier and A. Blin, "Widely linear MVDR beamformers for the reception of an unknown signal corrupted by noncircular interferences," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5323–5336, 2007.

[18] S. Zarei, W. H. Gerstacker, J. Aulin, and R. Schober, "I/Q imbalance aware widely-linear receiver for uplink multi-cell massive MIMO systems: Design and sum rate analysis," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3393–3408, May 2016.

[19] 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.

[20] 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.

[21] S. Leglaive, U. Şimşekli, 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.

[22] 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.

[23] N. Azzaoui and L. Clavier, "Statistical channel model based on $\alpha$-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.

[24] S. Leglaive, U. Şimşekli, A. Liutkus, L. Girin, and R. Horaud, "Speech enhancement with variational autoencoders and alpha-stable distributions," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 541–545, May 2019.

[25] G. Samorodnitsky and M. S. Taqqu, *Stable non-Gaussian random processes*. Chapman & Hall, 1994.

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

[27] 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.

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

[29] 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.

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

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

[32] S. P. Talebi, S. Werner, S. Li, and D. P. Mandic, "Tracking dynamic systems in $\alpha$-stable environments," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4853–4857, May 2019.

[33] K. K. . Delgado, *The Complex Gradient Operator and the CR-Calculus*. University of California, San Diego, Technical Report, 2005.

[34] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, August 1995.

[35] S. Javidi, M. Pedzisz, S. L. Goh, and D. P. Mandic, "The augmented complex least mean square algorithm with application to adaptive prediction problems," *In Proceedings of IARP Workshop on Cognitive Information Processing*, pp. 54–57, 2008.

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

[37] P. O. Amblard, M. Gaeta, and J. L. Lacoume, "Statistics for complex variables and signals - Part I: Variables," *Signal Processing*, vol. 53, no. 1, pp. 1–13, 1996.

[38] 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.

[39] ——, "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.

[40] D. Cafagna, "Fractional calculus: A mathematical tool from the past for present engineers," *IEEE Industrial Electronics Magazine*, vol. 1, no. 2, pp. 35–40, Summer 2007.

[41] 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.