Identification of improper processes by variable tap-length complex-valued adaptive filters

Bukhari Che Ujang, Clive Cheong Took and Danilo P. Mandic

***Abstract*—Variable tap-length is introduced into complex-valued adaptive filters in order to provide an additional degree of freedom, enhance tracking ability, and provide data-adaptive optimal modelling. This is achieved by extending the fractional tap-length (FT) algorithm from the real domain $\mathbb{R}$ and by accounting for some special properties of the complex domain $\mathbb{C}$. For generality, the augmented least mean square (ACLMS) and augmented complex nonlinear gradient descent(ACNGD) are equipped with the variable tap-length in order to cater for both the second order circular and noncircular signals. Simulations on model order selection and the identification of the noncircular nature of complex data support the approach.**

***Index Terms*—Widely linear modelling, complex circularity, fractional tap-length, augmented complex least mean square (ACLMS)**

I. Introduction

For optimal stochastic filtering of complex-valued signals, we need to consider three important factors: 1) the nature of the signal (circular or noncircular); 2) the signal generation mechanism (linear or nonlinear); 3) the order of the signal model. Although much work has been devoted to address the first two issues [1]–[3], model order selection in $\mathbb{C}$ is still an open problem. In $\mathbb{C}$ we also have more degrees of freedom, for instance, we differentiate between the standard and widely linear models and circular and noncircular signals.

One convenient way to perform the identification of time-varying parameters of a complex-valued system is to employ adaptive complex-valued filters with variable tap-length and optimise for both the filter length and filter coefficients. Amongst many such adaptive tap length algorithms [4], the fractional tap-length (FT) is considered in this work, due to its simplicity and robustness [5]. However, the FT algorithm was designed specifically for real-valued filters and therefore is not readily suited for processing complex-valued signals [5]. To this end, we extend the FT algorithm to the complex domain; this is achieved by considering specific features of $\mathbb{C}$, such as noncircularity of probability distributions. In this work, the fractional tap-length (FT) algorithm [5] will be incorporated within both linear and nonlinear adaptive complex-valued adaptive filters.

Bukhari Che-Ujang is with the Electrical and Electronic Engineering Department, Imperial College London, SW7 2BT, UK (email: che.che-ujang07@ic.ac.uk)

Clive Cheong-Took is with the Electrical and Electronic Engineering Department, Imperial College London, SW7 2BT, UK (email: c.cheong-took@ic.ac.uk)

Danilo Mandic is with the Electrical and Electronic Engineering Department, Imperial College London, SW7 2BT, UK (email: d.mandic@ic.ac.uk)

The aim of this paper is to: 1) introduce the variable tap-length into widely linear complex-valued adaptive filters; 2) provide a rigorous steady-state analysis to achieve optimal performance; 3) investigate the convergence properties of complex-valued adaptive algorithms for the identification of complex models (linear and widely linear); 4) identify the second order circular (proper) and second order noncircular (improper) complex processes.

We consider four algorithms, the standard complex LMS (CLMS) [6], the augmented complex LMS (ACLMS) [1], the complex nonlinear gradient descent (CNGD) [1] and the newly introduced augmented complex nonlinear gradient descent (ACNGD) [7]. The ACLMS is based on a widely linear stochastic moving average (MA) model, given by [1]

$$y(k) = \underbrace{\mathbf{x}^T(k)\mathbf{h}}_{\text{standard part}} + \underbrace{\mathbf{x}^H(k)\mathbf{g}}_{\text{augmented part}} + v(k) \tag{1}$$

where $\mathbf{x}(k)$ denotes the regressor vector, $\mathbf{h}$ and $\mathbf{g}$ are the coefficient vectors of the standard and 'augmented' part of the model, and $v(k)$ is circular white Gaussian noise. The symbols $(\cdot)^H$ and $(\cdot)^T$ denote respectively the Hermitian and vector transpose operator. The widely linear model (1) was also extended into the quaternion domain in [8].

The power of ACLMS and ACNGD stems from the so called augmented complex statistics, where for a centered complex random vector (RV) $\mathbf{x} \in \mathbb{C}^L$, the covariance $\mathcal{C}_{\mathbf{xx}}$ and pseudocovariance $\mathcal{P}_{\mathbf{xx}}$ matrices are defined as [9]

$$\mathcal{C}_{\mathbf{xx}} = E(\mathbf{x}\mathbf{x}^H) \qquad \mathcal{P}_{\mathbf{xx}} = E(\mathbf{x}\mathbf{x}^T) \tag{2}$$

A signal is called circular if it has a rotation invariant probability distribution. The distinguishing property of a circular signal is that its pseudocovariance vanishes, that is, $\mathcal{P}_{\mathbf{xx}} = \mathbf{0}$ [9]. Observe that the widely linear model (1) enables the ACLMS algorithm to operate on both the covariance $E(\mathbf{x}\mathbf{x}^H)$ and pseudocovariance $E(\mathbf{x}\mathbf{x}^T)$; making it suitable for the modelling of both second order circular (proper) and noncircular (improper) data, in contrast to CLMS and CNGD.

The paper is organized as follows. Section 2 presents an overview of $\mathbb{CR}$ calculus. The following section introduces the FT-ACLMS, FT-CLMS, FT-ACNGD and FT-CNGD fractional tap-length algorithms. This is followed by a steady-state analysis of all the algorithms considered. Section 4 provides performance comparisons between the four FT algorithms, through comprehensive simulations on Autoregressive (AR) and Nonlinear Autoregressive (NAR) systems.

978-1-4244-8126-2/10/$26.00 ©2010 IEEE

II. Overview of $\mathbb{CR}$ calculus

In order to extend the standard complex algorithms into their augmented complex (widely linear) counterparts, $\mathbb{CR}$ calculus can be applied in order to simplify the derivations [1]. When dealing with complex valued functions, it is required that the Cauchy-Riemann equation are satisfied (for the function to be analytic) in order to calculate the gradient.[1] The standard adaptive filtering cost function (error power) is not analytic and thus standard calculus in $\mathbb{C}$ is not adequate to evaluate its derivative. However, by using the $\mathbb{CR}$ calculus, it is possible to evaluate the derivative of the cost function directly in $\mathbb{C}$ [1], [10].

In this context, $\mathcal{J}(\mathbf{x}) : \mathbb{C}^N \mapsto \mathbb{R}$ can be rewritten as a function of complex vectors $\mathbf{x}$ and $\mathbf{x}^*$ such that $\mathcal{J}(\mathbf{x}, \mathbf{x}^*) : \mathbb{C}^N \mapsto \mathbb{R}$ where $\mathbf{x}$ and $\mathbf{x}^*$ are termed conjugate coordinates. Expanding the complex vectors $\mathbf{x}$ and $\mathbf{x}^*$ in terms of their real and imaginary components, $\mathbf{x_r}$ and $\mathbf{x_i}$, gives $\mathcal{J}(\mathbf{x_r}, \mathbf{x_i}) : \mathbb{R}^N \times \mathbb{R}^N \mapsto \mathbb{R}$. By using the duality between total differentials in $\mathbb{R}^2$ and $\mathbb{C}$, the $\mathbb{CR}$ derivatives are taken separately with respect to $\mathbf{x}$ and $\mathbf{x}^*$ while keeping the other variable constant, resulting in

$$\mathbb{R}-\text{derivative} \;:\; \frac{\partial \mathcal{J}}{\partial \mathbf{x}}\Big|_{\mathbf{x}^*=\text{const}} = \frac{1}{2}\left(\frac{\partial \mathcal{J}}{\partial \mathbf{x_r}} - \jmath \frac{\partial \mathcal{J}}{\partial \mathbf{x_i}}\right) \tag{4}$$

$$\mathbb{R}^*-\text{derivative} \;:\; \frac{\partial \mathcal{J}}{\partial \mathbf{x}^*}\Big|_{\mathbf{x}=\text{const}} = \frac{1}{2}\left(\frac{\partial \mathcal{J}}{\partial \mathbf{x_r}} + \jmath \frac{\partial \mathcal{J}}{\partial \mathbf{x_i}}\right) \tag{5}$$

It can then be shown that the direction of the steepest descent is given by the derivative with respect to $\mathbf{x}^*$ [10]. If $\mathbb{CR}$ calculus is applied to analytic functions, the $\mathbb{R}^*$ – derivative vanishes and we are only left with the $\mathbb{R}$ – derivative [10].

III. Model Order Identification

The proposed algorithms comprise of two parts: the finite impulse response (FIR) filter weights update which optimises the adaptive weight coefficients, followed by the FT algorithm that adapts the tap-length of the filter to an optimal value. We first review the existing approaches and then illustrate how the FT algorithm can be exploited within complex-valued adaptive systems.

A. Filter Weight Update Algorithms

The weight update of the standard CLMS algorithm is given by [6]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{x}^*(k) \tag{6}$$

where $\mathbf{w}(k)$ is the weight vector of the filter, $\mathbf{x}(k)$ is the filter input, $(\cdot)^*$ denotes the complex conjugate operator and μ is a real-valued learning rate.

[1]Given a complex function $f(x) = u(x_r, x_i) + v(x_r, x_i)\jmath$, the Cauchy-Riemann conditions are given by

$$\frac{\partial u}{\partial x_r} = \frac{\partial v}{\partial x_i}; \quad \frac{\partial v}{\partial x_r} = -\frac{\partial u}{\partial x_i} \tag{3}$$

The standard output of the nonlinear FIR filter trained by the complex nonlinear gradient descent (CNGD) is given by

$$y(k) = \Phi\big(\mathbf{w}^T(k)\mathbf{x}(k)\big) \tag{7}$$

where $\Phi(\cdot)$ is a nonlinear function chosen from a class of fully complex nonlinear functions [11]. The functions considered are analytic thus ensuring that the derivative of the nonlinear function exists.

Similar to the CLMS algorithm, the weight update of the CNGD algorithm is given by [1]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\Phi'^*\big(\mathbf{w}^T(k)\mathbf{x}(k)\big)\mathbf{x}^*(k) \tag{8}$$

where $\Phi'(\cdot)$ is the derivative of the nonlinear activation function.

In order to derive the ACLMS, based on the widely linear model in (1), we first rewrite the error signal as

$$e(k) = d(k) - y(k) = d(k) - \mathbf{h}^T(k)\mathbf{x}(k) - \mathbf{g}^T(k)\mathbf{x}^*(k) \tag{9}$$

The corresponding cost function $E(k) = e(k)e^*(k)$ is then minimised using a steepest descent adaptation given by

$$\mathbf{h}(k+1) = \mathbf{h}(k) - \mu\nabla_{\mathbf{h}^*}E(k) \tag{10}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) - \mu\nabla_{\mathbf{g}^*}E(k) \tag{11}$$

Recall from the $\mathbb{CR}$ calculus that the direction of steepest descent is given by $\mathbb{R}^*$–derivative, for both update equations. This yields the ACLMS weight update in the form [1]

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e(k)\mathbf{x}^*(k) \tag{12}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \mu e(k)\mathbf{x}(k) \tag{13}$$

where $\mathbf{g}(k)$ and $\mathbf{h}(k)$ are the weight vectors of the filter.

Proceeding in the same manner as with ACLMS, the ACNGD algorithm weight updates are given by [7]

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e(k)\Phi'^*\big(net(k)\big)\mathbf{x}^*(k) \tag{14}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \mu e(k)\Phi'^*\big(net(k)\big)\mathbf{x}(k) \tag{15}$$

where $net(k) = \mathbf{h}^T(k)\mathbf{x}(k) + \mathbf{g}^T(k)\mathbf{x}^*(k)$.

B. Fractional Tap Length Algorithm

The FT tap-length adaptation for complex-valued filters is governed by [5]

$$\eta_f(k+1) = (\eta_f(k) - \alpha) - \gamma\cdot\left[\left(E_N^{(N)}(k)\right) - \left(E_{N-\Delta}^{(N)}(k)\right)\right] \tag{16}$$

where η_f is the pseudo fractional tap-length which can take only positive real value, α and γ are the leaky factor and tap-length learning rate, which are small positive real values that satisfy $\alpha \ll \gamma$. Symbols $E_N^{(N)}(k)$ and $E_{N-\Delta}^{(N)}(k)$ denote respectively the instantaneous square errors for the tap-lengths of N and $N-\Delta$, symbol $N(k)$ denotes the "true" tap-length at discrete time instant 'k', and Δ is a real positive integer such that $\min\{N(k) - \Delta\} > 0$.

The instantaneous square output errors for filters of lengths N and $N - \Delta$ are given by

$$E_N^{(N)}(k) = \big(e_N^{(N)}(k)\big)\big(e_N^{(N)}(k)\big)^* \tag{17}$$

$$E_{N-\Delta}^{(N)}(k) = \big(e_{N-\Delta}^{(N)}(k)\big)\big(e_{N-\Delta}^{(N)}(k)\big)^* \tag{18}$$

based on the errors $e_N^{(N)}(k)$ and $e_{N-\Delta}^{(N)}(k)$ given by

$$e_M^{(N)}(k) = d(k) - y_M^{(N)}(k) \tag{19}$$

where $1 \leq M \leq N$, and $\mathbf{w}_M^{(N)}(k)$ and $\mathbf{x}_M^{(N)}(k)$ are vectors consisting of the first M coefficients of $\mathbf{w}^{(N)}(k)$ and $\mathbf{x}^{(N)}(k)$ respectively.
To calculate the optimal filter length, which also reflects the complexity of the system that generates the data, the tap-length parameter $N(k)$ is made adaptive according to [5]

$$N(k+1) = \begin{cases} \lfloor \eta_f(k) \rfloor, & |N(k) - \eta_f(k)| \geq \delta \\ N(k), & \text{otherwise} \end{cases} \tag{20}$$

where δ is a predefined integer threshold and $\lfloor \cdot \rfloor$ denotes the floor operator. This way, the true tap-length is robust to noise and remains unchanged until the fractional tap-length accumulates to the predefined integer threshold δ. The minimum value for $N(k)$ is defined as $\Delta + \delta$, to ensure that the lowest possible term $E_{N-\Delta}^{(N)}(k)$ in (16) is $E_1^{(N)}(k)$.

IV. Steady-State Analysis of FT Algorithms

In this section, we first provide a rigorous steady-state analysis of the FT algorithm in the context of the ACLMS and illustrate how this analysis also applies to the FT-CLMS, FT-ACNGD and FT-CNGD algorithms. We shall define the desired (teaching) signal $d(k)$ as

$$d(k) = \mathbf{x}_{Lopt}^T(k)\mathbf{h}_{Lopt}^o + \mathbf{x}_{Lopt}^H(k)\mathbf{g}_{Lopt}^o + v(k) \tag{21}$$

where $\mathbf{h}_{Lopt}^o$ and $\mathbf{g}_{Lopt}^o$ are the optimal weights coefficients of the optimal tap length.
Based on the widely linear FIR model (1), the ACLMS algorithm [1] makes the coefficient vectors adaptive, giving

$$y(k) = \underbrace{\mathbf{x}^T(k)\mathbf{h}(k)}_{\text{standard part [CLMS]}} + \underbrace{\mathbf{x}^H(k)\mathbf{g}(k)}_{\text{augmented part}} \tag{22}$$

Upon replacing (21) and (22) into (19), we can express the output error in terms of the optimal tap weights as

$$\begin{aligned} e_M^{(N)}(k) &= \mathbf{x}^T(k)\mathbf{h}_{Lopt}^o + \mathbf{x}^H(k)\mathbf{g}_{Lopt}^o \\ &\quad + v(k) - \mathbf{x}^T(k)\mathbf{h}(k) - \mathbf{x}^H(k)\mathbf{g}(k) \end{aligned} \tag{23}$$

A. Steady-State Analysis of the FT-CLMS and FT-ACLMS Algorithms

Proceeding in a manner similar to the analysis in [12], the optimal coefficients of the standard and conjugate part of the augmented weight vectors $\mathbf{h}_{Lopt}^o$ and $\mathbf{g}_{Lopt}^o$ can be split into three parts

$$\mathbf{h}_{Lopt}^o = \begin{bmatrix} \mathbf{h}'^o(k) \\ \mathbf{h}''^o(k) \\ \mathbf{h}'''^o(k) \end{bmatrix} \quad \mathbf{g}_{Lopt}^o = \begin{bmatrix} \mathbf{g}'^o(k) \\ \mathbf{g}''^o(k) \\ \mathbf{g}'''^o(k) \end{bmatrix} \tag{24}$$

where $\mathbf{h}'^o(k)$, $\mathbf{g}'^o(k)$ are the coefficients modelled by tap-length $1{:}N-\Delta$, $\mathbf{h}''^o(k)$, $\mathbf{g}''^o(k)$ are the coefficients modelled by the tap-length $N - \Delta + 1 : N$, and $\mathbf{h}'''^o(k)$, $\mathbf{g}'''^o(k)$ are the undermodelled coefficients.
The coefficient error vectors of ACLMS are denoted as

$$\begin{aligned} \tilde{\mathbf{h}}(k) &= \mathbf{h}_N^o - \mathbf{h}_N(k) \end{aligned} \tag{25}$$

$$\begin{aligned} \tilde{\mathbf{g}}(k) &= \mathbf{g}_N^o - \mathbf{g}_N(k) \end{aligned} \tag{26}$$

where $\mathbf{h}_N(k)$ and $\mathbf{g}_N(k)$ are the weight vectors of length N. Then, the weight error vectors $\tilde{\mathbf{g}}(k)$ and $\tilde{\mathbf{h}}(k)$ can also be split up into three parts

$$\tilde{\mathbf{h}}(k) = \begin{bmatrix} \tilde{\mathbf{h}}'(k) \\ \tilde{\mathbf{h}}''(k) \\ \tilde{\mathbf{h}}'''(k) \end{bmatrix} \quad \tilde{\mathbf{g}}(k) = \begin{bmatrix} \tilde{\mathbf{g}}'(k) \\ \tilde{\mathbf{g}}''(k) \\ \tilde{\mathbf{g}}'''(k) \end{bmatrix} \tag{27}$$

Substitute (24) and (27) into (23) to obtain the errors $e_N^{(N)}(k)$ and $e_{N-\Delta}^{(N)}(k)$ defined in (19) as (the time index 'k' has been dropped due to space limitations)

$$e_N^{(N)} = \begin{bmatrix} \mathbf{x}' \\ \mathbf{x}'' \\ \mathbf{x}''' \end{bmatrix}^T \begin{bmatrix} \tilde{\mathbf{h}}' \\ \tilde{\mathbf{h}}'' \\ \mathbf{h}'''^o \end{bmatrix} + \begin{bmatrix} \mathbf{x}' \\ \mathbf{x}'' \\ \mathbf{x}''' \end{bmatrix}^H \begin{bmatrix} \tilde{\mathbf{g}}' \\ \tilde{\mathbf{g}}'' \\ \mathbf{g}'''^o \end{bmatrix} + v \tag{28}$$

$$e_{N-\Delta}^{(N)} = \begin{bmatrix} \mathbf{x}' \\ \mathbf{x}'' \\ \mathbf{x}''' \end{bmatrix}^T \begin{bmatrix} \tilde{\mathbf{h}}' \\ \mathbf{h}''^o \\ \mathbf{h}'''^o \end{bmatrix} + \begin{bmatrix} \mathbf{x}' \\ \mathbf{x}'' \\ \mathbf{x}''' \end{bmatrix}^H \begin{bmatrix} \tilde{\mathbf{g}}' \\ \mathbf{g}''^o \\ \mathbf{g}'''^o \end{bmatrix} + v \tag{29}$$

To ensure mathematical tractability of the steady-state analysis, we shall make the following standard assumptions [12]:

- Both the input signal $\mathbf{x}(k)$ and the noise $v(k)$ are i.i.d. zero mean white jointly Gaussian with the respective variances σ_x^2 and σ_v^2;
- At the steady state, the input signal $\mathbf{x}(k)$ is independent of both the weight vectors $\mathbf{h}(k)$ and $\mathbf{g}(k)$;
- The tap-length parameter has converged at steady-state, hence $E\{\eta_f(k+1)\} = E\{\eta_f(k)\}$, leading to $\|\mathbf{h}'''^o(k)\|_2^2 = 0$ and $\|\mathbf{g}'''^o(k)\|_2^2 = 0$.

The MSE at the steady-state is obtained by applying the statistical expectation operator to (16) to give

$$E\left\{\left(E_N^{(N)}(k)\right) - \left(E_{N-\Delta}^{(N)}(k)\right)\right\} = \frac{-\alpha}{\gamma} \tag{30}$$

From the expectations of $E_N^{(N)}(k)$ and $E_{N-\Delta}^{(N)}(k)$ given in (18), we can substitute (28) and (29) into (30) to express the steady-state performance in terms of the leaky factor α and the stepsize γ, as

$$\begin{aligned} &E\{\|\mathbf{x}''^T(k)\tilde{\mathbf{h}}''(k)\|_2^2 + \|\mathbf{x}''^T(k)\tilde{\mathbf{g}}''(k)\|_2^2 \\ &-\|\mathbf{x}''^T(k)\mathbf{h}''^o(k)\|_2^2 - \|\mathbf{x}''^T(k)\mathbf{g}''^o(k)\|_2^2\} = -\frac{\alpha}{\gamma} \end{aligned} \tag{31}$$

The steady-state of the FT-CLMS algorithm is obtained by substituting $\tilde{\mathbf{g}}(k) = \mathbf{g}_N^o$ into (27) and $\mathbf{g}(k) = \mathbf{0}$, as it cannot model the augmented part into (22) while proceeding in a same manner as FT-ACLMS. This will yield the steady-state performance expectation in the form[2]

$$E\{\|\mathbf{x}''^T(k)\tilde{\mathbf{h}}''(k)\|_2^2 - \|\mathbf{x}''^T(k)\mathbf{h}''^o(k)\|_2^2\} = -\frac{\alpha}{\gamma} \tag{32}$$

[2] Convergence of ACLMS and CLMS for both circular and noncircular signals was addressed in [13]

Observe that the steady-state tap-length of the FT-ACLMS algorithm takes into consideration both the standard complex and augmented complex parts in the widely linear model (31), whereas the FT-CLMS only considers the standard complex parts (32). Therefore, the FT-ACLMS is suited for the processing of widely linear processes unlike the FT-CLMS which is optimal for only second order circular data. For second order linear circular processes, since $\mathbf{g}(k) = \mathbf{0}$, the steady-state performance of the FT-ACLMS degenerates into that of FT-CLMS.

B. Steady-State Analysis of the FT-CNGD and FT-ACNGD Algorithms

The output of the ACNGD algorithm $y(k)$ is given by [7]

$$y(k) = \Phi\Big(\underbrace{\mathbf{x}^T(k)\mathbf{h}(k)}_{\text{standard part [CNGD]}} + \underbrace{\mathbf{x}^H(k)\mathbf{g}(k)}_{\text{augmented part}} \Big) \tag{33}$$

Following the same approach to obtain (31) and replacing (22) with (33), the steady-state of the FT-ACNGD for the processing of widely linear processes is given by

$$\begin{aligned} &E\{\|\Phi\big(\mathbf{x}''^T(k)\tilde{\mathbf{h}}''(k) + \mathbf{x}''^T(k)\tilde{\mathbf{g}}''(k)\big)\|_2^2 \\ &-\|\mathbf{x}''^T(k)\mathbf{h}''^o(k)\|_2^2 - \|\mathbf{x}''^T(k)\mathbf{g}''^o(k)\|_2^2\} = -\frac{\alpha}{\gamma} \end{aligned} \tag{34}$$

Similarly, the steady-state for the FT-CNGD is obtained by substituting $\tilde{\mathbf{g}}(k) = \mathbf{g}_N^o$ into (27) and $\mathbf{g}(k) = \mathbf{0}$ into (33) to give

$$E\{\|\Phi\big(\mathbf{x}''^T(k)\tilde{\mathbf{h}}''(k)\big)\|_2^2 - \|\mathbf{x}''^T(k)\mathbf{h}''^o(k)\|_2^2\} = -\frac{\alpha}{\gamma} \tag{35}$$

Comparing (34) and (35), similar to the case of FT-ACLMS and FT-CLMS, the FT-ACNGD is more appropriate for the modelling of second order noncircular (improper) data than the FT-CNGD. The FT-ACNGD will degenerate into the FT-CNGD when processing second order linear circular processes.

V. SIMULATIONS

Simulations were conducted in the system identification setting. To generate the circular and noncircular test signals, a circular doubly white Gaussian noise was fed to the systems defined as

$$W_1 : z_k = 1.79z_{k-1} - 1.85z_{k-2} + 1.27z_{k-3} - 0.41z_{k-4} + w_k \tag{36}$$

$$W_2 : z_k = \Phi\big(1.79z_{k-1} - 1.85z_{k-2} + 1.27z_{k-3} - 0.41z_{k-4} + w_k\big) \tag{37}$$

$$\begin{aligned} W_3 : z_k = {} & 1.79z_{k-1} - 1.85z_{k-2} + 1.27z_{k-3} - 0.41z_{k-4} \\ & + w_k + 0.5w_k^* + 0.9w_{k-1}^* \end{aligned} \tag{38}$$

$$\begin{aligned} W_4 : z_k = {} & \Phi\big(1.79z_{k-1} - 1.85z_{k-2} + 1.27z_{k-3} - 0.41z_{k-4} \\ & + w_k + 0.5w_k^* + 0.9w_{k-1}^*\big) \end{aligned} \tag{39}$$

where W_1 is a circular linear system (AR4) [1], W_2 is a nonlinear system (NAR4), W_3 is a widely linear system, and W_4 is a nonlinear widely linear system. The nonlinearity $\Phi(\cdot)$ used to generate the signals is the $\tanh$ function which was the same as the nonlinearity used in the CNGD and ACNGD algorithms. System W_2 was obtained by applying a nonlinearity to W_1 and W_3 is constructed by extending W_1 with the augmented part of the widely linear system W, given by [14]

$$W : z_k = \exp(j)z_{k-1} + 2w_k + 0.5w_k^* + w_{k-1} + 0.9w_{k-1}^* \tag{40}$$

where j is the imaginary unit. System W_4 was obtained by applying a nonlinearity to W_3.

The following experiments were conducted in order to illustrate the usefulness of the proposed approach: 1) the optimal tap-length selection; 2) comparison of the proposed algorithms for system order identification of conventional and augmented complex systems.

A. Optimal Tap-Length

The optimal tap-lengths for both systems were determined by the steady-state MSE estimated by [5]

$$\bar{\varepsilon}(k) = \lambda\bar{\varepsilon}(k-1) + (1-\lambda)E(k) \tag{41}$$

where $\bar{\varepsilon}$ is the estimated steady-state MSE and $\lambda = 0.9$.

Figure 1 depicts the steady-state MSE for linear system W_1 using the CLMS, ACLMS, CNGD and ACNGD algorithms with $\mu = 10^{-4}$. All the four algorithms performed identically and the optimal tap-length for W_1 was found to be around $N_0 = 25$. Figure 2 shows the steady-state MSE for nonlinear noncircular system W_2 for all the four algorithms. Similar to Figure 1, all the algorithms performed identically and the optimal tap-length was approximately $N_0 = 25$.

Figure 3 illustrates the steady-state MSE for a widely linear noncircular system W_3 for the four algorithms. The two augmented algorithms performed similarly and were able to indicate the optimal tap-length, roughly $N_0 = 25$, whereas the standard techniques were not able to do so. Figure 4 depicts the steady-state MSE of a nonlinear widely linear system W_4; the optimal tap-length was found to be that of equivalent to the optimal tap-length of W_3. However, as desired, the ACNGD algorithm steady-state MSE was slightly lower than the ACLMS algorithm.

To summarize, the standard CLMS and CNGD algorithms were only able to estimate the optimal tap-length for conventional complex systems W_1 and W_2. In order to approximate the optimal tap-length for augmented complex linear systems W_3 and W_4, the augmented statistics must be considered. As a result, the ACLMS and ACNGD algorithms were able to find the optimal tap-length of the widely linear systems.

B. Modelling of Conventional Complex Systems

Figure 5 depicts the evolution of the optimal filter length parameter N for the FT-ACLMS, FT-CLMS, FT-CNGD and FT-ACNGD algorithms when employed for the modelling of linear system W_1. These algorithms were initialized with the following parameters: $\alpha = 0.03$, $\gamma = 10000$, δ=1, Δ=4, $N(0) = 7$ and $\mu = 5 \times 10^{-5}$. The large value of γ compensates for the scaling of the data prior to feeding them into the filters. The input data was scaled to the range [-0.8, 0.8]. From Figure 5, it is evident that the performances of all

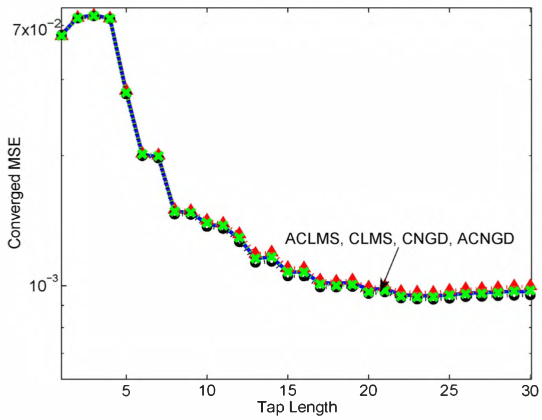

Fig. 1. The steady-state MSE for the linear circular system W_1 with respect to tap-length.

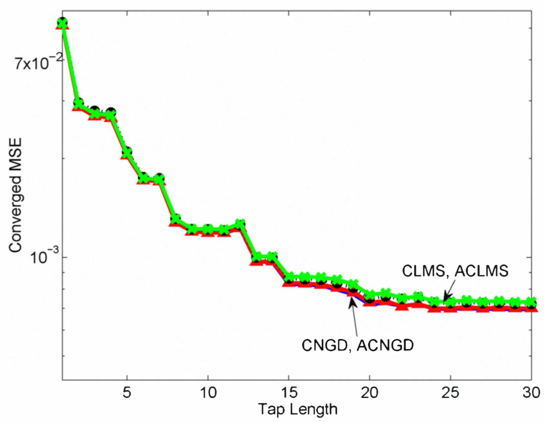

Fig. 2. The steady-state MSE for the nonlinear (noncircular) system W_2 with respect to tap-length.

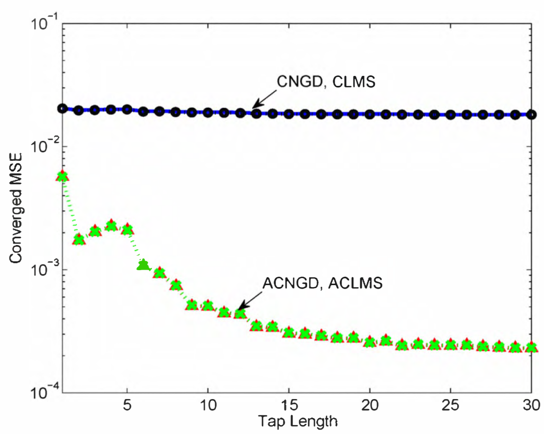

Fig. 3. The steady-state MSE for the widely linear (noncircular) system W_3 with respect to tap-length.

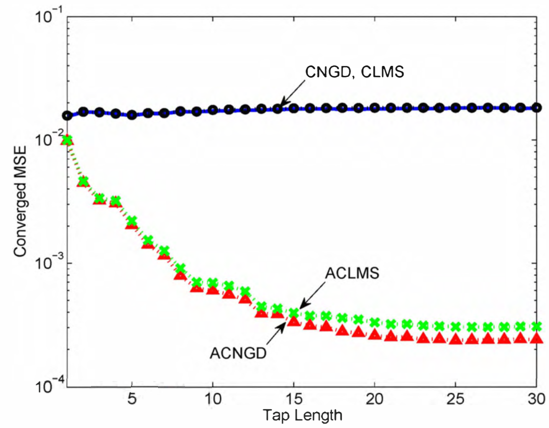

Fig. 4. The steady-state MSE for the nonlinear widely linear (noncircular) system W_4 with respect to tap-length.

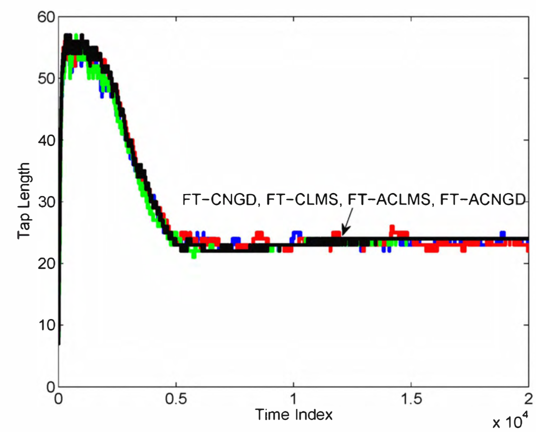

Fig. 5. The evolution of the optimal filter length parameter N for the modelling of the linear circular system W_1.

the algorithms considered were similar, as they all converged to the optimal tap-length at around the same time.

Figure 6 illustrates the evolution of the optimal filter length for the modelling of the nonlinear system W_2, initialized with the same parameters as previously. From Figure 6, it is apparent that the evolution of the filter length parameter is similar to that in Figure 5. Indeed, this is expected for both the linear and nonlinear systems, as the augmented algorithms (widely linear) degenerate into their conventional counterparts.

C. Modelling of Augmented Complex Systems

Figure 7 shows the evolution of the optimal filter length parameter N for FT-ACLMS, FT-CLMS, FT-ACNGD and FT-CNGD algorithms when employed for the modelling of the noncircular widely linear system W_3. These algorithms were initialized with the following parameters: $\alpha = 0.03$, $\gamma = 10000$, δ=1, Δ=4, $N(0) = 7$ and $\mu = 5 \times 10^{-5}$. As mentioned in the previous subsection, the reason behind the large value of γ is due to the input scaling. Observe that the augmented algorithms FT-ACNGD and FT-ACLMS were

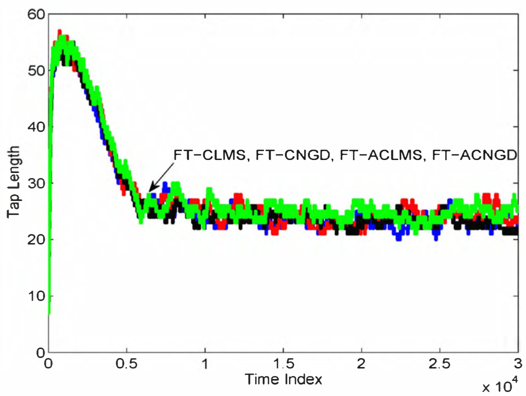

Fig. 6. The evolution of the optimal filter length parameter N for the modelling of the nonlinear system W_2.

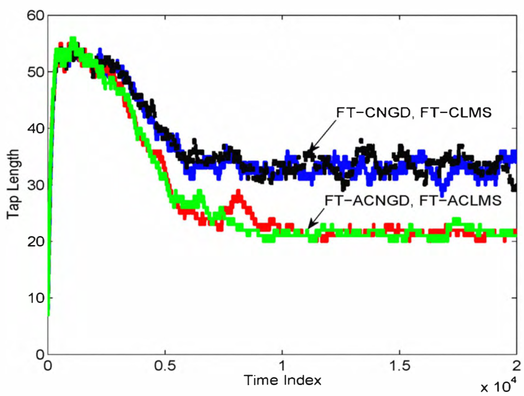

Fig. 7. The evolution of the optimal filter length parameter N for the modelling of the widely linear system W_3.

able to model W_3 correctly, in contrast to the FT-CLMS and FT-CNGD algorithms. In addition, the FT-CLMS and FT-CNGD algorithms were unstable at steady-state, whereas the FT-ACLMS and FT-ACNGD algorithms exhibited a smooth and stable convergence to the true tap-length. Notice also the slower convergence of FT-ACLMS and FT-ACNGD for the modelling of W_3 compared to W_1 and W_2, due to the additional information related to the widely linear models of W_3.

Figure 8 illustrates the evolution of the optimal filter length parameter N employed for the modelling of the nonlinear widely linear system W_4. As expected, the results were very similar to the modelling of W_3. The augmented algorithms were able to model W_4 accurately whereas their standard complex counterparts underperformed.

VI. Conclusions

We have introduced the fractional tap-length (FT) algorithm into complex-valued adaptive filters trained by the augmented least mean square (ACLMS) and augmented nonlinear gradient descent (ACNGD), and showed that the steady-state performance of the complex widely linear algorithms FT-ACLMS and FT-ACNGD can be used as a criterion

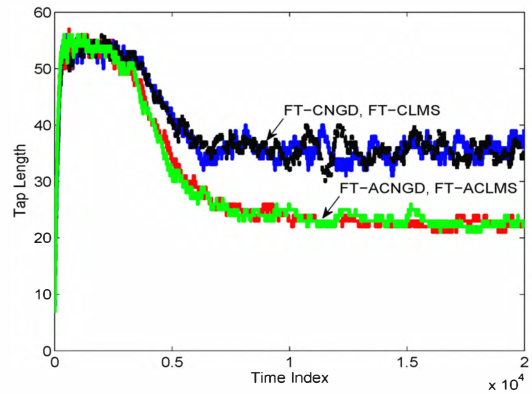

Fig. 8. The evolution of the optimal filter length parameter N for the modelling of the nonlinear widely linear system W_4.

for the identification of second order noncircular systems and model order selection. This complements the ability of ACLMS and ACNGD to model both second order circular (proper) and noncircular (improper) real world processes, and to track their nonlinear and nonstationary dynamics. Simulations on model order selection and the identification of complex impropriety support the approach.

References

[1] D. P. Mandic and S. L. Goh, *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*, Wiley, 2009.

[2] A. Hirose, *Complex-valued neural networks:theories and applications*, World Scientific Publishing, 2003.

[3] I. Aizenberg, N. N. Aizenberg, and J. P. L. Vandewalle, *Multi-valued and universal binary neurons*, Springer-Verlag New York, 2000.

[4] Z. Pritzker and A. Feuer, "Variable length stochastic gradient algorithm," *IEEE Transactions on Signal Processing*, vol. 39, no. 4, pp. 997–1001, 1991.

[5] Y. Gong and C. F. N. Cowan, "An LMS style variable tap-length algorithm for structure adaptation," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2400–2407, 2005.

[6] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719–720, 1975.

[7] Y. Xia, B. Jelfs, M. M. Van-Hulle, J. C. Principe, and D. P. Mandic, "An augmented echo state network for nonlinear adaptive filtering for complex noncircular signals," *sent to IEEE Transcations on Neural Networks December 2009*.

[8] C. Cheong-Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Transactions on Signal Processing [accepted for future publication]*, 2010.

[9] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, 1993.

[10] K. Kreutz-Delgado, "Complex gradient operator and the $\mathbb{CR}$ calculus," *Course Lecture Supplement ECE275A, Dept. of Electrical and Computer Engineering, UC San Diego*, 2006.

[11] T. Kim and T. Adali, "Approximation by fully complex multilayer perceptrons," *Neural Computation*, vol. 15, no. 7, pp. 1641–1666, 2003.

[12] Y. Zhang, N. Li, J. A. Chambers, and A. H. Sayed, "Steady-state performance analysis of variable tap-length LMS algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 839–845, 2008.

[13] S. C. Douglas and D. P. Mandic, "Performance analysis of the conventional complex LMS and augmented complex LMS algorithms," *In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3794–3797, 2010.

[14] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3061–3068, 2008.