The $N+1$ dimensional multiple integral can be replaced by a single zero mean Gaussian expectation with variance $\sigma_e^2(v/\sqrt{\beta}) = \sigma_o^2 + \sigma_x^2 v^T v/\beta$ as

$$E\left[g^2(x_1)\right] = \frac{1}{\sqrt{2\pi\sigma_e^2(v/\sqrt{\beta})}} \int_{-\infty}^{\infty} g^2(x_1) \exp\left(-\frac{x_1^2}{2\sigma_e^2(v/\sqrt{\beta})}\right) \mathrm{d}x_1. \tag{B4}$$

Equation (B4) can be evaluated using Price's Theorem [11]. Consider the joint Gaussian expectation

$$E\left[g(x_2)g(x_3)\right] = \frac{1}{2\pi|\mathrm{M}|^{1/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_2)g(x_3) \times \exp\left(-[x_2 x_3]\mathrm{M}^{-1}[x_2 x_3]^T/2\right) \mathrm{d}x_2 \mathrm{d}x_3 \tag{B5}$$

$$\text{with} \quad \mathrm{M} = E\left[x_1^2\right] \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \tag{B6}$$

Then [11]

$$E\left[\frac{\mathrm{d}g(x_2)}{\mathrm{d}x_2}\frac{\mathrm{d}g(x_3)}{\mathrm{d}x_3}\right] = \frac{\mathrm{d}E\left[g(x_2)g(x_3)\right]}{\mathrm{d}r} \tag{B7}$$

where $r = E[x_2 x_3] = \rho\sigma_e^2(v/\sqrt{\beta})$. Now

$$E\left[\frac{\mathrm{d}g(x_2)}{\mathrm{d}x_2}\frac{\mathrm{d}g(x_3)}{\mathrm{d}x_3}\right] = \frac{1}{2\pi|\mathrm{M}|^{1/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-x_2^2/2\sigma_y^2\right) \times \exp\left(-x_3^2/2\sigma_y^2\right) \exp\left(-[x_2 x_3]\mathrm{M}^{-1}[x_2 x_3]^T/2\right) \mathrm{d}x_2 \mathrm{d}x_3. \tag{B8}$$

Completing the square and performing the integration yields

$$E\left[\frac{dg(x_2)}{\mathrm{dx_2}}\frac{dg(x_3)}{\mathrm{dx_3}}\right] = \frac{\sigma_y^2}{\sqrt{\left(\sigma_y^2 + \sigma_e^2(v/\sqrt{\beta})\right)^2 - r^2}}. \tag{B9}$$

Thus, integrating on r,

$$E\left[g^2(x_1)\right] = \int_0^{\sigma_e^2(v/\sqrt{\beta})} \frac{\sigma_y^2 \mathrm{dr}}{\sqrt{\left(\sigma_y^2 + \sigma_e^2(v/\sqrt{\beta})\right)^2 - r^2}} = \sigma_y^2 \mathrm{Sin}^{-1}\left[\frac{\sigma_o^2 + \sigma_x^2 v^T v/\beta}{\sigma_y^2 + \sigma_o^2 + \sigma_x^2 v^T v/\beta}\right]. \tag{B10}$$

Finally, approximating $v^T v$ by K_{VV}, inserting (B11) in (B3) and integrating on β yields

$$E\left\{\frac{g^2\left(n_o(n) - X^T(n)v(n)\right)}{X^T(n)X(n)}\right\} = \frac{\sigma_y^2}{2\sigma_x^2} \int_1^{\infty} \frac{1}{\beta^{N/2}} \times \mathrm{Sin}^{-1}\left[\frac{\sigma_o^2 + \sigma_x^2 \mathrm{Tr}\left[K_{VV}(\mathrm{n})\right]/\beta}{\sigma_y^2 + \sigma_o^2 + \sigma_x^2 \mathrm{Tr}\left[K_{VV}(\mathrm{n})\right]/\beta}\right] \mathrm{d}\beta \tag{B11}$$

for the expectation in (18).

REFERENCES

[1] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 4, pp. 578–586, Aug. 1982.

[2] N. J. Bershad, "On the optimum data nonlinearity in LMS adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 2, pp. 69–78, Feb. 1986.

[3] N. J. Bershad, "On error saturation nonlinearities in LMS adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-36, no. 4, pp. 440–452, Apr. 1988.

[4] N. J. Bershad, "On weight update saturation nonlinearities in LMS adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-38, no. 4, pp. 623–630, Apr. 1990.

[5] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[6] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley-InterScience, 2003.

[7] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 4, pp. 793–806, Feb. 1986.

[8] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm for uncorrelated Gaussian data," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 680–691, Jul. 1988.

[9] D. Slock, "On the convergence behavior of the LMS and normalized LMS algorithms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2811–2825, Sep. 1993.

[10] S. C. Chan and Y. X. Zou, "A recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise: Fast algorithm and convergence performance analysis," *IEEE Trans. Signal Process.*, vol. SP-52, no. 4, pp. 975–991, Apr. 2004.

[11] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IEEE Trans. Inf. Theory*, vol. IT-4, no. 2, pp. 69–72, Jun. 1958.

[12] S. Marcos and O. Macchi, "Tracking capability of the least mean square algorithm: Applications to an asynchronous echo canceller," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 11, pp. 1570–1578, Nov. 1987.

Adaptive IIR Filtering of Noncircular Complex Signals

Clive Cheong Took and Danilo P. Mandic

***Abstract*—A recursive learning algorithm for the training of widely linear infinite impulse response complex valued adaptive filters is proposed. The use of so called augmented complex statistics makes this algorithm suitable for the processing of both second order circular (proper) and noncircular (improper) signals. A closed form solution for the bound on the stepsize is provided, and the small stepsize assumption in the derivation is used to reduce the computational complexity. Simulations for both synthetic and real-world circular and noncircular signals are provided in the prediction setting, illustrating the benefits of the proposed algorithm when modelling general complex signals.**

***Index Terms*—Adaptive prediction, augmented complex statistics, infinite impulse response filters, noncircular complex signals, wind modeling.**

I. INTRODUCTION

Complex-valued adaptive filtering algorithms are usually considered as generic extensions of their real-valued counterparts, with the second order statistics based on the covariance $\mathcal{C}_{\mathbf{xx}} = E\{\mathbf{x}(n)\mathbf{x}^H(n)\}$. However, recent advances in so-called augmented complex statistics show that for the modelling of general complex signals, this is not adequate, and the pseudocovariance $\mathcal{P}_{\mathbf{xx}} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$ (also known as the relation function [1] or the complementary covariance) should also be taken into account. This additional information proves crucial when processing second order non-circular (or improper) signals [2], for which the probability density functions are not rotation invariant. Thus, for instance, for second order circular (proper) signals $\mathcal{P}_{\mathbf{xx}} = 0$ and standard complex filters are adequate, whereas for

Manuscript received September 09, 2008; accepted April 05, 2009. First published May 05, 2009; current version published September 16, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alper Tunga Erdogan.

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: c.cheong-took@ic.ac.uk; d.mandic@ic.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2022353

improper signals $\mathcal{P}_{\mathbf{xx}} \neq 0$, and we need to design adaptive filters which account for this information. To ensure that both $\mathcal{C}_{\mathbf{xx}}$ and $\mathcal{P}_{\mathbf{xx}}$ are catered for, the complex-valued input $\mathbf{x}(n)$ is "augmented" to include its conjugate, that is, $\mathbf{x}^a(n) = [\mathbf{x}^T(n), \mathbf{x}^H(n)]^T$, and the second order statistics is based on the augmented covariance matrix $\mathcal{C}_{\mathbf{x^a x^a}}$, given by

$$\mathcal{C}_{\mathbf{x^a x^a}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{x}^* \end{bmatrix} [\mathbf{x}^H, \quad \mathbf{x}^T] = \begin{bmatrix} \mathcal{C}_{\mathbf{xx}} & \mathcal{P}_{\mathbf{xx}} \\ \mathcal{P}^*_{\mathbf{xx}} & \mathcal{C}^*_{\mathbf{xx}} \end{bmatrix}. \tag{1}$$

Algorithms based on augmented complex statistics have been used in communications [3], blind source separation [4], adaptive signal processing [5], and DOA (direction of arrival) estimation applications for general complex signals [6], [7].

Complex-valued processes are either complex by design (communications) or by convenience of representation (directional processes, radar, sonar, wind) [8]. The signals used in communications are usually complex circular (with rotation invariant probability density functions), whereas the class of signals made complex by convenience of representation is more general, and such signals are often noncircular. For the stochastic modelling of such signals, a widely linear moving average (MA) model was introduced by Picinbono, and is given by [9]

$$y(n) = \sum_{m=0}^{N} b_m x(n-m) + \sum_{m=0}^{N} h_m x^*(n-m) \tag{2}$$

where b_m and h_m are filter coefficients. Based on this widely linear model, an augmented complex least mean square (ACLMS) algorithm was recently developed in [10] and was employed in adaptive prediction of the improper wind signal.

As finite impulse response (FIR) widely linear models may not always be adequate, Moreno introduced an autoregressive moving average (ARMA) widely linear model with fixed coefficients, given by [11]

$$\begin{aligned} y(n) = &\sum_{m=1}^{p} a_m y(n-m) + \sum_{m=0}^{q} b_m x(n-m) \\ &+ \sum_{m=1}^{p} g_m y^*(n-m) + \sum_{m=0}^{q} h_m x^*(n-m) \end{aligned} \tag{3}$$

where p and q, respectively, are the order of the AR and MA part. This model serves as a theoretical background for the development of the proposed recursive algorithm for the training of adaptive infinite-impulse-response (IIR) filters.

To introduce a recursive algorithm for "augmented" complex adaptive IIR filtering, we extend the complex-valued adaptive IIR filtering algorithm introduced by Shynk [12] (referred to as CA-IIR), in order to cater for both proper and improper signals. This way, the coefficients of the widely linear ARMA model in (3) are made gradient adaptive, to give the output of a widely linear IIR filter in the form

$$\begin{aligned} y(n) = &\sum_{m=1}^{M} a_m(n) y(n-m) + \sum_{m=0}^{N} b_m(n) x(n-m) \\ &+ \sum_{m=1}^{M} g_m(n) y^*(n-m) + \sum_{m=0}^{N} h_m(n) x^*(n-m) \end{aligned} \tag{4}$$

where M is the order of the feedback and $(N+1)$ the length of the augmented tap input. This model can be written in a compact form as $y(n) = \mathbf{w}^{a^T}(n)\mathbf{z}^a(n)$, where $\mathbf{w}^a(n) = [a(n-1), \ldots, a(n-M), g(n-1), \ldots, g(n-M), b(n), \ldots, b(n-N), h(n), \ldots, h(n-N)]^T$, and $\mathbf{z}^a(n) = [\mathbf{y}^{a^T}(n), \mathbf{x}^{a^T}(n)]^T$. A stochastic gradient learning algorithm for the widely linear IIR filter will be referred to as the augmented CA-IIR (ACA-IIR).

The organization of the paper is as follows: in Section II, the ACA-IIR algorithm is derived. This is followed by the formulation of less computationally expensive gradient approximations. Next, an analysis on the bounds of the stepsize for convergence is provided in Section III. In Section IV, simulations are presented, based on a synthetic MA and ARMA processes and real world two-dimensional wind field measurements. Section V concludes the paper.

II. The ACA-IIR Algorithm

The aim of stochastic gradient based complex-valued adaptive filtering algorithms is to process the input signal $x(n)$ in order to estimate the desired response $d(n)$, based on the minimization of the cost function

$$\mathcal{J}(n) = \frac{1}{2}|d(n) - y(n)|^2 = \frac{1}{2}e(n)e^*(n). \tag{5}$$

Within a stochastic gradient optimization setting, the weights of the widely linear IIR filter in (4) are updated based on

$$\mathbf{w}^a(n+1) = \mathbf{w}^a(n) - \mu \nabla_{\mathbf{w}^a} \mathcal{J}(n). \tag{6}$$

The error gradients with respect to the filter coefficients can be computed as [12]

$$\begin{aligned} \nabla_{\mathbf{w}^a} \mathcal{J}(n) &= -\frac{1}{2}\left[e(n)\frac{\partial y^*(n)}{\partial \mathbf{w}^a(n)} + \frac{\partial y(n)}{\partial \mathbf{w}^a(n)} e^*(n) \right] \\ &= -\left[e(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) + \mathbf{\Psi}_{\mathbf{w}^a}(n)e^*(n) \right] \end{aligned} \tag{7}$$

where

$$\mathbf{\Phi}_{\mathbf{w}^a}(n) = \frac{1}{2}\left(\frac{\partial y^*(n)}{\partial \mathbf{w}^a_R(n)} + \imath \frac{\partial y^*(n)}{\partial \mathbf{w}^a_I(n)} \right) \tag{8}$$

$$\mathbf{\Psi}_{\mathbf{w}^a}(n) = \frac{1}{2}\left(\frac{\partial y(n)}{\partial \mathbf{w}^a_R(n)} + \imath \frac{\partial y(n)}{\partial \mathbf{w}^a_I(n)} \right). \tag{9}$$

The gradient vectors (8) and (9) can be evaluated as

$$\begin{aligned} \mathbf{\Phi}_{\mathbf{w}^a}(n) = &[\Phi_{a_1}(n), \ldots, \Phi_{a_M}(n), \Phi_{g_1}(n), \ldots, \Phi_{g_M}(n), \\ &\Phi_{b_0}(n), \ldots, \Phi_{b_N}(n), \Phi_{h_0}(n), \ldots, \Phi_{h_N}(n)]^T \\ \mathbf{\Psi}_{\mathbf{w}^a}(n) = &[\Psi_{a_1}(n), \ldots, \Psi_{a_M}(n), \Psi_{g_1}(n), \ldots, \Psi_{g_M}(n), \\ &\Psi_{b_0}(n), \ldots, \Psi_{b_N}(n), \Psi_{h_0}(n), \ldots, \Psi_{h_N}(n)]^T \end{aligned} \tag{10}$$

where the subscripts R and I denote the real and the imaginary parts of complex quantities, and $\imath = \sqrt{-1}$. To calculate the gradient in (7), each term in (8) and (9) has to be evaluated separately. Thus, for instance

$$\begin{aligned} \frac{\partial y^*(n)}{\partial a_{m_R}(n)} = &y^*(n-m) \\ &+ \sum_{\ell=1}^{M} a^*_\ell(n)\frac{\partial y^*(n-\ell)}{\partial a_{m_R}(n)} + \sum_{\ell=1}^{M} g^*_\ell(n)\frac{\partial y(n-\ell)}{\partial a_{m_R}(n)} \end{aligned} \tag{11}$$

$$\begin{aligned} \frac{\partial y^*(n)}{\partial a_{m_I}(n)} = &-\imath y^*(n-m) \\ &+ \sum_{\ell=1}^{M} a^*_\ell(n)\frac{\partial y^*(n-\ell)}{\partial a_{m_I}(n)} + \sum_{\ell=1}^{M} g^*_\ell(n)\frac{\partial y(n-\ell)}{\partial a_{m_I}(n)}. \end{aligned} \tag{12}$$

The presence of feedback within the IIR architecture gives rise to the recursions on the right-hand side of (11) and (12). These comprise the

derivatives of the past values of $y(n)$ with respect to the present value of the filter weight $a_m(n)$, and are not possible to compute. To circumvent this problem, for a small learning rate μ, the following standard approximation is adopted [12], [13]:

$$\mathbf{w}^a(n) \approx \mathbf{w}^a(n-1) \approx \cdots \approx \mathbf{w}^a(n-\tau)\tau = \max\{M, N+1\}. \tag{13}$$

Subsequently, the computation of (11) and (12) can be made mathematically tractable, to give the sensitivities

$$\frac{\partial y^*(n)}{\partial a_{m_R}(n)} \approx y^*(n-m) + \sum_{\ell=1}^{M} a_\ell^*(n)\frac{\partial y^*(n-\ell)}{\partial a_{m_R}(n-\ell)} + \sum_{\ell=1}^{M} g_\ell^*(n)\frac{\partial y(n-\ell)}{\partial a_{m_R}(n-\ell)} \tag{14}$$

$$\frac{\partial y^*(n)}{\partial a_{m_I}(n)} \approx -\imath y^*(n-m) + \sum_{\ell=1}^{M} a_\ell^*(n)\frac{\partial y^*(n-\ell)}{\partial a_{m_I}(n-\ell)} + \sum_{\ell=1}^{M} g_\ell^*(n)\frac{\partial y(n-\ell)}{\partial a_{m_I}(n-\ell)}. \tag{15}$$

The terms within the sums on the right-hand sides of (14) and (15) are now delayed versions of the corresponding terms on left-hand sides.

Thus, the gradients within the widely linear IIR filter in (8) have an additional term $\sum_{\ell=1}^{M} g_\ell^*(n)(\partial y(n-\ell))/(\partial a_m(n-\ell))$ as compared with the gradients within the standard CA-IIR algorithm [12]. These gradients are given by

$$\Phi_{a_m}(n) = y^*(n-m) + \sum_{\ell=1}^{M} a_\ell^*(n)\Phi_{a_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell^*(n)\Psi_{a_m}(n-\ell) \tag{16}$$

$$\Phi_{b_m}(n) = x^*(n-m) + \sum_{\ell=1}^{M} a_\ell^*(n)\Phi_{b_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell^*(n)\Psi_{b_m}(n-\ell) \tag{17}$$

$$\Phi_{g_m}(n) = y(n-m) + \sum_{\ell=1}^{M} a_\ell^*(n)\Phi_{g_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell^*(n)\Psi_{g_m}(n-\ell) \tag{18}$$

$$\Phi_{h_m}(n) = x(n-m) + \sum_{\ell=1}^{M} a_\ell^*(n)\Phi_{h_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell^*(n)\Psi_{h_m}(n-\ell) \tag{19}$$

and similarly for the gradient vector $\mathbf{\Psi}(n)$ in (9), we have

$$\Psi_{a_m}(n) = \sum_{\ell=1}^{M} a_\ell(n)\Psi_{a_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell(n)\Phi_{a_m}(n-\ell) \tag{20}$$

$$\Psi_{b_m}(n) = \sum_{\ell=1}^{M} a_\ell(n)\Psi_{b_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell(n)\Phi_{b_m}(n-\ell) \tag{21}$$

$$\Psi_{g_m}(n) = \sum_{\ell=1}^{M} a_\ell(n)\Psi_{g_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell(n)\Phi_{g_m}(n-\ell) \tag{22}$$

$$\Psi_{h_m}(n) = \sum_{\ell=1}^{M} a_\ell(n)\Psi_{h_m}(n-\ell) + \sum_{\ell=1}^{M} g_\ell(n)\Phi_{h_m}(n-\ell). \tag{23}$$

In the standard CA-IIR algorithm, gradients (20)–(21) represent unforced difference equations which vanish, as they do not depend on either the input $x(n)$ or the output $y(n)$. However, this is not the case with the augmented ACA-IIR algorithm due to the dependence of (20)–(21) on the terms $\mathbf{\Phi}(n)$, which are functions of either $x(n)$ or $y(n)$. Thus, no term in (16)–(23) can be considered negligible.

Finally, the filter coefficient update (6) can be expressed in a compact form as

$$\mathbf{w}^a(n+1) = \mathbf{w}^a(n) + \mu\big[e(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) + \mathbf{\Psi}_{\mathbf{w}^a}(n)e^*(n)\big]. \tag{24}$$

This completes the derivation of the recursive stochastic gradient based learning algorithm for widely linear adaptive IIR filters (ACA-IIR).

A. ACA-IIR as a Generalization of ACLMS

When feedback within the ACA-IIR algorithm is cancelled, that is, for a widely linear FIR filter, the partial derivatives on the right-hand side of (17), (19), (21), (21) and (23) vanish, yielding

$$\Phi_{b_m}(n) = x^*(n-m) \quad m = 0, \ldots, N \tag{25}$$

$$\Phi_{h_m}(n) = x(n-m) \quad m = 0, \ldots, N \tag{26}$$

$$\Psi_{b_m}(n) = \Psi_{h_m}(n) = 0 \quad m = 0, \ldots, N. \tag{27}$$

As desired the ACA-IIR algorithm (24) now simplifies into the ACLMS algorithm for FIR filters, given by [10]

$$\mathbf{w}^a(n+1) = \mathbf{w}^a(n) + \mu e(n)\mathbf{x}^{a^*}(n) \tag{28}$$

where the vector $\mathbf{x}^{a^*}(n)$ comprises the tap input vector and its complex conjugate.

B. Reducing the Computational Complexity of ACA-IIR

The weight update in ACA-IIR is computationally demanding, and it requires $4(M+N+1)$ recursions for the sensitivities $\mathbf{\Phi}_m(n)$ and $\mathbf{\Psi}_m(n)$, as seen from (16)–(23). However, by using the approximation (13), this can be simplified to updating only eight sensitivities [12]. For example, terms

$$\mathbf{\Phi}_a(n) = [\Phi_{a_1}(n), \Phi_{a_2}(n), \ldots, \Phi_{a_M}(n)]^T \tag{29}$$

can be replaced by

$$\mathbf{\Phi}_a^F(n) = \left[\Phi_{a_1}^F(n), \Phi_{a_2}^F(n), \ldots, \Phi_{a_M}^F(n)\right]^T. \tag{30}$$

Then, for a small learning rate, for example, the second term in the original gradient

$$\Phi_{a_2}(n) = y^*(n-2) + \sum_{\ell=1}^{M} a_\ell^*(n)\Phi_{a_2}(n-\ell) + \sum_{\ell=1}^{M} g_\ell^*(n)\Psi_{a_2}(n-\ell) \tag{31}$$

can be replaced by

$$\begin{aligned}\Phi_{a_2}^F(n) &= \Phi_{a_1}(n-1)\\ &= y^*(n-1-1) + \sum_{\ell=1}^{M} a_\ell^*(n-1)\Phi_{a_1}(n-1-\ell) + \sum_{\ell=1}^{M} g_\ell^*(n-1)\Psi_{a_1}(n-1-\ell)\\ &= y^*(n-2) + \sum_{\ell=1}^{M} \underbrace{a_\ell^*(n-1)}_{\text{time delayed version}}\ \Phi_{a_2}^F(n-\ell) + \sum_{\ell=1}^{M} \underbrace{g_\ell^*(n-1)}_{\text{time delayed version}}\ \Psi_{a_2}^F(n-\ell)\end{aligned} \tag{32}$$

and hence $\Phi_{a_2}(n) \approx \Phi_{a_2}^F(n)$. This approximation also applies for all the other sensitivities.

III. Convergence of ACA-IIR

In the analysis so far, it has been assumed that the stepsize μ is a small positive constant. We shall next examine the bounds on the stepsize, in order to provide insight into the convergence of the ACA-IIR algorithm. Consider the *a priori* and *a posteriori* estimation errors, given respectively by

$$e(n) = d(n) - \mathbf{w}^{a^T}(n)\mathbf{x}^a(n) \tag{33}$$

$$\tilde{e}(n) = d(n) - \mathbf{w}^{a^T}(n+1)\mathbf{x}^a(n). \tag{34}$$

To estimate the range of stepsize μ which ensures $|\tilde{e}(n)| < |e(n)|$ [14] and thus convergence, following the approach from [14, p. 344], perform a first-order Taylor series expansion of $|\tilde{e}(n)|$ around $|e(n)|$, given by

$$|\tilde{e}(n)|^2 = |e(n)|^2 + \Delta\mathbf{w}^{a^H}(n)\frac{\partial|e(n)|^2}{\partial\mathbf{w}^a(n)} \tag{35}$$

where $\Delta\mathbf{w}^a(n) = \mu\left[e(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) + e^*(n)\mathbf{\Psi}_{\mathbf{w}^{a(n)}}\right]$ is obtained from (24). The gradient term in (35) can be expressed as

$$\frac{\partial|e(n)|^2}{\partial\mathbf{w}^a(n)} = -2\left[e(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) + e^*(n)\mathbf{\Psi}_{\mathbf{w}^{a(n)}}\right] \tag{36}$$

to yield

$$\begin{aligned}
&|\tilde{e}(n)|^2 \\
&\quad = |e(n)|^2 - 2\mu\left[e(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) + e^*(n)\mathbf{\Psi}_{\mathbf{w}^{a(n)}}\right]^H \\
&\qquad \times \left[e(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) + e^*(n)\mathbf{\Psi}_{\mathbf{w}^{a(n)}}\right] \\
&\quad = |e(n)|^2 - 2\mu\left[|e(n)|^2\left(\|\mathbf{\Phi}_{\mathbf{w}^a}(n)\|^2 + \|\mathbf{\Psi}_{\mathbf{w}^a}(n)\|^2\right)\right. \\
&\qquad \left. + e^{*^2}(n)\mathbf{\Phi}_{\mathbf{w}^a}^H(n)\mathbf{\Psi}_{\mathbf{w}^a}(n) + e^2(n)\mathbf{\Psi}_{\mathbf{w}^a}^H(n)\mathbf{\Phi}_{\mathbf{w}^a}(n)\right].
\end{aligned} \tag{37}$$

Upon applying the statistical expectation operator, as shown in the Appendix, the cross-terms $e^*(n)\mathbf{\Phi}_{\mathbf{w}^a}^H(n)\mathbf{\Psi}_{\mathbf{w}^a}(n)$ and $e(n)\mathbf{\Psi}_{\mathbf{w}^a}^H(n)\mathbf{\Phi}_{\mathbf{w}^a}(n)$ can be neglected, to give

$$\begin{aligned}
|\tilde{e}(n)|^2 &= |e(n)|^2 - 2\mu|e(n)|^2\left(\|\mathbf{\Phi}_{\mathbf{w}^a}(n)\|^2 + \|\mathbf{\Psi}_{\mathbf{w}^a}(n)\|^2\right) \\
&= \left[1 - 2\mu\left(\|\mathbf{\Phi}_{\mathbf{w}^a}(n)\|^2 + \|\mathbf{\Psi}_{\mathbf{w}^a}(n)\|^2\right)\right]|e(n)|^2.
\end{aligned} \tag{38}$$

Since the error powers $|\tilde{e}(n)|^2$ and $|e(n)|^2$ are non-negative, so too

$$1 - 2\mu\left(\|\mathbf{\Phi}_{\mathbf{w}^a}(n)\|^2 + \|\mathbf{\Psi}_{\mathbf{w}^a}(n)\|^2\right) \geq 0 \tag{39}$$

and the range of the stepsize which preserves the stability of the ACA-IIR algorithm becomes

$$0 < \mu \leq \frac{1}{2\left(\|\mathbf{\Phi}_{\mathbf{w}^a}(n)\|^2 + \|\mathbf{\Psi}_{\mathbf{w}^a}(n)\|^2\right)}. \tag{40}$$

In practice, this upper bound is usually an order of magnitude lower [14]. The analysis of the mean-square-error surface is given in the Appendix.

IV. Simulations

The performance of the proposed ACA-IIR is evaluated based on three data sets: a proper MA process, an improper ARMA process, and a real-world improper complex wind field. The linear MA process was adopted from [11], and is given by

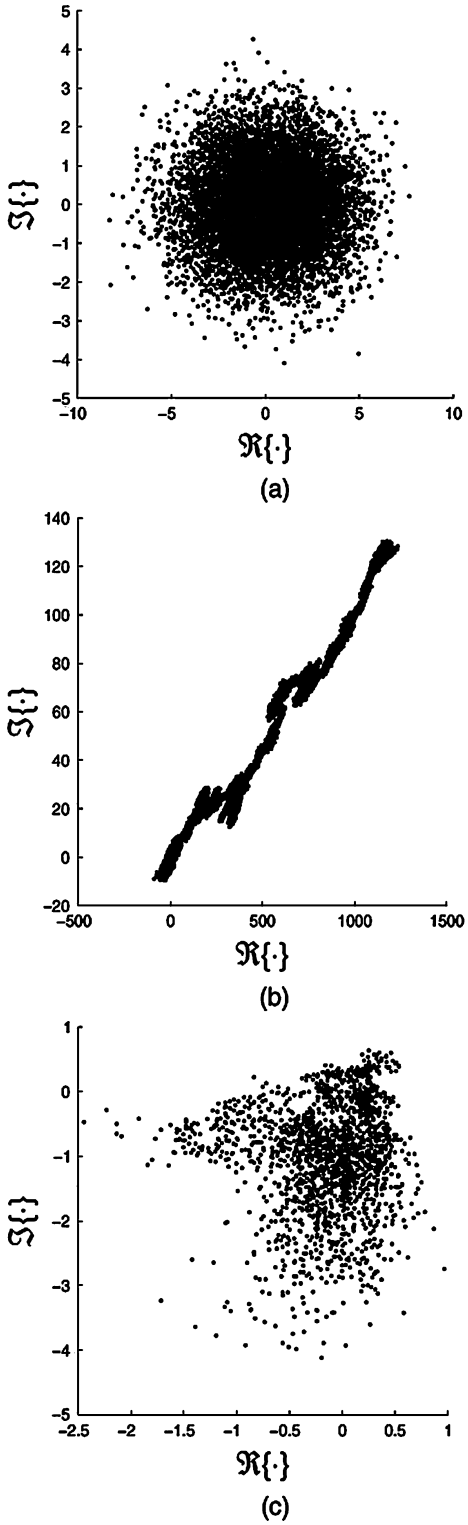

Fig. 1. Geometric view of circularity via "real-imaginary" scatter plot. (a) Proper MA process; (b) improper ARMA process; and (c) complex wind signal.

$$\begin{aligned}
&y(0) = 0 \\
&y(n) = 2w(n) + 0.5w^*(n) \\
&\qquad + w(n-1) + 0.9w^*(n-1), n \geq 1
\end{aligned} \tag{41}$$

whereas the widely linear ARMA process was a combination of the

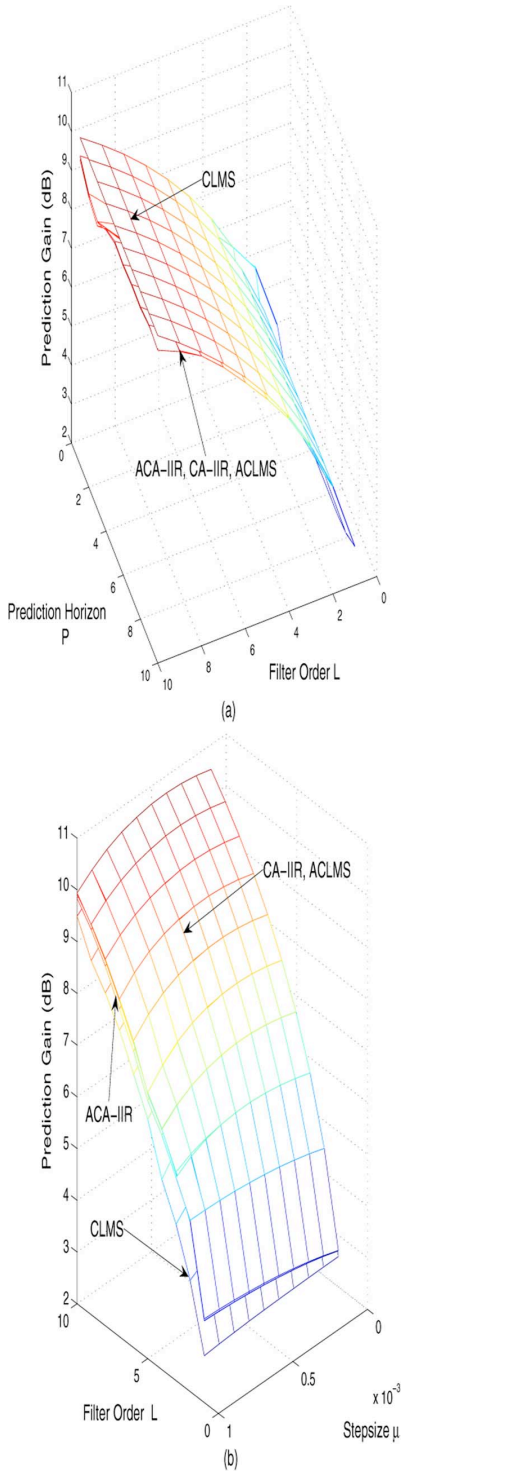

Fig. 2. Performance comparison between ACA-IIR, CA-IIR, ACLMS, and CLMS for the prediction of the circular MA signal, obtained from (41) for $K = 0$. (a) Performance dependence on P and L and (b) performance dependence on μ and L.

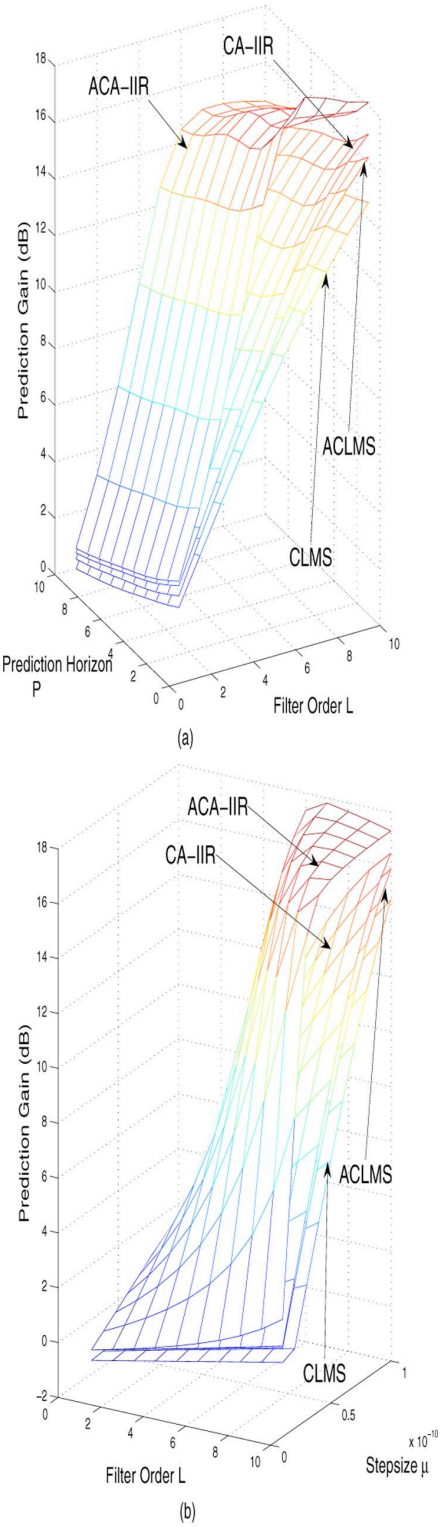

Fig. 3. Performance comparison between ACA-IIR, CA-IIR, ACLMS, and CLMS for the prediction of the noncircular ARMA model (42). (a) Performance dependence on P and L and (b) performance dependence on μ and L.

MA process (41) and the AR process from p. 155 [15], given by

$$\begin{aligned} y(0) &= 0 \\ y(n) &= 1.79y(n-1) - 1.85y(n-2) \\ &\quad + 1.27y(n-3) - 0.41y(n-4) \\ &\quad + 0.2y(n-5) + 2w(n) + 0.5w^*(n) \\ &\quad + w(n-1) + 0.9w^*(n-1),\ n \geq 1 \end{aligned} \tag{42}$$

with

$$\begin{aligned} E\{w(n-i)w(n-j)^*\} &= \delta(i-j) \\ E\{w(n-i)w(n-j)\} &= \mathrm{K}\delta(i-j) \end{aligned} \tag{43}$$

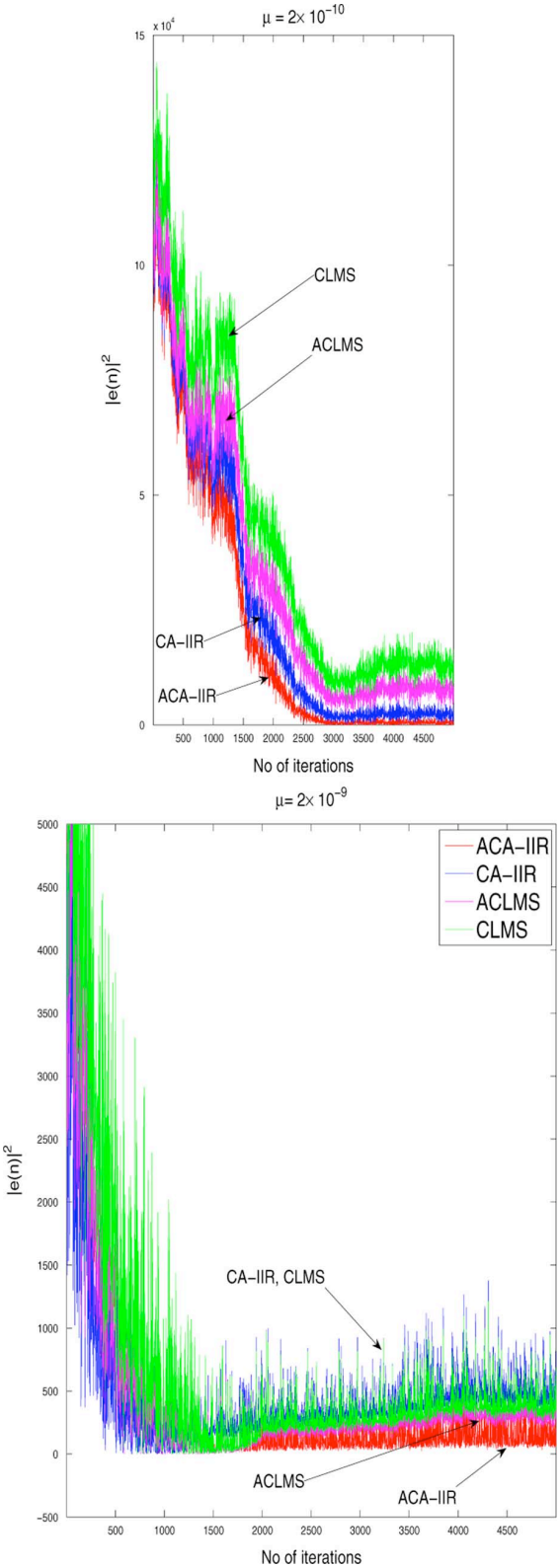

Fig. 4. Learning curves of ACA-IIR, CA-IIR, ACLMS, and CLMS for the prediction of the noncircular ARMA model (42).

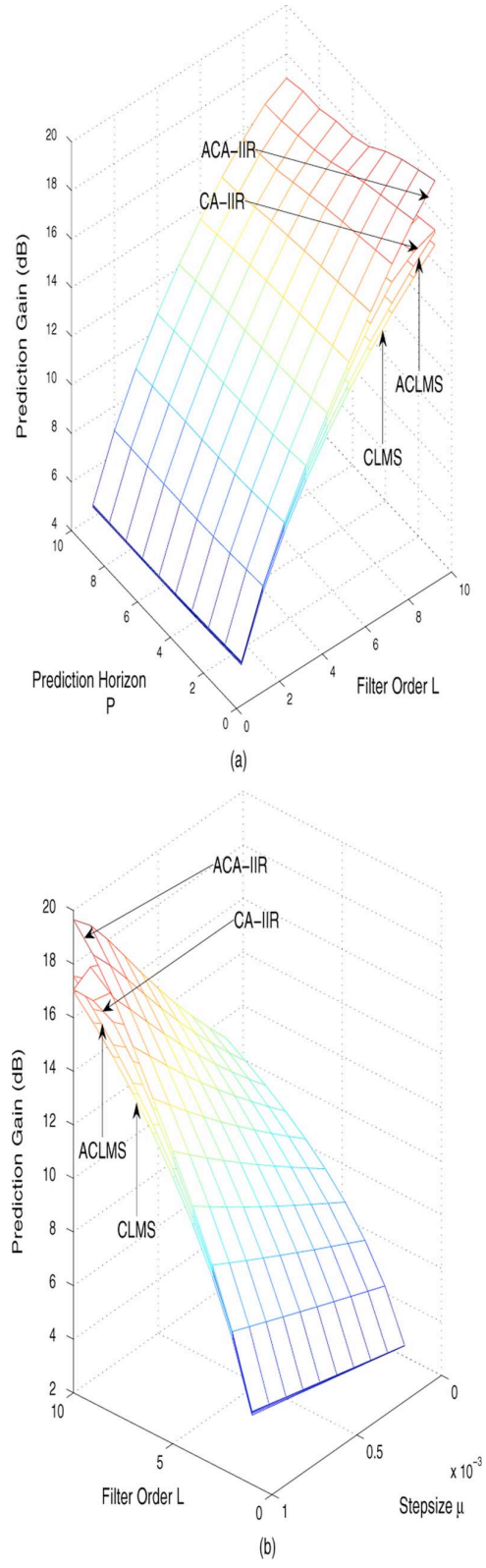

Fig. 5. Performance comparison between ACA-IIR, CA-IIR, ACLMS, and CLMS for the prediction of a 2-D wind field. (a) Performance dependence on P and L and (b) performance dependence on μ and L.

where $w(n)$ is doubly white circular noise [8]. The circular MA model is obtained for $K = 0$, whereas for the noncircular ARMA model $K = 0.95$. The 2-D wind field data[1] was collected from measurements of the wind speeds in the north-south and east-

[1]The wind data was provided by Prof. Aihara's team at the Institute of Industrial Science, University of Tokyo.

$$\begin{aligned}\mathcal{J} &= \frac{1}{2}E\{|E_C(q^{-1},n)x(n)+E_D(q^{-1},n)x^*(n)|^2\}\\ &= \frac{1}{2}E\{|E_C(q^{-1},n)|^2x(n)x^*(n)\\ &\quad +|E_D(q^{-1},n)|^2x(n)x^*(n)\\ &\quad +\underbrace{E_C(q^{-1},n)E_D^*(q^{-1},n)x(n)x(n)+E_D(q^{-1},n)E_C^*(q^{-1},n)x^*(n)x^*(n)\}}_{\text{cross-terms}}\end{aligned} \tag{48}$$

west direction. Fig. 1 shows the scatter plots of the complex signals considered; observe the circular symmetry (properness) for the circular MA signal and the non-circularity of the improper ARMA process and wind signal.

Simulations were conducted in a P-step ahead prediction setting. Thus, for instance, the P-step ahead prediction of the widely linear ARMA process (42), $y(n+P)$, was based on the delayed outputs $y(n-1),\ldots,y(n-L)$, and the corresponding filtered driving widely linear noise. The filter order L was set to $L=M=N$, where M is the order of the feedback and N is the order of the tap input of the filter (4). The performance of the ACA-IIR was compared with those for the widely linear augmented CLMS (ACLMS) [10], and the linear complex LMS (CLMS) [16] and the CA-IIR [12]. For a quantitative assessment of the performance, the standard prediction gain $R_p = 10\log(\sigma_{\mathbf{x}}^2/\sigma_e^2)$ was employed, where $\sigma_{\mathbf{x}}^2$ and σ_e^2 denote respectively the variance of the input and output error [15].

Fig. 2 compares the performances of all considered algorithms on the prediction of the circular MA process (41), over a range of stepsizes and prediction horizons. As the MA process exhibits a circularly symmetric distribution (Fig. 1), it is therefore expected that all the four algorithms would have similar performances, as illustrated in Fig. 2.

Fig. 3 illustrates the prediction performance for the improper ARMA process (42), whereas Fig. 4 shows the learning curves for all the algorithms considered. The ACA-IIR exhibited best performance, followed by the CA-IIR, ACLMS, and CLMS. Similar observations as for the improper ARMA case can be made for the simulations on the real-world complex wind signal, shown in Fig. 5.

V. Conclusion

We have introduced a recursive learning algorithm for widely linear adaptive IIR filtering. The proposed ACA-IIR algorithm has been derived based on augmented complex statistics, thus making it suitable for both circular and noncircular complex signals. Convergence analysis has provided the bound on the stepsize which preserves stability of the ACA-IIR. Further, computational complexity has been reduced by making use of the redundancy in the state vector of the filter. The proposed algorithm has been shown to exhibit superior performance over the ACLMS, CA-IIR, and CLMS on the multi-step ahead prediction of both synthetic and real world noncircular complex signals.

Appendix

To show that the cross-terms $(e^*)^2(n)\mathbf{\Phi}_{\mathbf{w}^a}^H(n)\mathbf{\Psi}_{\mathbf{w}^a}(n) + e^2(n)\mathbf{\Psi}_{\mathbf{w}^a}^H(n)\mathbf{\Phi}_{\mathbf{w}^a}(n)$ in (37) can be neglected, notice that $(e^*(n))^2 = (e^2(n))^*$ and $\mathbf{\Phi}_{\mathbf{w}^a}^H(n)\mathbf{\Psi}_{\mathbf{w}^a}(n) = [\mathbf{\Psi}_{\mathbf{w}^a}^H(n)\mathbf{\Phi}_{\mathbf{w}^a}(n)]^H$. Also, denote the products of the cross-term sensitivities by $\mathbf{\Phi}_{\mathbf{w}^a}^H(n)\mathbf{\Psi}_{\mathbf{w}^a}(n) = \mathbf{z}(n)$ and $\mathbf{\Psi}_{\mathbf{w}^a}^H(n)\mathbf{\Phi}_{\mathbf{w}^a}(n) = \mathbf{z}^*(n)$, and $\boldsymbol{\alpha} = (e^*)^2(n)\mathbf{z}(n)$ and $\boldsymbol{\beta} = e^2(n)\mathbf{z}^*(n)$. Then

$$\begin{aligned}\boldsymbol{\alpha} &= \left[\left(e_R^2(n)-e_I^2(n)\right)\mathbf{z}_R(n)+2e_R(n)e_I(n)\mathbf{z}_I(n)\right]\\ &\quad +\imath\left[\left(e_R^2(n)-e_I^2(n)\right)\mathbf{z}_I(n)-2e_R(n)e_I(n)\mathbf{z}_R(n)\right]\end{aligned}$$

$$\begin{aligned}\boldsymbol{\beta} &= \left[\left(e_R^2(n)-e_I^2(n)\right)\mathbf{z}_R(n)+2e_R(n)e_I(n)\mathbf{z}_I(n)\right]\\ &\quad -\imath\left[\left(e_R^2(n)-e_I^2(n)\right)\mathbf{z}_I(n)-2e_R(n)e_I(n)\mathbf{z}_R(n)\right]\end{aligned} \tag{44}$$

and the cross-terms in (37) are equal to $\boldsymbol{\alpha}+\boldsymbol{\beta}$. As the terms $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are complex conjugates, the imaginary part of $\boldsymbol{\alpha}+\boldsymbol{\beta}$ is always zero. In MMSE estimation, our aim is to obtain doubly white circular output error, for which $e_R^2 = e_I^2$ and $e_R \perp e_I$, and thus upon applying the statistical expectation operator the real part of $\boldsymbol{\alpha}+\boldsymbol{\beta}$ also vanishes. This can also be shown using the standard independence assumptions.

A. Analysis of the Mean-Square-Error Surface

The relationship between the mean square error and the coefficients of the widely linear ARMA input will be illustrated for the process (42) considered in the simulations, which has a general form

$$y(n) = C(q^{-1},n)x(n) + D(q^{-1},n)x^*(n) \tag{46}$$

where $C(q^{-1},n) = (B(q^{-1},n))/(A(q^{-1},n))$, $D(q^{-1},n) = (H(q^{-1},n))/(A(q^{-1},n))$ and the teaching signal

$$d(n) = C_{\text{opt}}(q^{-1})x(n) + D_{\text{opt}}(q^{-1})x^*(n). \tag{47}$$

The MSE now becomes (48), shown at the top of the page, where $E_C(q^{-1},n) = C_{\text{opt}}(q^{-1}) - C(q^{-1},n)$ and $E_D(q^{-1},n) = D_{\text{opt}}(q^{-1}) - D(q^{-1},n)$ are weight error vectors. Similarly to the analysis above, upon applying the statistical expectation operator the cross-terms in (48) vanish, and

$$\mathcal{J} = \frac{1}{2}E\{|E_C(q^{-1},n)|^2x(n)x^*(n)+|E_D(q^{-1},n)|^2x(n)x^*(n)\}. \tag{49}$$

The mean-square-error surface $\mathcal{J}$ is therefore expressed in terms of the weight error vectors and the covariance matrix of the input and can be analyzed using standard MSE analysis for IIR filters [17]

References

[1] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.

[2] P. J. Schreier, L. L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Process. Lett.*, vol. 13, no. 7, pp. 433–436, 2006.

[3] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, "Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 746–756, Mar. 2004.

[4] M. Novey and T. Adali, "Complex ICA by negentropy maximization," *IEEE Trans. Neural Netw.*, vol. 19, no. 4, pp. 596–609, 2008.

[5] S. L. Goh and D. P. Mandic, "An augmented ACRTRL for complex valued recurrent neural networks," *Elsevier Neural Netw.*, vol. 20, no. 10, pp. 1061–1066, 2007.

[6] J. P. Delmas, "Asymptotically minimum variance second-order estimation for noncircular signals with application to DOA estimation," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 12351241–, May 2004.

[7] H. Abeida and J. P. Delmas, "MUSIC-like estimation of direction of arrival for non-circular sources," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2678–2690, Jul. 2006.
[8] D. P. Mandic and V. S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. New York: Wiley, 2009.
[9] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
[10] D. P. Mandic, S. Javidi, S. L. Goh, A. Kuh, and K. Aihara, "Complex-valued prediction of wind profile using augmented complex statistics," *Renew. Energy*, vol. 34, pp. 196–201, 2009.
[11] J. Navarro-Moreno, "ARMA prediction of widely linear systems by using the innovations algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3061–3068, Jul. 2008.
[12] J. J. Shynk, "A complex adaptive algorithm for IIR filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 5, pp. 1342–1344, 1986.
[13] J. R. Treichler, C. Johnson, and M. G. Larimore, *Theory and Design of Adaptive Filters*. Englewood Cliffs, NJ: Prentice-Hall, 2001.
[14] P. A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. New York: Marcel Dekker, 1994.
[15] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction*. New York: Wiley, 2001.
[16] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," in *Proc. IEEE*, 1975, vol. 63, no. 4, pp. 719–720.
[17] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 3rd ed. New York: Springer, 2008.

On the Arbitrary-Length M-Channel Linear Phase Perfect Reconstruction Filter Banks

Zhiming Xu and Anamitra Makur

***Abstract*—This correspondence extends the theory and lattice factorizations for M-channel linear phase perfect reconstruction filter banks (LPPRFBs). We deal with FIR FBs with real-valued filter coefficients in which all filters have the same *arbitrary* length $L = KM + \beta\,(0 \le \beta < M)$ and same symmetry center, in contrast to traditional constrained length profile $L = KM$. First, refined existence conditions on this larger class of FBs are given. Then, lattice structures are developed for *both* odd and even-channel FBs, and their relationship with time-domain lapped transforms is explained. These structures are more general compared to conventional design methods, and cover them as special cases. Furthermore, we discuss how to structurally impose the regularity onto the lattice structure to ensure the smoothness of the basis function, which is very desirable in high performance image compression. Finally, these lattice structures are proven to be *minimal* in terms of the number of delay elements used in implementation, and to *completely* span the class of LPPRFBs with length $L \le 2M$ and the whole class of linear phase paraunitary filter banks (LPPUFBs).**

***Index Terms*—Completeness, filter bank, lattice structure, linear phase, minimality, perfect reconstruction, regularity.**

I. Introduction

The M-channel critically sampled uniform filter banks (FBs), which act as a powerful tool in linear time-frequency signal analysis, have been studied extensively and employed in various signal processing applications [1]. Denote the analysis and synthesis polyphase matrices by $\mathbf{E}(z)$ and $\mathbf{R}(z)$, respectively [1]. It is well known that the FB has perfect reconstruction (PR) property if $\mathbf{R}(z)$ is the inverse of $\mathbf{E}(z)$, i.e., $\mathbf{R}(z)\mathbf{E}(z) = \mathbf{I}_M$. In addition, if all the analysis and synthesis filters $H_k(z)$ and $F_k(z)$ have linear phase (LP) (very desirable in image/video processing), such FB is called LP FB. The theory, design and implementation of LPPRFBs were studied extensively [2]–[8]. However, most previous works imposed an *unnecessary a priori* condition on the filter length $L = KM$ for simplification of the factorization and design.

In this correspondence, we systematically investigate the factorizations of LPPRFBs with *arbitrary* filter lengths $L = KM+\beta\,(0 \le \beta < M)$ and the *same* symmetry center. The motivation for such work is to design more flexible FBs which can give more possible choices and offer better trade-off between the filter length and the performance of FBs than traditional methods. The existence conditions of such class of LPPRFBs in terms of filter lengths and filter symmetry polarities were explored thoroughly in [3]. However, most existing design methods [2], [4]–[8] concentrated only on the special case $\beta = 0$. As far as the case $\beta > 0$ is concerned, some works have been reported before. The general arbitrary-length cosine-modulated FBs were studied in [9] and in [10] for a class of even-channel LP FBs. Another design method was discussed in [3] for the class of restrictive even-channel LP paraunitary (PU) FBs. Later, it was also extended to even-channel LPPRFBs [11]. The case of even β has been further studied in [12] for both even and odd-channel systems. In addition, a subclass of such LPPRFBs, i.e., LPPUFBs with both odd M and β, was considered in [13]. Besides these, Tran *et al.* have contributed a lot for this class of LPPRFBs under the framework of time-domain lapped transforms (TDLTs) via pre and postfiltering [14]–[16]. Later, we will show that our derived structure is essentially equivalent to theirs. However, different from their presentation, we study such FBs in a *unifying* manner based on the systematic investigation of the necessary constraints held on $\mathbf{E}(z)$, which can lead us to derive lattice factorizations and show their connections with other works and some theoretical properties. For simplicity, we only consider the class of causal FIR FBs with anticausal FIR inverses (CAFACAFI) [17] in this correspondence.

Notations: Bold-faced quantities with uppercase and lowercase letters denote matrices and vectors, respectively. Subscripts indicate the size of a matrix if it is not clear from the context. $\mathbf{I}_M$, $\mathbf{J}_M$ and $\mathbf{0}_M$ denote the identity matrix, reversal matrix and null matrix, all with size $M \times M$. The rank of a matrix $\mathbf{A}$ is denoted by $\rho(\mathbf{A})$. For a real number x, $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the integer floor and ceiling of x, respectively. A special matrix useful for linear phase property of FBs is $\mathbf{I}_M^J = \mathrm{diag}(\mathbf{I}_{\lceil M/2 \rceil}, \mathbf{J}_{\lfloor M/2 \rfloor})$. In addition, $\mathbf{W}_{2m}$ and $\mathbf{W}_{2m+1}$ are $2m \times 2m$ and $(2m+1) \times (2m+1)$ butterfly-like orthogonal matrices, respectively, as follows:

$$\mathbf{W}_{2m} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_m & \mathbf{I}_m \\ \mathbf{I}_m & -\mathbf{I}_m \end{bmatrix}$$

$$\mathbf{W}_{2m+1} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I}_m & \mathbf{0} & \mathbf{I}_m \\ \mathbf{0} & \sqrt{2} & \mathbf{0} \\ \mathbf{I}_m & \mathbf{0} & -\mathbf{I}_m \end{bmatrix}.$$

II. Existence Conditions and Lattice Structures

In [3], some necessary conditions for LPPRFBs with lengths $L_i = K_i M + \beta\ (0 \le \beta < M)$ were derived. In particular, for the class of LPPRFBs with equal filter length, i.e., $K_i = K$ for $i = 0, 1, \ldots, M-1$, some further refined existence conditions can be obtained and listed in Table I, where n_s and n_a denote the number of symmetric and anti-

Manuscript received October 08, 2008; accepted April 22, 2009. First published May 27, 2009; current version published September 16, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Soontorn Oraintara.

The authors are with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore (e-mail: zmxu@ntu.edu.sg; eamakur@ntu.edu.sg).

Digital Object Identifier 10.1109/TSP.2009.2024026