THE AUGMENTED COMPLEX LEAST MEAN SQUARE ALGORITHM WITH APPLICATION TO ADAPTIVE PREDICTION PROBLEMS

[1]*Soroush Javidi,* [1]*Maciej Pedzisz,* [2]*Su Lee Goh and* [1]*Danilo P. Mandic*

[1]Imperial College London, UK
{d.mandic,soroush.javidi,m.pedzisz}@imperial.ac.uk
[2]Shell Exploration and Production, Assen, The Netherlands
vanessa.goh@shell.com

ABSTRACT

An augmented complex least mean square (ACLMS) algorithm for complex domain adaptive filtering which utilises the full second order statistical information is derived for adaptive prediction problems. This is achieved based on some recent advances in complex statistics and by using widely linear modelling in $\mathbb{C}$. This way, both circular and non–circular complex signals can be processed optimally, using the same model. Simulations on complex–valued wind field and on a complex autoregressive process show the effectiveness of this approach as compared to the standard Complex LMS algorithm.

Index Terms— Adaptive signal processing, Signal representations, Statistics, Least mean square methods, Wind forecasting

1. INTRODUCTION

The Least Mean Square (LMS) [1] algorithm is a workhorse of adaptive signal processing in $\mathbb{R}$. For adaptive filtering in the field of complex numbers $\mathbb{C}$, Widrow et al. introduced the Complex LMS (CLMS) in 1975 [2]. This algorithm benefits from the robustness and stability of the LMS, and enable simultaneous filtering of the real and imaginary parts of complex–valued data [3]. The need for extending and enhancing the performance of complex domain algorithms is due to the fact that $\mathbb{C}$ provides a natural processing platform. For instance, by processing real domain problems in $\mathbb{C}$, we can include the phase component and have a multidimensional solution exhibiting performance benefits over real domain solutions [4].

Unlike the usual assumption, statistics in the complex domain are not a straightforward extension from real domain statistics, the signal processing literature, however, usually deals with statistics in $\mathbb{C}$ as an extension of those in $\mathbb{R}$. For example, the covariance matrix of a zero mean complex vector $\mathbf{z}$ is $E\{\mathbf{z}\mathbf{z}^H\}$ is seen as an extension of the real covariance $E\{\mathbf{x}\mathbf{x}^T\}$, achieved by replacing the transpose operator $(\cdot)^T$ with the Hermitian operator $(\cdot)^H$ [5, 6]. However, strictly speaking, this is true only for circular data and does not accommodate for all the information contained within the complex data.

Some recent advances in complex domain statistics have enabled better modelling of complex data and therefore have opened the possibility for enhanced adaptive filtering algorithms [5, 7]. The use of so called augmented complex statistics leads to "widely" linear complex domain modelling [8, 9]. This is performed by considering both the covariance and pseudo-covariance matrices within the models. Based on this principle, the Widely Linear LMS was introduced in the communications field for use in a direct–sequence code division multiple access (DS–CDMA) receiver [10, 11]. It was shown that the algorithm has a lower complexity, while having an equally good performance to standard linear algorithms.

In this paper, the derivation of the widely linear LMS algorithm, or augmented CLMS (ACLMS), is provided in an adaptive prediction context, and illustrate the improvement in the performance of this algorithm as compared to the standard CLMS algorithm in an adaptive prediction setting for general complex signals. Our focus in on the forecasting of the complex wind profile, an important problem in renewable energy.

2. AUGMENTED COMPLEX STATISTICS

For a complex random vector (RV) $\mathbf{x} \in \mathbb{C}^m$, $E\{\mathbf{x}\} = \mathbf{0}$, we can define two covariance matrices

$$\mathcal{C}_{\mathbf{xx}} = E\{\mathbf{x}\mathbf{x}^H\}, \quad \mathcal{P}_{\mathbf{xx}} = E\{\mathbf{x}\mathbf{x}^T\} \tag{1}$$

where $\mathcal{C}_{\mathbf{xx}}$ and $\mathcal{P}_{\mathbf{xx}}$ are called respectively the *covariance matrix* and *pseudo-covariance matrix* [7]. In order to allow for all the information available within the complex RV to be used, the augmented $2m \times 1$ complex vector $\mathbf{x}^a$ can be defined as

$$\mathbf{x}^a = \begin{bmatrix} \mathbf{x} \\ \mathbf{x}^* \end{bmatrix} \tag{2}$$

where $(\cdot)^*$ denotes the complex conjugate operator. Then the covariance matrix $\mathcal{C}_{\mathbf{x}^a\mathbf{x}^a}$ of the augmented vector $\mathbf{x}^a \in$

$\mathbb{C}^{2m\times 2m}$, given by

$$\mathcal{C}_{\mathbf{x}^a\mathbf{x}^a} = \begin{bmatrix} \mathcal{C}_{\mathbf{xx}} & \mathcal{P}_{\mathbf{xx}} \\ \mathcal{P}^*_{\mathbf{xx}} & \mathcal{C}^*_{\mathbf{xx}} \end{bmatrix} \tag{3}$$

contains information from both the covariance and pseudo-covariance matrices of $\mathbf{x}^a$ [6]. When $\mathcal{P}_{\mathbf{xx}} = \mathbf{0}$, the RV is called *circular* (proper) [5, 6]. However in most applications, it is usually implicitly assumed that the pseudo-covariance matrix $\mathcal{P}_{\mathbf{xx}}$ is zero, thus resulting in undermodelling. We next show that by using the widely linear (WL) model [8], given by

$$y = \mathbf{h}^T\mathbf{x} + \mathbf{g}^T\mathbf{x}^* \tag{4}$$

it is possible to design adaptive filters suitable for general complex processes (both circular and non–circular).

3. THE AUGMENTED CLMS ALGORITHM

Using the augmented statistics, the output $y(k)$ of an FIR filter can be written as a widely linear process, given by

$$y(k) = \mathbf{h}^T(k)\mathbf{x}(k) + \mathbf{g}^T(k)\mathbf{x}^*(k) \tag{5}$$

where $\mathbf{h}(k)$ and $\mathbf{g}(k)$ are adaptive weight vectors, $\mathbf{x}(k)$ is the filter input, and the weights are updated by minimising the cost function

$$E(k) = \frac{1}{2}|e(k)|^2 = \frac{1}{2}|d(k) - y(k)|^2 = \frac{1}{2}e(k)e^*(k) \tag{6}$$

where $d(k)$ is the desired signal. Using the stochastic gradient based adaptation, we have

$$\mathbf{h}(k+1) = \mathbf{h}(k) - \mu\nabla E(k)\Big|_{h=h(k)} \tag{7}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) - \mu\nabla E(k)\Big|_{g=g(k)} \tag{8}$$

and

$$\nabla E(k)\Big|_{h=h(k)} = \frac{\partial E(k)}{\partial h_n^r(k)} + \jmath\frac{\partial E(k)}{\partial h_n^i(k)} \tag{9}$$

$$\nabla E(k)\Big|_{g=g(k)} = \frac{\partial E(k)}{\partial g_n^r(k)} + \jmath\frac{\partial E(k)}{\partial g_n^i(k)} \tag{10}$$

In this setting, μ is the step size, $(\cdot)^r$ and $(\cdot)^i$ denote respectively the real and imaginary part of a complex number and n denotes the n^{th} element of the weight vector. Since the input to the filter is complex, the error $e(k)$ is also complex and therefore the gradients from (9) and (10) should be evaluated as

$$\frac{\partial E(k)}{\partial h_n^r(k)} = \left\{e(k)\frac{\partial e^*(k)}{\partial h_n^r(k)} + e^*(k)\frac{\partial e(k)}{\partial h_n^r(k)}\right\} \tag{11}$$

$$\frac{\partial E(k)}{\partial h_n^i(k)} = \left\{e(k)\frac{\partial e^*(k)}{\partial h_n^i(k)} + e^*(k)\frac{\partial e(k)}{\partial h_n^i(k)}\right\} \tag{12}$$

$$\frac{\partial E(k)}{\partial g_n^r(k)} = \left\{e(k)\frac{\partial e^*(k)}{\partial g_n^r(k)} + e^*(k)\frac{\partial e(k)}{\partial g_n^r(k)}\right\} \tag{13}$$

$$\frac{\partial E(k)}{\partial g_n^i(k)} = \left\{e(k)\frac{\partial e^*(k)}{\partial g_n^i(k)} + e^*(k)\frac{\partial e(k)}{\partial g_n^i(k)}\right\} \tag{14}$$

Rewriting (6) in terms of its real and imaginary parts and substituting in (11)–(14) yields

$$\nabla E(k)\Big|_{h=h(k)} = -e(k)\mathbf{x}^*(k) \tag{15}$$

$$\nabla E(k)\Big|_{g=g(k)} = -e(k)\mathbf{x}(k) \tag{16}$$

The weight update equations (7) and (8) are now given as

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e(k)\mathbf{x}^*(k) \tag{17}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \mu e(k)\mathbf{x}(k) \tag{18}$$

In order to consolidate (17)–(18) into a compact vector form, we define the augmented weight vector $\mathbf{w}^a(k)$ as

$$\mathbf{w}^a(k) = [\mathbf{h}^T(k),\, \mathbf{g}^T(k)]^T \tag{19}$$

to give the augmented weight update

$$\mathbf{w}^a(k+1) = \mathbf{w}^a(k) + \mu e^a(k)(\mathbf{x}^a)^*(k) \tag{20}$$

where

$$e^a(k) = d(k) - \underbrace{(\mathbf{x}^a)^T(k)\mathbf{w}^a(k)}_{y(k)}, \tag{21}$$

$$\mathbf{x}^a(k) = [\mathbf{x}^T(k),\, \mathbf{x}^H(k)]^T \tag{22}$$

This concludes the derivation of the augmented CLMS (ACLMS) algorithm.

4. PERFORMANCE OF THE ACLMS

The advantage of the ACLMS algorithm over the standard CLMS is in the utilisation of the full second order statistical information available within the signal, achieved through WL modelling. For circular signals, where the pseudo-covariance is zero, it is anticipated that both algorithms will perform well, while ACLMS is expected to outperform the CLMS when applied to non–circular (improper) data. To demonstrate this, we used complex autoregressive $AR(4)$ process and real–world complex–valued wind signals. The performance was assessed based on the prediction gain R_p given by [12]

$$R_p \triangleq 10\log_{10}\left(\frac{\sigma_x^2}{\hat{\sigma}_e^2}\right) \text{[dB]} \tag{23}$$

where σ_x^2 denotes the variance of the input signal $\mathbf{x}(k)$, whereas $\hat{\sigma}_e^2$ denotes the estimated variance of the forward prediction error $\{e(k)\}$.

4.1. Prediction of Complex–Valued Autoregressive Signal

In the first experiment, we used a synthesised stable and circular complex–valued $AR(4)$ process used is given by

$$\begin{aligned} x(k) = 1.79x(k-1) - 1.85x(k-2) + 1.27x(k-3) \\ - 0.41x(k-4) + n(k) \end{aligned} \tag{24}$$

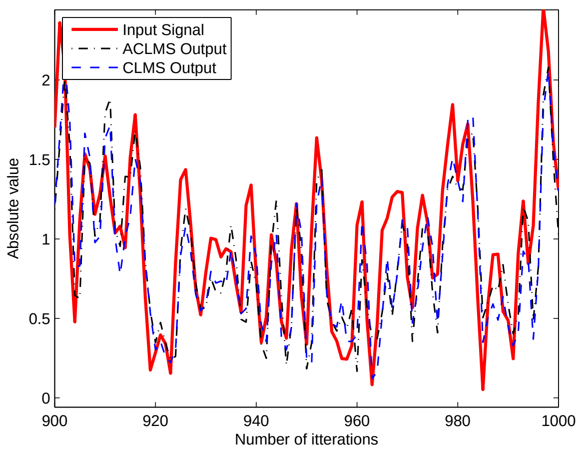

Fig. 1. The input and predicted signals obtained by using the CLMS (dash) and ACLMS (dot–dash).

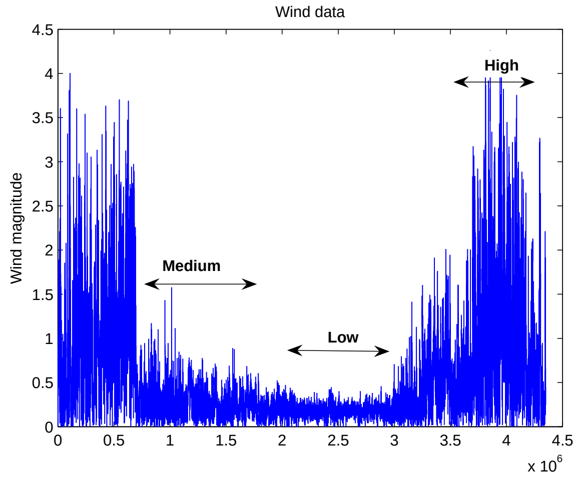

Fig. 2. Complex wind signal magnitude. Three wind speed regions have been identified as low, medium and high.

where $n(k) = n_r(k) + \jmath n_i(k)$ is a complex white Gaussian noise (CWGN), such that the real and imaginary parts are independent real WGN sequences $\sim \mathcal{N}(0,1)$ and $\sigma_n^2 = \sigma_{n_r}^2 + \sigma_{n_i}^2$.

The adaptive filter with $N = 10$ was trained using 1000 samples of $x(k)$, the step–size $\mu = 0.01$ was kept constant for both the algorithms. The obtained prediction gains were $R_{p,CLMS}$ = 3.22 dB and $R_{p,ACLMS}$ = 3.99 dB. Figure 1 demonstrates the convergence of the predicted signal to the original, which has been zoomed in for better clarity. The quantitative performances of both algorithms were adequate, with similar values of R_p. This was expected, since the $AR(4)$ signal is circular and there is no information available in the pseudo-covariance matrix to facilitate the performance of the ACLMS.

4.2. Prediction of Complex–Valued Wind Using ACLMS

Wind field was measured using an ultrasonic anemometer[1] over a period of 24 hours sampled at 50 Hz. A moving average filter was used to reduce the effects of high frequency noise; the signal was then resampled at 1 Hz. The window size w_F of the moving average filter varied according to

$$w_F = \{1, 2, 10, 20, 60\}, \tag{25}$$

where the window size is given in seconds.

The wind speed reading were taken in the north–south (V_N) and east–west (V_E) direction, which was used to create the complex wind signal $V = v \cdot e^{\imath\varphi}$, as

$$v = \sqrt{V_E^2 + V_N^2}\,, \quad \varphi = \arctan\left(\frac{V_N}{V_E}\right) \tag{26}$$

[1]Recorded in an urban environment at the Institute of Industrial Science, University of Tokyo, Japan

Based on the modulus of the complex wind data dynamics, we identified and labelled changes in the wind intensity as regions *high*, *medium* and *low* on Figure 2. To investigate the advantage of WL modelling for such intermittent and non–circular complex data, 5000 samples were taken from each region to train CLMS and ACLMS adaptive predictors for one step ahead prediction, and simulations results are shown in Figure 3.

It is evident that the ACLMS algorithm has provided better predictions compared to the CLMS algorithm in all the three considered regions. It is also seen that the best prediction was obtained for the *high* region where the wind speed had strongest variations, giving a maximum prediction gain of 16.20 dB. The performances of the two algorithms showed significant improvement for the *medium* region. This can also be seen in Figure 4 where after 5000 iterations, the ACLMS algorithm outperformed the CLMS algorithm, converged faster, and was able to track the dynamics of the input better.

Complex–valued wind is a non–circular signal, and clearly the use of augmented statistics helped to extract the full second order statistical information available within the data. The results of the ACLMS prediction clearly indicate the benefits of using augmented statistics for non–circular complex–valued data, resulting in faster convergence and better prediction performance.

5. CONCLUSION

A stochastic gradient–based algorithm for complex–valued adaptive filtering, which utilises some recent advances in complex statistics has been introduced. It has been shown that the second order statistics of an augmented random complex vector provides a mathematical model for enhancing the performance of adaptive filters. The so called augmented complex LMS (ACLMS) algorithm has been derived by considering

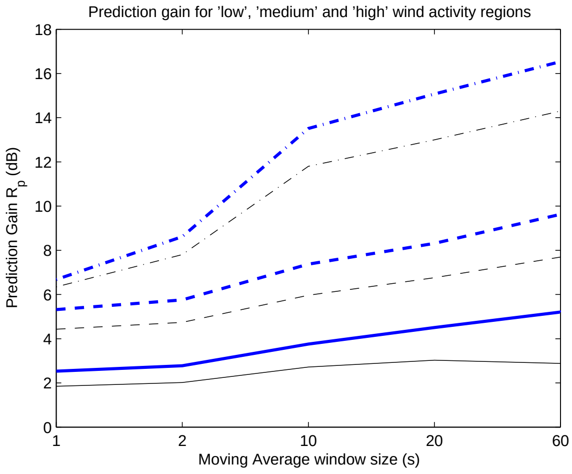

Fig. 3. Prediction gain of the ACLMS (thick lines) and CLMS (thin lines) algorithms in the *low* (solid), *medium* (dashed) and *high* (dot-dash) regions

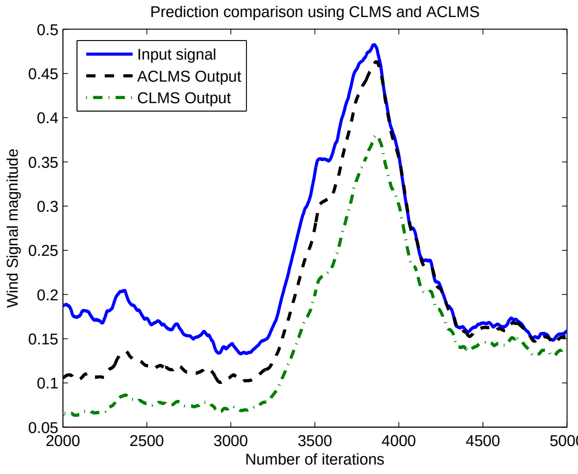

Fig. 4. Input and predicted signal of the *medium* region, comparing the performance of the ACLMS and CLMS after 5000 iterations (zoomed area).

both the pseudo-covariance matrix as well as the covariance matrix of the widely linear model.

Two sets of complex–valued data were used to illustrate the performance of the ACLMS, the circular $AR(4)$ process and a non–circular real–world wind. It has been shown that while the performances of the CLMS and ACLMS were relatively similar for circular data, the ACLMS outperformed the standard CLMS on non–circular data.

6. ACKNOWLEDGEMENT

We would like to thank Prof. Kazuyuki Aihara, Institute of Industrial Science, University of Tokyo, Japan for providing the wind data.

7. REFERENCES

[1] S Haykin, *Adaptive Filter Theory*, Prentice Hall, 1996.

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

[3] D. Mandic, P. Vayanos, C. Boukis, B. Jelfs, S.L. Goh, T. Gautama, and T. Rutkowski, “Collaborative adaptive learning using hybrid filters,” in *ICASSP 2007*, 2007, vol. 3, pp. 921–924.

[4] D.P. Mandic, S. Javidi, G. Souretis, and S.L. Goh, “Why a complex valued solution for a real domain problem,” in *Proceedings of the 17th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing*, 2007.

[5] B. Picinbono, “On circularity,” *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473–3482, 1994.

[6] P.J. Schreier and L.L. Scharf, “Second-order analysis of improper complex random vectors and processes,” *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 714–725, 2003.

[7] 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.

[8] B. Picinbono and P. Bondon, “Second-order statistics of complex signals,” *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 411–420, 1997.

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

[10] R. Schober, W.H. Gerstacker, and L.H.-J. Lampe, “A widely linear LMS algorithm for MAI suppression for DS-CDMA,” *IEEE International Conference on Communications, 2003. ICC ’03*, vol. 4, pp. 2520–2525, 2003.

[11] 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 Transactions on Signal Processing*, vol. 52, no. 3, pp. 746–756, 2004.

[12] S. Haykin and Liang Li, “Nonlinear adaptive prediction of nonstationary signals,” *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 526–535, 1995.