DUALITY BETWEEN WIDELY LINEAR AND DUAL CHANNEL ADAPTIVE FILTERING

Danilo P. Mandic[1], *Susanne Still*[2], *and Scott C. Douglas*[3]

[1]Electrical and Electronic Engineering Department, Imperial College London SW7 2AZ, UK,
[2]Department of Information and Computer Sciences Honolulu, HI 96822, USA
[3]Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275-0338, USA
E-mail: d.mandic@ic.ac.uk, sstill@hawaii.edu, douglas@engr.smu.edu

ABSTRACT

We address the duality between adaptive filtering in $\mathbb{C}$ and $\mathbb{R}^2$ and provide a comparison between the well understood dual channel real valued least mean square (DCRLMS) algorithm in $\mathbb{R}^2$ and the corresponding algorithms in $\mathbb{C}$. These include the complex LMS (CLMS) and the recently introduced augmented CLMS (ACLMS), a widely linear algorithm designed for the processing of noncircular complex valued signals. The analysis shows that the standard CLMS and DCRLMS in general provide different adaptive filtering solutions, whereas the ACLMS and DCRLMS are isomoprhic and can be made equivalent. The analysis is supported by simulations on noncircular real world signals.

***Index Terms*—** Widely linear modelling, augmented complex least mean square (ACLMS), bivariate least mean square.

1. INTRODUCTION

The generality of complex valued processes considered in statistical signal processing are

- Complex by design, such as the symbols used in communications, which are typically located equidistantly on the unit circle in the z plane;
- Complex by convenience of representation (radar, sonar, wind field) where two variables of different natures (speed and direction in the case of wind) are combined into a more compact representation.

Signals which are complex by design are usually circular, that is, with rotation invariant distributions, and for their processing there is a wealth of statistical signal processing tools. On the other hand, processes made complex by convenience of representation come from real world, and it is unlikely that their statistics will obey standard distributions. Tools for the processing of such noncircular signals are only emerging [1, 2]. It is therefore natural to ask whether it is more convenient to process such signals as two dimensional real valued vectors, as this is much better understood.

The duality between two-dimensional real valued vectors and complex numbers is usually addressed through the isomorphism between the fields $\mathbb{R}^2$ and $\mathbb{C}$. The one-to-one mapping between a point in the complex plane $x + \jmath y \in \mathbb{C}$ and a point $(x, y) \in \mathbb{R}^2$, can be expressed as

$$z=[1 \;\; \jmath]\begin{bmatrix} x \\ y \end{bmatrix} \;\; and \;\; \begin{bmatrix} z \\ z^* \end{bmatrix}=\begin{bmatrix} 1 & \jmath \\ 1 & -\jmath \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \qquad (1)$$

where the complex variable $z^a = [z, z^*]^T$ is called the "augmented" complex variable.

These mappings can be used to establish the relationship between the statistics in $\mathbb{R}^2$ and $\mathbb{C}$, and statistical signal processing techniques operating in $\mathbb{C}$ have been considered as generalizations of the corresponding techniques in $\mathbb{R}^2$ [3]. Thus, for instance, probability distributions of complex variables have been derived in terms of the corresponding distributions of "composite" real variables $w = (x, y) \in \mathbb{R}^2$, whereas for a (real or complex) column vector $\mathbf{z}$, its covariance is given by $E[\mathbf{z}\mathbf{z}^T]$ in the real case and $E[\mathbf{z}\mathbf{z}^H]$ in the complex case.

Recent development in the statistics of complex variable (called the augmented complex statistics) show that treating probability distributions in $\mathbb{C}$ as simple generalizations of the corresponding distributions in $\mathbb{R}$ is not adequate [4], and for complete second order statistical description both the covariance $\mathcal{C} = E[\mathbf{z}\mathbf{z}^H]$ and pseudocovariance $\mathcal{P} = E[\mathbf{z}\mathbf{z}^T]$ should be taken into account, to give the augmented covariance matrix [5, 6]

$$\mathcal{C}_a = cov\left(\mathbf{z}^a(k)\right) = E\left[\mathbf{z}^a(k)\mathbf{z}^{aH}(k)\right] = \begin{bmatrix} \mathcal{C} & \mathcal{P} \\ \mathcal{P}^* & \mathcal{C}^* \end{bmatrix} \qquad (2)$$

Consequently, linear stochastic models based on augmented complex statistics should be linear in both z and z^* and are termed widely linear stochastic models. One such model is the widely linear autoregressive (AR) model, given by

$$y(k) = \sum_{i=1}^{p} h_i z(k-i) + \sum_{i=1}^{p} g_i z^*(k-i) + n(k) \qquad (3)$$

where h_i and g_i are model coefficients and n is doubly white complex Gaussian noise. Based on this model the augmented complex least mean square (ACLMS) algorithm has been developed for linear adaptive filtering of noncircular signals, and has found applications in communications and wind forecasting [7, 8, 9, 10].

As multichannel LMS algorithms are a standard in multichannel adaptive filtering, our aim is to establish a correspondence between a dual channel real LMS (DCRLMS) and the ACLMS, and thus provide insight into the properties of ACLMS. The analysis also illustrates the duality between the processing in $\mathbb{R}^2$ and $\mathbb{C}$.

2. FILTERING OF TWO–DIMENSIONAL SIGNALS

For convenience, consider the operation of a dual channel linear adaptive filter in the prediction setting

$$\begin{aligned}\hat{x}(k) &= \mathbf{a}^T(k)\mathbf{x}(k) + \mathbf{b}^T(k)\mathbf{y}(k)\\ \hat{y}(k) &= \mathbf{c}^T(k)\mathbf{x}(k) + \mathbf{d}^T(k)\mathbf{y}(k)\end{aligned} \tag{4}$$

where $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^{L\times 1}$ are column vectors of filter coefficients, L denotes the filter length, $\hat{x}(k)$ and $\hat{y}(k)$ are the predictions of channels $x(k)$ and $y(k)$ and

$$\begin{aligned}\mathbf{x}(k) &= [x(k-1), \dots, x(k-L)]^T\\ \mathbf{y}(k) &= [y(k-1), \dots, y(k-L)]^T\end{aligned}$$

are the past samples from the x and y channel contained in the filter memory.

On the other hand, the operation of a complex linear adaptive filter is described by

$$\begin{aligned}\hat{x}(k) &= \mathbf{h}_r^T(k)\mathbf{x}(k) - \mathbf{h}_i^T(k)\mathbf{y}(k)\\ \hat{y}(k) &= \mathbf{h}_i^T(k)\mathbf{x}(k) + \mathbf{h}_r^T(k)\mathbf{y}(k)\end{aligned} \tag{5}$$

where $\mathbf{z}(k) = \mathbf{x}(k) + \jmath\mathbf{y}(k)$, $\mathbf{h} = \mathbf{h}_r(k) + \jmath\mathbf{h}_i(k) \in \mathbb{C}^{L\times 1}$ is a column vector of filter coefficients, and subscripts $(\cdot)_r$ and $(\cdot)_i$ denote respectively the real and imaginary part of a complex quantity.

Based on (3), the operation of a widely linear adaptive filter is described by $\hat{z}(k) = \mathbf{h}^T(k)\mathbf{z}(k) + \mathbf{g}^T(k)\mathbf{z}^*(k) = \mathbf{q}^T(k)\mathbf{z}^a(k)$ or in an expanded form

$$\begin{aligned}\hat{x}(k) &= \big(\mathbf{h}_r(k) + \mathbf{g}_r(k)\big)^T\mathbf{x}(k) + \big(\mathbf{g}_i(k) - \mathbf{h}_i(k)\big)^T\mathbf{y}(k)\\ \hat{y}(k) &= \big(\mathbf{h}_i(k) + \mathbf{g}_i(k)\big)^T\mathbf{x}(k) + \big(\mathbf{h}_r(k) - \mathbf{g}_r(k)\big)^T\mathbf{y}(k)\end{aligned} \tag{6}$$

From (4) and (5), the I/O relations of a dual channel adaptive filter and a standard complex adaptive filter are equivalent for

$$\begin{aligned}\mathbf{a}(k) &= \mathbf{h}_r(k) \qquad & \mathbf{b}(k) &= -\mathbf{h}_i(k)\\ \mathbf{c}(k) &= \mathbf{h}_i(k) \qquad & \mathbf{d}(k) &= \mathbf{h}_r(k)\end{aligned} \tag{7}$$

This is also clear from the isomorphism of $\mathbb{R}^2$ and $\mathbb{C}$ in (1). For fixed coefficient vectors, the standard complex valued filter can therefore be considered a constrained version of the dual channel real filter.

From (4) and (6), the I/O relations of a dual channel real adaptive filter and a widely linear complex adaptive filter are identical for

$$\begin{aligned}\mathbf{a}(k) &= \mathbf{h}_r(k) + \mathbf{g}_r(k) \qquad & \mathbf{b}(k) &= \mathbf{g}_i(k) - \mathbf{h}_i(k)\\ \mathbf{c}(k) &= \mathbf{h}_i(k) + \mathbf{g}_i(k) \qquad & \mathbf{d}(k) &= \mathbf{h}_r(k) - \mathbf{g}_r(k)\end{aligned} \tag{8}$$

or eqivalently

$$\begin{aligned}\mathbf{h}_r(k) &= \frac{1}{2}\left[\mathbf{a}(k) + \mathbf{d}(k)\right] \qquad & \mathbf{h}_i(k) &= \frac{1}{2}\left[\mathbf{c}(k) - \mathbf{b}(k)\right]\\ \mathbf{g}_r(k) &= \frac{1}{2}\left[\mathbf{a}(k) - \mathbf{d}(k)\right] \qquad & \mathbf{g}_i(k) &= \frac{1}{2}\left[\mathbf{c}(k) + \mathbf{b}(k)\right]\end{aligned} \tag{9}$$

We shall now establish the duality between the corresponding stochastic gradient learning algorithms for the real and complex adaptive filters considered.

3. DYNAMICS OF THE LEARNING ALGORITHMS

In the prediction setting, the teaching signals for the x and y channel are respectively $d_x(k) = x(k)$ and $d_y(k) = y(k)$ and the output errors at the x and y channel of the filter are defined as

$$\begin{aligned}e_x(k) &= x(k) - \mathbf{a}^T(k)\mathbf{x}(k) - \mathbf{b}^T(k)\mathbf{y}(k)\\ e_y(k) &= y(k) - \mathbf{c}^T(k)\mathbf{x}(k) - \mathbf{d}^T(k)\mathbf{y}(k)\end{aligned} \tag{10}$$

When it comes to the Wiener solution, the optimal weight vector for the dual channel real filter is calculated based on the $2L \times 2L$ correlation matrix

$$\mathbf{R} = E\left\{\begin{bmatrix} \mathbf{x}(k)\mathbf{x}^T(k) & \mathbf{x}(k)\mathbf{y}^T(k)\\ \mathbf{y}(k)\mathbf{x}^T(k) & \mathbf{y}(k)\mathbf{y}^T(k)\end{bmatrix}\right\} \tag{11}$$

The augmented correlation matrix $\mathcal{C}_a$ in (2) is also $2L \times 2L$ dimensional, but with complex coefficients, whereas the correlation matrix of a standard complex filter is $L \times L$ dimensional and with complex coefficients.

3.1. Weight updates

Within the stochastic gradient adaptive filtering setting, the cost function for the dual channel real valued least mean square (DCRLMS) algorithm is given by

$$J = J\left(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\right) = \frac{1}{2}\big(e_x^2(k) + e_y^2(k)\big) \tag{12}$$

and is equivalent to the cost function for the CLMS and ACLMS

$$J = J(\mathbf{h}, \mathbf{g}) = \frac{1}{2}e(k)e^*(k) = \frac{1}{2}\big(e_r^2(k) + e_i^2(k)\big) \tag{13}$$

where $e_x(k)$ and $e_y(k)$ are given in (10) and $e_r(k)$ and $e_i(k)$ are the real and imaginary components of the output error of complex filters $e(k) = e_r(k) + \jmath e_i(k)$. For the complex filters, for convenience denote $e_x(k) = e_r(k)$ and $e_y(k) = e_i(k)$. Since neither $\mathbf{x}(k)$ or $\mathbf{y}(k)$ are generated through the filter, coefficient updates of the dual channel real LMS (DCRLMS) algorithm are calculated similarly to the standard LMS and are given by [11]

$$\begin{aligned}\mathbf{a}(k+1) &= \mathbf{a}(k) + \Delta\mathbf{a}(k) = \mathbf{a}(k) + \mu e_x(k)\mathbf{x}(k)\\ \mathbf{b}(k+1) &= \mathbf{b}(k) + \Delta\mathbf{b}(k) = \mathbf{b}(k) + \mu e_x(k)\mathbf{y}(k)\\ \mathbf{c}(k+1) &= \mathbf{c}(k) + \Delta\mathbf{c}(k) = \mathbf{c}(k) + \mu e_y(k)\mathbf{x}(k)\\ \mathbf{d}(k+1) &= \mathbf{d}(k) + \Delta\mathbf{d}(k) = \mathbf{d}(k) + \mu e_y(k)\mathbf{y}(k)\end{aligned} \tag{14}$$

CLMS vs DCRLMS. We can express the standard complex least mean square algorithm, given by [12]

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \mu e(k)\mathbf{z}^*(k) \tag{15}$$

in the same form as the DCRLMS (14), to give

$$\begin{aligned}\mathbf{h}_r(k+1) &= \mathbf{h}_r(k) + \mu\left[e_x(k)\mathbf{x}(k) + e_y(k)\mathbf{y}(k)\right]\\ \mathbf{h}_i(k+1) &= \mathbf{h}_i(k) + \mu\left[e_y(k)\mathbf{x}(k) - e_x(k)\mathbf{y}(k)\right]\end{aligned} \tag{16}$$

that is, unlike the channels $\mathbf{a}(k), \ldots, \mathbf{d}(k)$ within DCRLMS, the real and imaginary parts of the filter coefficient vector, $\mathbf{h}_r(k)$ and $\mathbf{h}_i(k)$, are updated based on both the errors from the x and y channels and the tap inputs $\mathbf{x}(k)$ and $\mathbf{y}(k)$. Denote the learning rate for the DRCLMS by μ_R and the learning rate for CLMS by μ_C, then from (14)–(16) and by taking into account (7), we have[1]

$$\begin{aligned} \Delta\mathbf{h}_r(k) &= 2\frac{\mu_C}{\mu_R}\Delta\mathbf{h}_r(k) \\ \Delta\mathbf{h}_i(k) &= 2\frac{\mu_C}{\mu_R}\Delta\mathbf{h}_i(k) \end{aligned} \tag{17}$$

The CLMS and DCRLMS are therefore equivalent only when the rather stringent relation (7) is satisfied and the learning rate of CLMS is set to half the rate of DCRLMS.

ACLMS vs DCLMS. The weight updates for the ACLMS algorithm are given by [7, 9, 10]

$$\begin{aligned} \mathbf{h}(k+1) &= \mathbf{h}(k) + \mu_h e(k)\mathbf{z}^*(k) \\ \mathbf{g}(k+1) &= \mathbf{g}(k) + \mu_g e(k)\mathbf{z}(k) \end{aligned} \tag{18}$$

For $\mu_h = \mu_g = \mu_C$, the ACLMS update can be cast into the same form as the updates for DCRLMS (14), to give

$$\begin{aligned} \mathbf{h}_r(k+1) &= \mathbf{h}_r(k) + \mu_C\big[e_x(k)\mathbf{x}(k) + e_y(k)\mathbf{y}(k)\big] \\ \mathbf{h}_i(k+1) &= \mathbf{h}_i(k) + \mu_C\big[e_y(k)\mathbf{x}(k) - e_x(k)\mathbf{y}(k)\big] \\ \mathbf{g}_r(k+1) &= \mathbf{g}_r(k) + \mu_C\big[e_x(k)\mathbf{x}(k) - e_y(k)\mathbf{y}(k)\big] \\ \mathbf{g}_i(k+1) &= \mathbf{g}_i(k) + \mu_C\big[e_y(k)\mathbf{x}(k) + e_x(k)\mathbf{y}(k)\big] \end{aligned} \tag{19}$$

From (8), (9), (14) and (19), and by expressing the terms in the square brackets via the updates of the coefficient vectors within the DRCLMS, we can see that, for instance

$$\Delta\mathbf{h}_r(k) = \frac{\mu_C}{\mu_R}\big(\Delta\mathbf{a}(k) + \Delta\mathbf{d}(k)\big) = 2\frac{\mu_C}{\mu_R}\Delta\mathbf{h}_r(k)$$

The dual channel real adaptive filter and the widely linear complex adaptive filter, trained with the corresponding learning algorithms DCRLMS and ACLMS, are therefore equivalent when the stepsize of the DCRLMS is twice the stepsize of ACLMS, that is the ACLMS is twice faster than the DCRLMS. We can conclude that in the stochastic gradient setting, widely linear complex valued adaptive filters are isomorphic to dual channel real valued adaptive filters.

4. SIMULATIONS

To support the findings, simulations were conducted for a linear stable circular complex AR(4) process, noncircular and nonlinear complex Ikeda map, and real world noncircular complex wind signal. The AR(4) process was generated based on

$$\begin{aligned} z(k) &= 1.79z(k-1) - 1.85z(k-2) \\ &\quad +1.27z(k-3) - 0.41z(k-4) + n(k) \end{aligned} \tag{20}$$

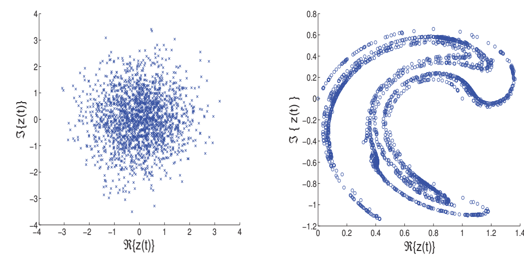
[1]From (14) we have $e_x(k)\mathbf{x}(k) = \Delta\mathbf{a}(k)/\mu_R$, $e_x(k)\mathbf{y}(k) = \Delta\mathbf{b}(k)/\mu_R$, $e_y(k)\mathbf{x}(k) = \Delta\mathbf{c}(k)/\mu_R$,and $e_y(k)\mathbf{y}(k) = \Delta\mathbf{d}(k)/\mu_R$.

where $n(k)$ is complex, doubly white Gaussian noise with variance $\sigma^2 = 1$. The Ikeda map chaotic signal is given by

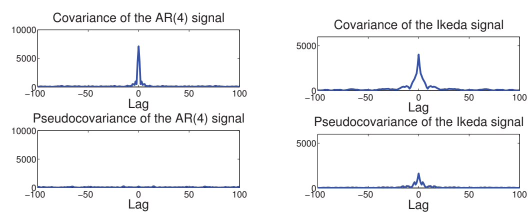
$$\begin{aligned} x(k+1) &= 1 + u\left(x(k)\cos[t(k)] - y(k)\sin[t(k)]\right) \\ y(k+1) &= u\left(x(k)\sin[t(k)] + y(k)\cos[t(k)]\right) \end{aligned} \tag{21}$$

where u is a parameter, typically $u = 0.8$, and $t(k) = 0.4 - \frac{6}{1+x^2(k)+y^2(k)}$. The wind signal was recorded by Windsonic, a 2D ultrasonic anemometer produced by Gill Instruments. All the test signals were made complex, for instance, in the case of wind $\mathbf{v} = ve^{\jmath\Phi}$, where v denotes the wind speed and Φ the direction.

Properties of the AR(4) and Ikeda test signals in terms of complex circularity are illustrated in Figure 1(a) and Figure 1(b), whereas their corresponding covariance and pseudocovariance functions are given in Figure 1(c) and Figure 1(d). Observe the circularly symmetric shape of the distribution



(a) Complex AR(4) process (b) Complex Ikeda map



(c) Complex AR(4) process (d) Complex Ikeda map

Fig. 1. Top: Circularity via "real–imaginary" scatter plots. *(a):* circular AR(4) signal; *(b):* noncircular chaotic Ikeda signal. Bottom: Noncircularity via pseudocovariance *(c)*: circular AR(4) signal; *(d):* noncircular chaotic Ikeda signal.

for the AR(4) signal or equivalently the vanishing pseudocovariance. The wind segment considered was also noncircular, with non–zero pseudocovariance; for more detail see [7].

Simulations were performed in a one step ahead prediction setting and the performance measure was the standard prediction gain

$$R_p = 10\log\frac{\sigma_z^2}{\sigma_e^2} \tag{22}$$

where σ_z^2 and σ_e^2 are respectively the estimated variance of the output and the prediction error. Table 1 summarizes the prediction gains for the above classes of signals. It can be seen that, since standard CLMS is designed for the adaptive

Table 1. Prediction gains R_p for the various classes of signals

Algorithm	AR4	Ikeda	Wind
DCRLMS	5.8423	3.9733	13.2604
CLMS	6.6380	2.4278	14.2941
ACLMS	6.6096	4.0330	14.8926
DCRLMS (double μ)	6.6096	4.0330	14.8926

filtering of circular complex signals, its performance for the AR(4) process was as good as that of ACLMS, whereas its performance for the noncircular Ikeda signal and the complex wind signal was worse than that of the widely linear ACLMS. When the same learning rate was used among all the algorithms, in general, the DCRLMS did not have advantage over the complex valued algorithms. However, when the DCRLMS had the learning rate twice the size of the learning rate of the complex algorithms, its performance was identical to that of ACLMS. The ALCMS outperformed CLMS for the noncircular Ikeda and wind signals, as by design, it accounts for complex noncircularity.

5. DISCUSSION AND CONCLUSIONS

The analysis and simulations have shown that

- The dual channel real least mean square (DCRLMS) and the augmented complex LMS (ACLMS) are isomorphic and provide a different filtering solution to that obtained by CLMS. The adaptive widely linear complex filter and the dual channel real filter are identical when the learning rate of DCRLMS is twice the size of the learning rate of ACLMS;

- The CLMS has half the number of coefficients as compared to DCRLMS and ACLMS; it therefore converges faster, however, its optimal Wiener solution is different from that for ACLMS and DCRLMS;

- DCRLMS simplifies into CLMS when the constraints $\mathbf{a} = \mathbf{d}$ and $\mathbf{b} = -\mathbf{c}$ are imposed on the parameters (see equation (7)), while ACLMS degenerates into CLMS for circular data when $\mathbf{g} = \mathbf{0}$.

The overall conclusion of this work is that although standard adaptive filtering algorithms in $\mathbb{R}^2$ and $\mathbb{C}$, that is CLMS and DCRLMS, provide the same solutions only as a special case, the stochastic gradient adaptive filtering algorithm based on augmented complex statistics (ACLMS) gives effectively the same solution as the DCRLMS, the difference being the scaling between the corresponding learning rates.

Acknowledgements: We wish to thank Mr Yili Xia for helping out with the simulations. The wind measurements were conducted with the help of Gill Instruments.

6. REFERENCES

[1] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*. Cambridge University Press, 2009.

[2] J. Eriksson and V. Koivunen, “Complex random vectors and ICA models: Identifiability, uniqueness, and separability,” *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1017–1029, 2006.

[3] S. Haykin, *Adaptive Filter Theory*, 4th ed. Prentice-Hall, 2002.

[4] A. V. D. Bos, “The multivariate complex normal distribution: A generalization,” *IEEE Transactions on Information Theory*, vol. 41, no. 2, pp. 537–539, 1995.

[5] B. Picinbono and P. Chevalier, “Widely Linear Estimation with Complex Data,” *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, 1995.

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

[8] R. Schober, W. H. Gerstacker, and L. H. J. Lampe, “Data–aided and blind stochastic gradient algorithms for widely linear MMSE MAI supression for DS–CDMA,” *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 746–756, 2004.

[9] S. Javidi, M. Pedzisz, S. P. Goh, and D. P. Mandic, “The augmented complex least mean square algorithm with application to adaptive prediction problems,” in *Proceedingss of the I Cognitive Information Processing Systems Conference*, 2008, p. on CD.

[10] D. P. Mandic, S. Javidi, S. L. Goh, A. Kuh, and K. Aihara, “Complex–valued prediction of wind profile using augmented complex statistics,” *Renewable Energy*, vol. 34, no. 1, pp. 196–210, 2009.

[11] J. Benesty, T. Gansler, Y. Huang, and M. Rupp, “Adaptive algorithms for MIMO acoustic echo cancellation,” in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y. Huang and J. Benesty, Eds. Springer, 2007, pp. 119–147.

[12] B. Widrow, J. McCool, and M. Ball, “The complex LMS algorithm,” *Proceedings of the IEEE*, vol. 63, no. 3, pp. 712–720, 1975.