

# A Regularised Normalised Augmented Complex Least Mean Square Algorithm

Yili Xia <sup>#1</sup>, Soroush Javidi <sup>#2</sup>, Danilo P. Mandic <sup>#3</sup>

<sup>#</sup> *Department of Electrical and Electronic Engineering, Imperial College London  
Exhibition Road, London, SW7 2BT, U.K.*

<sup>1</sup>yili.xia06@imperial.ac.uk

<sup>2</sup>soroush.javidi@imperial.ac.uk

<sup>3</sup>danilo.mandic@imperial.ac.uk

**Abstract**—A Regularised Normalised Augmented Complex Least Mean Square (RNACLMS) algorithm is proposed for widely linear adaptive filtering in the complex domain  $\mathbb{C}$ . Based on augmented complex statistics, the RNACLMS is shown to utilise complete second order information in  $\mathbb{C}$ , thus being suitable to deal with both circular and noncircular complex signals. Furthermore, a gradient adaptive regularisation term makes the proposed algorithm exhibit enhanced robustness and convergence over the NACLMS algorithm. Simulations on circular and noncircular benchmark signals and on real-world noncircular wind signals support the analysis.

## I. INTRODUCTION

The normalised least mean square (NLMS) algorithm is a standard method used in adaptive filtering due to its stability and low computational complexity [1]. The idea behind the NLMS algorithm is to employ the norm of the input vector in the denominator of the weight update term, and hence, provide decorrelation of the input and faster convergence than least mean square (LMS) algorithm. To deal with the situations where the value of the weight update term becomes extremely large due to close-to-zero input vectors, a compensation term  $\varepsilon$ , a small positive constant, is routinely employed. However, in practice, input signals with unknown and possibly very large dynamical range, and an ill-conditioned input autocorrelation matrix make it difficult to find an optimal value of  $\varepsilon$ , especially when required to meet the conflicting requirements of fast convergence and low misadjustment. To this end, the generalized normalized gradient descent (GNGD) algorithm, a variant of the NLMS with a gradient adaptive regularisation term  $\varepsilon$ , has been proposed [2]. It has been shown that the GNGD exhibits excellent stability and is robust to initialization of the compensation term  $\varepsilon$  and learning rate  $\mu$ , and its sign variants have been introduced to reduce computational complexity [3].

In the design of complex-valued adaptive filtering algorithms, standard adaptive algorithms in complex domain  $\mathbb{C}$  are normally considered as generic extensions of the corresponding algorithms in  $\mathbb{R}$ ; this implicitly results from the second order statistics in  $\mathbb{C}$  which is normally based on the covariance matrix  $C_{zz} = E\{zz^H\}$  [4], [5], where  $\mathbf{z}$  is a complex-valued random vector with zero mean. However, recent results in the so-called augmented complex statistics show that the covariance matrix  $C_{zz}$  cannot completely describe

the second order behaviour of  $\mathbf{z}$ , and the pseudo-covariance matrix  $P_{zz} = E\{zz^T\}$  should also be taken into account when processing second order noncircular signals, for which  $P_{zz} \neq \mathbf{0}$ . To ensure that both  $C_{zz}$  and  $P_{zz}$  are taken into account, the adaptive filtering algorithm should be based on the augmented complex input vector  $\mathbf{z}^a = [\mathbf{z}^T, \mathbf{z}^H]^T$ , and its augmented covariance matrix  $C_{z^a z^a}$ , given by

$$C_{z^a z^a} = E \begin{bmatrix} \mathbf{z} \\ \mathbf{z}^* \end{bmatrix} [\mathbf{z}^H \mathbf{z}^T] = \begin{bmatrix} C_{zz} & P_{zz} \\ P_{zz}^* & C_{zz}^* \end{bmatrix} \quad (1)$$

To this end, the widely linear model [4] has been introduced into design adaptive filtering algorithms to make them suitable for processing both circular and noncircular signals. Examples include the augmented CLMS algorithm for signal prediction [6], widely linear recursive least squares (WL-RLS) for adaptive beamforming [7], and the augmented affine projection algorithm [8]. These augmented widely linear algorithms offer improvement in performance for general complex signals.

Recent analyses show that the convergence of ACLMS algorithm is critically dependent on the properties of the augmented covariance matrix and the degree of circularity of the input [9]. To enhance its convergence and stability, we employ the GNGD [2] in the ACLMS framework. We first derive the normalised ACLMS (NACLMS) algorithm and next make the regularisation term  $\varepsilon$  within NACLMS gradient adaptive. This way, the RNACLMS algorithm adapts its learning rate according to the dynamics of the input signal, with the additional adaptive term also compensating for the simplifications in the derivation of NCLMS and ill-conditioned inputs. It is shown that the performance is bounded from below by the performance of the NACLMS, whereas the proposed algorithm exhibits enhanced convergence in various adverse environments where the NACLMS diverges. Simulations in the prediction setting for signals with varying degrees of circularity support the analysis.

## II. PROPOSED RNACLMS ALGORITHM

The output  $y(k)$  of a finite impulse response (FIR) adaptive filter based on the widely linear model can be written as

$$y(k) = \underbrace{\mathbf{x}^T(k)\mathbf{h}(k)}_{\text{conventional update}} + \underbrace{\mathbf{x}^H(k)\mathbf{g}(k)}_{\text{conjugate update}} \quad (2)$$

where  $\mathbf{h}(k)$  and  $\mathbf{g}(k)$  are the  $L \times 1$  weight vectors,  $\mathbf{x}(k)$  denotes the  $L \times 1$  input vector  $\mathbf{x}(k) = [x(k-1), \dots, x(k-L)]^T$  and the error  $e(k)$  is obtained as  $e(k) = d(k) - y(k)$ , where  $d(k)$  is the desired signal. To update the conventional weight vector  $\mathbf{h}(k)$  and the conjugate weight vector  $\mathbf{g}(k)$ , based on the minimum disturbance principle, the standard constrained minimisation problem for normalised CLMS (NCLMS) algorithm is modified to cater for the widely linear model, that is

$$J(k) = \|\mathbf{h}(k+1) - \mathbf{h}(k)\|_2^2 + \|\mathbf{g}(k+1) - \mathbf{g}(k)\|_2^2 + \Re[(d(k) - \mathbf{x}^T(k)\mathbf{h}(k+1) - \mathbf{x}^H(k)\mathbf{g}(k+1))^* \cdot \lambda] \quad (3)$$

where  $\lambda$  is a Lagrange multiplier operator,  $\Re[\cdot]$  denotes the real part of complex numbers, and  $J(k)$  is the real-valued cost function of complex variable. Note that in case of widely linear modelling, the constraint becomes

$$d(k) = \mathbf{x}^T(k)\mathbf{h}(k+1) + \mathbf{x}^H(k)\mathbf{g}(k+1) \quad (4)$$

For the cost function  $J(k)$ , the maximum change in the gradient is in the direction of  $\frac{\partial J(k)}{\mathbf{g}^*(k+1)}$ . Since  $J(k)$  is real-valued and so is not directly differentiable in  $\mathbb{C}$ , we resort to the  $\mathbb{C}\mathbb{R}$  calculus. Following the so-called  $\mathbb{C}\mathbb{R}$  derivatives [10], we obtain

$$\frac{\partial J(k)}{\partial \mathbf{g}^*(k+1)} = \frac{1}{2} \left( \frac{\partial J(k)}{\partial (\mathbf{g}^*(k+1))^r} - j \frac{\partial J(k)}{\partial (\mathbf{g}^*(k+1))^i} \right) \quad (5)$$

where  $(\cdot)^r$  and  $(\cdot)^i$  denote respectively the real and imaginary part of complex gradient. To calculate the update for the conjugate weight vector  $\mathbf{g}$ , the cost function  $J(k)$  in (3) can be rewritten as

$$\begin{aligned} J(k) &= [\mathbf{h}(k+1) - \mathbf{h}(k)]^H [\mathbf{h}(k+1) - \mathbf{h}(k)] \\ &+ [\mathbf{g}(k+1) - \mathbf{g}(k)]^H [\mathbf{g}(k+1) - \mathbf{g}(k)] \\ &+ \frac{1}{2} [(d^*(k) - \mathbf{h}^H(k+1)\mathbf{x}^*(k) - \mathbf{g}^H(k+1)\mathbf{x}(k))\lambda] \\ &+ \frac{1}{2} [(d(k) - \mathbf{h}^T(k+1)\mathbf{x}(k) - \mathbf{g}^T(k+1)\mathbf{x}^*(k))\lambda^*] \end{aligned}$$

and therefore

$$\frac{\partial J(k)}{\partial (\mathbf{g}^*(k+1))^r} = \mathbf{g}(k+1) - \mathbf{g}(k) - \frac{1}{2}\mathbf{x}(k)\lambda \quad (6)$$

$$\frac{\partial J(k)}{\partial (\mathbf{g}^*(k+1))^i} = j(\mathbf{g}(k+1) - \mathbf{g}(k)) - \frac{j}{2}\mathbf{x}(k)\lambda \quad (7)$$

Finally, the derivative of cost function  $J(k)$  with respect to  $\mathbf{g}^*(k+1)$  becomes

$$\frac{\partial J(k)}{\partial \mathbf{g}^*(k+1)} = \mathbf{g}(k+1) - \mathbf{g}(k) - \frac{1}{2}\mathbf{x}(k)\lambda \quad (8)$$

Setting this derivative to zero, the update of conjugate weight vector  $\mathbf{g}(k)$  is obtained as

$$\mathbf{g}(k+1) - \mathbf{g}(k) = \frac{1}{2}\mathbf{x}(k)\lambda \quad (9)$$

In the same way, the update of conventional weight vector  $\mathbf{h}(k)$  is given by

$$\mathbf{h}(k+1) - \mathbf{h}(k) = \frac{1}{2}\mathbf{x}^*(k)\lambda \quad (10)$$

Premultiplying both sides of (9) by  $\mathbf{x}^H(k)$  and both sides of (10) by  $\mathbf{x}^T(k)$  yields

$$\begin{aligned} \mathbf{x}^H(k)(\mathbf{g}(k+1) - \mathbf{g}(k)) &= \|\mathbf{x}(k)\|_2^2 \lambda \\ \mathbf{x}^T(k)(\mathbf{h}(k+1) - \mathbf{h}(k)) &= \|\mathbf{x}(k)\|_2^2 \lambda \end{aligned}$$

Upon adding these two equations together and using the constraint (3) and the widely linear model (2), the Lagrange multiplier is obtained as

$$\lambda = \frac{e(k)}{\|\mathbf{x}(k)\|_2^2} \quad (11)$$

Thus, the NACLMS updates of the conventional weight vector  $\mathbf{h}(k)$  and the conjugate weight vector  $\mathbf{g}(k)$  become

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \frac{\mu e(k)\mathbf{x}^*(k)}{2\|\mathbf{x}(k)\|_2^2}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \frac{\mu e(k)\mathbf{x}(k)}{2\|\mathbf{x}(k)\|_2^2}$$

To preserve stability for close-to-zero input vectors, the compensation term  $\varepsilon$  is added to the above updates, giving the final weight updates of the NACLMS algorithm in the form

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \frac{\mu e(k)\mathbf{x}^*(k)}{2\|\mathbf{x}(k)\|_2^2 + \varepsilon} \quad (12)$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \frac{\mu e(k)\mathbf{x}(k)}{2\|\mathbf{x}(k)\|_2^2 + \varepsilon} \quad (13)$$

Combining  $\mathbf{h}(k)$  and  $\mathbf{g}(k)$  into an augmented weight vector  $\mathbf{w}^a(k)$ , we have

$$\mathbf{w}^a(k+1) = \mathbf{w}^a(k) + \frac{\mu e(k)\mathbf{x}^a(k)}{\|\mathbf{x}^a(k)\|_2^2 + \varepsilon} \quad (14)$$

where  $\mathbf{x}^a(k) = [\mathbf{x}^T(k), \mathbf{x}^H(k)]^T$  is the augmented input vector.

#### A. Gradient Adaptive Regularisation Factor

When dealing with signals with unknown and possibly very large dynamical range, a fixed regularisation parameter  $\varepsilon$  may lead to divergence and poor performance of the NACLMS algorithm. To this end, following the standard gradient adaptive approach [2], [11], we introduce an gradient adaptive  $\varepsilon$ , updated as

$$\varepsilon(k) = \varepsilon(k-1) - \rho \nabla_{\varepsilon(k-1)} E(k) \quad (15)$$

where  $E(k) = \frac{|e^2(k)|}{2} = \frac{e(k)e^*(k)}{2}$  is the standard cost function,  $\rho$  is step size parameter, and  $\nabla_{\varepsilon(k-1)} E(k)$  can be derived as

$$\begin{aligned} \nabla_{\varepsilon(k-1)} E(k) &= \frac{1}{2} \left[ e(k) \frac{\partial e^*(k)}{\partial \varepsilon(k-1)} + e^*(k) \frac{\partial e(k)}{\partial \varepsilon(k-1)} \right] \\ &= \Re \left[ e^*(k) \frac{\partial e(k)}{\partial \varepsilon(k-1)} \right] \end{aligned} \quad (16)$$

Since  $e(k) = d(k) - y(k)$ , using the chain rule, we have

$$\frac{\partial e(k)}{\partial \varepsilon(k-1)} = -\frac{\partial y(k)}{\partial \varepsilon(k-1)} = -\left[ \mathbf{x}^{aT}(k) \frac{\partial \mathbf{w}^a(k)}{\partial \varepsilon(k-1)} \right] \quad (17)$$

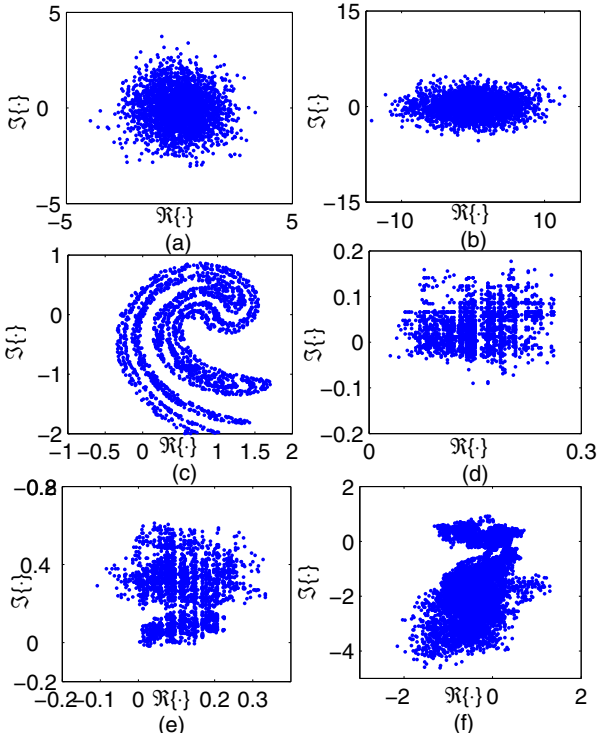


Fig. 1. Geometric view of circularity (a) AR(1); (b) Noncircular ARMA; (c) Ikeda; (d) Wind (*low*); (e) Wind (*medium*); (f) Wind (*high*).

From (14), we then obtain

$$\frac{\partial \mathbf{w}^a(k)}{\partial \varepsilon(k-1)} = -\frac{\mu e(k-1) \mathbf{x}^{a*}(k)}{[\|\mathbf{x}^a(k)\|_2^2 + \varepsilon(k-1)]^2} \quad (18)$$

to give

$$\frac{\partial e(k)}{\partial \varepsilon(k-1)} = \frac{\mu e(k-1) \mathbf{x}^{aT}(k) \mathbf{x}^{a*}(k-1)}{[\|\mathbf{x}^a(k)\|_2^2 + \varepsilon(k-1)]^2}$$

and

$$\varepsilon(k) = \varepsilon(k-1) - \frac{\rho \mu \Re[e^*(k) e(k-1) \mathbf{x}^{aT}(k) \mathbf{x}^{a*}(k-1)]}{[\|\mathbf{x}^a(k)\|_2^2 + \varepsilon(k-1)]^2}$$

The conventional NCLMS algorithm and the derivation of RNCLMS algorithm are given in the Appendix.

### III. SIMULATIONS

In the simulations, the order of the FIR adaptive filters was  $L = 10$ . To avoid close-to-zero input vector,  $\varepsilon(0)$  was set to be 0.1, whereas its step-size parameter  $\rho = 0.1$ . For generality, both circular and noncircular benchmark and real-world nonstationary and noncircular wind signals were used in simulations. The performance of algorithms on synthetic signals was produced by averaging 100 independent trails, whereas single trial experiments were performed on real-world wind signals<sup>1</sup>.

<sup>1</sup>Wind is usually measured as bivariate process of its speed and direction. In  $\mathbb{C}$ , the wind vector  $w(k)$  can be represented as  $w(k) = |v(k)|e^{j\Phi(k)}$ , where  $v(k)$  denotes the speed and  $\Phi(k)$  the direction.

The benchmark circular complex signal used in simulations was an auto regressive AR(1) complex process, given by

$$r(k) = 0.7r(k-1) + n(k) \quad (19)$$

The driving input  $n(k)$  was complex-valued doubly white Gaussian noise, with zero mean and unit variance.

The benchmark widely linear noncircular complex signal was an auto regressive moving average (ARMA) complex process, which adopt the MA model in [12] and the stable AR(1) model, given by

$$r(k) = 0.7r(k-1) + 2n(k) + 0.5n^*(k) + n(k-1) + 0.9n^*(k-1) \quad (20)$$

where  $n(k)$  was complex valued doubly white Gaussian noise.

The nonlinear and noncircular chaotic Ikeda map signal is given by [13]

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

where typically  $u = 0.7$  and  $t(k) = 0.4 - \frac{6}{1+x^2(k)+y^2(k)}$ .

The noncircular wind signals were with different dynamical characteristics, identified as regions *high*, *medium* or *low* based on the changes in the wind intensity.

Fig. 1 shows the scatter plots of the complex signals considered in simulations. Observe the circular symmetry for the AR(1) signal and the noncircularity of the ARMA model, Ikeda map and wind signals. For a quantitative measurement of the degree of noncircularity of a complex vector  $\mathbf{z}$ , we used the index  $s$  proposed in [14], defined as  $s = 1 - \det(\mathbf{C}_{\mathbf{z}^a \mathbf{z}^a}) \det^{-2}(\mathbf{C}_{\mathbf{z}\mathbf{z}})$ , where  $\det(\cdot)$  is the matrix determinant operator, and the degree of noncircularity  $s$  is normalised to within  $[0, 1]$  with the value of 0 indicating perfect circularity, and 1 maximum noncircularity. Table I illustrates the degrees of noncircularity  $s$  for the various classes of signals. Observe the excellent match between index of noncircularity in Table I and scatter plot descriptions in Fig. 1; for instance, the wind (*low*) region was least noncircular, whereas the wind (*high*) region exhibited strong noncircularity.

To illustrate the enhanced robustness and stability of employing a gradient adaptive regularisation factor, a performance comparison between the NACLMS and RNACLMS algorithms was performed on prediction of the noncircular ARMA signal (20) and Ikeda map (21). Fig. 2 shows the corresponding learning curves for various values of  $\mu$ . By design, the performance of RNACLMS was similar to that of NACLMS when NACLMS was stable, whereas RNACLMS did not diverge in cases when NACLMS did, as in the case of  $\mu = 2$ . Fig. 3 shows the learning curves of the conventional RNCLMS and RNACLMS algorithms on noncircular ARMA and Ikeda map with  $\mu = 0.1$ . Observe that both algorithms exhibited comparable rate of convergence, however, in the steady state, the RNACLMS algorithm achieved superior performance, indicating the suboptimality of standard RNCLMS algorithm when dealing with noncircular signals.

To further illustrate the advantage of using the widely linear model, in the next experiment, the prediction gain  $R_p$

TABLE I  
COMPARISON OF DEGREES OF NONCIRCULARITY  $s$  FOR THE VARIOUS CLASSES OF SIGNALS

	Circular AR(1)	Noncircular ARMA	Ikeda map	Wind ( <i>low</i> )	Wind ( <i>medium</i> )	Wind ( <i>high</i> )
$s$	0.0031	0.7220	0.9654	0.1155	0.6962	0.9009

TABLE II  
COMPARISON OF PREDICTION GAINS  $R_p$  FOR THE VARIOUS CLASSES OF SIGNALS

$R_p$ [dB]	Circular AR(1)	Noncircular ARMA	Ikeda map	Wind ( <i>low</i> )	Wind ( <i>medium</i> )	Wind ( <i>high</i> )
CLMS	2.4372	5.2089	3.0157	3.5182	6.8953	10.0822
ACLMS	2.3940	5.3920	5.1448	4.1213	7.2127	10.3517
RNCLMS	2.5943	5.2618	3.5325	3.8854	7.4442	10.1501
RNACLMS	2.5620	5.4152	5.2600	4.4291	7.7391	10.4543

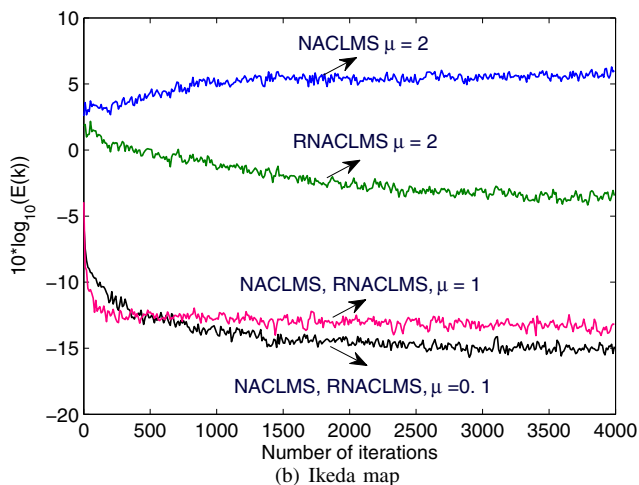
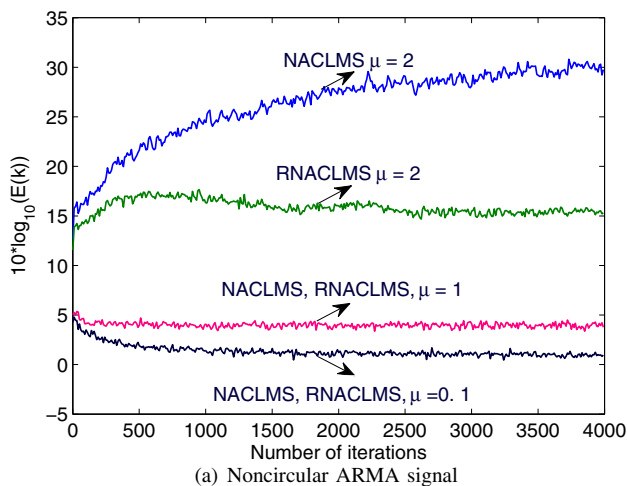


Fig. 2. Learning curves for the NACLMS and RNACLMS algorithms. (a) noncircular ARMA signal (20) and (b) Ikeda map (21) with  $\mu = 0.1, 1$  and  $2$ .

was used as an overall performance measurement, defined as  $R_p = 10 \log_{10} (\sigma_x^2 / \hat{\sigma}_e^2)$  [dB], where  $\sigma_x^2$  denotes the variance of the input signal  $x(k)$ , and  $\hat{\sigma}_e^2$  denotes the variance of the prediction error  $e(k)$  [15]. Table II shows the prediction gains  $R_p$  of the proposed RNACLMS, standard RNCLMS, ACLMS [6] and standard CLMS [16] for  $\mu = 0.1$ . As expected, for

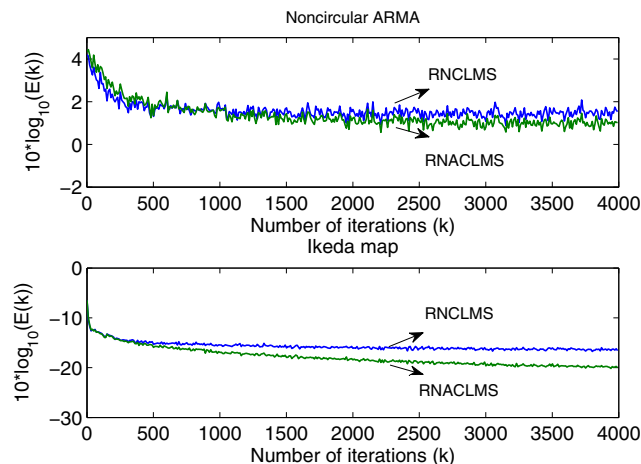


Fig. 3. Learning curves for standard RNCLMS and the proposed RNACLMS algorithms on one step ahead prediction of the noncircular ARMA signal (20) and Ikeda map (21) with  $\mu = 0.1$ .

all the noncircular signals considered, there was significant improvement when augmented algorithms were employed over that of standard algorithms, and RNACLMS always achieved best performance.

#### IV. CONCLUSIONS

A Regularised Normalised ACLMS (RNACLMS) algorithm has been introduced for adaptive filtering of general complex valued signals. It has been derived by utilising the widely linear model and a gradient adaptive regularisation factor within the NACLMS. Simulations on both synthetic and real world complex-valued signals have illustrated the advantage of RNACLMS algorithm over standard RNCLMS algorithm for adaptive filtering of noncircular signals. It also has been shown that by employing a gradient adaptive compensation term, the performance of RNACLMS algorithm is bounded from below by the performance of the NACLMS algorithm, whereas it achieves enhanced convergence and stability when NACLMS algorithm diverges.

#### ACKNOWLEDGMENT

We wish to thank Gill Instruments for providing their ultrasonic anemometers used in our wind measurements.

## V. APPENDIX

The conventional NCLMS algorithm is summarised as [17]

$$\begin{aligned} y(k) &= \mathbf{x}^T(k)\mathbf{w}(k) \\ e(k) &= d(k) - y(k) \\ \mathbf{w}(k+1) &= \mathbf{w}(k) + \frac{\mu e(k)\mathbf{x}^*(k)}{\|\mathbf{x}(k)\|_2^2 + \varepsilon} \end{aligned} \quad (22)$$

where  $\mathbf{w}(k)$  is  $L \times 1$  weight vector, and RNCLMS algorithm can be achieved by making  $\varepsilon(k)$  gradient adaptive as

$$\varepsilon(k) = \varepsilon(k-1) - \rho \nabla_{\varepsilon(k-1)} E(k) \quad (23)$$

where  $\nabla_{\varepsilon(k-1)} E(k) = \Re[e^*(k) \frac{\partial e(k)}{\partial \varepsilon(k-1)}]$ . Note that

$$\frac{\partial e(k)}{\partial \varepsilon(k-1)} = - \left[ \mathbf{x}^T(k) \frac{\partial \mathbf{w}(k)}{\partial \varepsilon(k-1)} \right] \quad (24)$$

and

$$\frac{\partial \mathbf{w}(k)}{\partial \varepsilon(k-1)} = - \frac{\mu e(k-1)\mathbf{x}^*(k-1)}{[\|\mathbf{x}(k-1)\|_2^2 + \varepsilon(k-1)]^2} \quad (25)$$

Then we can obtain

$$\nabla_{\varepsilon(k-1)} E(k) = \frac{\mu \Re[e^*(k)e(k-1)\mathbf{x}^T(k)\mathbf{x}^*(k-1)]}{[\|\mathbf{x}(k-1)\|_2^2 + \varepsilon(k-1)]^2} \quad (26)$$

and

$$\varepsilon(k) = \varepsilon(k-1) - \frac{\rho \mu \Re[e^*(k)e(k-1)\mathbf{x}^T(k)\mathbf{x}^*(k-1)]}{[\|\mathbf{x}(k)\|_2^2 + \varepsilon(k-1)]^2}$$

## REFERENCES

- [1] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Prentice Hall, 1985.
- [2] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 115–118, 2004.
- [3] Y. S. Choi, H. C. Shin, and W. J. Song, "Robust regularization for normalized LMS algorithm," *IEEE Transactions on Circuits and Systems II*, vol. 53, no. 8, pp. 627–631, 2006.
- [4] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, 1995.
- [5] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. John Wiley & Sons, 2009.
- [6] 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.
- [7] S. C. Douglas, "Widely-linear Recursive Least-Squares algorithm for adaptive beamforming," *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, pp. 2041–2044, 2009.
- [8] Y. Xia, C. C. Took, and D. P. Mandic, "An augmented affine projection algorithm for the filtering of noncircular complex signals," *Signal Processing*, vol. 90, no. 6, pp. 1788–1799, 2010.
- [9] S. C. Douglas and D. P. Mandic, "Performance analysis of the conventional Complex LMS and Augmented Complex LMS algorithms," *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, pp. 3794–3797, 2010.
- [10] K. Kreutz-Delgado, "The complex gradient operator and the CR calculus," *Lecture Supplement ECE275A*, pp. 1–74, 2006.
- [11] D. P. Mandic, A. I. Hanna, and M. Razaz, "A normalized gradient descent algorithm for nonlinear adaptive filters using a gradient adaptive step size," *IEEE Signal Processing Letters*, vol. 8, no. 11, pp. 295–297, 2001.
- [12] 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.
- [13] K. Aihara, *Applied Chaos and Applicable Chaos*. Science-Sha, 1994.
- [14] P. J. Schreier, "The degree of impropriety (noncircularity) of complex random vectors," *Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, pp. 3909–3912, 2008.
- [15] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. John Wiley & Sons, 2001.
- [16] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of IEEE*, vol. 63, pp. 719–720, 1975.
- [17] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Springer, 2008.