# An Adaptive Approach for the Identification of Improper Complex Signals

Beth Jelfs[a,*], Danilo P. Mandic[a], Scott C. Douglas[b]

[a] *Department of Electrical & Electronic Engineering, Imperial College London, SW7 2BT UK*
[b] *Department of Electrical Engineering, Southern Methodist University, Dallas, Texas 75275 USA*

**Abstract**

A real-time approach for the identification of second order noncircularity (improperness) of complex valued signals is introduced. This is achieved based on a convex combination of a standard and widely linear complex adaptive filter, trained by the corresponding complex least mean square (CLMS) and augmented CLMS (ACLMS) algorithms. By providing a rigorous account of widely linear autoregressive modelling the analysis shows that the monitoring of the evolution of the adaptive convex mixing parameter within this structure makes it possible to both detect and track the complex improperness in real time, unlike current methods which are block based and static. The existence and uniqueness of the solution is illustrated through the analysis of the convergence of the convex mixing parameter. The analysis is supported by simulations on representative datasets, for a range of both proper and improper inputs.

*Keywords:* Complex Circularity, Widely Linear Autoregressive Modelling, Collaborative Filter, Augmented Complex Least Mean Square (ACLMS), Improper Complex Signals, Wind Modelling

*Corresponding author
*Email address:* beth.jelfs05@imperial.ac.uk (Beth Jelfs)

## 1. Introduction

Complex valued statistical processing is a well established area; it deals with detection, estimation and adaptive signal processing, and has found a large number of applications across the engineering disciplines. For instance, in signal processing for communications the data symbols are complex by design, and many problems related to arrays and multipath processing in wireless communications are conveniently represented by both the amplitude and phase, thus complex. Directional processes (radar, sonar, vector fields, bearings only estimation), where both the "intensity" (amplitude) and "direction" (phase) components carry the information, are also most conveniently analysed as complex valued [1].

Statistics in $\mathbb{C}$ are typically treated as a straightforward extension of real valued statistics, leading to the same generic solutions for most classic estimators. For instance, the covariance matrix $\mathcal{C}_{\mathbf{zz}} = E\{\mathbf{z}\mathbf{z}^H\}$ of a zero mean complex vector $\mathbf{z} \in \mathbb{C}^{N\times 1}$ is obtained by replacing the vector transpose operator $(\cdot)^T$ in the real covariance matrix $E\{\mathbf{x}\mathbf{x}^T\}$ with the Hermitian transpose in $\mathbb{C}$. Whereas most practical algorithms have been developed based on this assumption [2, 3], the statistics of complex variables show that this approach is optimal only for second order *circular (or proper)* complex random processes [4, 5], for which the probability distribution is rotation invariant, thus limiting the number of applications [6].

Recently, "augmented" complex statistics have established that for optimal second order statistical modelling of the generality of complex signals we need to take into account both the *covariance matrix* $\mathcal{C}_{\mathbf{zz}}$ and *pseudocovariance matrix* $\mathcal{P}_{\mathbf{zz}}$, defined as $\mathcal{P}_{\mathbf{zz}} = E\{\mathbf{z}\mathbf{z}^T\}$. These two matrices are conveniently combined into the "augmented covariance matrix", calculated as

$$\mathcal{R}_{\mathbf{zz}} = E\{\mathbf{z}_a\mathbf{z}_a^H\} = \begin{bmatrix} \mathcal{C}_{\mathbf{zz}} & \mathcal{P}_{\mathbf{zz}} \\ \mathcal{P}_{\mathbf{zz}}^* & \mathcal{C}_{\mathbf{zz}}^* \end{bmatrix}, \tag{1}$$

where the augmented complex vector $\mathbf{z}_a$ is given by

$$\mathbf{z}_a = \begin{bmatrix} \mathbf{z}, & \mathbf{z}^* \end{bmatrix}^T, \tag{2}$$

and the symbol $(\cdot)^*$ denotes the complex conjugation.

The rotation invariant distributions associated with circular signals imply equal powers in the real and imaginary part, and thus a vanishing pseudocovariance matrix of a *proper* zero mean random complex signal[1], that is, $\mathcal{P}_{\mathbf{zz}} = \mathbf{0}$. Notice that complex circularity is a wider notion than properness - it is a property of the probability density function (rotation-invariance) while properness is a second order property indicating a vanishing pseudocovariance matrix [7]. Thus, the modelling of noncircular data based on the covariance matrix only is generally inadequate[2] [5], whereas the modelling of proper complex processes based on augmented statistics (1) is second order optimal but involves additional computational complexity. Therefore, the identification and tracking of the nature of complex valued signals (degree of noncircularity) is a key to efficient statistical signal processing and should ideally be performed in real time.

Signal modality characterisation has been developed in Physics to reveal changes in the nature of real world data (nonlinear, sparse, deterministic, stochastic [8]), and is only just being adopted in signal processing [9, 10]. Statistical hypothesis testing based measures for the validity of complex representation do exist [11, 6], however, such tests are difficult to generalise for measuring the degree of noncircularity. Existing measures of the degree of noncircularity are typically block-based deterministic functions, based e.g. on multivariate associations in real valued vectors as a measure of the linear dependence between $\mathbf{z}$ and $\mathbf{z}^*$. The circularity index in [12] was introduced as a function of canonical correlations; it was subsequently made more flexible based upon likelihood measures, such as the generalised likelihood ratio tests (GLRT) in [13] and [14]. Whereas such block based tests are accurate and intuitive for off-line processing of stationary signals, they are unsuitable in real time processing and for real

[1]For illustration, consider a complex number $z = x + \jmath y$. Then $zz^T = x^2 - y^2$, which for $\sigma_x^2 = \sigma_y^2$ vanishes upon applying the statistical estimation operator.

[2]Unless we have a special case of a standard autoregressive process driven by a doubly white noise with different powers in the real and imaginary part, as shown in Section 2.

world applications.

To this end, we extend our earlier work in [15, 16] to propose a flexible method for the identification and tracking of the degree of (non)circularity of the generality of complex valued signals. The proposed approach employs a collaborative adaptive filter, based on a convex combination of the complex least mean square (CLMS) [3] and augmented CLMS (ACLMS) [17], thus facilitating real time adaptive mode of operation. The evolution of the convex mixing parameter within this structure is shown to quantify the time varying degree of improperness of a complex signal. The analysis is supported by illustrative simulations on both synthetic signals and real world wind modelling.

## 2. Widely Linear Estimation and Autoregressive Modelling

Consider the standard mean square error (MSE) estimator of a real valued signal $y$ in terms of another observation $x$, that is

$$\hat{y} = E[y|x]. \tag{3}$$

For zero mean, jointly normal $y$ and $x$, the solution is the linear estimator

$$\hat{y} = \mathbf{h}^T\mathbf{x}, \tag{4}$$

where $\mathbf{h}$ is a coefficient vector and $\mathbf{x}$ the regressor vector. By continuation, in standard MSE in the complex domain (based on only the covariance matrix), it is assumed that[3]

$$\hat{y} = E\{y|z\} \quad \rightarrow \quad \hat{y} = \mathbf{h}^H\mathbf{z}.$$

However, it is important to realise that both the real and imaginary parts of the complex number $z = z_r + \jmath z_i$ can be estimated using a real MSE estimator

[3]Both $y = \mathbf{h}^T\mathbf{z}$ and $y = \mathbf{h}^H\mathbf{z}$ are correct, yielding the same output and the mutually conjugate coefficient vectors. The latter form is more common and the former was used in the original CLMS paper [3]; in this work we will use the first form.

in (3), thus giving [18]

$$\hat{y}_r = E[y_r|z_r, z_i] \quad \& \quad \hat{y}_i = E[y_i|z_r, z_i],$$
$$\text{thus} \quad \hat{y} = E[y_r|z_r, z_i] + \jmath E[y_i|z_r, z_i]. \tag{5}$$

Upon employing the identities $z_r = (z + z^*)/2$ and $z_i = (z - z^*)/2\jmath$, we arrive at

$$\hat{y} = E[y_r|z, z^*] + \jmath E[y_i|z, z^*], \tag{6}$$

yielding the "widely linear" estimator for general complex signals (both proper and improper) in the form

$$\hat{y} = \mathbf{h}^T\mathbf{z} + \mathbf{g}^T\mathbf{z}^*, \tag{7}$$

where $\mathbf{h}$ and $\mathbf{g}$ are the coefficient vectors of the widely linear model. For more detail on augmented complex statistics, widely linear models, and their applications in adaptive signal processing see [6, 19, 20, 21].

### *2.1. Performance Bounds of the Standard Autoregressive Model*

The standard autoregressive AR(n) model of order $n$ (in $\mathbb{R}$ or $\mathbb{C}$) is described by

$$z(k) = a_1 z(k-1) + \cdots + a_n z(k-n) + q(k) = \mathbf{a}^T\mathbf{z}(k) + q(k), \tag{8}$$

where $\mathbf{a} = [a_1, \ldots, a_n]^T$ are the fixed AR coefficients, $\mathbf{z}(k) = [z(k-1), \ldots, z(k-n)]^T$ the regressor vector, and $q(k)$ the driving doubly white Gaussian noise (proper or improper). Using the Yule-Walker equations, the AR coefficients can be found from [22]

$$\begin{aligned} \mathbf{a}^* &= \mathcal{C}_{\mathbf{zz}}^{-1}\mathbf{c} \\ \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_n^* \end{bmatrix} &= \begin{bmatrix} c(0) & c^*(1) & \ldots & c^*(n-1) \\ c(1) & c(0) & \ldots & c^*(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ c(n-1) & c(n-2) & \ldots & c(0) \end{bmatrix}^{-1} \begin{bmatrix} c(1) \\ c(2) \\ \vdots \\ c(n) \end{bmatrix} \end{aligned} \tag{9}$$

where $\mathbf{c} = [c(1), c(2), \ldots, c(n)]^T$ is the time shifted correlation vector. Observe that using the model in (8) it is possible to generate both proper and improper linear processes, depending on the nature of the doubly white driving noise $q(k)$.

To evaluate the advantage of *widely linear* over *strictly linear* stochastic modelling, consider again the linear and widely linear estimates of a random process $z(k)$ given respectively by [23]

$$\hat{z}_l(k) = \mathbf{a}^T \mathbf{z}(k), \tag{10}$$

$$\hat{z}_{wl}(k) = \mathbf{h}^T \mathbf{z}(k) + \mathbf{g}^T \mathbf{z}^*(k). \tag{11}$$

Then, the corresponding optimal AR estimation errors are given by

$$\begin{aligned} e_l^2 &= E[|z(k)|^2] - E[|\hat{z}_l(k)|^2] = \mathbf{c}^T \mathcal{C}_{\mathbf{zz}}^{*^{-1}} \mathbf{c}^* + \sigma_q^2 - \mathbf{c}^T \mathcal{C}_{\mathbf{zz}}^{*^{-1}} \mathbf{c}^* = \sigma_q^2 \\ e_{wl}^2 &= E[|z(k)|^2] - E[|\hat{z}_{wl}(k)|^2] = \mathbf{c}^T \mathcal{C}_{\mathbf{zz}}^{*^{-1}} \mathbf{c}^* + \sigma_q^2 - \mathbf{r}^T \mathcal{R}_{\mathbf{zz}}^{*^{-1}} \mathbf{r}^* \end{aligned} \tag{12}$$

where $\sigma_q^2$ denotes the driving noise variance and $\mathbf{r} = [c(1), \ldots, c(n), p^*(1), \ldots, p^*(n)]^T$ is the corresponding time shifted correlation vector of the augmented covariance matrix $\mathcal{R}_{\mathbf{zz}}$, defined in (1).

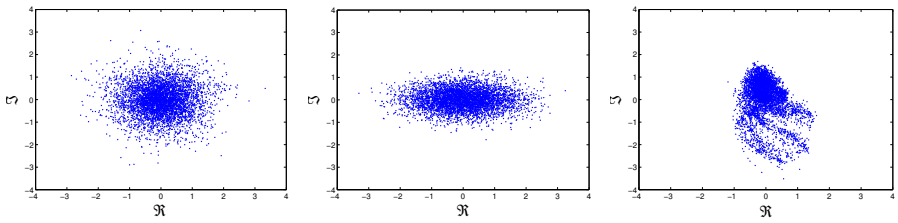
### *2.2. Processes Generated by Standard Strictly AR Models*

In standard autoregressive modelling in the complex domain, the only requirement on the driving noise $q(k)$ is its double whiteness, that is, the real and imaginary part are jointly white and uncorrelated - there are no requirements on the particular distributions or variances in the noise channels. Depending on the nature of the driving noise (doubly white circular, doubly white noncircular, general noncircular) there are three possible scenarios for the circularity properties of the processes generated by the standard AR model.

- Doubly white circular noise (Fig. 1(a)): In this case $\mathcal{P}_{\mathbf{qq}} = \mathbf{0}$, $q_r(k) \perp q_i(k)$, and $\sigma_{q_r}^2 = \sigma_{q_i}^2$; The resulting AR process is also circular (Fig. 2(a)), and from (12) we have

  $$e_l^2 = e_{wl}^2, \tag{13}$$

  and thus such process can be optimally modelled by a standard AR model.



(a) AR(4) driven by circular noise (b) AR(4) driven by doubly white noncircular noise (c) AR(4) driven by general noncircular noise

Figure 1: Distributions of an AR(4) process for different realisations of the driving noise

- Doubly white improper noise (Fig. 1(b)): In this case $\mathcal{P}_{\mathbf{qq}} \neq \mathbf{0}$, $q_r(k) \perp q_i(k)$ and $\sigma_{q_r}^2 > \sigma_{q_i}^2$. The resulting AR process is therefore second order noncircular, as shown in Fig. 2(b). From (10)–(12) the error of the standard linear AR process driven by $q(k)$ is

$$e(k) = z(k) - \hat{z}_l(k) = q(k). \tag{14}$$

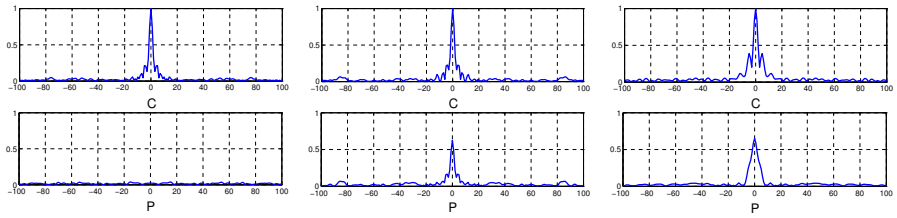
  Since the innovation $e(k)$ is uncorrelated with $\mathbf{z}(k)$ and $\mathbf{z}^*(k)$, then the standard AR model is second order optimal and, therefore, the errors of the standard and widely linear estimator are equal

$$e_l^2 = e_{wl}^2 = \sigma_q^2, \tag{15}$$

  illustrating that when noncircular linear processes are generated by a doubly white noncircular noise which drives standard AR models, there is no advantage in using the widely linear model [24].

- General noncircular driving noise (Fig. 1(c)): In this case $\mathcal{P}_{\mathbf{qq}} \neq \mathbf{0}$ and the resulting standard AR process is noncircular (Fig. 2(c)). The advantage of the widely linear model over the linear model is then assessed from

$$\begin{aligned}
\delta e^2 =& e_l^2 - e_{wl}^2 \\
=& E[|z(k)|^2] - E[|\hat{z}_l(k)|^2] - \left(E[|z(k)|^2] - E[|\hat{z}_{wl}(k)|^2]\right) \\
=& E[|\hat{z}_{wl}(k)|^2] - E[|\hat{z}_l(k)|^2] \\
=& \mathbf{r}^T \mathcal{R}_{\mathbf{zz}}^{*^{-1}} \mathbf{r}^* - \mathbf{c}^T \mathcal{C}_{\mathbf{zz}}^{*^{-1}} \mathbf{c}^*,
\end{aligned} \tag{16}$$




(a) AR(4) driven by doubly white circular noise

(b) AR(4) driven by doubly white noncircular noise

(c) AR(4) driven by general noncircular noise

Figure 2: Covariance (top) and pseudocovariance (bottom) of a standard AR(4) model driven by different driving noises

Following the approach in [18], as shown in Appendix A, this can be rewritten as

$$\delta e^2 = \left[\mathbf{p} - \mathcal{P}_{\mathbf{zz}}^{*}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{c}^{*}\right]^{H}\left[\mathcal{C}_{\mathbf{zz}} - \mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^{*}\right]^{-1}\left[\mathbf{p} - \mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{c}^{*}\right]. \quad (17)$$

Figures 1 and 2 illustrate the properties of distributions and correlation structures for the processes generated by a standard AR model, driven by the three classes of noises discussed.

### *2.3. Widely Linear Autoregressive Modelling*

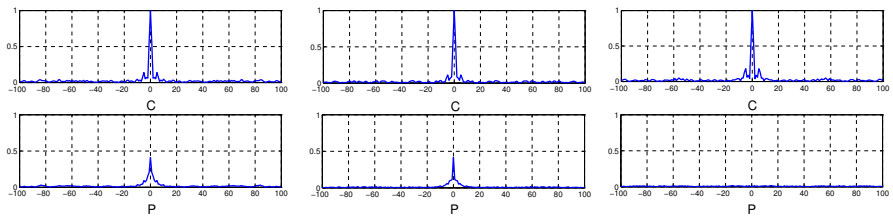
The widely linear AR (WLAR) model caters for the complete (augmented) second order complex statistics, and is given by (based on (7))

$$z(k) = h_1 z(k-1) + g_1 z^*(k-1) + \cdots + h_n z(k-n) + g_n z^*(k-n) + q(k) \quad (18)$$

The Yule-Walker equations for the coefficients of the widely linear model are then given by

$$\begin{bmatrix} \mathbf{h}^* \\ \mathbf{g}^* \end{bmatrix} = \begin{bmatrix} \mathcal{C}_{\mathbf{zz}} & \mathcal{P}_{\mathbf{zz}} \\ \mathcal{P}_{\mathbf{zz}}^* & \mathcal{C}_{\mathbf{zz}}^* \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{c} \\ \mathbf{p}^* \end{bmatrix}, \quad (19)$$

where $\mathbf{h}$ denotes the coefficient vector of the standard complex AR model and $\mathbf{g}$ the coefficient vector of the conjugate part of the WLAR. To illustrate the second order optimality of this method, we shall apply the widely linear normal




(a) Original Ikeda map (b) Widely linear AR(4) model of the Ikeda map (c) Standard linear AR(4) model of the Ikeda map

Figure 3: Covariance (top) and pseudocovariance (bottom) of the Ikeda map models

equations (19) to estimate parameters of a WLAR(4) model generated from the nonlinear and noncircular Ikeda map, described by

$$x(k+1) = 1 + u\left[x(k)\cos t(k) - y(k)\sin t(k)\right], \tag{20}$$

$$y(k+1) = u\left[x(k)\sin t(k) + y(k)\cos t(k)\right], \tag{21}$$

where $u$ is a parameter and

$$t(k) = 0.4 - \frac{6}{1 + x^2(k) + y^2(k)}. \tag{22}$$

The resulting covariance and pseudocovariance functions are shown in Fig. 3; observe that unlike the standard AR model, the WLAR model caters for improper signals, as indicated its ability to model both the covariance and pseudocovariance.

## 3. Collaborative Adaptive Filter for the Tracking of Noncircularity

The proposed approach for the assessment of the signal noncircularity is based on a collaborative combination of CLMS and ACLMS trained subfilters, and is an extension of our earlier work in [16]. All the filter parameters are updated by minimising the cost function

$$\mathcal{J}(k) = \frac{1}{2}|e(k)|^2 = \frac{1}{2}|d(k) - y(k)|^2. \tag{23}$$

The CLMS is described by [3]

$$\begin{aligned} e_c(k) =& d(k) - y_c(k), \\ y_c(k) =& \mathbf{h}_c^T(k)\mathbf{z}(k), \\ \mathbf{h}_c(k+1) =& \mathbf{h}_c(k) + \mu_c e_c(k)\mathbf{z}^*(k), \end{aligned} \tag{24}$$

where $\mathbf{h}(k) = [h_1(k), h_2(k), \ldots, h_N(k)]^T$ is the filter coefficient vector, $d(k)$ and $e_c(k)$ are the desired response and output error at time instant $k$ and $\mu$ is the learning rate. The ACLMS utilises the full second order statistical information available by using the widely linear model and is given by [17, 25]

$$\begin{aligned} e_a(k) =& d(k) - y_a(k), \\ y_a(k) =& \mathbf{h}_a^T(k)\mathbf{z}(k) + \mathbf{g}_a^T(k)\mathbf{z}^*(k), \\ \mathbf{h}_a(k+1) =& \mathbf{h}_a(k) + \mu_a e_a(k)\mathbf{z}^*(k), \\ \mathbf{g}_a(k+1) =& \mathbf{g}_a(k) + \mu_a e_a(k)\mathbf{z}(k). \end{aligned} \tag{25}$$

The collaborative filter, shown in Figure 4, consists of two independently adapted subfilters, operating in the prediction setting, sharing the common input $z(k)$ and the desired signal $d(k)$. The convex combination of the subfilter outputs

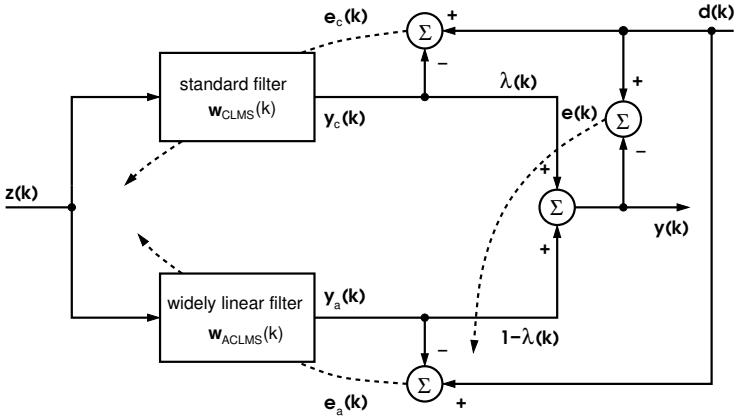



Figure 4: Hybrid filter structure

$y_c(k)$ and $y_a(k)$ forms the overall output $y(k)$, given by

$$y(k) = \lambda(k)y_c(k) + \big(1 - \lambda(k)\big)y_a(k), \quad 0 \leq \lambda(k) \leq 1 \tag{26}$$

where $\lambda(k)$ is the real-valued convex mixing parameter, whose update is obtained from

$$\lambda(k+1) = \lambda(k) - \mu_\lambda \nabla_\lambda \mathcal{J}(k)|_{\lambda=\lambda(k)}, \tag{27}$$

where $\mu_\lambda$ is the step size. Since the input to the filters is complex, the error $e(k)$ is also complex, and therefore the gradient

$$\nabla_\lambda \mathcal{J}(k)|_{\lambda=\lambda(k)} = \left\{ e(k)\frac{\partial e^*(k)}{\partial \lambda(k)} + e^*(k)\frac{\partial e(k)}{\partial \lambda(k)} \right\}. \tag{28}$$

can be evaluated as

$$\frac{\partial e(k)}{\partial \lambda(k)} = y_c(k) - y_a(k), \tag{29}$$

$$\frac{\partial e^*(k)}{\partial \lambda(k)} = \big(y_c(k) - y_a(k)\big)^*, \tag{30}$$

yielding the update of the mixing parameter in the form

$$\lambda(k+1) = \lambda(k) + \mu_\lambda \Big[ e(k)\big(y_c(k) - y_a(k)\big)^* + e^*(k)\big(y_c(k) - y_a(k)\big) \Big]. \tag{31}$$

Due to the convex nature of the collaborative filter, providing at least one of the subfilters converges the collaborative filter is guaranteed to converge [26], provided the mixing parameter remains within the range $[0, 1]$. Several approaches have been proposed for this purpose, however, as our aim is to track the behaviour of the mixing parameter, we cannot interfere with the evolution of $\lambda$, and a hard bound[4] on the values of $\lambda$ is used when $\lambda > 1$ or $\lambda < 0$.

### *3.1. Convergence of the Mixing Parameter*

The convergence of the CLMS and ACLMS for both proper and improper inputs has been analysed in [27, 28], and a rigorous account of the convergence of collaborative filters in [26]. Since the linear and widely linear adaptive filter are adapted independently, results in [27, 28] also apply to the CLMS and ACLMS subfilters within the collaborative filter in Fig. 4. We therefore only need to illustrate the ability of the collaborative filter to identify the noncircularity of the

[4] In practice this is seldom necessary.

input in real time; for the structure in Fig. 4 this means that for proper signals the output of a collaborative filter is dominated by the strictly linear CLMS and $\lambda \to 1$, and for improper signals the ACLMS trained subfilter prevails and $\lambda \to 0$.

Without loss in generality assume that the desired response can be expressed as [29]

$$d(k) = \mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k) + q(k), \tag{32}$$

where $\mathbf{h}_o$ and $\mathbf{g}_o$ are the optimal Wiener filter weights and $q(k)$ is doubly white noise, so that the minimum mean squared error $\mathcal{J}_{min} = \sigma_q^2$. In the steady state both the subfilters converge towards the optimal values of their respective coefficient vectors $\mathbf{h}_{co}$, $\mathbf{h}_{ao}$ and $\mathbf{g}_{ao}$.

We next illustrate the principle of the proposed solution for the two extremes - purely circular and purely noncircular inputs.

- Circular input signal. In this case $\mathbf{h}_{co} = \mathbf{h}_{ao} = \mathbf{h}_o$ and $\mathbf{g}_{ao} = \mathbf{g}_o = \mathbf{0}$, and from (26) in the steady state, the overall instantaneous output error becomes

$$\begin{aligned} e(k) =& \mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k) + q(k) - \lambda(k) y_c(k) - \big(1 - \lambda(k)\big) y_a(k) \\ =& \mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k) + q(k) - \lambda(k) \mathbf{h}_o^T \mathbf{z}(k) - \big(1 - \lambda(k)\big) \mathbf{h}_o^T \mathbf{z}(k) \\ =& q(k). \end{aligned} \tag{33}$$

  Substitute into the update for the mixing parameter in (31) to yield the evolution of the mixing parameter in the form

$$\begin{aligned} \lambda(k+1) =& \lambda(k) + \mu_\lambda \Big[ e(k) \big(y_c(k) - y_a(k)\big)^* + e^*(k) \big(y_c(k) - y_a(k)\big) \Big] \\ =& \lambda(k) + \mu_\lambda q(k) \big(\mathbf{h}_o^H \mathbf{z}^*(k) - \mathbf{h}_o^H \mathbf{z}^*(k)\big) \\ & + \mu_\lambda q^*(k) \left(\mathbf{h}_o^T \mathbf{z}(k) - \mathbf{h}_o^T \mathbf{z}(k)\right) \\ =& \lambda(k). \end{aligned} \tag{34}$$

  Since the CLMS initially converges faster than the ACLMS [28], and therefore in the beginning of adaptation the collaborative filter favours the

CLMS trained subfilter, the convex mixing parameter $\lambda \to 1$, thus correctly reflecting the circular nature of the input.

- Noncircular input signal. For noncircular inputs, upon convergence $\mathbf{h}_{ao} = \mathbf{h}_o \neq \mathbf{h}_{co}$ and $\mathbf{g}_{ao} = \mathbf{g}_o$, and the overall output error becomes

$$\begin{aligned} e(k) =& \mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k) + q(k) - \lambda(k) y_c(k) - \big(1 - \lambda(k)\big) y_a(k) \\ =& \mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k) + q(k) - \lambda(k) \mathbf{h}_{co}^T \mathbf{z}(k) \\ & - \big(1 - \lambda(k)\big)\big(\mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k)\big) \\ =& q(k) + \lambda(k)\big(\mathbf{h}_o^T \mathbf{z}(k) + \mathbf{g}_o^T \mathbf{z}^*(k) - \mathbf{h}_{co}^T \mathbf{z}(k)\big). \end{aligned} \tag{35}$$

Substituting this result into the update of the mixing parameter (31) gives

$$\begin{aligned} \lambda(k+1) =& \lambda(k) + \mu_\lambda \Big[ e(k)\big(y_c(k) - y_a(k)\big)^* + e^*(k)\big(y_c(k) - y_a(k)\big) \Big] \\ =& \lambda(k) + \mu_\lambda e(k) y_c^*(k) - \mu_\lambda e(k) y_a^*(k) + \mu_\lambda e^*(k) y_c(k) \\ & - \mu_\lambda e^*(k) y_a(k) \\ =& \lambda(k) + \mu_\lambda \Big[ q(k) \mathbf{h}_{co}^H \mathbf{z}^*(k) + q^*(k) \mathbf{h}_{co}^T \mathbf{z}(k) - q(k) \mathbf{h}_o^H \mathbf{z}^*(k) \\ & - q^*(k) \mathbf{h}_o^T(k) \mathbf{z}(k) - q(k) \mathbf{g}_o^H \mathbf{z}(k) - q^*(k) \mathbf{g}_o^T(k) \mathbf{z}^*(k) \\ & + 2\lambda(k) \Big( \mathbf{h}_{co}^T \mathbf{z}(k) \mathbf{h}_o^H \mathbf{z}^*(k) + \mathbf{h}_{co}^T \mathbf{z}(k) \mathbf{g}_o^H \mathbf{z}(k) - \mathbf{h}_{co}^T \mathbf{z}(k) \mathbf{h}_{co}^H \mathbf{z}^*(k) \\ & + \mathbf{h}_{co}^H \mathbf{z}^*(k) \mathbf{h}_o^T \mathbf{z}(k) + \mathbf{h}_{co}^H \mathbf{z}^*(k) \mathbf{g}_o^T \mathbf{z}^*(k) - \mathbf{h}_o^T \mathbf{z}(k) \mathbf{h}_o^H \mathbf{z}^*(k) \\ & - \mathbf{h}_o^T \mathbf{z}(k) \mathbf{g}_o^H \mathbf{z}(k) - \mathbf{g}_o^T \mathbf{z}^*(k) \mathbf{h}_o^H \mathbf{z}^*(k) - \mathbf{g}_o^T \mathbf{z}^*(k) \mathbf{g}_o^H \mathbf{z}(k) \Big) \Big]. \end{aligned} \tag{36}$$

On applying the statistical expectation operator and employing the stan-

dard independence assumptions[5], we have

$$\begin{aligned}E[\lambda(k+1)] =& E[\lambda(k)]\Big[1+2\mu_\lambda\Big(E[\mathbf{h}_{co}^T\mathbf{z}(k)\mathbf{z}^H\mathbf{h}_o^*]+E[\mathbf{h}_{co}^T\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{g}_o^*]\\ &-E[\mathbf{h}_{co}^T\mathbf{z}(k)\mathbf{z}^H(k)\mathbf{h}_{co}^*]+E[\mathbf{h}_{co}^H\mathbf{z}^*(k)\mathbf{z}^T(k)\mathbf{h}_o]\\ &+E[\mathbf{h}_{co}^H\mathbf{z}^*(k)\mathbf{z}^H(k)\mathbf{g}_o]-E[\mathbf{h}_o^T\mathbf{z}(k)\mathbf{z}^H(k)\mathbf{h}_o^*]\\ &-E[\mathbf{h}_o^T\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{g}_o^*]-E[\mathbf{g}_o^T\mathbf{z}^*(k)\mathbf{z}^H(k)\mathbf{h}_o^*]\\ &-E[\mathbf{g}_o^T\mathbf{z}^*(k)\mathbf{z}^T(k)\mathbf{g}_o^*]\Big)\Big]. \end{aligned} \tag{37}$$

For clarity[6], define the inverse Schur complement of the conjugate augmented covariance as $\mathbf{A} = \left(\mathcal{C}_{\mathbf{zz}} - \mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\right)^{-1}$ and $\mathbf{A}^* = \mathcal{C}_{\mathbf{zz}}^{*^{-1}} + \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}$, then from (A.4) and (A.5) we have

$$\begin{aligned}\mathbf{h}_o =& \left[\mathcal{C}_{\mathbf{zz}}^{*^{-1}} + \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\right]\left[\mathbf{c}^* - \mathcal{P}_{\mathbf{zz}}^*\mathcal{C}_{\mathbf{zz}}^{-1}\mathbf{p}\right]\\ =& \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{c}^* - \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathbf{p} + \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{c}^*, \end{aligned} \tag{38}$$

$$\mathbf{h}_o^* = \mathbf{A}\left[\mathbf{c} - \mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{p}^*\right], \tag{39}$$

$$\mathbf{g}_o = \mathbf{A}\left[\mathbf{p} - \mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{c}^*\right], \tag{40}$$

$$\begin{aligned}\mathbf{g}_o^* =& \left[\mathcal{C}_{\mathbf{zz}}^{*^{-1}} + \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\right]\left[\mathbf{p}^* - \mathcal{P}_{\mathbf{zz}}^*\mathcal{C}_{\mathbf{zz}}^{-1}\mathbf{c}\right]\\ =& \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{p}^* - \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathbf{c} + \mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{p}^*. \end{aligned} \tag{41}$$

Substitute into (37) to give the expression for the evolution of the mixing

[5]Namely that the input signal and filter coefficient vectors are zero mean, stationary, jointly normal and with finite moments; the successive increments of filter weights are independent of one another and the error and input vector sequences are statistically independent of one another [30, 22, 31].

[6]For more detail see Appendix A.

parameter $\lambda$ for improper signals, in the form

$$\begin{aligned}
E[\lambda(k+1)] =& E[\lambda(k)]\Big(1+2\mu_\lambda\Big[-\mathbf{c}^H\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}^T\mathcal{P}_{\mathbf{zz}}\mathbf{A}\mathbf{c}+\mathbf{p}^T\mathbf{A}^T\mathcal{P}_{\mathbf{zz}}\mathbf{A}\mathbf{c}\\
&-\mathbf{p}^T\mathbf{A}^T\mathbf{p}^*+\mathbf{c}^H\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}^T\mathbf{p}^*\\
&+\mathbf{c}^H\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}^T\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathbf{c}\\
&-\mathbf{p}^T\mathbf{A}^T\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathbf{c}\Big]\Big)\\
=& E[\lambda(k)]\Big(1+2\mu_\lambda\Big[\big(\mathbf{p}^T-\mathbf{c}^H\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}^*\big)\mathbf{A}^T\big(\mathcal{P}_{\mathbf{zz}}\mathbf{A}\mathbf{c}-\mathbf{p}^*\\
&-\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\mathbf{A}\mathbf{c}\big)\Big]\Big)\\
=& E[\lambda(k)]\Big(1+2\mu_\lambda\Big[\big(\mathbf{p}^T-\mathbf{c}^H\mathcal{C}_{\mathbf{zz}}^{-1}\mathcal{P}_{\mathbf{zz}}^*\big)\mathbf{A}^T\big(\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{-1}\mathbf{c}-\mathbf{p}^*\big)\Big]\Big)\\
=& E[\lambda(k)]\Big(1+2\mu_\lambda\Big[\big(\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{-1}\mathbf{c}-\mathbf{p}^*\big)^T\big(\mathcal{C}_{\mathbf{zz}}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathcal{P}_{\mathbf{zz}}^*\big)^{-1}\\
&\big(\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}_{\mathbf{zz}}^{*^{-1}}\mathbf{c}^*\big)\Big]\Big)\\
=& E[\lambda(k)]\Big(1-2\mu_\lambda\delta e^2\Big), \qquad (42)
\end{aligned}$$

where $\delta e^2$ denotes the performance advantage of the widely linear model over the standard linear model, as shown in (17). Since both the learning rate $\mu_\lambda$ and $\delta e^2$ are positive, the mixing parameter $\lambda(k)$ converges towards zero (favouring the ACLMS trained subfilter) whenever the widely linear filter outperforms the standard filter, that is, for second order noncircular (improper) inputs.

These two cases reflect the ability of a collaborative CLMS-ACLMS filter to identify and track the circular/noncircular nature of real-world inputs. Due to the convexity of the mixing parameter $\lambda$, by continuity the analysis is also valid for any degree of noncircularity and for processes with time varying statistics. The usefulness of this approach is illustrated through representative simulation studies.

## 4. Simulations

For all the simulations $\mu_a = \mu_c = 0.001$, $\mu_\lambda = 0.05$ and the initial value of $\lambda = 0.5$ (neither proper or improper). Recall that $\lambda \to 1$ corresponds to the filter being dominated by the CLMS subfilter (indicating a circular signal), whereas $\lambda \to 0$ indicates the collaborative filter is dominated by the ACLMS subfilter, indicating a noncircular input. All the simulations based on synthetic data were averaged over 100 independant trials while the real world wind example was analysed over a single trial.

Figure 5 shows the evolution of the mixing parameter $\lambda$ for an AR(4) process, a WLAR(4) model of the Ikeda signal, and for the original noncircular Ikeda signal. The linear circular AR(4) process was a standard AR model driven by doubly white noise, while the WLAR(4) process was calculated from (18)-(19). As desired, for the linear AR(4) signal the value of the mixing parameter $\lambda$ was between 0.8 and 0.9, indicating its circular nature, whereas for the original Ikeda process, the value of the mixing parameter approached zero indicating its second order noncircular (improper), also illustrated in Fig. 3. For the WLAR(4) model of the Ikeda signal (improper), the mixing parameter moved initially upwards, indicating the faster convergence of the CLMS, but then settled to approximately $\lambda = 0.1$, illustrating that the widely linear model of the improper Ikeda process is better modelled by the ACLMS. This behaviour of $\lambda$ is also in line with the convergence analysis in Section 3.1.

In the next set of simulations, the effect of the nature of the driving noise on the proper/improper nature of the standard AR(4) model was investigated, in order to verify the proposed solution on the three classes of data elaborated in Section 2.2. Figure 6 shows that both the cases driven by the doubly white noise were optimally modelled by a standard AR model, conforming with the analysis in (13)–(17). Although, theoretically such inputs are equally modelled by CLMS and ALCMS (in the steady state), the CLMS is initially faster converging and was thus the dominant subfilter, as indicated by $\lambda$ evolving towards unity. The improper AR signal driven by a general noncircular noise was correctly better

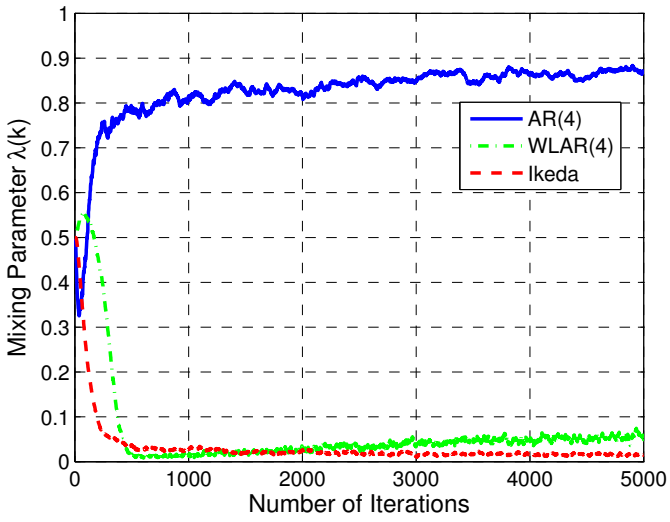



Figure 5: Evolution of the mixing parameter $\lambda$ for the linear circular AR(4) process, the WLAR(4) model of the Ikeda map (improper) and the original noncircular Ikeda signal

modelled by the widely linear ACLMS algorithm, as reflected in the the value of the convex mixing parameter $\lambda$ approaching zero, indicating the improper nature of the output.

The previous two sets of simulations were conducted for static processes. To illustrate the ability of the collaborative filter to track changes in the circularity of a signal, in a dynamically changing environment, in the next experiment the input was alternated between the proper AR(4) process and the improper Ikeda map in the first setting, and between the same AR(4) signal and the improper WLAR(4) model of the Ikeda signal in the second setting with the segment lenght fixed to 1000 samples. The results shown in Fig. 7 illustrate that the proposed approach was able to accurately track these changes. Finally, Fig. 8 shows the behaviour of the convex mixing parameter $\lambda$ when the segment length of the alternating proper/improper inputs changed over time, illustrating flexibility and real time tracking ability of the proposed approach.

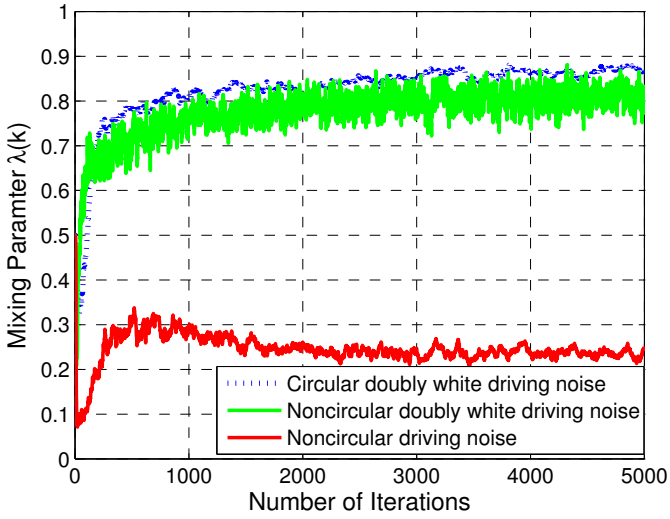



Figure 6: Evolution of the mixing parameter $\lambda$ for a standard linear AR(4) model driven by noises of different nature

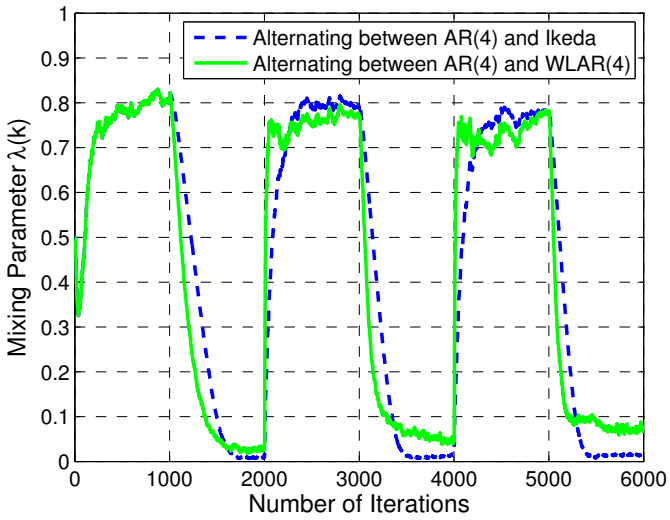



Figure 7: Evolution of the mixing parameter $\lambda$ for signals alternating from linear AR(4) to either widely linear AR(4) model of the Ikeda process or the original Ikeda signal

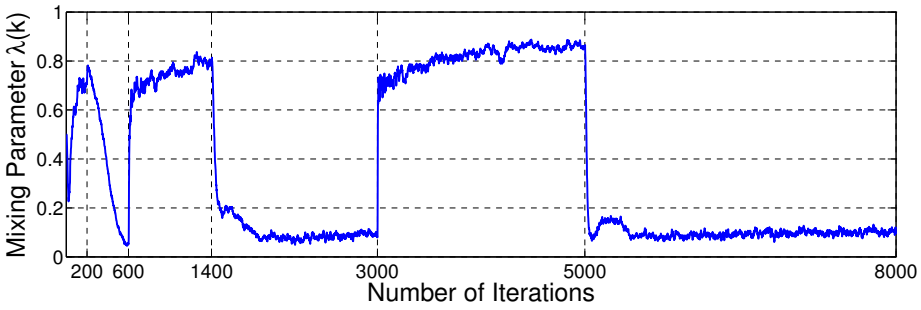



Figure 8: Evolution of the mixing parameter $\lambda$ for a signal alternating from a proper AR(4) to an improper linear AR(4) process with a varying segment duration

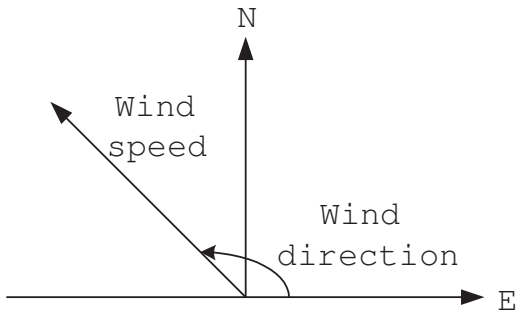


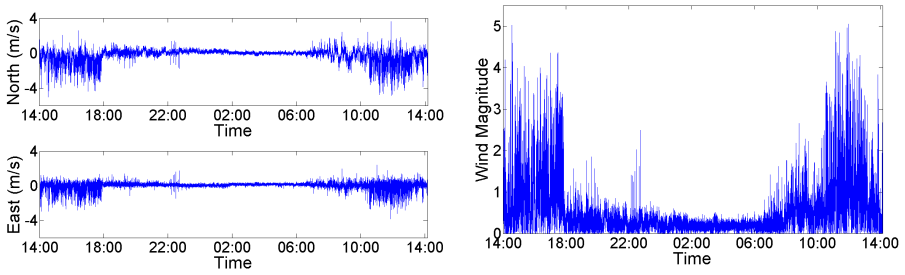

Figure 9: Wind recordings in a complex [speed,direction] representation

### *4.1. Improperness of Wind Data*

Wind modelling is a key in several renewable energy applications; wind is normally measured as a bivariate process of direction and speed [32]. From Fig. 9 it is clear that wind can be represented as a vector of speed and direction components in the North – East coordinate system; the wind speed $\mathbf{v}$ and direction $\theta$ are then combined to form a complex signal

$$\mathbf{V} = \mathbf{v} \cdot e^{\jmath\theta}. \tag{43}$$

The wind data used was measured over a 24 hour period sampled at 50Hz




(a) Wind speed readings

(b) Wind magnitude

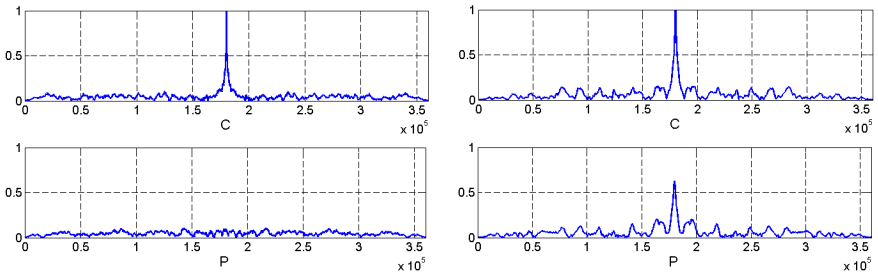
Figure 10: Wind over a 24 hour period in an urban environment

in an urban environment[7]. The wind speed readings (Fig. 10(a)) were taken in the north–south ($\mathbf{V}_N$) and east–west ($\mathbf{V}_E$) directions, where

$$\mathbf{v} = \sqrt{\mathbf{V}_E^2 + \mathbf{V}_N^2} \quad \text{and} \quad \theta = \arctan\left(\frac{\mathbf{V}_N}{\mathbf{V}_E}\right), \tag{44}$$

were used to give the complex signal (43). The magnitude of the complex signal obtained from combining the two speed readings $\mathbf{V}$ is shown in Fig. 10(b). It can be seen from the recordings that there is a distinct 'calm' period in the wind diagrams in the late evening and the early morning between 18:00 and 08:00 compared to the rapid fluctuations in the wind at other times. To best assess the performance of the collaborative filter for detecting the level of noncircularity of the wind data, two periods of 1 hour duration each were assessed, one from the 'calm' period between 04:00-05:00 and one from the 'high' wind dynamics between 16:00-17:00. The covariance and pseudocovariance plots for both sections, shown in Fig. 11, show that the 'calm' wind section has a close to zero pseudocovariance indicating the near-circular nature of 'calm' wind and an improper nature of 'high' wind. The comparison of the evolution of the mixing parameter $\lambda$ for the 'calm' and 'high' sections of wind are shown in Fig. 12. As desired, for the 'calm' wind the value of $\lambda \to 1$, indicating its proper nature, whereas for the 'high' wind, the value of $\lambda$ indicated its rapidly changing

[7]The wind data was provided by Prof. Kazuyuki Aihara and Dr Yoshito Hirata from the Institute of Industrial Science, University of Tokyo.




(a) Calm wind (b) High wind

Figure 11: Normalised covariance (top) and pseudocovariance (bottom) for 1 hour sections of wind data

and predominantly improper nature.

## 5. Conclusions

We have proposed an online adaptive test for the identification and assessment of the proper/improper nature of complex valued signals. By revisiting widely linear AR modelling this has been achieved based on a collaborative adaptive filtering approach, whereby each subfilter has been chosen so as to be optimal for either proper or improper input processes. It has been shown that for circular data the convex mixing parameter within this structure favours the standard, strictly linear, subfilter whereas for noncircular data it favours the widely linear subfilter. The analysis has addressed the convergence of the solution when identifying such processes, and the simulations illustrate that the evolution of the mixing parameter correctly reflects the proper/improper nature of the data. It has also been shown that unlike the existing static, block based, approaches the proposed method has the ability to both identify and track the degree of circularity of a signal in real time, a crucial feature in time-varying scenarios, such as in wind modelling for renewable energy applications.

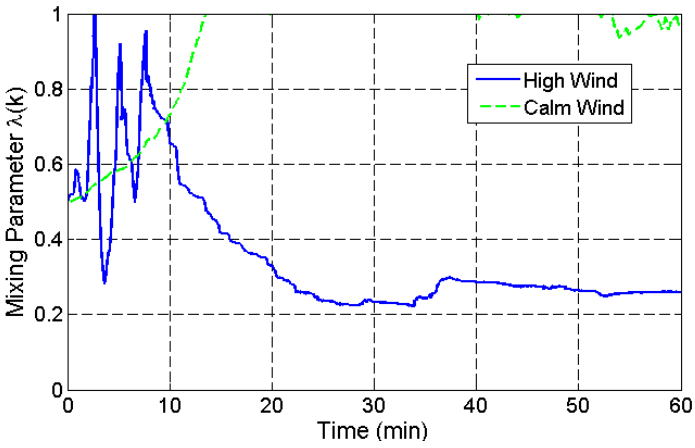



Figure 12: Evolution of the mixing parameter $\lambda(k)$ for 1hr of 'calm' wind and 1 hour of 'high' wind

## Appendix A. Advantage of Widely Linear Model Over Standard Linear Model

To quantify the advantage of the WLAR model over the standard linear AR model when modelling improper processes, we start from the normal equations for the widely linear AR model and find the augmented covariance matrix $\mathcal{R}_{\mathbf{zz}}$ from

$$\mathcal{R}^*_{\mathbf{zz}} \begin{bmatrix} \mathbf{h} \\ \mathbf{g} \end{bmatrix} = \mathbf{r}^*$$

$$\begin{bmatrix} \mathcal{C}^*_{\mathbf{zz}} & \mathcal{P}^*_{\mathbf{zz}} \\ \mathcal{P}_{\mathbf{zz}} & \mathcal{C}_{\mathbf{zz}} \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} \mathbf{c}^* \\ \mathbf{p} \end{bmatrix}, \tag{A.1}$$

$$\mathcal{C}^*_{\mathbf{zz}}\mathbf{h} + \mathcal{P}^*_{\mathbf{zz}}\mathbf{g} = \mathbf{c}^* \quad \longrightarrow \quad \mathbf{h} = \mathcal{C}^{*^{-1}}_{\mathbf{zz}}\left[\mathbf{c}^* - \mathcal{P}^*_{\mathbf{zz}}\mathbf{g}\right], \tag{A.2}$$

$$\mathcal{P}_{\mathbf{zz}}\mathbf{h} + \mathcal{C}_{\mathbf{zz}}\mathbf{g} = \mathbf{p} \quad \longrightarrow \quad \mathbf{g} = \mathcal{C}^{-1}_{\mathbf{zz}}\left[\mathbf{p} - \mathcal{P}_{\mathbf{zz}}\mathbf{h}\right]. \tag{A.3}$$

Solving this system of equations gives

$$\begin{aligned}
\mathcal{C}^{*}_{\mathbf{zz}}\mathbf{h}+\mathcal{P}^{*}_{\mathbf{zz}}\left[\mathcal{C}^{-1}_{\mathbf{zz}}\left[\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathbf{h}\right]\right] =&\mathbf{c}^{*}\\
\left[\mathcal{C}^{*}_{\mathbf{zz}}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{-1}_{\mathbf{zz}}\mathcal{P}_{\mathbf{zz}}\right]\mathbf{h} =&\mathbf{c}^{*}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{-1}_{\mathbf{zz}}\mathbf{p}\\
\mathbf{h} =&\left[\mathcal{C}^{*}_{\mathbf{zz}}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{-1}_{\mathbf{zz}}\mathcal{P}_{\mathbf{zz}}\right]^{-1}\left[\mathbf{c}^{*}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{-1}_{\mathbf{zz}}\mathbf{p}\right], \quad \text{(A.4)}\\
\mathcal{P}_{\mathbf{zz}}\left[\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\left[\mathbf{c}^{*}-\mathcal{P}^{*}_{\mathbf{zz}}\mathbf{g}\right]\right]+\mathcal{C}_{\mathbf{zz}}\mathbf{g} =&\mathbf{p}\\
\left[\mathcal{C}_{\mathbf{zz}}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathcal{P}^{*}_{\mathbf{zz}}\right]\mathbf{g} =&\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\\
\mathbf{g} =&\left[\mathcal{C}_{\mathbf{zz}}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathcal{P}^{*}_{\mathbf{zz}}\right]^{-1}\left[\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\right]. \quad \text{(A.5)}
\end{aligned}$$

Using these values of $\mathbf{h}$ and $\mathbf{g}$, the difference between the squared estimation errors in (16) becomes

$$\begin{aligned}
\delta e^{2} =&\mathbf{c}^{T}\left[\mathcal{C}^{*}_{\mathbf{zz}}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{-1}_{\mathbf{zz}}\mathcal{P}_{\mathbf{zz}}\right]^{-1}\left[\mathbf{c}^{*}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{-1}_{\mathbf{zz}}\mathbf{p}\right]\\
&+\mathbf{p}^{H}\left[\mathcal{C}_{\mathbf{zz}}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathcal{P}^{*}_{\mathbf{zz}}\right]^{-1}\left[\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\right]-\mathbf{c}^{T}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}, \quad \text{(A.6)}
\end{aligned}$$

which can be rewritten as

$$\delta e^{2}=\left[\mathbf{p}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\right]^{H}\left[\mathcal{C}_{\mathbf{zz}}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathcal{P}^{*}_{\mathbf{zz}}\right]^{-1}\left[\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\right]. \quad \text{(A.7)}$$

Thus, as the matrix $\left[\mathcal{C}_{\mathbf{zz}}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathcal{P}^{*}_{\mathbf{zz}}\right]$ is positive definite, the term $\delta e^{2}=0$ only when either $\left[\mathbf{p}-\mathcal{P}^{*}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\right]=\mathbf{0}$ or $\left[\mathbf{p}-\mathcal{P}_{\mathbf{zz}}\mathcal{C}^{*^{-1}}_{\mathbf{zz}}\mathbf{c}^{*}\right]=\mathbf{0}$.

## References


[1] D. P. Mandic, S. Javidi, G. Souretis, S. L. Goh, Why a complex valued solution for a real domain problem, in: Proceedings IEEE International Workshop on Machine Learning for Signal Processing, 2007, pp. 384–389.

[2] M. H. Hayes, Statistical Digital Signal Processing and Modeling, Wiley, 1996.

[3] B. Widrow, J. McCool, M. Ball, The complex LMS algorithm, Proceedings of the IEEE 63 (4) (1975) 719–720.

[4] B. Picinbono, On circularity, IEEE Transactions on Signal Processing 42 (12) (1994) 3473–3482.

[5] F. Neeser, J. Massey, Proper complex random processes with application to information theory, IEEE Transactions on Information Theory 39 (4) (1993) 1293–1302.

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

[7] P. Schreier, L. Scharf, Second-order analysis of improper complex random vectors and processes, IEEE Transactions on Signal Processing 51 (3) (2003) 714–725.

[8] H. Kantz, T. Schreiber, Nonlinear Time Series Analysis, Cambridge University Press, 1997.

[9] T. Schreiber, A. Schmitz, On the discrimination power of measures for nonlinearity in a time series, Physical Review E 55 (5) (1997) 5443–5447.

[10] D. P. Mandic, M. Chen, T. Gautama, M. M. Van Hulle, A. Constantinides, On the characterisation of the deterministic/stochastic and linear/nonlinear nature of time series, Proceedings of the Royal Society A 464 (2008) 1141–1160.

[11] T. Gautama, D. P. Mandic, M. V. Hulle, A non–parametric test for detecting the complex–valued nature of time series, International Journal of Knowledge–Based Intelligent Engineering Systems 8 (2) (2004) 99–106.

[12] P. J. Schreier, The degree of impropriety (noncircularity) of complex random vectors, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2008, pp. 3909–3912.

[13] P. J. Schreier, L. L. Scharf, A. Hanssen, A generalized likelihood ratio test for impropriety of complex signals, IEEE Signal Processing Letters 13 (7) (2006) 433.

[14] E. Ollila, On the circularity of a complex random variable, IEEE Signal Processing Letters 15 (2008) 841–844.

[15] B. Jelfs, S. Javidi, P. Vayanos, D. Mandic, Characterisation of signal modality: Exploiting signal nonlinearity in machine learning and signal processing, Journal Signal Processing Systems 61 (1) (2010) 105–115.

[16] B. Jelfs, Y. Xia, D. Mandic, S. Douglas, Collaborative adaptive filtering in the complex domain, in: Proceedings IEEE International Workshop on Machine Learning for Signal Processing, 2008, pp. 421–425.

[17] S. Javidi, M. Pedzisz, S. L. Goh, D. P. Mandic, The augmented complex least mean square algorithm with application to adaptive prediction problems, in: Proceedings of the IAPR Workshop on Cognitive Information Processing, 2008, pp. 54–57.

[18] B. Picinbono, P. Chevalier, Widely linear estimation with complex data, IEEE Transactions On Signal Processing 43 (8) (1995) 2030–2033.

[19] B. Picinbono, Random Signals and Systems, Englewood Cliffs, Prentice Hall, 1993.

[20] J. Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm, IEEE Transactions on Signal Processing 56 (7) (2008) 3061–3068.

[21] P. J. Schreier, L. L. Scharf, Statistical Signal Processing of Complex-Valued Data, Cambridge University Press, 2010.

[22] A. H. Sayed, Fundamentals of Adaptive Filtering, Wiley, 2003.

[23] B. Picinbono, M. Bouvet, Complex white noises and autoregressive signals, in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 9, 1984, pp. 596–599.

[24] B. Picinbono, Wide-sense linear mean square estimation and prediction, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 3, 1995, pp. 2032–2035.

[25] R. Schober, W. H. Gerstacker, L. H.-J. Lampe, A widely linear LMS algorithm for MAI suppression for DS-CDMA, in: IEEE International Conference on Communications, Vol. 4, 2003, pp. 2520–2525.

[26] J. Arenas-Garcia, A. R. Figueiras-Vidal, A. H. Sayed, Mean-square performance of a convex combination of two adaptive filters, IEEE Transactions on Signal Processing 54 (3) (2006) 1078–1090.

[27] S. C. Douglas, D. P. Mandic, Mean and mean–square analysis of the complex LMS algorithm for non–circular Gaussian signals, in: Proceedings of the IEEE Digital Signal Processing Workshop, 2009, pp. 101–106.

[28] S. C. Douglas, D. P. Mandic, Performance analysis of the conventional complex LMS and augmented complex LMS algorithms, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2010, pp. 3794–3797.

[29] S. C. Douglas, W. Pan, Exact expectation analysis of the LMS adaptive filter, IEEE Transactions on Signal Processing 43 (12) (1995) 2863–2871.

[30] S. Haykin, Adaptive Filter Theory, 4th Edition, Prentice Hall, 2002.

[31] B. Widrow, S. Stearns, Adaptive Signal Processing, Prentice–Hall, 1985.

[32] S. L. Goh, M. Chen, D. Popović, K. Aihara, D. Obradovic, D. P. Mandic, Complex-valued forecasting of wind profile, Renewable Energy 31 (11) (2006) 1733–1750.