

# WHY A COMPLEX VALUED SOLUTION FOR A REAL DOMAIN PROBLEM

*Danilo P. Mandic, Soroush Javidi, George Souretis, and Vanessa S. L. Goh*

Department of Electrical and Electronic Engineering, Imperial College, London, UK  
{d.mandic,soroush.javidi,g.souretis,s.goh}@imperial.ac.uk

## ABSTRACT

An insight into the potential benefits of using complex valued models for real valued data is provided. The problem itself is not new; it is however timely and important to revisit this issue, due to a plethora of modern applications based on multidimensional and multichannel measurements which can be cast into an equivalent problem in the field of complex numbers  $\mathbb{C}$ . The analysis and simulations highlight the duality between several classes of real domain problems and their complex valued representations. This is supported by case studies on image processing, modelling of point processes for brain prosthetics, and forecasting of vector fields.

## 1. INTRODUCTION

Recent breakthroughs in technology and biomedicine have highlighted applications where nonlinearity, nonstationarity, multidimensional data natures, and uncertainty play major roles. Subsequently novel signal processing and machine learning theories have been rapidly developed in order to cater for these new classes of problems.

When it comes to nonlinear models, over the past decade, much research effort has been dedicated to the development of new approaches suitable for nonlinear phenomena, the landmark in this direction arguably being the Santa Fe prediction competition [1]. On the other hand, linear models, although often suboptimal, are well established and understood, and widely applied in practice. From the practical viewpoint, it is natural to ask ourselves whether nonlinear solutions are general enough to be applicable to both nonlinear and linear scenarios. One important result in that direction is by T. Adali [2], where it was shown that even though the problem in hand may be linear, nonlinear solutions can exhibit far superior performance, especially in the presence of noise. The use of nonlinear models for linear noisy problems was also proven beneficial in the radar and sonar literature, and more recently in blind source separation [3].

Apart from the problems with nonlinearity, nonstationarity and noise, novel applications in brain science, communications, and data and sensor fusion are dealing with mul-

tidimensional (multichannel) measurements, for which the online processing algorithms are essential [4]. If possible (data dimensionality, computational power), the processing of such data should be performed directly in the spaces where the data vectors reside ( $\mathbb{R}^N, \mathbb{C}^N$ ). Machine learning theories for multidimensional domains are still being developed; this is especially in the presence of nonlinearity and noise.

Following a related study for nonlinear models [2], it is natural to ask ourselves whether it is beneficial to use multidimensional solutions for lower dimensional problems. The simplest, yet extremely important, case to consider is whether complex valued solutions for real valued problems provide a theoretical and performance advantage over standard real valued solutions.

In applications of adaptive systems it is the signal magnitude that is used as the main source of information [5]. Whereas this facilitates the established algorithms in  $\mathbb{R}$ , it is important to realise that this way the full information potential (phase information) within the signal is not utilised. Indeed real world processes with the “intensity” and “direction” component (radar, sonar, vector fields) require the consideration of such phase information. Our question then boils down to whether the processing of such real-valued data in the complex domain  $\mathbb{C}$ , where the phase information can be accounted for naturally, has advantages over straightforward processing in  $\mathbb{R}$ . For instance our recent results on the modelling of wind profile [6], indicate that the processing in  $\mathbb{C}$  (simultaneously speed and direction as a complex vector) has major benefits over the direct processing in  $\mathbb{R}$  (speed only), or in  $\mathbb{R}^2$  (speed and direction as independent processes).

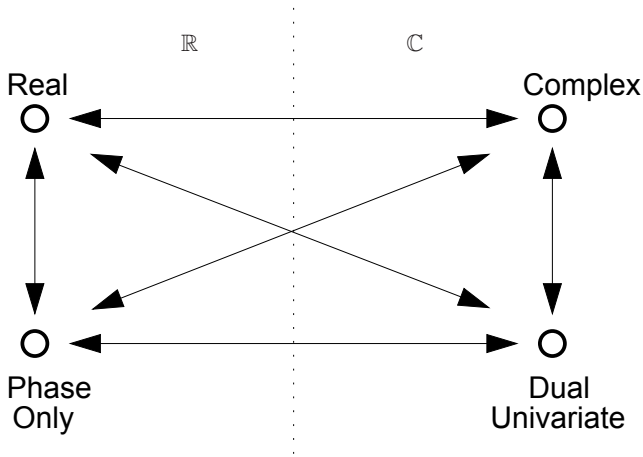
Our aim is to provide a comprehensive account of the advantages of using complex valued solutions for several classes of real valued problems. It is shown that these potential benefits arise not only from the full utilisation of the phase information, but also from the use of different algebra; this provides rigorous compact framework for the processing of both complex and real signals. The analysis is supported by experimental evidence on case studies including examples from image processing, processing of point processes, and wind forecasting for renewable energy applications. These

help towards answering our underlying question “Why complex valued solutions for real valued problems”.

## 2. REAL VS COMPLEX VALUED PROCESSING

When it comes to the duality between real and complex processes and highlight two aspects of this duality:- the importance of phase (direction) information, and the advantages of the simultaneous modelling of the “intensity” and “direction” component of vector field processes in  $\mathbb{C}$  over the so called “dual univariate” modelling (where the component of such processes are treated as independent random processes).

Complex valued data are complex either by their nature (e.g. communications) or by convenience of representation (e.g. phasors). Figure 1 illustrates such duality between the processing in  $\mathbb{R}$  and  $\mathbb{C}$ . The nature of purely *real* and *complex* sig-



**Fig. 1.** Duality between signal processing in  $\mathbb{R}$  and  $\mathbb{C}$ . Arrows show the direction in which data can be transformed.

nals is obvious – *real* signals are magnitude-only whereas *complex* signals comprise both magnitude and phase components. *Phase only* signals are real signals formed from the phase of a complex signal. *Dual univariate* signals are the components of a complex signal that are processed separately as real valued quantities.

In order to transform a real signal into its complex valued counterpart, it is convenient to make use of the delay or phase associated with the time of arrival of the real valued signal (or vector field) at sensors. When transforming a complex signal into its dual univariate representation, we need to identify the heterogeneous (e.g. wind speed and direction) components within the complex representation. Another interesting scenario is the “complex to phase only” transformation for phase only modelling (tracking); this is very practical for cases where the magnitude of the signal has small or no variation [7]. There are a variety of other ways of transforming signals from  $\mathbb{R}$  into their  $\mathbb{C}$  domain

counterparts, most of them are application specific. We shall now illustrate these transformations using examples in image processing and in the processing of wind field.

### 2.1. Phase Information in Imaging

Oftentimes in machine learning our aim is to preserve or enhance the phase of a signal(s); this plays major part in perception based modelling (video, speech, haptic, multi-modal). While the phase information in 1-D signals is subtly hidden, in 2-D signals such as images, the role of the phase of a signal is more obvious [5]. The top panel in Figure 2 shows two greyscale images denoted by  $I_1$  (wheel) and  $I_2$  (child); the 2-D Fourier transform is applied and the phase spectra of  $I_1$  and  $I_2$  are swapped to give

$$S_1(x, y) = |\mathcal{F}(I_1)| \cdot \exp(j \cdot \angle(\mathcal{F}(I_2))) \quad (1)$$

$$S_2(x, y) = |\mathcal{F}(I_2)| \cdot \exp(j \cdot \angle(\mathcal{F}(I_1))) \quad (2)$$

Thus,  $S_1$  has the magnitude spectrum of  $I_1$  and phase spectrum of  $I_2$ , whereas  $S_2$  has the magnitude spectrum of  $I_2$  and phase spectrum of  $I_1$ . Finally, the inverse Fourier transform (IVT) is applied to spectra  $S_1$  and  $S_2$  and the resulting images are displayed in the bottom panel of Figure 2. Observe that  $\hat{I}_1 = \mathcal{F}^{-1}(S_1)$  which has the amplitude spectrum of  $I_1$  and phase spectrum of  $I_2$  appears as a noisy version of  $I_2$  (child) but in the background similar to that of  $I_1$ . Similar observations apply to  $\hat{I}_2 = \mathcal{F}^{-1}(S_2)$  which is recognised as a wheel, making it clear that as far as the human visual system is concerned, it is the phase which is the information carrier.

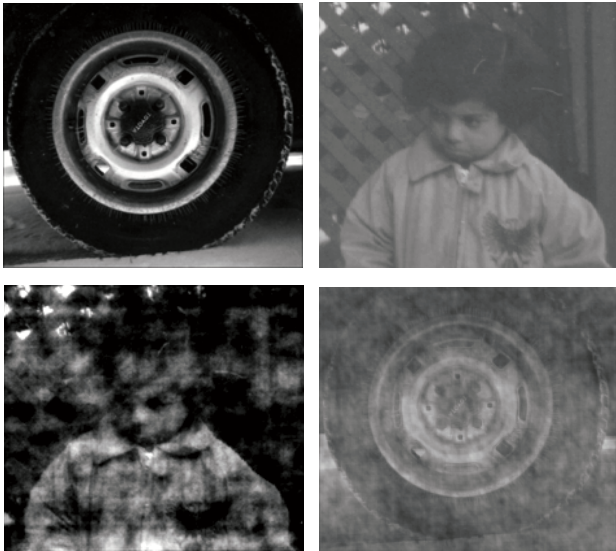
It is therefore natural and convenient to conduct processing of images based solely on their complex valued representation [8]. One way to achieve this would be to consider a greyscale or “intensity” image

$$I = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \quad (3)$$

where for an image with  $n$  intensity levels, elements  $a_{ij}$  take values from  $\{0, 1, \dots, n\}$ . Alternatively, these elements can be modelled as  $n^{th}$  (complex) roots on the unit circle in the  $z$ -plane, that is  $\exp(j \frac{2\pi k}{n})$ ,  $k = 1, 2, \dots, n$ , thus obtaining a phase-only complex representation of the image. The subsequent image processing can be now performed directly in  $\mathbb{C}$  [9].

### 2.2. Modelling of Directional Processes

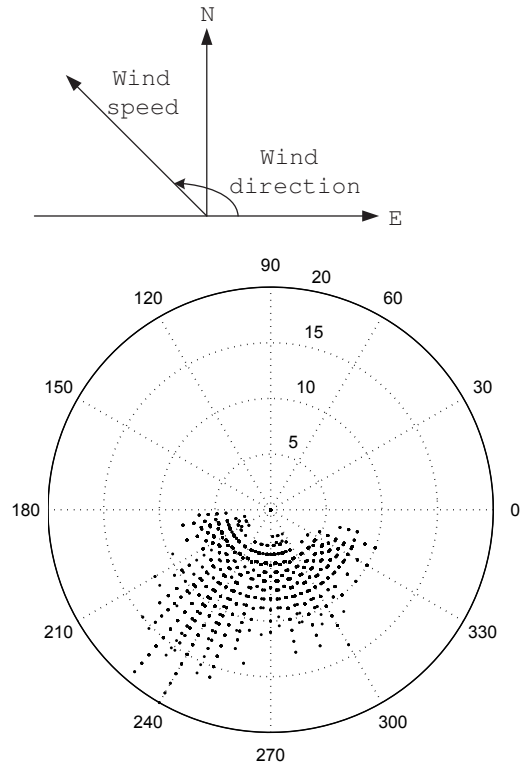
Consider a class of processes with the “intensity” and “direction” components, such as wind, radar, sonar, or sensor



**Fig. 2.** Surrogate images. *Top:* Original images  $I_1$  and  $I_2$ ; *Bottom:* Images  $\hat{I}_1$  and  $\hat{I}_2$  generated by exchanging the amplitude and phase spectrum of the original images

array measurements. Figure 3 represents a wind measurement<sup>1</sup> as a vector  $\mathbf{w}(t) = v(t)e^{jd(t)}$  of its speed  $v(t)$  and direction  $d(t)$  components, in the  $N - E$  coordinate system, together with the distribution of wind speeds over various directions. Notice that this way we have inherently introduced *fusion* of the heterogeneous speed and direction wind components, and represented them as a single quantity in the complex domain  $\mathbb{C}$  [10]. Despite the clear interdependence between wind signal components (wind lattice in Figure 3), the existing models treat these as dual *univariate* time series, introducing error in both the wind dynamics and associated forecasts. We can therefore exploit the natural coupling between the real and imaginary part in  $\mathbb{C}$ , in order to simultaneously estimate the changes in the dynamics of both the speed and direction component of wind. This gives us the opportunity to rigorously develop novel architectures and algorithms for greatly improved and computationally efficient modelling and robust forecasting [11]. Figure 4 illustrates the performance gain obtained by using the complex wind model. Clearly the dual univariate approach (see Figure 1) using two real models (which were combined into a complex quantity at the output) for the signal components was not able to capture the wind dynamics, whereas the fully complex approach exhibited excellent per-

<sup>1</sup>The wind signal used was obtained from readings from the Iowa (USA) Department of Transport <http://mesonet.agron.iastate.edu/request/awos/1min.php> database. The complex wind vector is estimated as  $\mathbf{v} = ve^{jd}$ , where  $v$  denotes the speed and  $d$  its direction.



**Fig. 3.** Wind recordings: *Top:* a complex-valued representation, *Bottom:* wind lattice

formance.<sup>2</sup>

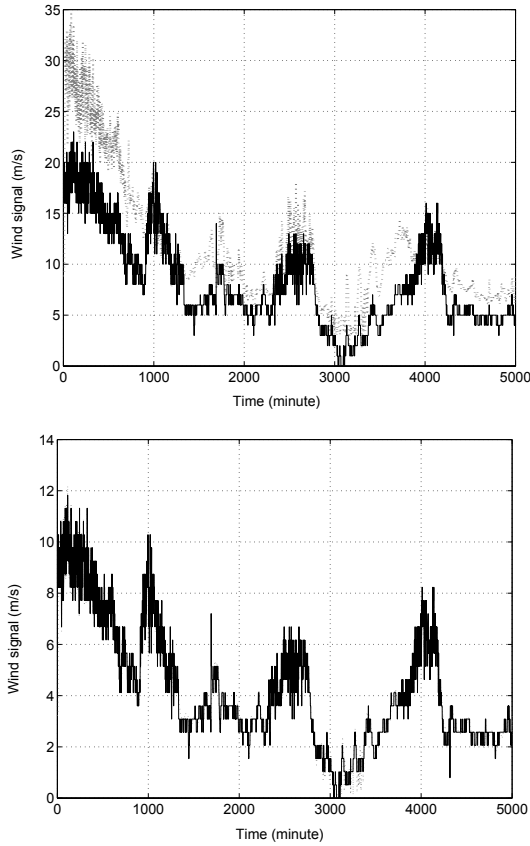
### 3. EXPLOITING THE PHASE INFORMATION

There are several ways to obtain phase from real world signals. One convenient way is to apply the Hilbert transform to the signal in hand and obtain an analytic signal, from which both the phase and instantaneous frequency can be obtained [12]. In online adaptive signal processing applications it is also possible to consider a “mixed norm” type approach, with one term within the cost function representing the minimum mean squared error (MMSE) and the other some measure of the phase [7].

#### 3.1. Complex processing of spiking neurons

The processing of spiky neuronal recordings presents considerable difficulties due to the underlying point process statistics, and nonlinear and nonstationary nature of the data. In addition, the information is typically coded through synchronised or phase-locked events. One major problem when obtaining information from an implanted microarray of microelectrodes is spike synchronisation. This problem is very difficult to solve in  $\mathbb{R}$ , however, time delays between the

<sup>2</sup>The simulations were based on LMS and complex LMS.



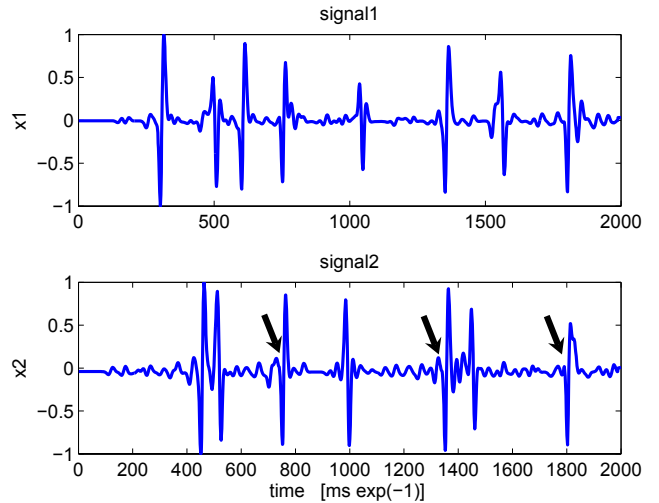
**Fig. 4.** Complex vs dual univariate modelling, dashed line represents the original signal and solid line the one step ahead prediction. *Top:* Dual univariate (real valued) prediction. *Bottom:* Fully complex prediction

spiky events are straightforward to model in  $\mathbb{C}$ .

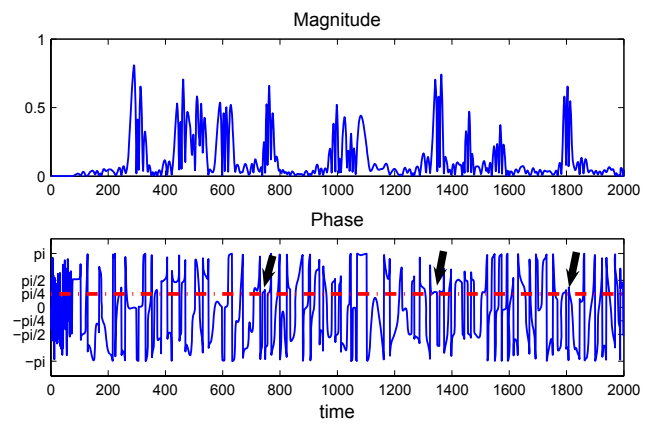
Consider two simultaneous recordings from two artificially generated point processes  $x_1$  and  $x_2$ , depicted in Figure 5(a). Since the spike synchronisation problem boils down to the issue of phase matching, let us form a complex valued signal

$$z = x_1 + jx_2 \quad (4)$$

for which the real part is  $x_1$  and the imaginary part  $x_2$ . To detect the synchronised spike events, we may employ a complex version of Empirical Mode Decomposition (EMD) [13], or the Rotation Invariant EMD [14], where the synchronisation of the spike events within  $x_1$  and  $x_2$  can be examined straightforwardly from  $z$ , when its phase equals  $\pi/4$ , as depicted in Figure 5(b). A rigorous approach for converting real spike sequences into their complex counterparts is proposed in [15] by means of a new spatiotemporal neural algebra. The main underlying idea is to code spike events as complex numbers, whereby the magnitude of a complex number represents the “magnitude” of the spike and the phase of the complex number codes the interspike interval.



(a) Point processes  $x_1$  and  $x_2$ .



(b) Synchronised spike events detected at the phase values of  $\pi/4$ . (compare with Figure 5(a)).

**Fig. 5.** Synchronised spike event detection. The arrows indicate the instances of the synchronisation between the two series of spiky events.

### 3.2. Incorporating Phase Within the Cost Function

Standard adaptive filtering algorithms are based on the minimisation of MMSE, which is perfectly reasonable for real valued signals. However, in several classes of complex signals the magnitude information is not an information carrier (MPSK, QPSK), and it is the phase that conveys the information. In those cases, when designing an adaptive filtering algorithm it is important to consider both the error in magnitude and phase of the received symbols in order to provide efficient learning for fast channel variations. One such algorithm is the recently introduced Least Mean Phase–Least Mean Square (LMP–LMS) [7] algorithm for which the cost function is given by

$$J(\mathbf{w}) = k_1 \underbrace{E |d - \mathbf{x}^T \mathbf{w}|^2}_{LMS} + k_2 \underbrace{E |\angle d - \angle \mathbf{x}^T \mathbf{w}|^m}_{LMP} \quad (5)$$



where  $\mathbf{x}$  is the input signal,  $d$  is the teaching signal,  $\mathbf{w}$  are the filter coefficients,  $m \in \{1, 2\}$ , and  $k_1$  and  $k_2$  are mixing coefficients. From (5), both the phase and magnitude error term are present within the cost function, for which the weight update is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu_1 e(k) \mathbf{x}^*(k) + \mu_2 \angle e(k) \frac{j \mathbf{x}^*(k)}{y^*(k)}, \quad (6)$$

where  $y$  denotes the output of the filter,  $(\cdot)^*$  is the complex conjugation operator and  $\mu_1$  and  $\mu_2$  are positive learning rates.

### 3.3. Connection Between Real Valued and Complex Valued Adaptive Filtering

Consider the adaptive denoising configuration shown in Figure 6. The primary input is the useful signal corrupted by

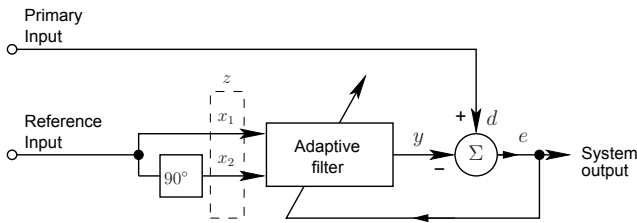


Fig. 6. The noise cancelling adaptive filtering configuration.

additive noise, whereas the reference input is any noise correlated with the noise within the primary input. In [16] it was shown that the scheme from Figure 6 outperforms the standard noise cancellation scheme in ECG type of applications. In the spirit of this paper, it is clear that by introducing a phase shift of  $\pi/2$  within the filter, inputs  $x_1$  and  $x_2$  can be combined within an equivalent complex input  $z$ , to be processed by the complex LMS [17], given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k) \mathbf{x}^*(k) \quad (7)$$

## 4. BENEFITS OF COMPLEX DOMAIN PROCESSING

Not only based on the examples of complex domain processing for real valued signals introduced so far, the processing in  $\mathbb{C}$  possesses several distinctive features which are not present in  $\mathbb{R}$ . These include:-

- Different statistics: although it is often assumed that the statistics in  $\mathbb{C}$  is a straightforward extension of the statistics in  $\mathbb{R}$  (replace  $(\cdot)^T$  by the Hermitian  $(\cdot)^H$ ), some recent results in complex statistics have introduced the notions of proper ( $E\{\mathbf{x}\mathbf{x}^T\} = \mathbf{0}$ ) and improper ( $E\{\mathbf{x}\mathbf{x}^T\} \neq \mathbf{0}$ ) complex random variables

[18]. This allows us to account for a wide range of real-world signals which are processed in  $\mathbb{R}$ ;

- A variety of ways to “complexify” the signal in hand [19]. Apart from the dual univariate approach (Figure 1) which applies mainly to linear processes, ways to introduce nonlinearity include complex and fully complex approach [19]. Within the split complex approach, a pair of nonlinear activation functions is employed to separately process real and imaginary components of the net input. A split-complex activation function is given by

$$f(x) = f_R(\text{Re}(x)) + j f_I(\text{Im}(x)) \quad (8)$$

where  $f_R = f_I$  are real functions of real variable.

Fully complex functions are standard complex functions of complex variables, such as the complex tanh.

The issue of the second order statistics in  $\mathbb{C}$  is particularly important and ought to be further clarified. Since the covariance matrix  $E[\mathbf{x}\mathbf{x}^H]$  will not completely describe the second-order statistical behavior of  $\mathbf{x}$ , we need to consider both variable  $\mathbf{x}$  and its complex conjugate  $\mathbf{x}^*$  in order to make full use of the available statistical information. This additional information is contained in the cross-moments [18].

To derive so called augmented learning algorithms, consider a complex vector  $\mathbf{x}$  and its conjugate to produce a  $(2n \times 1)$  vector  $\mathbf{x}^a = [\mathbf{x}, \mathbf{x}^*]$ . Consequently, it is the augmented  $(2n \times 2n)$  complex covariance matrix  $\mathbf{P}_{\mathbf{x}^a \mathbf{x}^a} = E[\mathbf{x}^a (\mathbf{x}^a)^T]$  (rather than just the  $n \times n$  matrix  $\mathbf{P}_{\mathbf{x}\mathbf{x}} = E[\mathbf{x}\mathbf{x}^H]$ ) that contains the complete second-order statistical information. Such an augmented covariance matrix is given by [20]

$$\mathbf{P}_{\mathbf{x}^a \mathbf{x}^a} = E \begin{bmatrix} \mathbf{x} \\ \mathbf{x}^* \end{bmatrix} [\mathbf{x}^T \mathbf{x}^H] = \begin{bmatrix} \mathbf{P}_{\mathbf{x}\mathbf{x}} & \mathbf{P}_{\mathbf{x}\mathbf{x}}^\xi \\ \mathbf{P}_{\mathbf{x}\mathbf{x}}^{\xi*} & \mathbf{P}_{\mathbf{x}\mathbf{x}}^* \end{bmatrix} \quad (9)$$

Recently derived augmented algorithms include the augmented Extended Kalman filter (AEKF) and augmented CRTRL (ACRTRL) algorithm [21]. To illustrate the power of augmented complex statistics Table 1 illustrates the improved quantitative performance of the augmented CRTRL over the standard CRTRL for the task of one step ahead prediction of a stable  $AR(4)$  process and the complex radar data<sup>3</sup>.

## 5. CONCLUSIONS

A comprehensive account of the benefits of using complex valued models for real valued processes has been provided. Unlike in standard transform domain processing, where the DCT or DFT are employed to transform the data into  $\mathbb{C}$  in

<sup>3</sup>Radar (high) is referred to as “high sea state data” and radar (low) is referred to as “low sea state data”. Publicly available from “<http://soma.ece.mcmaster.ca/ipix/>”.

**Table 1.** Comparison of prediction gains  $R_p$  [dB] for the various classes of signals

$R_p$ [dB]	CRTRL	ACRTRL
Linear $AR(4)$ process	3.22	4.10
Radar (low)	11.40	13.57
Radar (high)	4.56	5.41

order to decorrelate the input and hence improve convergence, our focus has been on direct time domain processing. It has been shown that in many cases where the phase information within the signal is important, complex valued processing represents not only a viable alternative to real domain processing, but also a natural processing domain.

## 6. REFERENCES

- [1] A. Weigend and N. Gershenfeld, *Tim Series Prediction: Forecasting the Future and Understanding the Past*, Wesley, 1998.
- [2] T. Adali, "Why a nonlinear solution for a linear problem?," in *Proceedings of the IEEE Workshop on Neural Networks for Signal Processing*, 1999, vol. 9, pp. 157–165.
- [3] W. Y. Leong and D. P. Mandic, "Blind extraction of noisy event using nonlinear predictor," in *Proceedings of the IEEE ICASSP 2007*, 2007, p. tba.
- [4] D. P. Mandic, M. Golz, A. Kuh, D. Obradovic, and T. Tanaka, *Signal processing techniques for Knowledge Extraction and Information Fusion*, Springer, 2007.
- [5] A.V. Oppenheim and J.S. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.
- [6] D. P. Mandic and S. L. Goh, *Complex Valued Nonlinear Adaptive Filters*, Wiley, 2007.
- [7] A. Tarighat and A.H. Sayed, "Least mean-phase adaptive filters with application to communications systems," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 220–223, 2004.
- [8] H. Aoki, M. R. Azimi-Sadjadi, and Y. Kosugi, "Image association using a complex-valued associative memory model," *Transactions of the IEICE*, vol. 83A, pp. 1824–1832, 2000.
- [9] A. Hirose, *Complex valued neural networks, Theories and Applications*, World Scientific, 2003.
- [10] Danilo P. Mandic, Su Lee Goh, and Kazuyuki Aihara, "Sequential data fusion via vector spaces: Fusion of heterogeneous data in the complex domain," *The Journal of VLSI Signal Processing*, Oct. 2006.
- [11] S.L. Goh, M. Chen, D.H. Popovic, K. Aihara, D. Obradovic, and D.P. Mandic, "Complex-valued forecasting of wind profile," *Renewable Energy*, vol. 31, no. 11, pp. 1733–50, 2006.
- [12] Tristan Needham, *Visual Complex Analysis*, Oxford University Press, 2000.
- [13] N.E. Huang, Z. Shen, S.R. Long, M. L. Wu, H. H. Shih, Z. Quanan, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A*, vol. 454, pp. 903–995, 1998.
- [14] M. Umair Bin Altaf, T. Gautama, T. Tanaka, and D. P. Mandic, "Rotation invariant complex empirical mode decomposition," *accepted for publication in IEEE ICASSP*, 2007.
- [15] G. Vaucher, "A complex valued spiking machine," *Lecture Notes in Computer Science*, pp. 967–976, 2003.
- [16] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, 1985.
- [17] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719–720, 1975.
- [18] F. D. Neeser and J. L. Massey, "Proper complex random processes with application to information theory," *IEEE Transactions on Information Theory*, vol. 39, pp. 1293–1302, 1992.
- [19] Taehwan Kim and Tulay Adali, "Approximation by Fully Complex Multilayer Perceptrons," *Neural Comp.*, vol. 15, no. 7, pp. 1641–1666, 2003.
- [20] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Transactions on Signal Processing*, vol. 51, pp. 714–725, 2003.
- [21] S. L. Goh and D. P. Mandic, "An augmented extended Kalman filter algorithm for complex-valued recurrent neural networks," *Neural Computation*, vol. 19, no. 4, pp. 1–17, 2003.