1

Collaborative Adaptive Filters for Online Knowledge Extraction and Information Fusion

Beth Jelfs[1], Phebe Vayanos[1], Soroush Javidi[1], Vanessa Su Lee Goh[2], and Danilo P. Mandic[2]

Imperial College London, UK[1], Shell Exploration and Production, Assen, The Netherlands[2]

We present a method for extracting information (or knowledge) about the nature of a signal, this is achieved by employing recent developments in signal characterisation for online analysis of the changes in signal modality. We show that it is possible to use the fusion of the outputs of adaptive filters to produce a single collaborative hybrid filter and that by tracking the dynamics of the mixing parameter of this filter rather than the actual filter performance, a clear indication as to the nature of the signal is given. Implementations of the proposed hybrid filter in both the real $\mathbb{R}$ and complex $\mathbb{C}$ domains are analysed and the potential of such a scheme for tracking signal nonlinearity in both domains is highlighted. Simulations on linear and nonlinear signals in a prediction configuration support the analysis; real world applications of the approach have been illustrated on electroencephalogram (EEG), radar and wind data.

1.1 Introduction

Signal modality characterisation is becoming an increasingly important area of multidisciplinary research and large effort has been put into devising efficient algorithms for this purpose. Research in this area started in physics in the mid 1990s but its applications in machine learning and signal processing are only recently becoming apparent. Before discussing characterisation of signal modalities certain key properties for defining the nature of a signal should be outlined [8, 21]:

1. Linear (strict definition) – a linear signal is generated by a linear time-invariant system, driven by white Gaussian noise;
2. Linear (commonly adopted) – definition 1. is relaxed somewhat by allowing the distribution of the signal to deviate from the Gaussian one, which can be interpreted as a linear signal from 1. measured by a static (possibly nonlinear) observation function;

3. Nonlinear – a signal that cannot be generated in the above way is considered nonlinear;
4. Deterministic (predictable) – a signal is considered deterministic if it can be precisely described by a set of equations;
5. Stochastic – a signal that is not deterministic.

Figure 1.1 (modified from [19]) illustrates the range of signals spanned by the characteristics of nonlinearity and stochasticity. Whilst signals with certain characteristics are well defined, for instance chaotic signals (nonlinear and deterministic) or those produced by autoregressive moving average (ARMA) models (linear and stochastic signals), these represent only the extremes in signal nature and do not highlight the majority of signals which do not fit into such classifications. Due to the presence of such factors as noise or uncertainty, any real world signals are represented in the areas (a), (b), (c) or '?'; these are significant areas about which we know little or nothing. As changes in the signal nature between linear and nonlinear and deterministic and stochastic can reveal information (knowledge) which is critical in certain applications (e.g. health conditions) the accurate characterisation of the nature of signals is a key prerequisite prior to choosing a signal processing framework.
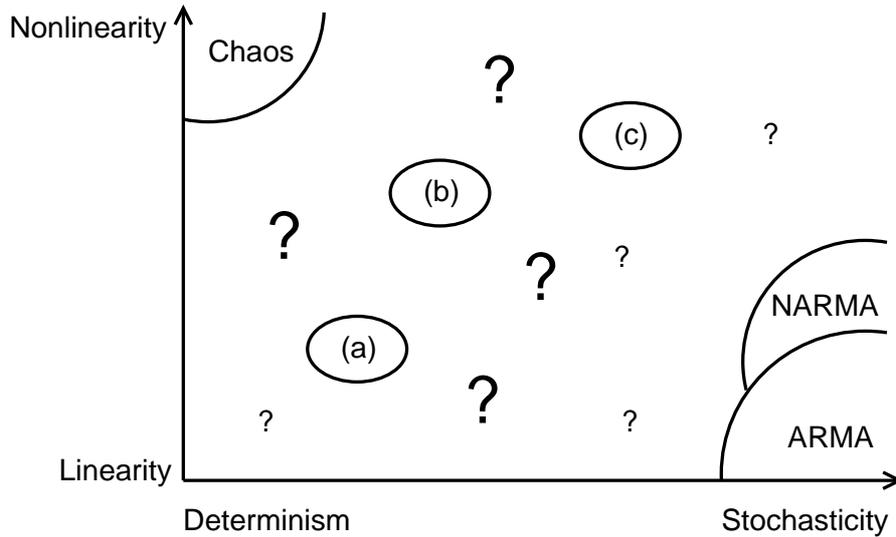
Fig. 1.1. Deterministic vs. stochastic nature or linear vs. nonlinear nature

The existing algorithms in this area are based on hypothesis testing [6, 7, 20] and describe the signal changes in a statistical manner. However, there are very few online algorithms which are suitable for this purpose. The purpose of the approach described in this chapter is to introduce a class of online

algorithms which can be used not only to identify, but also to track changes in the nature of the signal (signal modality detection).

One intuitive method to determine the nature of a signal has been to present the signal as input to two adaptive filters with different characteristics, one nonlinear and the other linear. By comparing the responses of each filter, this can be used to identify whether the input signal is linear or not. Whilst this is a very useful simple test for signal nonlinearity, it does not provide an online solution. There are additional ambiguities due to the need to choose many parameters of the corresponding filters and this approach does not rely on the "synergy" between the filters considered.

1.1.1 Previous Online Approaches

In [17] an online approach is considered which successfully tracks the degree of nonlinearity of a signal using adaptive algorithms, but relies on a parametric model to effectively model the system in order to provide a true indication of the degree of nonlinearity. Figure 1.2 shows an implementation of this method using a third order Volterra filter and the normalised least mean square (NLMS) algorithm with a step size $\mu = 0.008$ to update the system parameters. The system input and output can be described by

$$u[k] = \sum_{i=0}^{I} a_i x[k-i] \text{ where } I = 2 \text{ and } a_0 = 0.5, a_1 = 0.25, a_2 = 0.125$$
$$y[k] = F(u[k]; k) + \eta[k] \tag{1.1}$$

where $x[k]$ are i.i.d uniformly distributed over the range $[-0.5, 0.5]$ and $\eta[k] \sim \mathcal{N}(0, 0.0026)$. The function $F(u[k]; k)$ varies with k,

$$F(u[k]; k) = \begin{cases} u^3[k], \text{ for } 10000 < k \leq 20000 \\ u^2[k], \text{ for } 30000 < k \leq 40000 \\ u[k], \text{ at all other times} \end{cases} \tag{1.2}$$

The output $y[k]$ can be seen in the first trace of Fig. 1.2, the second and third traces show the residual estimation errors of the optimal linear system and Volterra system respectively, the final trace is the estimated degree of signal nonlinearity. Whilst these results show that this approach can detect changes in nonlinearity and is not affected by the presence of noise, this may be largely due to the nature of the input signal in question being particularly suited to the Volterra model.

This type of method relies on the nature of the nonlinearity under observation being suited to the actual signal model; in real world situations it is not always possible to know the nonlinearity in advance, therefore their application is limited. To overcome these limitations we propose a much more flexible method based on collaborative adaptive filtering.
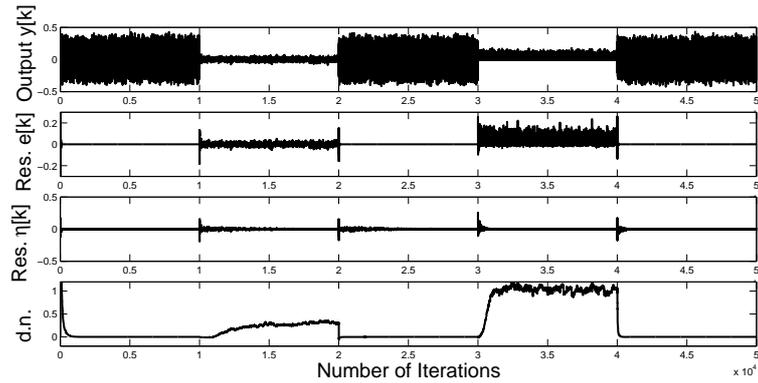
Fig. 1.2. Estimated degree of signal nonlinearity for an input alternating from linear to nonlinear

1.1.2 Collaborative Adaptive Filters

Developing on the well established tracking capabilities of adaptive filters using combinations of adaptive subfilters in a more natural way produces a single hybrid filter without the need for any knowledge of underlying signal generation models. Hybrid filters consist of multiple individual adaptive subfilters operating in parallel and all feeding into a mixing algorithm which produces the single output of the filter [4, 13]. The mixing algorithms are also adaptive and combine the outputs of each subfilter based on the estimate of their current performance on the input signal from their instantaneous output error.

Many previous applications of hybrid filters have focused mainly on the improved performance they can offer over the individual constituent filters. Our aim is to focus on one additional effect of the mixing algorithm that is, to show whether it can give an indication of which filter is currently responding to the input signal most effectively. Therefore intuitively by selecting algorithms which are particularly suited to one type of input signals, it is possible to cause the mixing algorithm to adapt according to fundamental properties of the input signal.

A simple form of mixing algorithm for two adaptive filters is a convex combination. Convexity can be described as [5]

$$\lambda x + (1 - \lambda)y \text{ where } \lambda \in [0, 1] \tag{1.3}$$

For x and y being two points on a line, as shown in Fig. 1.3, their convex mixture (1.3) will lie on the same line between x and y.

For convex mixing of the outputs of adaptive filters, it is intuitively clear that initially λ will adapt to favour the faster filter (that is the filter with faster learning rate) and following convergence it will favour the filter with
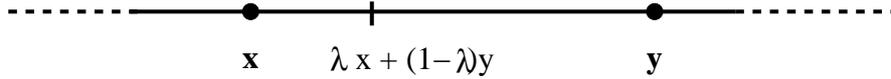
Fig. 1.3. Convexity

better steady state properties[1]; should one of the subfilters fail to converge the values of λ adapt such that the hybrid filter follows the stable subfilter [16]. The approach in this chapter focuses on observing the dynamics of mixing parameter λ, to allow conclusions to be drawn about the current nature of the input signal.

1.2 Derivation of The Hybrid Filter

Unlike the existing approaches to hybrid adaptive filters which focus on the quantitative performance of such filters, in this case the design of the hybrid filters is such that it should combine the characteristics of two distinctly different adaptive filters. Signal modality characterisation is achieved by making the value of the "mixing" parameter λ adapt according to the fundamental dynamics of the input signal. In this chapter we illustrate applications of this method for characterisation of nonlinearity and complexity on both synthetic and real world data, but this method can be equally well applied to any other signal characteristics. With that in mind we start from the general derivation of the convex hybrid filter before moving on to specific implementations.

Figure 1.4 shows the block diagram of a hybrid filter consisting of two adaptive filters combined in a convex manner. At every time instant k, the output of the hybrid filter, $y(k)$, is an adaptive convex combination of the output of the first subfilter $y_1(k)$ and the output of the second subfilter $y_2(k)$, and is given by

$$y(k) = \lambda(k)y_1(k) + (1 - \lambda(k))\, y_2(k), \qquad (1.4)$$

where $y_1(k) = \mathbf{x}^T(k)\mathbf{w}_1(k)$ and $y_2(k) = \mathbf{x}^T(k)\mathbf{w}_2(k)$ are the outputs of the two subfilters with corresponding weight vectors $\mathbf{w}_1(k) = [w_{1,1}(k), \ldots, w_{1,N}(k)]^T$ and $\mathbf{w}_2(k) = [w_{2,1}(k), \ldots, w_{2,N}(k)]^T$ which are dependent on the algorithms used to train the subfilters based on the common input vector $\mathbf{x}(k) = [x_1(k), \ldots, x_N(k)]^T$ for filters of length N.

To preserve the inherent characteristics of the subfilters, which are the basis of our approach, the constituent subfilters are each updated by their own errors $e_1(k)$ and $e_2(k)$, using a common desired signal $d(k)$, whereas the parameter λ is updated based on the overall error $e(k)$. The convex mixing parameter $\lambda(k)$ is updated based on minimisation of the quadratic cost function $E(k) = \frac{1}{2}e^2(k)$ using the following gradient adaptation

---

[1] Unlike traditional search then converge approaches this method allows for potentially nonstationary data.
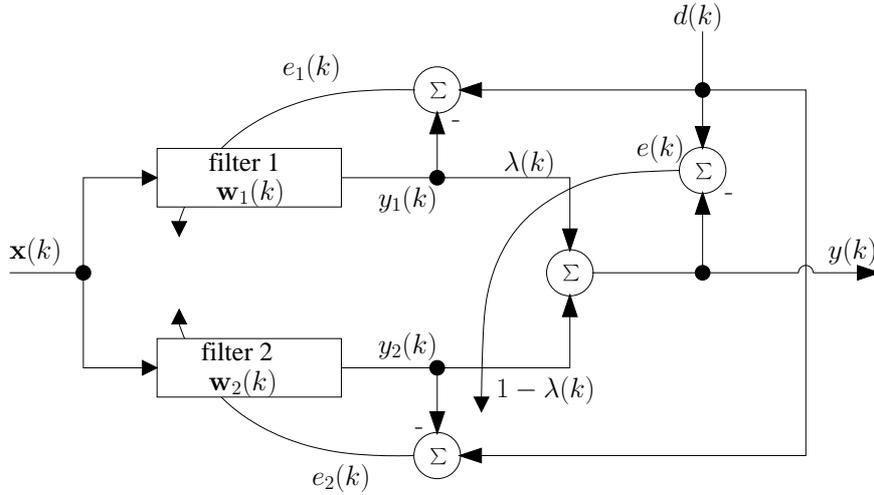
Fig. 1.4. Convex combination of adaptive filters (hybrid filter)

$$\lambda(k+1) = \lambda(k) - \mu_\lambda \nabla_\lambda E(k)_{|\lambda=\lambda(k)}, \tag{1.5}$$

where μ_λ is the adaptation step-size. From (1.4) and (1.5), using an LMS type adaptation, the λ update can be obtained as

$$\lambda(k+1) = \lambda(k) - \frac{\mu_\lambda}{2}\frac{\partial e^2(k)}{\partial \lambda(k)} = \lambda(k) + \mu_\lambda e(k)(y_1(k) - y_2(k)). \tag{1.6}$$

To ensure the combination of adaptive filters remains a convex function, it is critical λ remains within the range $0 \leq \lambda(k) \leq 1$. In [4] the authors obtained this through the use of a sigmoid function as a post-nonlinearity to bound $\lambda(k)$. Since, in order to determine the changes in the modality of a signal, we are not interested in the overall performance of the filter but in the dynamics of parameter λ, the use of a sigmoid function would interfere with true values of $\lambda(k)$ and was therefore not appropriate. In this case a hard limit on the set of allowed values for $\lambda(k)$ was therefore implemented.

1.3 Detection of the Nature of Signals - Nonlinearity

Implementations of the hybrid filter described above using the LMS algorithm [23] to train one of the subfilters and the generalised normalised gradient descent (GNGD) algorithm [15] for the other, have been used to distinguish the linearity/nonlinearity of a signal [11]. The LMS algorithm was chosen as it is widely used, known for its robustness and excellent steady state properties whereas the GNGD algorithm has a faster convergence speed and better tracking capabilities. By exploiting these properties it is possible to show that

due to the synergy and simultaneous mode of operation, the hybrid filter has excellent tracking capabilities for signals with extrema in their inherent linearity and nonlinearity characteristics.

The output of the LMS trained subfilter y_{LMS} is generated from [23]

$$y_{\mathrm{LMS}}(k) = \mathbf{x}^T(k)\mathbf{w}_{\mathrm{LMS}}(k)$$
$$e_{\mathrm{LMS}}(k) = d(k) - y_{\mathrm{LMS}}(k)$$
$$\mathbf{w}_{\mathrm{LMS}}(k+1) = \mathbf{w}_{\mathrm{LMS}}(k) + \mu_{\mathrm{LMS}}e_{\mathrm{LMS}}(k)\mathbf{x}(k) \tag{1.7}$$

and y_{GNGD} is the corresponding output of the GNGD trained subfilter given by [15]

$$y_{\mathrm{GNGD}}(k) = \mathbf{x}^T(k)\mathbf{w}_{\mathrm{GNGD}}(k) \tag{1.8}$$
$$e_{\mathrm{GNGD}}(k) = d(k) - y_{\mathrm{GNGD}}(k)$$
$$\mathbf{w}_{\mathrm{GNGD}}(k+1) = \mathbf{w}_{\mathrm{GNGD}}(k) + \frac{\mu_{\mathrm{GNGD}}}{\|\mathbf{x}(k)\|_2^2 + \varepsilon(k)}e_{\mathrm{GNGD}}(k)\mathbf{x}(k)$$
$$\varepsilon(k+1) = \varepsilon(k) - \rho\mu_{\mathrm{GNGD}}\frac{e_{\mathrm{GNGD}}(k)e_{\mathrm{GNGD}}(k-1)\mathbf{x}^T(k)\mathbf{x}(k-1)}{\left(\|\mathbf{x}(k-1)\|_2^2 + \varepsilon(k-1)\right)^2}$$

where the step-size parameters of the filters are μ_{LMS} and μ_{GNGD}, and in the case of the GNGD ρ is the step-size adaptation parameter and ε the regularisation term.

By evaluating the resultant hybrid filter in an adaptive one step ahead prediction setting with the length of the adaptive filters set to $N = 10$, it is possible to illustrate the ability of the hybrid filter to identify the modality of a signal of interest. The behaviour of λ has been investigated for benchmark synthetic linear and nonlinear inputs. Values of λ were averaged over a set of 1000 independent simulation runs, for the inputs described by a stable linear AR(4) process:

$$x(k) = 1.79x(k-1) - 1.85x(k-2) + 1.27x(k-3) - 0.41x(k-4) + n(k) \tag{1.9}$$

and a benchmark nonlinear signal [18]:

$$x(k+1) = \frac{x(k)}{1 + x^2(k)} + n^3(k) \tag{1.10}$$

where $n(k)$ is a zero mean, unit variance white Gaussian process. The values of the step-sizes used were $\mu_{\mathrm{LMS}} = 0.01$ and $\mu_{\mathrm{GNGD}} = 0.6$. For the GNGD filter $\rho = 0.15$ and the initial value of the regularisation parameter was $\varepsilon(0) = 0.1$. Within the convex combination of the filters, filter 1 corresponds to the GNGD trained subfilter and filter 2 to the LMS trained subfilter, the step-size for the adaptation of $\lambda(k)$ was $\mu_\lambda = 0.05$ and the initial value[2] of $\lambda(0) = 1$.

---

[2] Since GNGD exhibits much faster convergence than LMS, it is natural to start the adaptation with $\lambda(0) = 1$. This way, we avoid possible artefacts that may arise due to the slow initial response to the changes in signal modality.

From the curves shown in Fig. 1.5 it can be seen the value of $\lambda(k)$ for both inputs moves towards zero as the adaptation progresses. As expected, the output of the convex combination of adaptive filters approaches the output of the LMS filter y_{LMS} predominately. This is due to the better steady state properties of the LMS filter when compared to the GNGD filter, which due to its constantly 'alert' state does not settle in the steady state as well as the LMS. In the early stages of adaptation the nonlinear input (1.10) adapts to become dominated by the LMS filter much faster than the linear input and rapidly converges, whereas the linear input (1.9) changes much more gradually between the two filters[3].
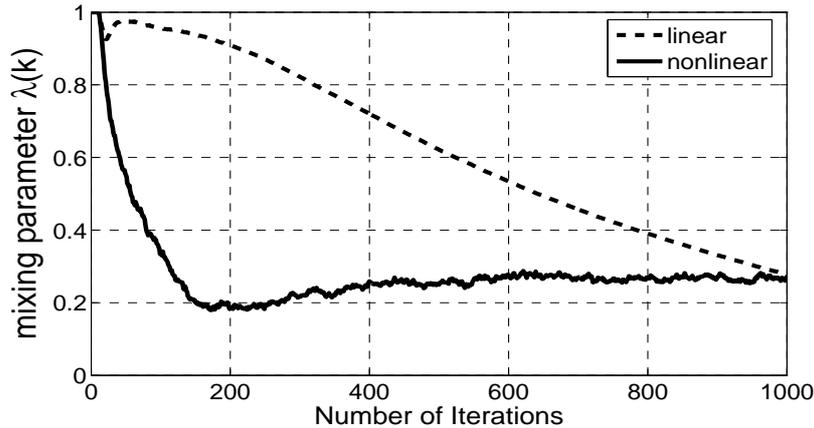

Fig. 1.5. Comparison of the mixing parameter λ for linear and nonlinear inputs

1.3.1 Tracking Changes in Nonlinearity of Signals

It is also possible to use changes in λ along the adaptation to track the changes in signal modality. Since the behaviour of λ as a response to the different inputs is clearly distinct, especially in the earliest stages of adaptation, the convex combination was presented with an input signal which alternated between linear (1.9) and nonlinear (1.10). The input signal was alternated every 200 samples and the corresponding dynamics of the mixing parameter $\lambda(k)$ are shown in Fig. 1.6. From Fig. 1.6 it is clear that the value of $\lambda(k)$ adapts in a way which ensures that the output of the convex combination is dominated by the filter most appropriate for the input signal characteristics.

To illustrate the discrimination ability of the proposed approach, the next set of simulations show the results of the same experiment as in Fig. 1.6, but

---

[3] Both filters perform well on a linear input and are competing along the adaptation.
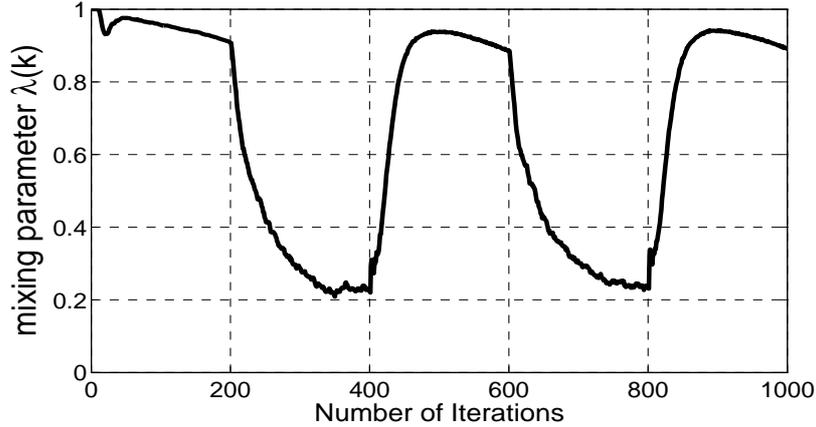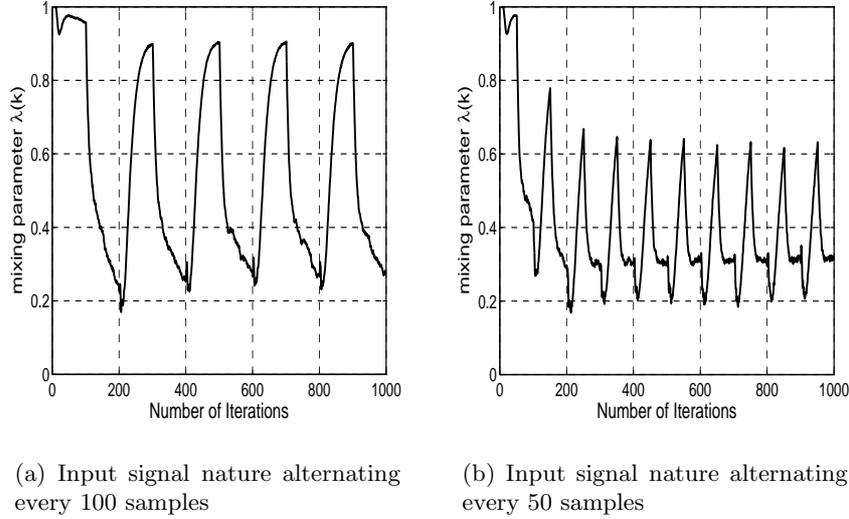
Fig. 1.6. Evolution of the mixing parameter λ for input nature alternating every 200 samples

for a decreased number of samples between the alternating segments of data. Figure 1.7 shows the response of $\lambda(k)$ to the input signal alternating every 100 and 50 samples respectively. There is a small anomaly in the values of λ immediately following the change in input signal from nonlinear to linear, which can be clearly seen in Fig. 7(b) around sample numbers $100i$, $i = 1, 2, \ldots$, where the value of λ exhibits a small dip before it increases. This is due to the fact that the input to both the current AR process (1.9) and the tap inputs to both filters use previous nonlinear samples where we are in fact predicting the first few "linear" samples. This does not become an issue when alternations between the input signals occur less regularly or if there is a more natural progression from "linear" to "nonlinear" in the the input signal.

Real World Applications

To examine the usefulness of this approach for the processing of real world signals, a set of EEG signals has been analysed. Following the standard practice, the EEG sensor signals were averaged across all the channels and any trends in the data were removed. Figure 1.8 shows the response of λ when applied to two different sets of EEG data from epileptic patients, both showing the onset of a seizure as indicated by a sudden change in the value of λ. These results show that the this approach can effectively detect changes in the nature of the EEG signals which can be very difficult to achieve otherwise.

(a) Input signal nature alternating every 100 samples

(b) Input signal nature alternating every 50 samples

Fig. 1.7. Evolution of the mixing parameter λ for a signal with input nature alternating between linear to nonlinear

1.4 Detection of the Nature of Signals - Complex Domain

For generality, building upon identification and tracking of nonlinearity in the real domain $\mathbb{R}$ we shall extend this to the complex domain $\mathbb{C}$. In order to facilitate this, the update of λ (1.6) was extended to the complex domain, resulting in

$$\lambda(k+1) = \lambda(k) - \frac{\mu_\lambda}{2} \left\{ e(k) \frac{\partial e^*(k)}{\partial \lambda(k)} + e^*(k) \frac{\partial e(k)}{\partial \lambda(k)} \right\}$$
$$= \lambda(k) + \mu_\lambda \Re \left\{ e(k)(y_1(k) - y_2(k))^* \right\}, \tag{1.11}$$

where $(\cdot)^*$ denotes the complex conjugation operator. For the complex version of the hybrid convex combination, the subfilters previously discussed were substituted with the complex NLMS and complex normalised nonlinear gradient descent (NNGD) [14], in this case the normalised versions were used as opposed to the standard complex LMS (CLMS) and NGD (CNGD) to overcome problems with the convergence of the individual subfilters and hence dependence on the combination of input signal statistics. The CLMS update is given by [22]

$$\mathbf{w}_{\text{CLMS}}(k+1) = \mathbf{w}_{\text{CLMS}}(k) + \eta_{\text{CLMS}}(k) e_{\text{CLMS}}(k) \mathbf{x}^*(k) \tag{1.12}$$
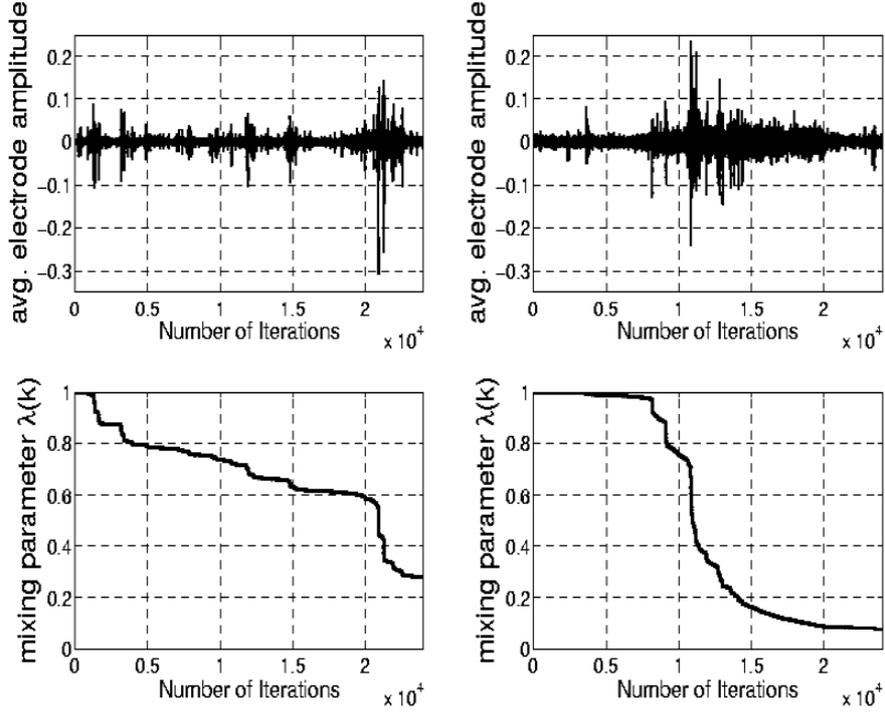
Fig. 1.8. Top panel: EEG signals for two patients showing epileptic seizures; Bottom panel: corresponding adaptations of λ

where η denotes the learning rate which for the CLMS is $\eta_{\text{CLMS}}(k) = \mu_{\text{CLMS}}$ and for the CNLMS and $\eta_{\text{CLMS}}(k) = \mu_{\text{CLMS}} / \left(\|\mathbf{x}(k)\|_2^2 + \varepsilon \right)$.

The CNGD is described by

$$
\begin{aligned}
e_{\text{CNGD}}(k) &= d(k) - net(k) \\
net(k) &= \Phi\left(\mathbf{x}^T(k)\mathbf{w}_{\text{CNGD}}(k)\right) \\
\mathbf{w}_{\text{CNGD}}(k+1) &= \mathbf{w}_{\text{CNGD}}(k) - \eta_{\text{CNGD}}(k)\nabla_{\mathbf{w}} E(k)
\end{aligned}
\tag{1.13}
$$

where $net(k)$ is the net input, $\Phi(\cdot)$ denotes the complex nonlinearity and $E(k)$ is the cost function given by

$$
E(k) = \frac{1}{2}|e(k)|^2.
\tag{1.14}
$$

Following the standard complex LMS derivation [22] for a fully complex non-linear activation function, Φ, the weight update is expressed as

$$
\mathbf{w}_{\text{CNGD}}(k+1) = \mathbf{w}_{\text{CNGD}}(k) + \eta_{\text{CNGD}}(k)e_{\text{CNGD}}(k)\left(\Phi'\left[net(k)\right]\right)^* \mathbf{x}^*(k) \tag{1.15}
$$

where $(\cdot)'$ denotes the complex differentiation operator and $\eta_{\mathrm{CNGD}} = \mu_{\mathrm{CNGD}}$ for the CNGD and $\eta_{\mathrm{CNGD}} = \mu_{\mathrm{CNGD}} / \left(C + [\Phi'(net(k))]^2 \|\mathbf{x}\|_2^2 \right)$ for the CN-NGD.

For the purposes of tracking changes in the nonlinearity of signals the hybrid filter was again presented with an input signal alternating between linear and nonlinear. The process $n(k)$ in the linear AR(4) signal (1.9) was replaced with a complex white Gaussian process again with zero mean and unit variance,

$$n(k) = n_r(k) + jn_i(k),$$

where the real and imaginary components of n are mutually independent sequences having equal variances so that $\sigma_n^2 = \sigma_{n_r}^2 + \sigma_{n_i}^2$. The complex benchmark nonlinear signal [18] was

$$x(k) = \frac{x^2(k-1)\left(x(k-1)+2.5\right)}{1+x^2(k-1)+x^2(k-2)} + n(k-1) \tag{1.16}$$

and the nonlinearity used was the sigmoid function

$$\Phi(z) = \frac{1}{1+e^{-z}}, \ \text{ where } z \in \mathbb{C} \tag{1.17}$$

Figure 1.9 shows the response of λ to the input signal alternating every 200 and every 100 samples and again the hybrid filter was clearly capable of tracking such changes in the nonlinearity of the input signal.
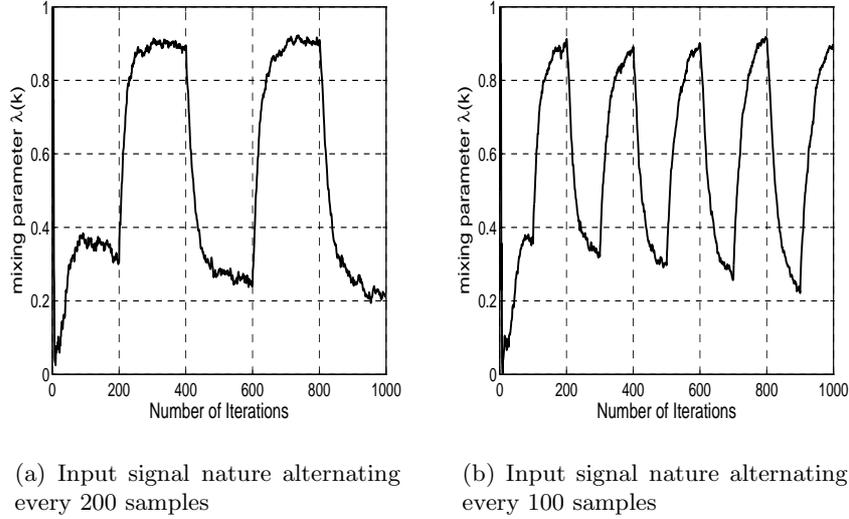
1.4.1 Split-Complex vs. Fully-Complex

Whilst being able to identify the nonlinearity of a signal is important and can give key knowledge about the signal under observation, within nonlinear adaptive filtering in $\mathbb{C}$ one of the biggest problems is the choice of nonlinear complex activation function (AF). There are three main methods to deal with this:-

- processing the real and imaginary components separately using a real non-linearity;
- processing in the complex domain using a so called "split-complex" non-linearity;
- or using a so called "fully-complex" nonlinearity.

A fully-complex nonlinearity is a function $f : \mathbb{C} \to \mathbb{C}$ and are the most efficient in using higher order statistics within a signal [12]. For a split-complex function the real and imaginary components of the input are separated and fed through the dual real valued activation function $f_R(x) = f_I(x)$, $x \in \mathbb{R}$. A split complex AF can be represented as

$$\Phi_{split}(z) = f_R(z^r) + jf_I(z^i) = u(z^r) + jv(z^i) \tag{1.18}$$

(a) Input signal nature alternating every 200 samples

(b) Input signal nature alternating every 100 samples

Fig. 1.9. Evolution of the mixing parameter λ for a signal with input nature alternating between linear to nonlinear

Whilst algorithms using split-complex AFs have been shown to give good results, when the real and imaginary components of a signal are strongly correlated algorithms using the split-complex AFs are not suitable as they rely on the real and imaginary weight updates being mutually exclusive to calculate the true gradient.

Consider the Ikeda map, a well known benchmark signal in chaos theory [2], given by

$$x(k+1) = 1 + u\left[x(k)\cos t(k) - y(k)\sin t(k)\right]$$
$$y(k+1) = u\left[x(k)\sin t(k) + y(k)\cos t(k)\right] \tag{1.19}$$

where u is a parameter and

$$t(k) = 0.4 - \frac{6}{1 + x^2(k) + y^2(k)}. \tag{1.20}$$

Figure 10(a) illustrates that the hybrid filter can clearly identify the Ikeda map as nonlinear when presented with it as an input. It is natural however to expect that as the signal generation mechanism is in the form of coupled difference equations, by representing the pair $[x(n), y(n)]$ as a vector in $\mathbb{C}$, the Ikeda map (1.19) will represent a fully complex signal. This is indeed confirmed by the simulation results shown in Fig. 10(b) where to test the application of the hybrid filter method for detection of the nature of nonlinear complex signals,

the hybrid filter consisted of a combination of a fully-complex and a split-complex subfilter trained by the CNGD algorithm with $\lambda = 1$ corresponding to the fully-complex subfilter. As expected (by design), from Fig. 1.10, the Ikeda map is a nonlinear signal (Fig. 10(a)) which exhibits fully-complex nonlinear properties (Fig. 10(b)).

To illustrate this further, Fig. 1.11 shows the performance of the complex real time recurrent learning (CRTRL) algorithm [9] for both split- and fully-complex learning on the prediction of the Ikeda map; observe that the split-complex CRTRL did not respond as well as the fully complex version to prediction of the Ikeda map.

(a) Nonlinear vs. Linear   (b) Fully- vs. Split-Complex
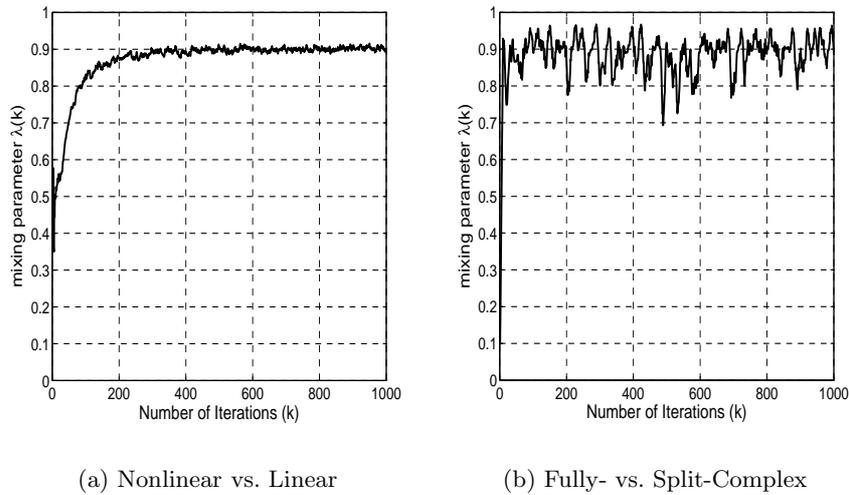
Fig. 1.10. Evolution of the mixing parameter λ for prediction of Ikeda map

Knowledge of the complex nonlinearity that describes a real-world complex signal is critical, as it can help us to understand the nature of the dynamics of the system under observation (radar, sonar, vector fields). To illustrate the tracking capabilities of this hybrid filter, the filter was presented with real world radar data. The radar data comes from a maritime radar (IPIX, publicly available from [1]), for different sea states, "low" (calm sea) and "high" (turbulent) states. Whilst there are off-line statistical tests for radar data [10] and it has been shown that radar data is predominantly fully complex in nature when the target is in the beam [8], on-line estimation algorithms are lacking and it is clear that it is important to track the onset of changes in the nature whilst recording. Figure 1.12 shows the evolution of λ when predicting radar data that was alternated every 50 samples from the low sea state to
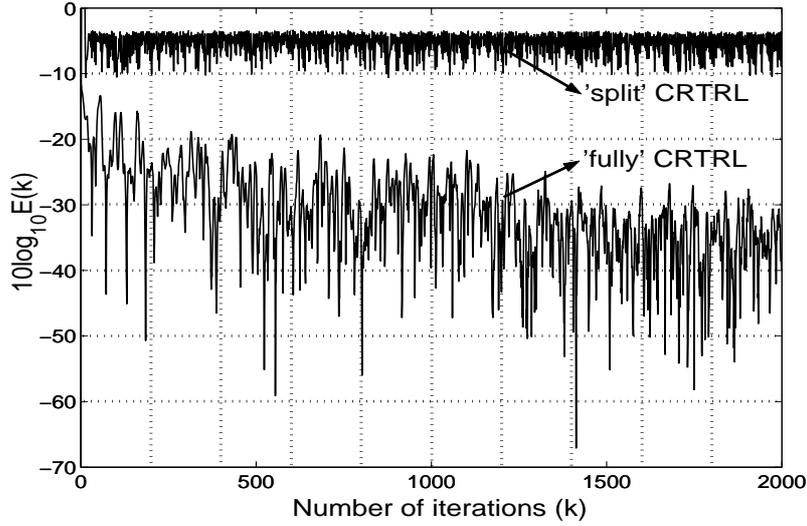
Fig. 1.11. Learning curves for fully-complex vs. split-complex prediction of Ikeda map

the high sea state. All the data sets were standardised so all the magnitudes were in the range $[-1, 1]$. The initial weight vectors were set to zero and the filter order $N = 10$. Figure 1.12 shows that the modality of the high sea state was predominantly fully complex and similarly the low sea state was predominantly split complex.
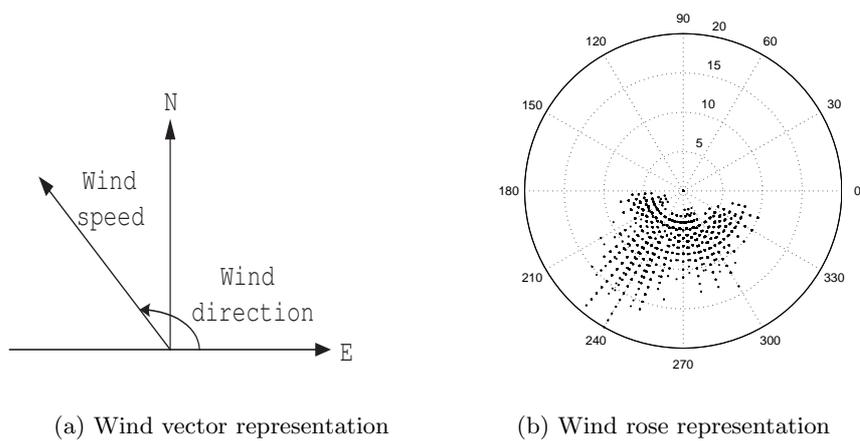
1.4.2 Complex Nature of Wind

Wind modelling is an illustration for the need for complex valued techniques; Fig. 13(a) represents a wind rose plot of direction versus magnitude and shows the need for wind to be modelled based on both direction and speed. Wind is normally measured either as a bivariate process of these measurements [3] or, despite the clear interdependence between the components, only the speed component in taken into account. From Fig. 13(b) it is clear that wind could also be represented as a vector of speed and direction components in the North – East coordinate system. Following this, the wind vector $\mathbf{v}(k)$ can be represented in the complex domain $\mathbb{C}$, as

$$\mathbf{v}(k) = |\mathbf{v}(k)|e^{j\theta(k)} = v_E(k) + jv_N(k) \tag{1.21}$$

where v is the speed component and θ the direction, modelled as a single complex quantity.
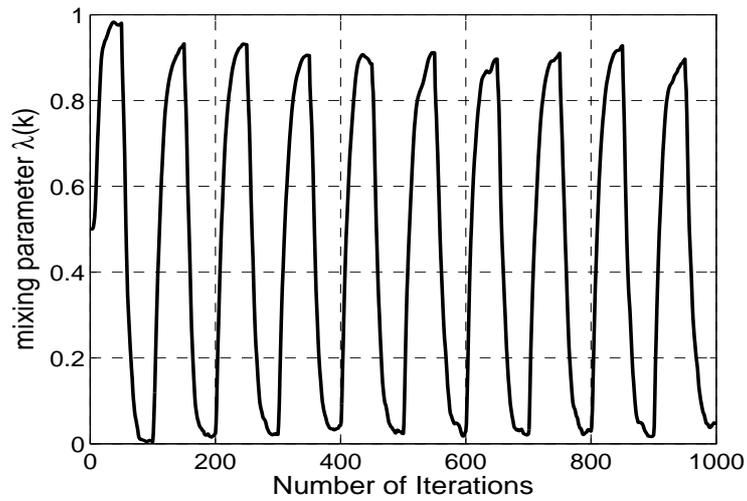
Fig. 1.12. Evolution of the mixing parameter λ for alternating blocks of 400 data samples of radar data from the "low" to "high" sea state


(a) Wind vector representation

(b) Wind rose representation

Fig. 1.13. Wind recordings as complex [speed,direction] quantities

Complex Surrogate Data Testing for the Nature of Wind

Following the approach from [8], to support the complex-valued modelling of wind we first employ a statistical test for the complex nature of wind. The test is based on the complex-valued surrogate data analysis and is set within the framework of hypothesis testing and the statistical testing methodology from [7] is adopted. The signals are characterised by the Delay Vector Variance (DVV) method and the null hypothesis was that the original signal is complex-valued. Figure 1.14 shows the results of this test indicating there is a significant component dependence within the complex-valued wind signal representation. This is indicated by the rejection ratio of the null hypothesis of fully complex wind data being significantly greater than zero.

The results from Fig. 1.14 show the proposed test repeated 100 times, with the number of times the null hypothesis was rejected computed. The wind signal was averaged over either one hour intervals or six hour intervals and as can be seen from the rejection ratios the signal averaged over one hour showed a stronger indication of having a complex nature than those averaged over six hours. Therefore the components of complex-valued wind signals become more dual univariate and linear when averaged over longer intervals, this is in line with results from probability theory where a random signal becomes more Gaussian (and therefore linear) with the increase in the degree of averaging.
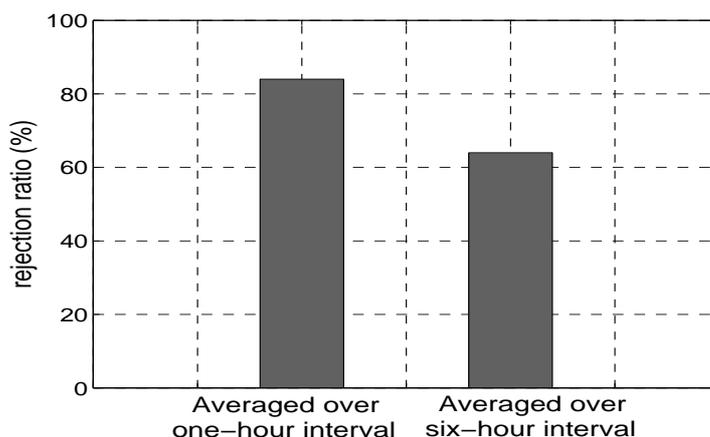
Fig. 1.14. Complex valued surrogate data test for the complex nature of wind signal [speed and direction]

Tracking Modality of Wind Data

As it has been shown that wind can be considered a complex valued quantity and that it is possible to track the nature of complex signals using the hybrid filter combination of a fully complex and a split complex adaptive filter, the hybrid filter was used to predict a set of wind data. The data used was measurements of the wind in an urban area over one day. The filter length was set to $N = 10$, the learning rates of the split and fully complex NGD algorithms were $\mu_{split} = 0.01$ and $\mu_{fully} = 0.01$ and the step size of the learning parameter was $\mu_\lambda = 0.5$. The results of this can be seen in Fig. 1.15, as for the majority of the time the value of λ is around 0.9 the wind signal can be considered mainly fully complex. It is also clear that the first and last measurements are more unstable in nature as λ oscillated in the range $[0.5, 0.9]$, this indicates the intermittent wind nature was mainly fully complex but does at times become more split complex. Since these measurements were taken from a 24 hour period starting from 14:00 these sections correspond to the recordings taken between 14:00–18:00 and 08:00–14:00 the next day, this is to be expected as during these times the wind is changing rapidly compared to the "calm" period in the late evening and the early morning.

1.5 Conclusions

We have proposed a novel approach to identify changes in the modality of a signal. This is achieved by a convex combination of two adaptive filters for which the transient responses are significantly different. By training the two filters with different algorithms, it is possible to exploit the difference in the performance capabilities of each. The evolution of the adaptive convex mixing parameter λ, helps determine which filter is more suited to the current input signal dynamics, and thereby gain information about the nature of the signal. This way, information fusion is achieved by collaborative modular learning, suitable for the online mode of operation. The analysis and simulations illustrate that there is significant potential for the use of this method for online tracking of some fundamental properties of the input signal. Both synthetic and real world examples on EEG, radar and wind data support the analysis. The extension to the simultaneous tracking of several parameters follows naturally; this can be achieved by a hierarchical and distributed structure of hybrid filters.
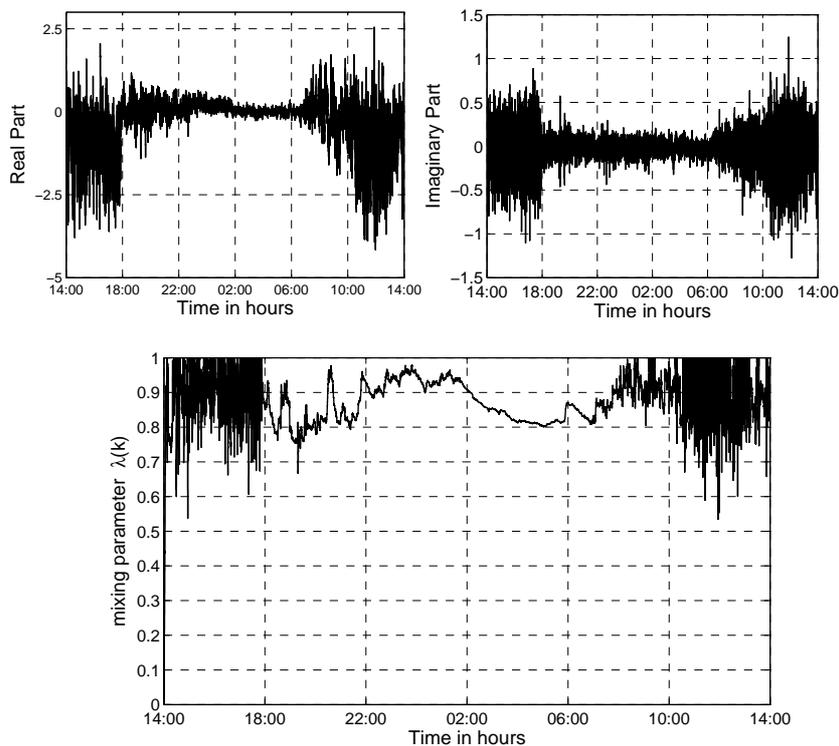
Acknowledgements

Fig. 1.15. Evolution of the mixing parameter λ for the prediction of wind

References

1. The McMaster IPIX radar sea clutter database. URL `http://soma.crl.mcmaster.ca/ipix/`
2. Aihara, K. (ed.): Applied Chaos and Applicable Chaos. Tokyo: Science–Sha (1994)
3. Alexiadis, M.C., Dokopoulos, P.S., Sahsamanoglou, H.S., Manousaridis, I.M.: Short-term forecasting of wind speed and related electrical power. Solar Energy **63**(1), 61–68 (1998)
4. Arenas-Garcia, J., Figueiras-Vidal, A., Sayed, A.: Steady state performance of convex combinations of adaptive filters. In: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '05), vol. 4, pp. 33–36 (2005)
5. Cichocki, A., Unbehauen, R.: Neural Networks for Optimisation and Signal Processing. Wiley (1993)
6. Gautama, T., Mandic, D., Hulle, M.V.: Signal nonlinearity in fMRI: a comparison between BOLD and MION. IEEE Transactions on Medical Imaging, **22**(5), 636–644 (2003)

7. Gautama, T., Mandic, D., Hulle, M.V.: The delay vector variance method for detecting determinism and nonlinearity in time series. Physica D **190**(3–4), 167–176 (2004)
8. Gautama, T., Mandic, D., Hulle, M.V.: On the characterisaion of the deterministic/stochastic and linear/nonlinear nature of time series. Technical report dpm-04-5, Imperial College (2004)
9. Goh, S.L., Mandic, D.P.: A general complex valued RTRL algorithm for nonlinear adaptive filters. Neural Computation **16**(12), 2699–2731 (2004)
10. Haykin, S., Principe, J.: Making sense of a complex world [chaotic events modeling]. IEEE Signal Processing Magazine **15**(3), 66–81 (1998)
11. Jelfs, B., Vayanos, P., Chen, M., Goh, S.L., Boukis, C., Gautama, T., Rutkowski, T., Kuh, T., Mandic, D.: An online method for detecting nonlinearity within a signal. In: Proceedings 10th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems KES2006, vol. 3, pp. 1216–1223 (2006)
12. Kim, T., Adali, T.: Approximation by fully complex multilayer perceptrons. Neural Computation **15**(7), 1641–1666 (2003)
13. Kozat, S., Singer, A.: Multi-stage adaptive signal processing algorithms. In: Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, pp. 380–384 (2000)
14. Mandic, D.: NNGD algorithm for neural adaptive filters. Electronics Letters **36**(9), 845–846 (2000)
15. Mandic, D.: A generalized normalized gradient descent algorithm. IEEE Signal Proc. Lett. **11**(2), 115–118 (2004)
16. Mandic, D., Vayanos, P., Boukis, C., Jelfs, B., Goh, S., Gautama, T., Rutkowski, T.: Collaborative adaptive learning using hybrid filters. In: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2007, vol. 3, pp. 921–924 (2007)
17. Mizuta, H., Jibu, M., Yana, K.: Adaptive estimation of the degree of system nonlinearity. In: Proceedings IEEE Adaptive Systems for Signal Processing and Control Symposium (AS-SPCC), pp. 352–356 (2000)
18. Narendra, K., Parthasarathy, K.: Identification and control of dynamical systems using neural networks. IEEE Transactions on Neural Networks **1**(1), 4–27 (1990)
19. Schreiber, T.: Interdisciplinary application of nonlinear time series methods. Physics Reports **308**(1), 1–64 (1999)
20. Schreiber, T., Schmitz, A.: Discrimination power of measures for nonlinearity in a time series. Physical Review E **55**(5), 5443–5447 (1997)
21. Schreiber, T., Schmitz, A.: Surrogate time series. Physica D **142**, 346–382 (2000)
22. Widrow, B., McCool, J., Ball, M.: The complex LMS algorithm. Proceedings of the IEEE **63**(4), 719–720 (1975)
23. Widrow, B., Stearns, S.D.: Adaptive Signal Processing. Prentice-Hall (1985)