COMPLEX-VALUED RADIAL BASIS FUNCTION NETWORKS

S. Chen, P.M. Grant, S. McLaughlin and B. Mulgrew

University of Edinburgh, UK

Abstract

We propose a novel complex radial basis function (RBF) network. The network has complex centres and weights but the response of its hidden nodes remains real. Several learning algorithms for the existing real RBF network are extended to this complex network. The proposed network is capable of generating complicated nonlinear decision surface or approximating an arbitrary nonlinear function in multi-dimensional complex space and it provides a powerful tool for nonlinear signal processing involving complex signals. This is demonstrated using two practical applications to communication systems. The first case considers the equalisation of time-disperisve communication channels, and we show that the underlying Bayesian solution has an identical structure to the complex RBF network. In the second case, we use the complex RBF network to model nonlinear channels, and this application is typically found in channel estimation and echo cancellation involving nonlinear distortion.

1. Introduction

Many signal processing applications are performed in multi-dimensional complex space. For example, in many communication systems, information is transmitted in the form of complex digital symbols, and the channel distortion can be modelled as a finite impulse response filter with complex taps. Removing channel distortion and interference can be formulated concisely and elegantly as a signal processing problem in multi-dimensional complex space. Most available artificial neural networks however are real-valued and are suitable for signal processing applications in multi-dimensional real space. Some research has been done on complex multilayer perceptron (MLP) and on extension of the backpropagation algorithm to complex form (e.g. [1]-[3]). Like its real counterpart, the complex MLP is highly nonlinear in the parameters and suffers the same drawbacks of slow convergence and unpredictable solutions during learning.

The present paper proposes a complex RBF network, which is a logic extension of the existing real RBF network [4]-[6]. The inputs and outputs of the network are both complex. Each node in the hidden layer has a real radially symmetric response around a complex vector called the node centre which has the same dimension as the network input vector. The output layer of the network contains a set of linear combiners with complex weights. The existing real RBF network can be viewed as a special case of this complex network. In fact, if the desired outputs for the network are to be real the network weights also become real and, similarly, if the network inputs are real the RBF centres become real. When both the network inputs and desired outputs are reduced to real, this complex RBF network degenerates naturally into the real RBF network. The intrinsic Bayesian interpretation for the real RBF network is preserved in this complex RBF network. Since the response of a hidden node can be interpreted as some kind of generalized potential function, it should be real. It will be shown later in the application to channel equalisation that a hidden node actually realizes the conditional probability density function for a given channel state. Furthermore we show that existing learning algorithms for real RBF networks [4],[6]-[10] can easily be extended to the complex RBF network.

This novel network provides a powerful means for nonlinear signal processing in multi-dimensional complex space, and this is demonstrated using two applications to communication systems. In the first application, equalisation of time-dispersive channels is formulated as a decision problem. It is shown that the Bayesian solution for this decision problem has an identical structure to the complex RBF network, and therefore the latter provides an ideal means to realize the Bayesian performance. Superiority of the Bayesian approach over the conventional equalisation approach is highlighted using both the stationary and multi-path fading channels [11]. In some equalisation schemes, it is necessary to identify a channel model. Echo cancellation [12] can also be viewed as an identification problem. When nonlinear distortion is present in transmission path, nonlinear approximation capability is required. The second application considers modelling nonlinear complex channels based on the complex RBF network.

2. The complex radial basis function network

The topology of the complex RBF network is similar to the real RBF network and is depicted in Fig.1. By convention a complex quantity y is defined as

$$y = \mathrm{Re}[y] + j\mathrm{Im}[y] = y_R + jy_I, \qquad (1)$$

where y_R and y_I are the real and imaginary parts of y respectively; $j = \sqrt{-1}$. The network input and output spaces are both complex and have dimensions n_i and n_o respectively. Assume that the network has n_h hidden

This work was supported by the UK Science and Engineering Research Council under award GR/G/53095.

nodes. The outputs of the hidden nodes are defined by

$$\phi_i = \phi((\mathbf{x} - \mathbf{c}_i)^H (\mathbf{x} - \mathbf{c}_i)/\rho_i), \quad 1 \le i \le n_h, \tag{2}$$

where $\mathbf{x}$ is a n_i-dimensional complex network input vector; $\mathbf{c}_i$ are n_i-dimensional complex vectors called RBF centres; ρ_i are real positive scalars known as widths; the operator $(\cdot)^H = ((\cdot)^T)^*$, $(\cdot)^T$ denotes vector or matrix transpose and $(\cdot)^*$ represents complex conjugate; $\phi(\cdot)$ is a real nonlinear function. Each output node is a complex linear combiner given by

$$f_l(\mathbf{x}) = \sum_{i=1}^{n_h} \phi_i w_{il}, \quad 1 \le l \le n_o, \tag{3}$$

where w_{il} are the complex connection weights.

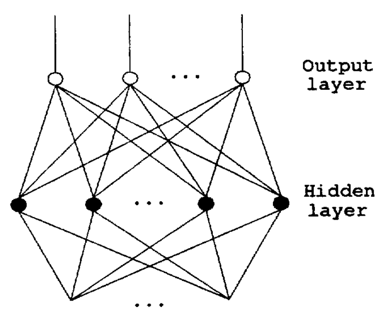

Fig.1. Topology of radial basis function network.

This network realizes a mapping from the n_i-dimensional complex space onto the n_o- dimensional complex space. The real response of the hidden nodes bear close resemblance to some probability density functions and this complex RBF network preserves the Bayesian interpretation of the real RBF network. As in the case of real RBF networks, many real nonlinear functions can be chosen for the hidden nodes. Two examples are the thin-plate-spline function

$$\phi(\chi) = \chi^2 \log(\chi), \tag{4}$$

and the Gaussian function

$$\phi(\chi) = \exp(-\chi^2/\rho), \tag{5}$$

where $\log(\cdot)$ and $\exp(\cdot)$ are both real functions; χ is real and non-negative.

3. Learning algorithms

An advantage of the RBF network is that linear learning laws can be derived. Many such learning algorithms have been developed for real RBF networks. These algorithms can easily be extended to the complex RBF network.

(i) Least squares algorithm with fixed centres [4].

Extension of this learning algorithm to complex RBF networks is simple. As in the real case, complex RBF centres are randomly selected from network input data or from the region where input data exist. Once the centres have been fixed, the least squares algorithm is used to identify weights. Except that the weight vector and the error signal between the desired output and the network output are complex, the rest of the variables in the least squares algorithm remain real.

(ii) Orthogonal least squares algorithm [6],[9].

This is a powerful constructive algorithm based on a block of training data. The algorithm identifies appropriate RBF centres from the training data and estimates the corresponding weights simultaneously in an efficient manner. With some obvious modifications to take into account the complex nature of the centres and the weights, the algorithm remains the same and can readily be applied to complex RBF networks.

(iii) Recursive clustering and LS algorithm [7],[8].

In this algorithm, the RBF centres are adjusted using a recursive clustering algorithm and the weights are updated using the recursive least squares algorithm. To extend the clustering algorithm to the complex RBF network, we only need to define the squared distance between the network input vector $\mathbf{x}$ and a centre $\mathbf{c}_i$ as $(\mathbf{x} - \mathbf{c}_i)^H (\mathbf{x} - \mathbf{c}_i)$. The use of the recursive least squares algorithm to adjust the complex weights is staightforward as in the case (i).

(iv) Dynamic complexity learning algorithm [10].

In this recursive learning procedure, whether to add a new basis function to the network is based on the angle formed between a new basis function and the existing basis functions and the prediction error. With minor notation modifications, it can readily be applied to complex RBF networks.

4. Application to channel equalisation

Many digital communications channels are subject to intersymbol interference (ISI) and can be characterised by a finite impulse response filter and an additive noise source [11]. The relationship between the transmitted complex symbol $s(k)$ and the channel output $y(k)$ can be summarized as follows

$$y(k) = \sum_{l=0}^{n_c-1} a_l s(k-l) + e(k), \tag{6}$$

where n_c is the length of the channel impulse response; a_l are the complex channel taps; $e(k)$ is a complex Gaussian white noise. Without the loss of generality, we will assume that $s(k)$ are 4QAM symbols, that is, the constellation of $s(k)$ is given by the set $\{s^{(1)} = 1 + j, s^{(2)} = -1 + j, s^{(3)} = 1 - j, s^{(4)} = -1 - j\}$.

At the receiver, the ISI must be compensated in order to reconstruct the transmitted symbols, and this is known as the equalisation. A generic equalisation structure is shown in Fig.2, where the integers m, n and d are known as the feedforward order, feedback order and decision delay respectively. The equaliser uses the information present in the observed output vector $\mathbf{y}(k) = [y(k) \cdots y(k-m+1)]^T$ and the detected symbol vector $\hat{\mathbf{s}}_f(k-d) = [\hat{s}(k-d-1) \cdots \hat{s}(k-d-n)]^T$ to produce an estimate $\hat{s}(k-d)$ of $s(k-d)$. The transmitted symbols that influence equaliser decision at k are $\mathbf{s}(k)=[s(k)\cdots s(k-m-n_c+2)]^T$, and this vector has $N_s=4^{n_c+m-1}$ combinations. This gives rise to N_s channel states of $\hat{\mathbf{y}}(k)=[\hat{y}(k)\cdots\hat{y}(k-m+1)]^T$, where $\hat{y}(k)=y(k)-e(k)$. The set of these N_s states is denoted as $Y_{m,d}$. It is sufficient to consider a feedback order of $n=n_c+m-d-2$. The feedback vector $\hat{\mathbf{s}}_f(k-d)$ has $N_f=4^n$ states, and these states are denoted as $\mathbf{s}_{f,i}$, $1\le i\le N_f$. The set of channel states $Y_{m,d}$ can be divided into N_f subsets conditioned on $\hat{\mathbf{s}}_f(k-d)=\mathbf{s}_{f,i}$

$$Y_{m,d}=\bigcup_{1\le i\le N_f} Y_{m,d,i}. \tag{7}$$

Each $Y_{m,d,i}$ can further be divided into 4 subsets according to the value of $s(k-d)$

$$Y_{m,d,i}=\bigcup_{1\le l\le 4} Y_{m,d,i}^{(l)}, \tag{8}$$

where $Y_{m,d,i}^{(l)}=\{\hat{\mathbf{y}}(k)|s(k-d)=s^{(l)}\cap\hat{\mathbf{s}}_f(k-d)=\mathbf{s}_{f,i}\}$ and the number of states in $Y_{m,d,i}^{(l)}$ is $N_{s,i}^{(l)}=4^d$, $1\le l\le 4$.

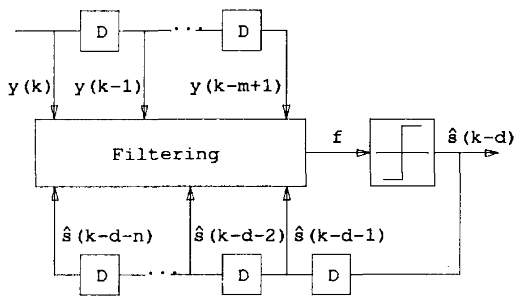

Fig.2. Schematic of decision feedback equaliser.

By applying the Bayes decision theory, we can derive the optimal solution which minimizes the average error probability in symbol detection for the structure of Fig.2. This is known as the Bayesian decision feedback equaliser (DFE) and it takes the form

$$\hat{s}(k-d)=\mathrm{sgn}(f_B(\mathbf{y}(k)|\hat{\mathbf{s}}_f(k-d)=\mathbf{s}_{f,i})), \tag{9}$$

where sgn(·) is the complex signum function, and the conditional Bayesian decision function given $\hat{\mathbf{s}}_f(k-d)=\mathbf{s}_{f,i}$ is

$$f_B(\mathbf{y}(k)|\hat{\mathbf{s}}_f(k-d)=\mathbf{s}_{f,i})=$$

$$\sum_{q=1}^{4}\sum_{l=1}^{N_{s,i}^{(q)}} h_l^{(q)}\exp(-(\mathbf{y}(k)-\mathbf{y}_l^{(q)})^H(\mathbf{y}(k)-\mathbf{y}_l^{(q)})/\rho), \tag{10}$$

where $\rho=2\sigma_e^2$ is the noise variance; the inner sums are over the subset states $\mathbf{y}_l^{(q)}\in Y_{m,d,i}^{(q)}$, $1\le q\le 4$, respectively; the coefficients $h_l^{(1)}=\alpha_l^{(1)}+j\alpha_l^{(1)}$, $h_l^{(2)}=-\alpha_l^{(2)}+j\alpha_l^{(2)}$, $h_l^{(3)}=\alpha_l^{(3)}-j\alpha_l^{(3)}$, $h_l^{(4)}=-\alpha_l^{(4)}-j\alpha_l^{(4)}$, $\alpha_l^{(q)}$ are related to the a-priori probabilities of $\mathbf{y}_l^{(q)}$. Since the a-priori probabilities of $\mathbf{y}_l^{(q)}$ are all equal, all the $\alpha_l^{(q)}$ can be set to one. Compared the Bayesian DFE with the complex RBF network, we can readily draw the following equivalent relationships. The RBF centres correspond to the channel states; the network weights can be interpreted as the generalized a-priori probabilities of channel states; the nonlinearity of hidden nodes is equal to the noise probability density function (p.d.f.) which is Gaussian; the response of hidden nodes realizes the conditional p.d.f. for given states; the response of the network has an identical form to the Bayesian decision function.

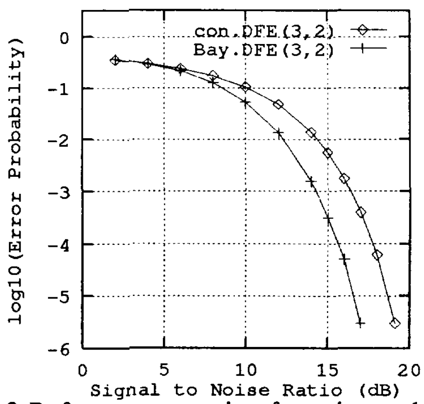

Fig.3. Performance comparison for stationary channel.

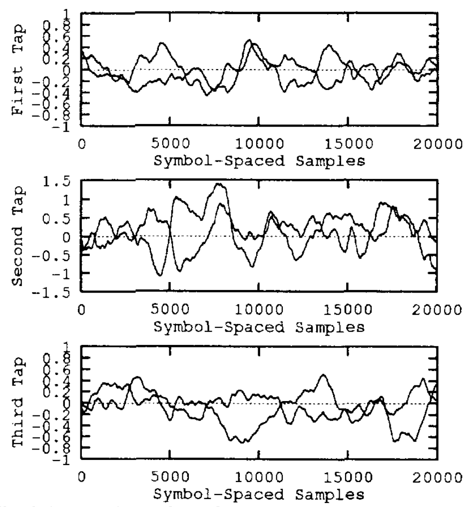

Fig.4. Tap trajectories of three-path fading channel.

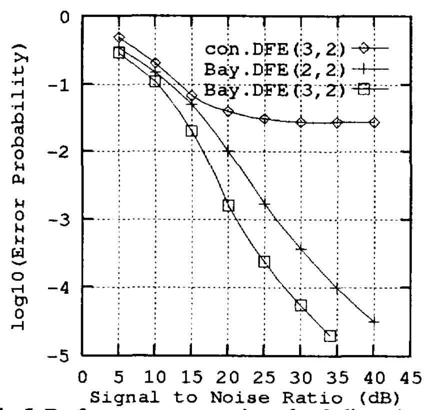

Fig.5. Performance comparison for fading channel.

Better performance of the Bayesian DFE over the conventional DFE [11] is demonstrated using two examples. The first example is a three-tap stationary channel with a transfer function $A(z) = (0.4313 + j0.4311)(1-(0.5+j)z^{-1})(1-(0.35+j0.7)z^{-1})$. Both equalisers had a structure of $d=2$, $m=3$ and $n=2$. The error rate curves of these two equalisers are plotted in Fig.3. The second example is a three-path Rayleigh fading channel [11], and the trajectories of the channel taps are shown in Fig.4. The least mean square algorithm was used to track this time-varying channel and the channel estimate was then employed to design equaliser. The error rates of the conventional and Bayesian DFEs are depicted in Fig.5.

5. Nonlinear modelling and prediction

In some communication systems, significant nonlinear distortion is present in transmission path and it is necessary to represent these channels by the model

$$y(k) = f_c(s(k), \cdots, s(k-n_c+1)) + e(k), \tag{11}$$

where $f_c(\cdot)$ is some complex nonlinear function. In this situation, it is required to identify the nonlinear channel (11) in order to implement the maximum likelihood Viterbi detection [11]. Another application of nonlinear modelling and prediction is the echo cancellation in the communication network [12]. In this scenario, the received signal contains three components

$$y(k) = y_1(k) + y_2(k) + e(k), \tag{12}$$

where $y_1(k)$ is the far-end signal, and $y_2(k)$ is the echo. The aim of echo cancellation is to produce an accurate estimate of the echo and to subtract this estimate from the received signal. When the echo path contains significant nonlinear elements, nonlinear modelling capability is required. In the current application, we use the complex RBF network to model the nonlinear channel (11).

Assume that digital symbols are 4QAM and the nonlinear channel contains three elements in cascade. The symbols $s(k)$ are first distorted by a static nonlinearity

$$u(k) = \frac{2s(k)}{1+|s(k)|^2}\exp\left(j\frac{\pi}{3}\frac{|s(k)|^2}{1+|s(k)|^2}\right), \tag{13}$$

where $\exp(j\phi) = \cos(\phi) + j\sin(\phi)$. This is followed by a finite impulse response filter with the transfer function

$$A(z) = (0.3725 + j0.2172)(1- (0.5+j)z^{-1})(1-(0.35+j0.7)z^{-1}). \tag{14}$$

The output of this linear filter $v(k)$ is further distorted by a third-order complex Voterra nonlinearity and an additive Gaussian white noise

$$y(k) = v(k) + 0.2v^2(k) - 0.1v^3(k) + e(k). \tag{15}$$

We first fitted a linear model $\hat{y}(k) = \hat{a}_0 s(k) + \hat{a}_1 s(k-1) + \hat{a}_2 s(k-2)$ to this nonlinear channel using the complex RLS algorithm. At every recursion, a separate block of test data was used to calculate the mean square error (MSE). The MSE trajectories for different noise powers are plotted in Fig.6. Since $\mathbf{s}(k) = [s(k)\ s(k-1)\ s(k-2)]^T$ has 64 combinations, the channel has 64 states. The errors between the channel states and the linear model states are depicted in Fig.7. Next we identified a RBF network model of 30 centres using the recursive clustering and least squares algorithm. The MSE trajectories are shown in Fig.8 and the modelling errors are given in Fig.9. Finally we used a RBF network model of 64 centres. The centres were fixed to the 64 states of $\mathbf{s}(k)$, and we only estimated the weights using the RLS algorithm. The MSE plots are shown in Fig.10 and the modelling errors are given in Fig.11.

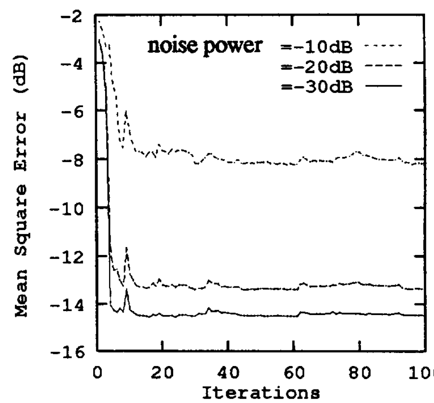

Fig.6. Mean square error plot of linear model using recursive least squares algorithm.

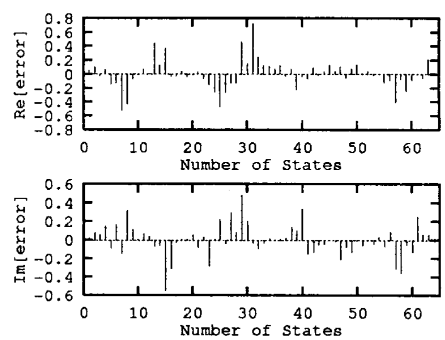

Fig.7. State errors of linear model. Noise power is -30dB.

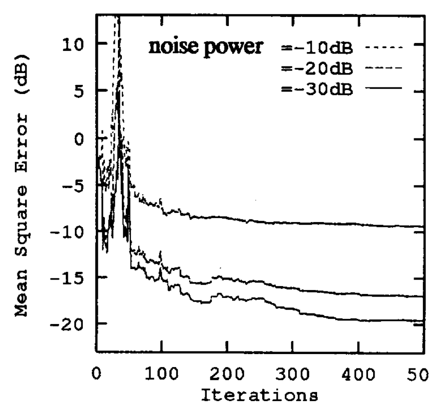

Fig.8. Mean square error plot of RBF model with 30 centres using recursive clustering and LS algorithm.

6. Conclusions

A complex RBF network has been proposed for nonlinear signal processing in multi-dimensional complex space. This network contains the usual real RBF network as a special case. Furthermore it has been shown that the existing learning algorithms for the real RBF network can easily be extended to the complex RBF network. The application of this complex RBF network to communication systems has been demonstrated using the equalisation of time-dispersive channels and the modelling of nonlinear channels.

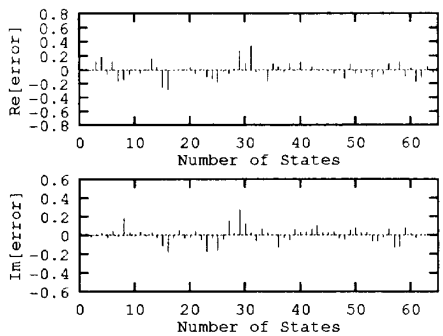

Fig.9. State errors of RBF model with 30 centres. Noise power is -30dB.

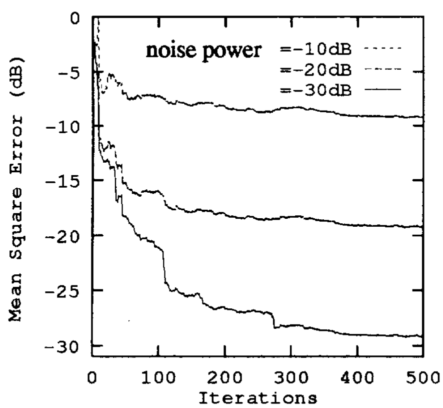

Fig.10. Mean square error plot of RBF model with 64 centres using recursive LS algorithm.

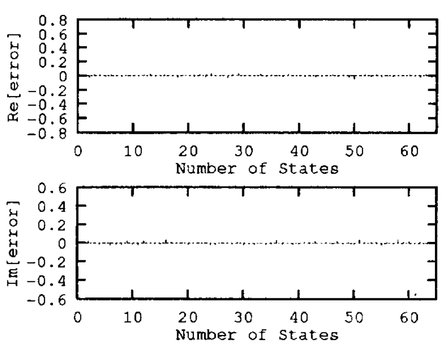

Fig.11. State errors of RBF model with 64 centres. Noise power is -30dB.

7. References

[1] T.L. Clarke, "Generalization of neural networks to the complex plane," in *Proc. Int. Joint Conf. Neural Networks* (San Diego), 1990, Vol.II, pp.435-440.

[2] N. Benvenuto, M. Marchesi, F. Piazza and A. Uncini, "Complex neural networks for equalizing nonlinear digital radio links," in *Adaptive Algorithms: Applications and Non-classical Schemes*, D. Docampo and A.R. Figueiras, eds., Universidada de Vigo, 1991, pp.162-170.

[3] M. Peng, C.L. Nikias and J.G. Proakis, "Adaptive equalization with neural networks: new multilayer perceptron structures and their evaluation," in *Proc. ICASSP* (San Francisco), 1992, Vol.II, pp.301-304.

[4] D.S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," *Complex Systems*, Vol.2, pp.321-355, 1988.

[5] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, Vol.78, No.9, pp.1481-1497, 1990.

[6] S. Chen, C.F.N. Cowan and P.M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Networks*, Vol.2, No.2, pp.302-309, 1991.

[7] J. Moody and C.J. Darken, "Fast-learning in networks of locally-tuned processing units," *Neural Computation*, Vol.1, No.2, pp.281-294, 1989.

[8] S. Chen, S.A. Billings, and P.M. Grant, "Recursive hybrid algorithm for non-linear system identification using radial basis function networks," *Int. J. Control*, Vol.55, No.5, pp.1051-1070, 1992.

[9] S. Chen and S.A. Billings, "Neural networks for nonlinear dynamic system modelling and identification," *Int. J. Control*, Vol.56, No.2, pp.319-346, 1992.

[10] V. Kadirkamanathan, M. Niranjan and F. Fallside, "Models of dynamic complexity for time-series prediction," in *Proc. ICASSP* (San Francisco), 1992, Vol.II, pp.269-272.

[11] J.G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

[12] D.G. Messerschmitt, "Echo cancellation in speech and data transmission," *IEEE J. Selected Areas Communications*, Vol.SAC-2, No.2, pp.283-297, 1984.