RF Power Amplifier Behavioral Modeling using a Globally Recurrent Neural Network

Bill O'Brien, John Dooley, and Thomas J. Brazil

School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Ireland.

***Abstract* — In this paper it is shown that a globally recurrent time delay neural network can accurately model a nonlinear RF power amplifier having significant memory. The recognized difficulty of training a recurrent neural network is overcome by reducing it initially to a feedforward network, training that network, and then using the weights established by this training sequence in a restructured recurrent network. The training of the recurrent network thus reduces to the training of a feedforward network and a simple restructuring. The required maximum input delay is established by examination of the temporal profile of the energy contained in the amplifier impulse response. The model was successfully trained with an RF passband time domain multi-sine signal and subsequently validated with another multi-sine signal composed of different sine components at different amplitudes. A second model trained with a wider bandwidth multi-sine was successfully validated with a W-CDMA signal.**

***Index Terms* — Behavioral modeling, neural networks, power amplifiers, recurrent neural networks.**

I. INTRODUCTION

In recent years increasing attention has been given to the application of neural networks to behavioral modeling of RF power amplifiers, see for example [1]. It has been shown that feedforward neural networks with sufficient neurons are 'universal approximators' and can model any linear or non-linear system without memory [2]. Difficulties have arisen due to the 'curse of dimensionality', which refers to the amount of training data required to fully characterise a system. Another problem is the sometimes lengthy training time required to reach an acceptable level of performance.

With the development of newer communication systems such as W-CDMA, wider power amplifier bandwidths have become necessary and the memory effects of the PA start to become more critical to performance, making a simple feedforward model inadequate. Recent approaches have used feedforward time delay networks (TDNN) [3] to enable a feedforward network to model a system's dynamics, but as this method is limited to an input moving average representation, it may be expected to fail when the system being modelled has significant memory effects [4].

By the addition of feedback connections (recurrency), a neural network can fully model dynamic systems. Fully recurrent networks have been proposed to model power amplifiers [5]. However recurrent networks can be very slow to converge using standard training algorithms such as the back propagation through time algorithm [6] due to the problem of vanishing gradients.

The addition suggested here of a global feedback connection to the TDNN network, but with no local feedback, allows the implementation of a full non-linear autoregressive moving average model with external inputs (NARMAX) and can give a more effective and flexible behavioral modeling method with potentially fewer weights needed in the model, leading to faster execution. In addition, a globally recurrent network can be trained using an approach that is simpler and faster than the back propagation through time method. This approach is not possible in fully recurrent networks as the locally recurrent signals are unknown.

The remainder of this paper is organised as follows: Section II outlines the method used to train the globally recurrent network. In Section III we demonstrate the validation of a globally recurrent time delay neural network model for an amplifier with significant memory. Conclusions are contained in Section IV.

II. TRAINING RECURRENT NEURAL NETWORKS

The conventional approach to training recurrent neural networks is the back propagation through time method [6]. In this method the network is 'unfolded in time' or effectively duplicated at each time step so that the dependency of the output on previous outputs can be calculated at each step. This means that the training effort is increased by a multiple of the number of feedback delays used relative to training of feedforward networks. Convergence using this or any of the standard recurrent training methods can be slow, resulting in very long training times. Also, the error performance may not be very good. These methods must be used when the network is locally recurrent as the local feedback signals are unknown. However, if recurrency is restricted to global feedback from output to input with no local feedback connections, the feedback signal is known, as it is actually the desired output signal or signals. A feedforward network can be constructed with inputs including the actual and delayed inputs and also the delayed desired output signals. This network can then be trained using any of the standard training methods used for feedforward networks. When trained, the network can be restructured into a recurrent network using the weight matrices learned under the feedforward training. This provides a fast training technique with good error performance.

Linear activation functions are used in the output layer. A direct connection of the input to the output layer allows the network to learn the linear component of the amplifier behavior more easily, essentially saving the hidden layer for the non-linear part [7]. The structure of the final network is shown in Fig.1, and the equation describing the input-output mapping of the NARMAX model is presented in (1).

$$y(t) = g(\sum_{k=1}^{h} w_2^k (f(\sum_{l=0}^{p} w_1^l u(t-l) + \sum_{m=1}^{q} w_1^{p+m} y(t-m)) + w_2^0 u(t)) \quad (1)$$

where u(t) is the input
y(t) is the output
h is the number of hidden layer neurons
p is the number of input delays
q is the number of feedback delays
w_1 is the weight vector of the hidden layer
w_2 is the weight vector of the output layer
f is the sigmoid activation function of the hidden layer
g is the linear activation function of the output layer.

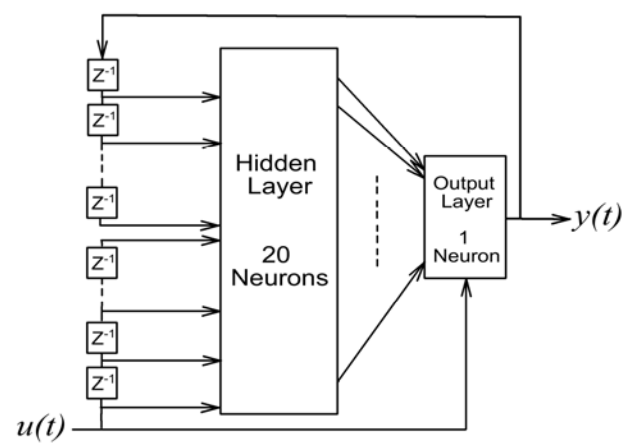

Fig. 1 Recurrent Neural Network

III. Model Validation

The key issues to be addressed in this model are the number of input and feedback delays, the maximum delay length and finally the training technique. The number of neurons needed in the hidden layer can be established by trial and error, or by using a growth or pruning algorithm and will depend on the degree of nonlinearity of the amplifier and the complexity of the training signal.

It can be difficult to establish what length of input delay to apply and how many individual delay states to use within this overall delay time. The Lipschitz criterion [8] gives an indication of the order of delay required, but does not indicate how many or which delay states to use. A trial and error approach can be used, but is time consuming. The approach taken here was to look at the amplifier forward transmission impulse response and use a maximum delay length equal to the time during which significant energy is present in the impulse response. A simple pruning algorithm was then used to remove any delay states within this delay period that had little effect on the overall system performance.

The feedback delays are not as critical as the input delays. Even one feedback delay can result in a very significant improvement in the training time and the error performance of the model. As just one feedback delay signal contains all previous output information from the start of the simulation, one delay should be sufficient [2]. Additional feedback delays reduce the initialisation time required for the trained network to converge on the solution. Too many feedback delays can encourage instability.

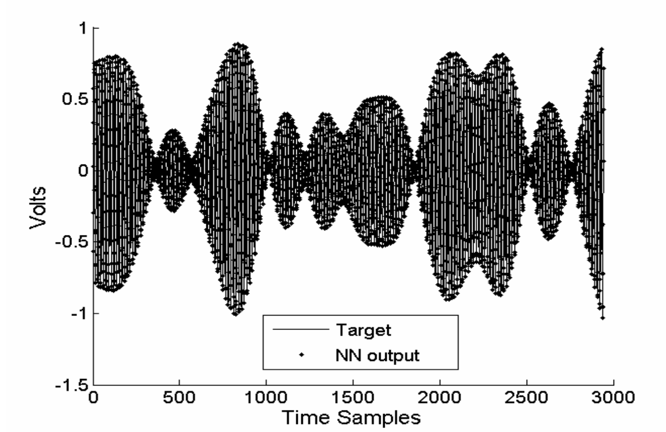

Fig. 2 Training signal (20 sine wave components); target output and NN model output

The signal sampling rate has a direct impact on the number of delay states needed. If too high a sampling rate is used, the network becomes very large and therefore slow to train and execute. For best performance the sampling rate should be kept close to the minimum needed to ensure fidelity to the original waveform. This is critical when simulating at passband frequencies as the required number of delays can become very large. A sampling rate of 16 times the centre frequency was used here to model at passband frequencies, but this could be reduced.

A multi-sine training signal consisting of 20 randomly phased sine waves of equal amplitudes with frequencies covering a 5MHz band centred at 2GHz was first constructed. A sequence of several different amplitudes was used to drive an in-house convolution-based impulse solver of the PA device model and the simulated output saved. A second signal with 14 sine wave components in the same frequency band and with the same amplitudes as the first was then run on the circuit model to generate a test signal. A second test signal consisting of 14 sine waves but with lower amplitudes than the previous signals was also run. The first 20-component input and output signals were then used as training and target signals to train a feedforward neural network with the same dimensions as the recurrent network shown in Fig. 1. The weights generated were then used to build a globally recurrent network. The training algorithm used was the Levenberg-Marquardt algorithm.

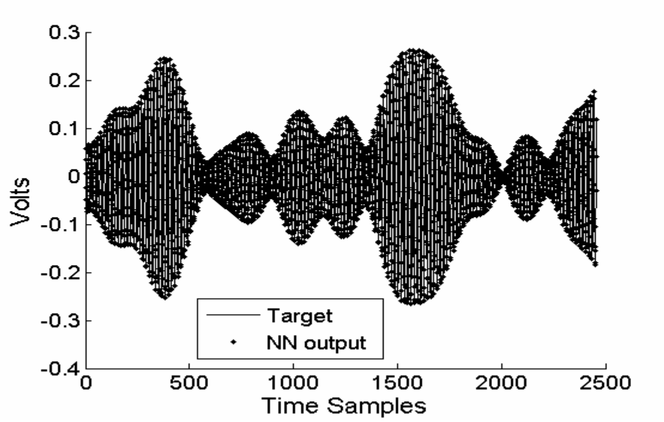

Fig.3 Test signal (14 sine wave components, lower amplitude than training signal); target output and NN model output

A further multi-sine training signal consisting of 70 sine components covering a bandwidth of 60 MHz was used to train another network. A single carrier W-CDMA signal was used as the test signal for this model.

IV. Validation Test Results

The network was tested first with the training signal, then with the other two multi-sine test signals. The resulting plots are shown in Fig.2, Fig. 3, Fig.4 and Fig. 5. The normalised mean square error performance is given in table 1. The NMSE on the training signal shows that the recurrent network constructed using the weights from the feedforward network is behaving as would be expected for a network actually trained using the training signal, with a training error of 3e-9.

Signal	NMSE
Training Signal	-55.5 dB
14 Sine Test Signal	- 49.3 dB
14 Sine Test Signal, Reduced Amplitude	- 31.5 dB

Table 1 Multi-sine error performance of the recurrent network

Signal	NMSE
70-Sine Training Signal	-45 dB
W-CDMA Test Signal	-38 dB

Table 2 W-CDMA error performance when trained on 70 component multi-sine

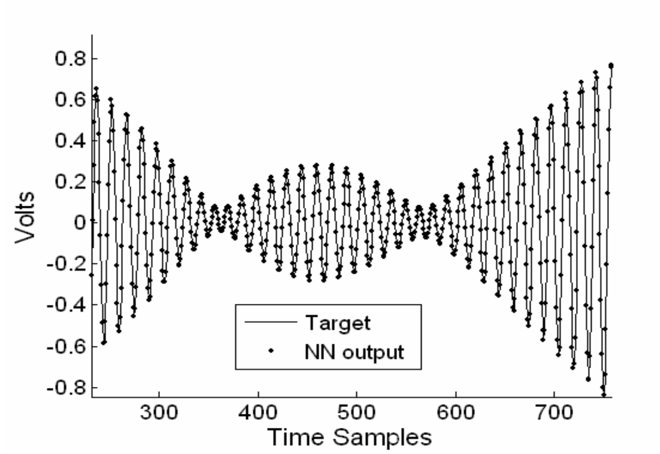

Fig. 4 Enlarged part of training signal in Fig. 2

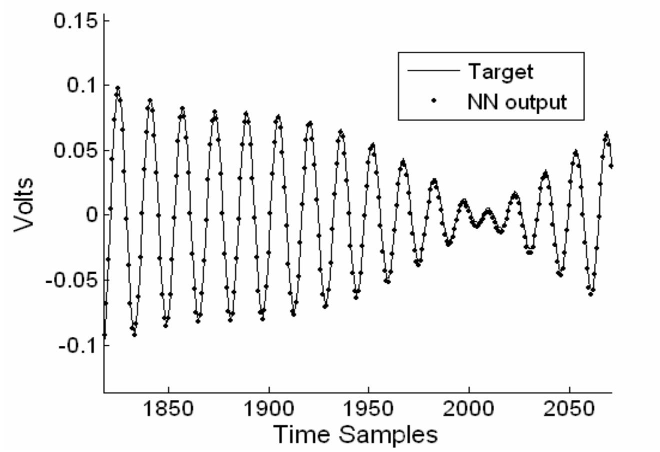

Fig. 5 Enlarged part of test signal in Fig. 3

The performance on the other signals which were not used in training shows that the constructed recurrent network has excellent generalisation capability.

As a further test, a 25-neuron network was trained with 8000 samples of a 70-component multi-sine signal and then tested with 250k samples of a single carrier W-CDMA signal. Figs. 6 and 7 show the training signal. The results are given in Table 2.

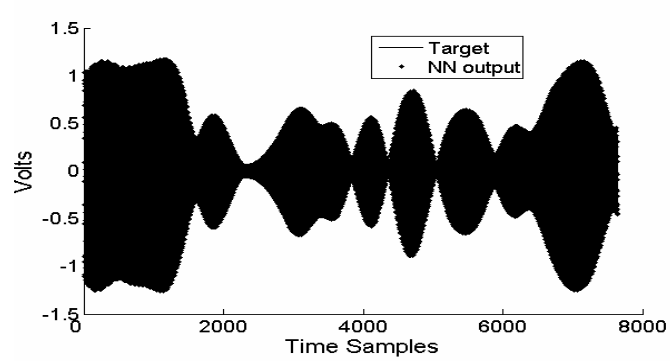

Fig. 6 Target and NN model output response to 70-component multi-sine training signal

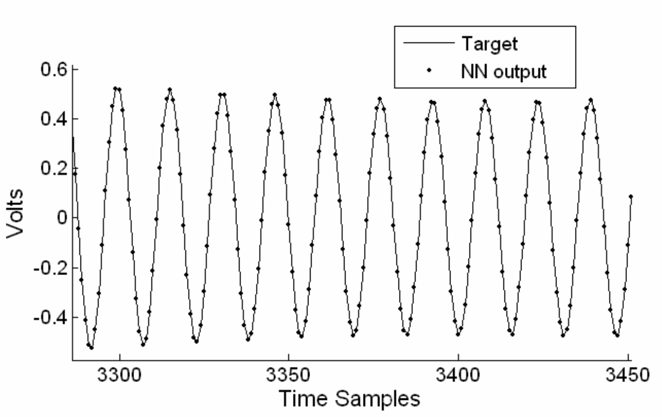

Fig. 7 Enlarged part of training signal in Fig. 6

The test signal is shown in Figs. 8 and 9.

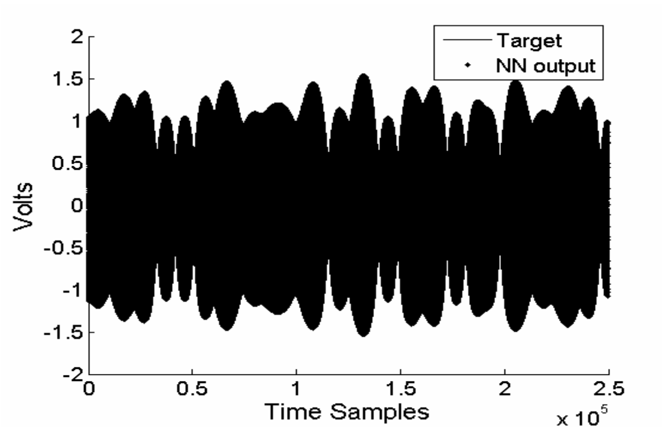

Fig. 8 Independent validation of NN model; Output response to W-CDMA test signal

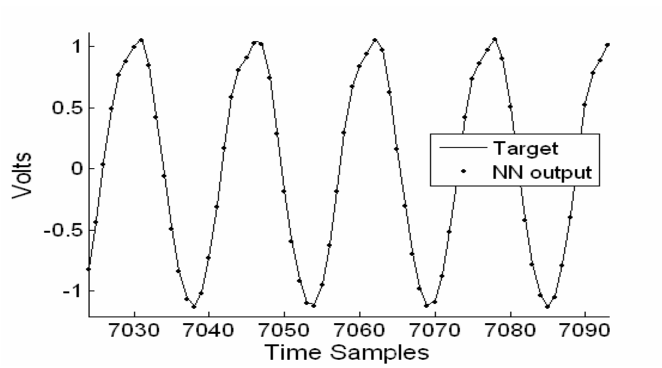

Fig. 9 Enlarged part of W-CDMA test signal in Fig. 8

As can be seen, the model was able to accurately simulate the amplifier when driven by the W-CDMA signal even though the training was carried out using the 70-component multi-sine signal.

V. Conclusion

In this paper it was shown that a globally recurrent neural network can provide a highly accurate RF passband behavioral model of a microwave power amplifier with significant memory and nonlinearity. A simple technique was also shown that allowed the recurrent network to be trained as efficiently as a feedforward network. Additionally, an approach was given to estimate the maximum required input delay, and which delays to include based on the amplifier impulse response. It was also shown that a network trained with a multi-sine signal could accurately simulate the amplifier when driven by a W-CDMA signal.

Acknowledgement

This work was supported by the Enterprise Ireland "Advanced Technology Research Program" and Science Foundation Ireland.

References

[1] Q.J. Zhang and K.C. Gupta, *Neural Networks for RF and Microwave Design*, Norwood, MA: Artech House, 2000.

[2] S. Haykin, *Neural Networks: a comprehensive foundation*, 2nd ed, Prentice Hall, 1999.

[3] Taijun Liu, Slim Boumaiza, and Fadhel M. Ghannouchi, "Dynamic Behavioral Modeling of 3G Power Amplifiers Using Real-Valued Time-Delay Neural Networks," *IEEE Trans. Microwave Theory Tech.,* vol. 52, pp. 1025-1032, March 2004.

[4] Danilo Mandic and Jonathon A. Chambers, *Recurrent Neural Networks for Prediction*, Wiley, 2001.

[5] Dahn Luongvinh, Youngwoo Kwon, "Behavioral Modeling of Power Amplifiers Using Fully Recurrent Neural Networks", *Microwave Symposium Digest, 2005 IEEE MTT-S International,* June 2005

[6] Paul J. Werbos, "Back propagation through time: what it is and how to do it," in *Proc. IEEE*, vol. 78, pp. 1550-1560, Oct. 1990.

[7] Gerard Dreyfus, *Neural Networks Methodology and Applications*, Springer-Verlag, 2005.

[8] M. Norgaard, O. Ravn, N.K. Poulson and L.K. Hansen, *Neural Networks for Modeling and Control of Dynamic Systems,* Springer-Verlag, 2000.