Utilizing Sparse-Aware Volterra for Power Amplifier Behavioral Modeling

Keith Finnerty, John Dooley, Ronan Farrell

National University of Ireland Maynooth, Callan Institute, Maynooth, Kildare Ireland

Abstract — This paper presents a method for reducing the number of weights in a time series behavioral model for a power amplifier. The least-absolute shrinkage and selection operator (Lasso) algorithm is used to reduce the kernel size, preserving the important kernels, while eliminating the less important kernels. The algorithm is evaluated on a behavioral model for a class AB amplifier, the algorithm reduces the number of weights by greater than 70% without degrading model performance by a significant amount.

Index Terms — Behavioral modeling, time series, Volterra, system identification, Lasso.

I. Introduction

Behavioral modeling of power amplification systems is an important step in RF system modeling, providing a computationally efficient structure with an accurate description of the non-linear effects can greatly reduce time required to model the entire system. It also provides a platform to analyze and develop a solution for the non-linear effects, the reduced computational time increases the number of possible iterations that can be preformed, greatly increasing the speed with which the optimum solution can be found.

Volterra series offers a powerful time series estimator, capable of modeling complex non-linear systems with memory. One group of nonlinear dynamic systems are high power radio frequency amplifiers. The Volterra series calculates every interaction between the given inputs up to a defined order of non-linearity, as the number of inputs or the order of non-linearity increases the number of coefficients rapidly increase. This expansion described as the "curse of dimensionality" can greatly reduce the computational efficiency of the model as the nonlinear order or memory depth of the system to be modeled grows.

Modern amplification systems such as Doherty [**1**], envelope tracking (ET) [**2**] and outphasing amplifiers [**3**] have increasingly complex characteristic behavior as they exploit more efficient modes of operation of the active devices. In some cases a large number of non-linear orders are required to accurately identify the system, especially in cases where the devices are operated in switch mode. In such systems, higher order terms may be required, however using a classical Volterra series as the model will generate all intermediate coefficients, some of which are not required.

In [**4**] a method of generating sparse Volterra kernels is presented. A Lasso algorithm introduced in [**5**] is a method of variable selection and shrinkage in linear models. Comparable to ridge regression as a method to shrink the set of weights, it simultaneously has the ability to set some of the remaining coefficients to zero in order to generate a sparse estimation. The estimation removes the least important weights, leaving a set of core weights. This reduced kernel space is what is used to derive the Volterra behavioral model. In this paper we will investigate the effectiveness of the method in reducing the number of kernels required to identify the behavioral model of a power amplification system.

II. LASSO Algorithm

Parameter reduction is an important tool for data analysis, the two main reasons being prediction accuracy and interpretation. With a large number of weights the predicted system can have a low bias or systematic error, but the individual weights can have a large variance, the weights derived have a large deviation from their true value. The aim of parameter reduction is to reduce the variance while maintaining a reasonable bias relationship.

With a large number of weights it can be difficult to identify the most influential ones, though using parameter reduction techniques we can remove the least important weights to leave core subset which can accurately represent the system behavior. For behavioral modeling, parameter reduction is employed to achieve the smallest subset of weights which will maintain the model accuracy but provide a computational saving.

The Lasso definition for a time series is outlined as follows, $\boldsymbol{x}^{\boldsymbol{t}}=(x_{t,1}, \ldots ,x_{tp})^T$ where $\boldsymbol{x}^{\boldsymbol{t}}$ are the computed kernels of the input signal and y_t are a vector of training data. The following assumptions, common to regression methods are made: the observations (y_t) are independent and we assume that x_{tp} is normalised so that the mean of x_{tp} is zero and the mean of x^2_{tp} is one. The Lasso estimate [**5**] is defined by,

$$(\hat{\alpha},\hat{\beta}) = \arg min\left\{\textstyle\sum_{t=1}^{N}\left(y_t - \alpha - \sum_p \beta_p x_{tp}\right)^2\right\} \quad (1)$$

$$subject\ to \sum_p |\beta_p| \le k \quad (2)$$

Where $\hat{\beta}$ is the regression estimation of β that minimizes the cost function, $k \geq 0$ is a tuning parameter and α is a scalar

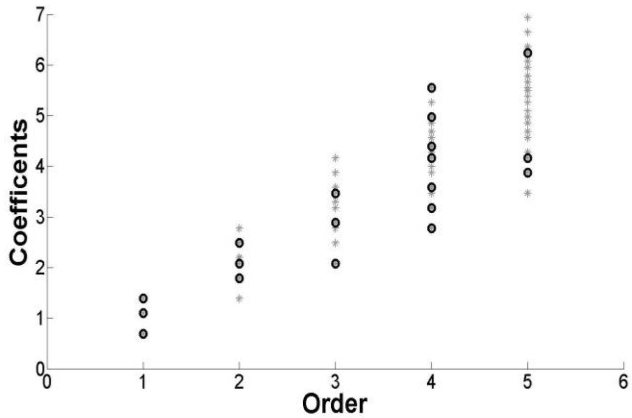

Fig 1 . Kernel Reduction utilizing Lasso function. Grey is the full Volterra series, Black are the chosen parameters for reduced model.

used to determine an offset in the dataset y_t, given that the mean of y_t is zero we can omit α and $\hat{\alpha}$ without loss of generality. By tuning the value of k the reduction function can be controlled to find a balance between model accuracy and number of remaining weights. In behavioral modeling this value of k will be dependent on the characteristics of the system being identified, several iterations of the algorithm are run with varying values of k, and the best balance is chosen. Due to the fact that a power amplifier is a largely deterministic system and the input and output data is highly correlated, the Lasso method can be very effective in reducing the number of weights in the time series.

III. Volterra Series

Volterra series behavioral modeling is a powerful time series estimator for non-linear systems. It has been extensively used to model and pre-distort power amplification systems [**6**]. The Volterra series is outlined in equation (3), where P is the order of non-linearity and M is the number of memory terms.

$$y(n) = h_0 + \sum_{p=1}^{P} h_p[x(n), \ldots, x(n-M)] \tag{3}$$

The Volterra series is composed of all possible combinations of non-linear order and memory terms and as the terms P and M increase the number of weights increase rapidly. The computational complexity for the Volterra series is outlined in equation (4) [**7**]. The number of operations calculated is the total number of complex computations. It is clear that increasing the number of either memory or non-linear polynomial coefficients has a dramatic impact on the total number of weights.

To combat the rate of expansion various reduced Volterra kernels have been proposed such as Memory polynomial (MP) [**8**], Cross Memory polynomial (CMP) [**9**], and sparse Volterra series (SVS) [**10**]. Each method aims to reduce the kernel size to increase the efficiency of the behavioral model. In many cases such as MP and CMP the resulting kernel structure is static. As a result there exists the possibility of missing important weights specific to a system. Alternatively including unnecessary weights can result in an inefficient model.

$$C(P,M) = \sum_{p=1}^{P} \frac{(M-1+p)!}{(p-1)!\,(M-1)!} \tag{4}$$

Utilizing the Lasso [**5**] function, parameter estimation is carried out on the full set of Classical Volterra weights, an m by n matrix where m is the total number of Volterra weights and n is the length of the input signal and the target signal, an 1 by n vector, to derive a sparse-aware Volterra kernel [**4**], derived specifically for the measured system. This estimation is carried out in advance of extracting the model weights. Depending on the number of Volterra weights the Lasso algorithm can be carried out on a subset of the total measurement data, reducing the total computation load in the parameter reduction function. The system identification process is then carried out on the reduced Volterra kernel. Through reduction of both kernel calculation and weight estimation the total computation complexity of model extraction and implementation can be significantly reduced. An example of the kernel reduction is presented in Fig 1. The data is separated by orders of non-linearity for clarity. The reduced number of unknown coefficients in the model prevents the identification process from under determining the system due to insufficient measurement data.

IV. Results

To evaluate the performance of the model it was applied to measurement data of a NXP Class AB power amplifier. The amplifier capable of a 10W peak output power was operated close to saturation in order to extract a non-linear signal. The signal was provided by a Rhode and Schwarz SMU200A signal generator and captured by a Rhode and Schwarz FSQ vector signal analyzer. In order to minimize measurement noise a series of measurements were taken, the individual measurements were aligned and an average of six measurements was taken, this significantly lowers the noise floor of the measurement system. In order to operate the amplifier close to saturation for non-linear operation, a linear driver amplifier was used, allowing peak output power from the Class AB amplifier to be achieved.

The extraction of the behavioral model was carried out in Matlab, the performance of the model was evaluated using the normalized mean square error (NMSE) of equation (5). The captured data was arranged in separate training and test data sets, allowing the model to be independently evaluated. Each data set has at least 16000 samples. The Lasso method estimated the sparse kernels before system identification,

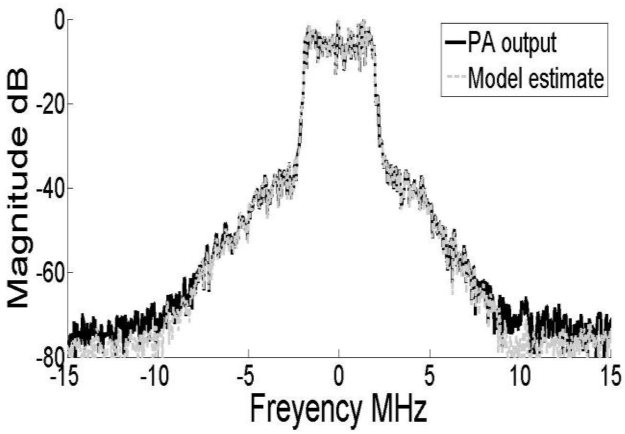

Fig 2 presents the measured non-linear signal and the reduced Volterra model superimposed. The model accurately predicts the amplifier non-linearity to -60 dB spectral power.

therefore it does not limit the system identification to a particular algorithm. In this paper the least squares method with QR decomposition is used for model extraction [**11**]. The model is extracted for a series of Volterra kernel sizes and compared to the corresponding Lasso reduced model. From the table of results it can be seen that the Lasso is capable of reducing the number of Volterra kernels on average by 78% while maintaining model accuracy.

$$\text{NMSE} = \frac{|d_k(n) - y_k(n)|^2}{|d_k|^2} \tag{5}$$

Table 1. Results from power amplifier modeling, *P* is the order of non-linearity, *M* is the number of memory terms, and the reduction of terms is presented as a percentage.

		Full Volterra		Reduced Volterra		
P	**M**	**NMSE**	**Kernels**	**NMSE**	**Kernels**	**Reduction**
5	**4**	**-45.0**	**125**	**-44.8**	**27**	**78%**
5	**5**	**-44.8**	**251**	**-44.9**	**57**	**77%**
7	**4**	**-44.6**	**329**	**-45.2**	**86**	**74%**
7	**5**	**-43.7**	**791**	**-45.1**	**113**	**86%**

V. Conclusion

This paper evaluates the effectiveness of the Lasso function as a method for parameter reduction in power amplifier behavioral models. A 10W class AB power amplifier is used in the experimental validation of this approach . It is shown in Table 1 that a reduction in the number of weights greater than 70% could be achieved. This represents a significant reduction in computation load. The results also show that model accuracy was maintained and in some cases the accuracy increased as a result of the reduction of variance in the core weight values by removing the least important weights.

VI. Acknowledgement

This material is based upon works supported by Science foundation Ireland under Grant No. 10/CE/I1853 and by HEA under PRTLI5 and is being co-funded by the Irish Government and the EU under Ireland's Structural Funds Programmes 2007-2013: Investing in your future. The authors would also like to acknowledge NXP Semiconductors for providing the Class AB power amplifier used in this paper. The authors gratefully acknowledge this support.

References

[1] Kim, Bumman; Jangheon Kim; Ildu Kim; Jeonghyeon Cha, "The Doherty power amplifier," *Microwave Magazine*, vol. 7, no. 5, pp. 42,50, Oct 2006.

[2] Staudinger, J.; Gilsdorf, B.; Newman, D.; Norris, G.; Sadowniczak, G.; Sherman, R.; Quach, T, "High efficiency CDMA RF power amplifier using dynamic envelope tracking technique," in *Microwave Symposium Digest*, 2000, pp. 873,876.

[3] H Chireix, "High Power Outphasing Modulation," in *Radio Engineers, Proceedings of the Institute of*, 1935, pp. 1370,1392.

[4] Kekatos, V.; Angelosante, D.; Giannakis, G.B, "Sparsity-aware estimation of nonlinear Volterra kernels," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2009, pp. pp.129,132.

[5] Robert Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society*, vol. 58, no. Series B, pp. 267 -- 288, 1994.

[6] V. John Mathews and Giovanni L. Sicuranza, *Polynomial Signal Processing* , 1st ed., Wiley & Sons, Ed., 2000.

[7] Tsimbinos, J.; Lever, K.V., "Computational complexity of Volterra based nonlinear compensators," , 1996, pp. 852,854.

[8] Wang Huadong; Bao Jingfu; Wu Zhengde; Huang Jingfu, "An Memory polynomial model for power amplifiers," in *Communications, Circuits and Systems*, 2008, pp. 1346,1349.

[9] Farsi, S.; Dooley, J.; Finnerty, K.; Schreurs, D.; Nauwelaers, B.; Farrell, R., "Cross-memory polynomial modeling for RF circuits," in *Microwave Integrated Circuits Conference* , 2012, pp. 262,265.

[10] Kalouptsidis, N.; Mileounis, G.; Babadi, B.; Tarokh, Vahid, "Adaptive algorithms for sparse nonlinear channel estimation," in *Statistical Signal Processing*, 2009, pp. 221,224.

[11] David C. Lay, "Linear Algebra and its applications," in *Linear Algebra and its applications*., 2004, pp. 364-365.