

On the Wiener and Hammerstein Models for Power Amplifier Predistortion

P.L. Gilibert, G. Montoro and E. Bertran

Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), Spain

Abstract — This paper presents a comparative study on the suitability of using Hammerstein or Wiener models to identify the power amplifier (PA) nonlinear behavior considering memory effects. This comparative takes into account the operational complexity regarding the identification process as well as their accuracy to follow the PA behavior. Both identified PA models will be used to estimate a Hammerstein based predistorter in order to see which model combination provides better linearization results. In addition, two adaptive algorithms for predistorting both PA models are compared in terms of accuracy and converge speed.

I. INTRODUCTION

The significant growth of digital predistortion solutions to linearize power amplifiers (PA's) has to do with the fact that nowadays, most of the communication equipment already incorporates some digital processor for mandatory issues within the wireless standards (i.e, coding, interleaving, OFDM,...). So then, such digital processors can also be used to implement digital-based linearizers avoiding the need for including a specific digital device with its consequent cost increment and energy consumption [1]. Most of the digital predistorters proposed in the recent past years have been modeled using memoryless techniques, aiming at compensate the instantaneous nonlinear behavior of the PA, characterized by the AM/AM and AM/PM static curves. This PA memoryless model might be an acceptable approximation for narrowband signals (e.g. nearly-constant envelope modulations). However, memoryless predistortion has shown insufficient cancellation performance since new multilevel (and multicarrier) modulation formats claim for power handling capability and higher baseband bandwidth.

Current wideband multilevel modulation formats demand better PA models, all considering memory effects. Some of the proposed PA and predistorter (PD) models in literature are based on Volterra series (or pruned Volterra)[2][3], memory polynomials [4], Wiener-Hammerstein models [5] [6], or neural networks [7]. Leaving apart the Volterra or neural networks models for being more computational complex, this paper studies the suitability in terms of complexity and accuracy of using the Wiener and Hammerstein models for the PA identification and predistortion.

II. PROBLEM STATEMENT

In order to carry out the predistortion learning process there are two basic configurations aimed at estimating the predistorter coefficients. These basic configurations are the direct learning (predistortion) and the indirect

learning, also called postdistortion and translation method. In the indirect learning, see Fig. 1, the inverse PA model is estimated by using its input and output data, as described in eq. (1),

$$F_{post}(G(x)) = G_0 \cdot x \quad (1)$$

Once the learning process has finished, the postdistorter coefficients are copied to an identical model that is used to predistort the input. With this configuration there is no need to calculate the PA model, only its input and output data is needed to estimate the postdistortion model. But applying the commutative property (translation method) in nonlinear systems is not properly formal in the mathematical sense.

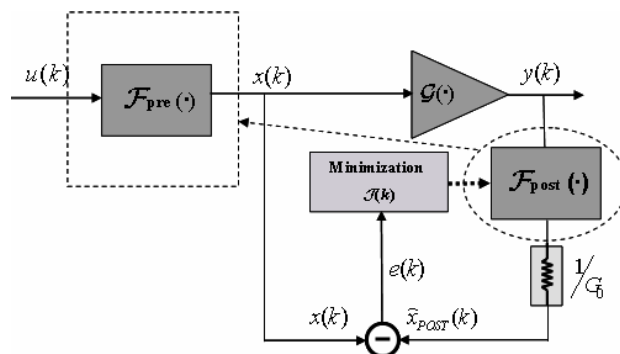


Fig. 1. Indirect learning (postdistortion and translation method).

The direct learning approach, depicted in Fig. 2, permits a direct (more formal) predistorter estimation but requires the real PA or a good model of it to train their coefficients, as described in eq. (2):

$$G(F_{pre}(u)) = G_0 \cdot u \quad (2)$$

Thus for obtaining a good predistorter estimation, first it is necessary to obtain a reliable PA model that respects the trade-off between accuracy and simplicity.

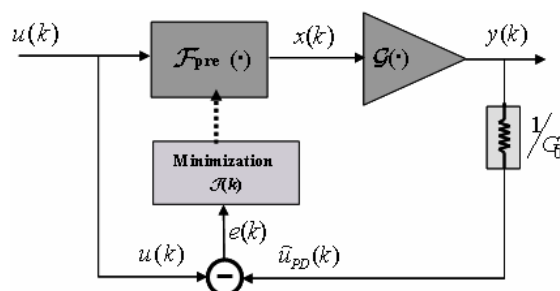


Fig. 2. Direct learning predistortion.

III. POWER AMPLIFIER AND PREDISTORTER MODEL

Volterra series and Neural Networks show good performance in estimating general nonlinear models with memory, but their predistortion is complex and its real-time implementation difficult.

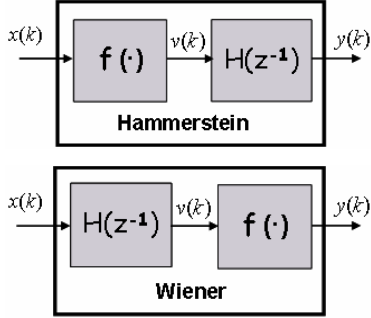


Fig. 3. Block diagram of the Hammerstein and Wiener models.

Apart from those accurate but complex models, the Hammerstein and Wiener models are two simple but effective possibilities for describing the PA and/or the predistorter nonlinear behavior taking into account memory effects. Hammerstein models are composed by a memoryless nonlinearity followed by a linear time-invariant system, as it is shown in Fig. 3. While the Wiener model (Fig. 3) consists of the same subsystems but connected in the reverse order. The use of memory polynomials has been also considered in literature to model the PA or the predistorter, but memory polynomials can be seen as a particular configuration of a more generic Hammerstein model.

A. Hammerstein Models

In this paper we have considered to use an IIR filter in the linear time invariant block, since IIR filters present better resolution than FIR filters at the same filter order, so then, reducing computational load in the DSP. For the nonlinear system a general power series is used. Thus, the Hammerstein model depicted in Fig. 3 can be defined by the following equations:

$$v(k) = \sum_{p=0}^P \gamma_p \cdot x(k) \cdot |x(k)|^p \quad (3)$$

$$y(k) = \sum_{d=1}^D \alpha_d \cdot y(k-d) + \sum_{n=0}^N \beta_n \cdot v(k-n) \quad (4)$$

Being $y(k)$ the PA (or the predistorter) output and $v(k)$ the output of the memoryless non-linear block. P is the order of the memoryless polynomial, while N and D are the general zero/pole order respectively. By combining eq. (3) and (4) we obtain the following closed expression,

$$y(k) = \sum_{d=1}^D \alpha_d \cdot y(k-d) + \sum_{n=0}^N \beta_n \cdot \left(\sum_{p=0}^P \gamma_p \cdot x(k-n) \cdot |x(k-n)|^p \right) \quad (5)$$

Rewriting eq. (5) in a more compact matrix notation, it results:

$$y(k) = \mathbf{c}_k^H \cdot \mathbf{q}_k \quad (6)$$

$$\text{where } \mathbf{c}_k = \begin{bmatrix} d_{01}, d_{02}, \dots, d_{0P}, d_{11}, d_{12}, \dots, d_{1P} \\ \dots, d_{N1}, d_{N2}, \dots, d_{NP}, \dots, \alpha_1, \dots, \alpha_D \end{bmatrix}^T,$$

$$\mathbf{q}_k = \begin{bmatrix} x(k), \dots, x(k) |x(k)|^p, x(k-1), \dots, x(k-1) |x(k-1)|^p \\ \dots, x(k-N), \dots, x(k-N) |x(k-N)|^p \\ y(k-1), y(k-2), \dots, y(k-D) \end{bmatrix}^T,$$

and $d_{np} = \beta_n \cdot \gamma_p$. Vectors are noted in bold and matrices are noted in bold and with a hat. The superindex ^H denotes hermitian transpose.

Note that the Hammerstein model permits linear regression, since the predicted output $y(k)$, is a linear combination of known data (\mathbf{q}_k). The parameter estimation (\mathbf{c}_k) is easy to calculate by minimizing a cost function (defined as the real output minus the predicted output) using gradient techniques.

B. Wiener Models

On the contrary to Hammerstein models, we have considered to use a FIR filter in the linear time invariant block for simplicity, since Wiener models are more complicated to estimate than Hammerstein ones. For the nonlinearity we have considered power series as well as in the Hammerstein model. So then, eq. (7) and eq. (8) describe the relation between the input ($x(k)$) and output ($y(k)$) in a Wiener model (see Fig. 3) by means of an unknown intermediate variable $v(k)$,

$$v(k) = \sum_{n=0}^N \beta_n \cdot x(k-n) \quad (7)$$

$$y(k) = \sum_{p=0}^P \gamma_p \cdot v(k) \cdot |v(k)|^p \quad (8)$$

If we now combine eq. (7) and eq. (8) we obtain,

$$y(k) = \sum_{p=0}^P \gamma_p \cdot \left(\sum_{n=1}^N \beta_n \cdot x(k-n) \right) \cdot \left| \sum_{n=1}^N \beta_n \cdot x(k-n) \right|^p \quad (9)$$

But as it can be deduced from eq. (9), an easy linear regression of its coefficients it is not possible since the filter coefficients are integrated in the power series. In order to solve this problem, it is possible to first estimate the intermediate variable $v(k)$, as it is proposed in [8], for later divide the estimation problem into two steps. To estimate the intermediate variable, it is necessary to use the following assumptions [8]: a) the linear subsystem (the FIR filter) is stable, b) the nonlinear function (power series) is invertible, and c) there is no noise in the system. Assuming this, it is possible to calculate the intermediate variable as described in eq. (10),

$$v(k) = \sum_{n=1}^N \beta_n \cdot x(k-n) = \sum_{p=0}^P \gamma_p^{-1} \cdot y(k) \cdot |y(k)|^p \quad (10)$$

So then, eq. (10) describes an equation where the parameters come in linearly and thus can be estimated by linear regression. The error that wants to be minimized is described in eq. (11), where to avoid the trivial solution ($\gamma, \beta = 0$), it is possible to fix one parameter ($\beta_1 = 1$)

without loss of generality due to the over-parameterization ([8]).

$$e(k) = x(k-1) - \left(\sum_{p=0}^P \gamma_p^{-1} \cdot y(k) \cdot |y(k)|^p - \sum_{n=2}^N \beta_n \cdot x(k-n) \right) \quad (11)$$

where $(\beta_1 = 1)$

Rewriting eq. (11) in matrix notation, we obtain eq. (12)

$$e(k) = x(k-1) - \boldsymbol{\eta}_k^H \cdot \mathbf{g}_k \quad (12)$$

where $\boldsymbol{\eta}_k = [\gamma_0^{-1}, \gamma_1^{-1}, \dots, \gamma_P^{-1}, \beta_2, \beta_3, \dots, \beta_N]^T$, and

$$\mathbf{g}_k = [y(k), \dots, y(k) |y(k)|^p, -x(k-2), \dots, -x(k-D)]^T.$$

Once the quadratic criterion $\left(\min \sum_{k=1}^L (e(k))^2 \right)$ is minimized, the intermediate variable $v(k)$ is available, permitting a two steps estimation of the Wiener coefficients by linear regression by using eq. (7) and eq. (8).

Since the estimation of Wiener models is reasonably more complicated than the Hammerstein ones, in this paper we have considered, for simplicity, that the predistorter (PD) is based on a Hammerstein model.

IV. PREDISTORTION ADAPTIVE ALGORITHMS

In order to estimate the predistorter (PD) coefficients in Fig. 2, we have considered a PD based on a Hammerstein model. The cost function to be minimized ($J(k) = |e(k)|^2$), is defined as the mean square error between the output samples of the PA (previously divided by the desired linear gain G_0) and the input samples of the PD, as is described in eq. (13)

$$e(k) = u(k) - \hat{u}_{PD}(k) = u(k) - \left(\frac{1}{G_0} G(\mathbf{c}_k^H \cdot \mathbf{q}_k) \right) \quad (13)$$

being $G(\cdot)$ the PA response.

To minimize this cost function two algorithms are proposed: the Least Mean Square (LMS) algorithm and the Fast-Kalman Filter. Other algorithms such Least Squares or Recursive Least Squares are suitable for training the predistorter or for estimating the PA model, but result more computational complex if real-time (or quasi real-time) adaptation is pretended.

A. Least Mean Square Algorithm

The Least Mean Square algorithm is described in eq. (15).

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mu \cdot \mathbf{q}_k \cdot e(k)^* \quad (14)$$

being μ the error step (trade-off between speed of convergence and accuracy), and its bounds are:

$$0 < \mu < \frac{2}{\text{Tr}[\mathbf{q}_k \cdot \mathbf{q}_k^H]} \quad (15)$$

where $\text{Tr}[\cdot]$ denotes the trace of the data (signal) matrix.

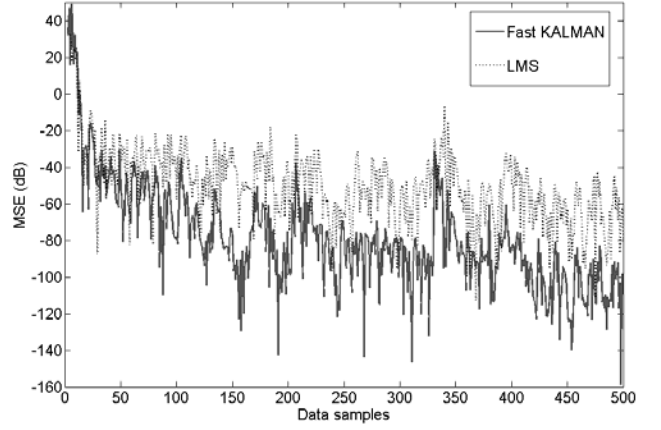


Fig. 4. Comparison between the Fast Kalman and the LMS algorithms in the predistorter adaptation process

B. Fast Kalman Filter

Fast-Kalman algorithms ([9]) use the optimum Kalman filtering technique to adaptively estimate the predistorter coefficients without the need of knowing *a priori* any transition matrix (unlike the conventional Kalman filter). The Fast-Kalman equations describing the algorithm are [9] the following:

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mathbf{h}_k \cdot e(k)^* \quad (16)$$

$$\mathbf{h}_k = \frac{\bar{\boldsymbol{\lambda}}_{k-1} \mathbf{q}_k}{\mathbf{q}_k^H \bar{\boldsymbol{\lambda}}_{k-1} \mathbf{q}_k + Q_M} \quad (17)$$

$$\bar{\boldsymbol{\lambda}}_k = \bar{\boldsymbol{\lambda}}_{k-1} - \mathbf{h}_k \mathbf{q}_k^H \bar{\boldsymbol{\lambda}}_{k-1} + \bar{Q}_p \quad (18)$$

where Q_M y \bar{Q}_p are related to the estimation and measure error variance [9]. The $\bar{\boldsymbol{\lambda}}_k$ matrix is recursively updated as defined in eq. (19).

Fig. 4 shows the mean square error (MSE) in dB of the Fast Kalman and the LMS algorithm in the adaptation process of the predistorter coefficients. It is possible to observe how the Fast Kalman algorithm provides a faster and more accurate estimation than the LMS algorithm. For that reason, in the following the Fast Kalman algorithm will be used to estimate and adapt the PD coefficients.

V. RESULTS

A. PA model estimation

The PA model has been identified using modulated input and output data of a class-A PA designed with the Agilent ATF-54143 PHEMT transistor. The PA has been estimated by means of the Narendra-Gallman (NG) algorithm [10], which is based in a Least Squares estimation. After some empirical simulations of the optimum order for identifying the PA (trade-off between accuracy and complexity), the Hammerstein model consists in a 7th order memoryless polynomial and a second order IIR filter, while the Wiener model is composed by a FIR filter with 15 delays and a 7th order memoryless polynomial.

Fig. 5 shows the MSE resulting from the PA model estimation for both Hammerstein and Wiener models.

The PA identification with the Hammerstein model provides better accuracy (MSE around -90 dB) in comparison to the Wiener model (MSE around -60 dB).

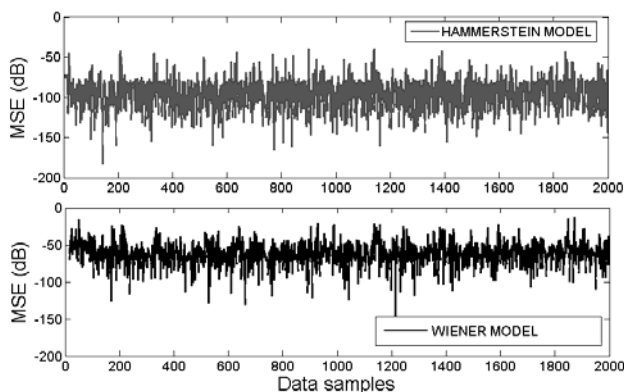


Fig. 5. MSE (dB) of a Hammerstein model ($P=7$, $N=2$ and $D=2$) and Wiener model ($P=7$, $N=15$) respectively in the identification of the PA.

As it can be observed in Fig. 6, the output spectrum of the 16-QAM modulated signal provided by the Hammerstein model resembles the real PA output power spectrum much more than the Wiener model does.

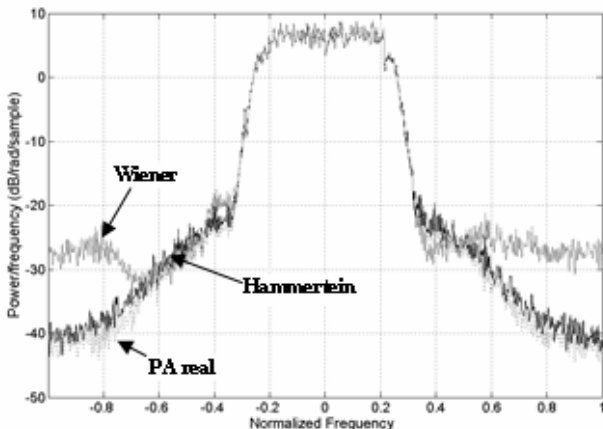


Fig. 6. Output power spectra of a 16-QAM modulated signal for: a real PA, a Hammerstein and a Wiener based PA model

B. Predistortion results

The 1-dB compression point of our particular PA is at 7.5 dBm of input power and 23 dBm of output power. A 256-QAM Gray constellation ordered modulation, filtered by a root raised cosine (RRC) filter with a roll-off factor of $\alpha=0.35$ has been considered in the predistorter adaptation process.

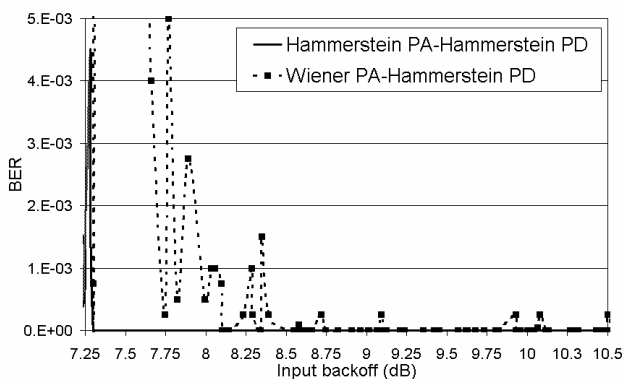


Fig. 7. BER for different IBO's considering both Wiener and Hammerstein PA models.

In order to see the performance achieved by the Hammerstein based predistorter when considering both PA models, Fig. 7 shows the BER at reception for both Wiener PA-Hammerstein PD and Hammerstein PA-Hammerstein PD combinations. An additive white Gaussian noise (AWGN) channel with SNR=10 dB has been considered. The PAPR of a 256-QAM RRC filtered ($\alpha=0.35$) is around 7.5 dB which means that an IBO bigger than 7.5 dB is necessary for having linear amplification. Fig. 7 shows how the Hammerstein PA-Hammerstein PD combination permits operating less backed off than Wiener one, what corresponds to a power efficiency improvement.

VI. CONCLUSION

This paper has presented an study on the suitability of using Wiener or Hammerstein models to identify the power amplifier nonlinear behavior taking into account memory effects. Results has shown that a PA identified using a Hammerstein model presents better performance than using a Wiener model, since it is easier to estimate, needs less parameters for the identification and achieves better results when predistorting the PA with a Hammerstein based predistorter. Finally, the fast Kalman algorithm has demonstrated quality skills to be considered for real time adaptation in memory limited embedded signal processors.

ACKNOWLEDGEMENT

This work was partially supported by Spanish Government (MCYT) under project TIC2002-04084-C03-01, by the EU network TARGET "Top Amplifier Research Group in a European Team" (IST-1-507893-NOE) and by the EU STREP CAPANINA "Communications from Aerial Platform Networks delivering 'Broadband for All'" (FP6-IST-2003-506745).

REFERENCES

- [1] Steve C. Cripps, "RF Power 2005", *Microwave Journal*, vol. 48, pp. 22-36, April 2005.
- [2] C. Euns, , and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture", *IEEE Tran. on Signal Processing*, vol. 45 , pp.223-227, Jan. 1997
- [3] A. Zhu, and T. J. Brazil, "An adaptive Volterra predistorter for the linearization of RF high power amplifiers", *MTT-IMS 2002*, vol. 1, pp. 461-464, 2002.
- [4] J. Kim, and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory", *Electronics Letters*, vol. 37, n. 23, pp: 1417-1418, Nov. 2001.
- [5] Y. Ding, H. Ohmori and A. Sano, "Adaptive Predistortion for High Power Amplifier with Linear Dynamics," *IEEE Int. Midwest Symp. on Circuits and Sys.*, pp.121-124, Japan, July 2004.
- [6] T. Wang and J. Ilow, "Compensation of Nonlinear Distortions with Memory Effects in Digital Transmitters," *IEEE Conf. On Comm. Networks and Services Research (CNSR'04)*, pp. 3-9, Canada, May 2004.
- [7] N. Benvenuto, F. Piazza, A. Uncini, "A neural network approach to data predistortion with memory in digital radio systems", *IEEE Int. Conf. Comm. ICC 93*, Vol 1 , pp. 232 - 236, Geneva, May 1993.
- [8] A. Hagenblad, "Aspects of the Identification of Wiener Models," PhD Thesis Division of Automatic Control, Department of Electrical Engineering Linköping universitet, Sweden, 1999.
- [9] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, N. J., 1991.
- [10] Narendra K. S. and Gallman P. G. "An iterative method for the identification of nonlinear systems using a Hammerstein model," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 546-550, July 1966.