Unconditionally Convergent pre-Distortion of non-Linear High Power Amplifiers

Josep Sala – Hugo Durney
Signal Processing Group
Department of Signal Theory and Communications
Universitat Politècnica de Catalunya
alvarez@gps.tsc.upc.es - durney@gps.tsc.upc.es

Abstract

The use of High Power Amplifiers (HPA) in digital communications is plagued with non-linear (AM/AM and AM/PM) distortion. This is specially acute for those modulations that are not constant amplitude, such as OFDM or MQAM. We present a pre-distortion algorithm that, in contrast to previously researched architectures, is unconditionally convergent.

1. Introduction

The identification of the transfer characteristic of non-linear systems has been widely dealt with in the literature. It is also interesting for some applications to provide methods for identification of the corresponding non-linear inverse transfer function. As such, Pre-distortion of High Power Amplifiers (HPA) for the transmission and relay of communication signals constitutes an active area of research[1] motivated by the utilization of non-constant amplitude modulations [1] [2]. Very frequently in transmission applications featuring HPA's, it is chosen either to utilize constant-amplitude modulations which are immune to the HPA non-linear distortion (as in satellite communications), or to operate the amplifier in back-off mode, thus reducing the transmission power budget. Alternative techniques rely on improving the peak to average power ratio of the modulation [3]. Non-constant amplitude modulations which are potentially advantageous in the context of trellis coding (such a M-QAM) or in the context of channel equalization (such as Orthogonal Frequency Division Multiplexing: OFDM), are not robust under non-linear distortion.

[1] This work has been financed by grants TIC98-0703, TIC99-0849 and TIC2000-1025 of the Spanish Science and Technology Commission (CICYT), by FIT-070000-2000-649 and by CIRIT's grant 2000SGR-00083 (Generalitat de Catalunya).

Adaptive pre-distortion algorithms prove a feasible alternative where this type of modulations is concerned. Nevertheless, the formulation of cost functions for the implementation of reliable pre-distortion algorithms must consider absolute convergence a key point of the design. Many pre-distortion cost functions are plagued by the appearance of local minima that result in a reduced pre-distortion capacity. In this paper we present a unimodal pre-distortion algorithm that guarantees absolute convergence.

The formulation of the pre-distortion algorithm will require first the characterization of the non-linear distortion characteristic. High Power Amplifiers introduce two different effects [4]: AM/AM and AM/PM distortion on the modulus and phase of the output pass-band signal. These non-linear characteristics are solely dependent on the modulus of the input pass-band signal.

Let us consider first the signal and HPA models in continuous time (t). Let $b_x(t)$ denote the complex base band equivalent of the real HPA input pass-band signal $x(t)$, with $x(t)=\text{Re}\{b_x(t)\exp(j\omega_c t)\}$, and ω_c the angular carrier frequency, and let $b_y(t)$ denote the complex base band equivalent of the real HPA output pass-band signal $y(t)$. Then, the HPA output is modeled as,

$$b_y(t) = b_x(t)\frac{\Psi(u_x)}{u_x}e^{j\Phi(u_x)} \qquad (1)$$

with $u_x(t) = |b_x(t)|$ and $\Psi(u_x)$ and $\Phi(u_x)$ the AM/AM and AM/PM curves, respectively. $g(u_x)= \Psi(u_x) / u_x$ is the modulus-dependent gain of the non-linear HPA. The role of pre-distortion is to introduce a controlled non-linearity that inverts that subsequently introduced by the HPA. The characteristic of the pre-distorter can be easily expressed from the AM/AM and AM/PM curves. Let $\Psi^{-1}(u)$ and $\Phi^{-1}(u)$ denote the inverse AM/AM and AM/PM curves, and let,

$$b_z(t) = u_z(t)\exp[\alpha_z(t)]$$
$$b_y(t) = u_y(t)\exp[\alpha_y(t)]$$

It then holds that,

$$u_y(t) = \Psi\left(u_x(t)\right)$$

$$\alpha_y(t) = \alpha_x(t) + \Phi\left(u_x(t)\right)$$

Hence, the equivalent pre-distortion characteristics $\Psi'(u)$ and $\Phi'(u)$ are given by,

$$\Psi'(u) = \Psi^{-1}(u)$$

$$\Phi'(u) = -\Phi\left(\Psi^{-1}(u)\right)$$

The following complex functions determine the AM/AM and AM/PM characteristics of the HPA and its equivalent pre-distorter, respectively,

$$A(u) = \Psi(u)e^{j\Phi(u)}$$

$$A'(u) = \Psi'(u)e^{j\Phi'(u)} = \Psi^{-1}(u)e^{-j\Phi\left(\Psi^{-1}(u)\right)}$$

We will seek to express $A'(u)$ in terms of moments of the input and output of the HPA. The phase term α_x is but a nuisance parameter that can be canceled with the following product,

$$b_y^*(t)\, b_x(t) = u_x(t)u_y(t)e^{j\Phi(u_x(t))}$$

and hence,

$$\begin{aligned} b_y^*(t)\, b_x(t) &= u_y(t)e^{-j\Phi(u_x(t))}u_x(t) \\ &= \Psi^{-1}\left(u_y(t)\right)e^{-j\Phi\left(\Psi^{-1}(u_y(t))\right)}u_y(t) \\ &= A'\left(u_y(t)\right)u_y(t) \end{aligned}$$

Whence we obtain the final direct relationship for the identification on the inverse non-linearity,

$$A'\left(u_y(t)\right) = e^{-j\alpha_y(t)}b_x(t) \qquad (2)$$

2. Pre-Distortion Algorithm

The pre-distortion algorithm is formulated in discrete time where we assume avaliability of the sampled versions of $b_x(t)$ and $b_y(t)$: $b_x(kT_s)$ and $b_y(kT_s)$, with $f_s = 1/T_s$ the sampling frequency. As both $b_x(kT_s)$ and $b_y(kT_s)$ are available[2], it becomes possible to obtain samples of the curve $A'(u)$ at distinct points. For subsequent operation, an estimate of $A'(u)$ valid for the whole dynamic range of the modulus is required. This can be achieved approximating $A'(u)$ in terms of a set of basis functions $\{f_n(u)\}$,

$$\hat{A}'(u) = \sum_{n=0}^{L-1} a'^*_n f_n(u) = \mathbf{a}'^H \mathbf{f}(u)$$

where each $f_n(u)$ will be henceforth called the n-th activation function and a'_n its associated coefficient. Both are grouped into vectors $\mathbf{f}(u)$ and $\mathbf{a}'$, respectively. The optimum set of coefficients is derived according to minimization of the following square error cost function,

$$J(\mathbf{a}') = \sum_{\kappa=0}^{N-1} K_\kappa \left| e^{-j\alpha_y(t_k)}b_x(t_k) - \mathbf{a}'^H \mathbf{f}\left(u_y(t_k)\right)\right|^2$$

with $K_k \geq 0$ any suitably defined kernel, and with $t_k = kT_s$. In the scope of this paper we constraint ourselves to $K_k=1/N$.

Equating to the **0**-vector the gradient of this cost function, $\nabla_{\mathbf{a}'^H} J(\mathbf{a}') = \mathbf{0}$, will provide a closed-form expression for the coefficients of the expansion. This is given by,

$$\nabla_{\mathbf{a}'^H} J(\mathbf{a}') = -\frac{1}{N}\sum_{k=0}^{N-1}\left(e^{-j\alpha_y(t_k)}b_x(t_k) - \mathbf{a}'^H \mathbf{f}\left(u_y(t_k)\right)\right)^* \mathbf{f}\left(u_y(t_k)\right)$$

and the optimum coefficient vector, $\mathbf{a}'_{opt}$, resulting from $\nabla_{\mathbf{a}'^H} J(\mathbf{a}') = \mathbf{0}$, becomes,

$$\mathbf{a}'_{opt} = \boldsymbol{M}^{-1}\sum_{k=0}^{N-1} e^{j\alpha_y(t_k)}b_x^*(t_k)\mathbf{f}\left(u_y(t_k)\right) \qquad (3)$$

$$\boldsymbol{M} = \sum_{k=0}^{N-1} \mathbf{f}\left(u_y(t_k)\right)\mathbf{f}^H\left(u_y(t_k)\right) \qquad (4)$$

We may define several types of activation functions. We should consider:

1. piece-wise linear (triangular) expansions.

2. positive/negative slope functions.

The positive/negative slope functions $f_n^{(0)}(u)$ and $f_n^{(1)}(u)$ are defined by,

$$f_n^{(1)}(u) = \begin{cases} 0 & , \quad 0 \leq u/\Delta \leq n \\ u/\Delta - n & , \quad n < u/\Delta < n+1 \\ 0 & , \quad n+1 \leq u/\Delta \end{cases}$$

$$f_n^{(0)}(u) = \begin{cases} 0 & , \quad 0 \leq u/\Delta \leq n \\ 1 - f_n^{(1)}(u) & , \quad n < u/\Delta < n+1 \\ 0 & , \quad n+1 \leq u/\Delta \end{cases}$$

[2] Note that due to the non-linear nature of distortion and pre-distortion, the involved signals are not band-limited. Thus, and for the sake of brevity, when we speak of sampled versions of the continuous-time signals, it should be understood that they are subject to residual aliasing. This residual effect, and also quantization noise, are negligible compared with the magnitude of non-linear distortion.

with Δ the length of each resolution bin. The triangular activation functions $f_n^{\Delta}(u)$ are defined by,

$$f_n^{\Delta}(u) = f_n^{(1)}(u) + f_{n+1}^{(0)}(u)$$

Both sets of activation functions appear depicted in figure (1). The structure of matrix ***M*** is tri-diagonal for triangular activation functions, diagonal for rectangular non-overlapping activation functions and block-diagonal for positive/negative slope activation functions (with 2 × 2 sub-blocks).

It will be possible in this context to work not on the output modulus but on an invertible function of it: $u'_y = \chi(u_y)$, the motivation being that matrix ***M*** and its inverse are highly sensitive to the probability distribution of the activated bins in **f**(*u*) and that the evaluation of the modulus requires a square-root which is not a trivial operation. In general, this function can be expressed as $u'_y = \chi(u_y) = \chi'(u_y^2)$, with u_y^2 much easier to compute than u_y and $\chi'(u_y^2)$ any simple function, not necessarily the square-root.

The solution $u'_y = \chi'(u_y^2) = u_y^2$, although simple is not valid if the bins activated by are equi-spaced. In this case the probability density function of u'_y is highly concentrated in the low region of the u_y^2 range so that the higher bins of u_y^2 are seldom activated and matrix ***M*** is worse conditioned. The approach is to use any simple compansion function similar to the square-root where the following cost function is minimized,

$$J_2(\mathbf{a}'') = \frac{1}{N}\sum_{k=0}^{N-1}\left| e^{-j\alpha_y(t_k)} b_x(t_k) - \mathbf{a}''^H \mathbf{f}\left(u'_y(t_k)\right)\right|^2$$

Hence the optimum coefficient vector $\mathbf{a}''_{opt}$ is now provided by,

$$\mathbf{a}''_{opt} = \left[\sum_{k=0}^{N-1} \mathbf{f}\left(u'_y(t_k)\right)\mathbf{f}^H\left(u'_y(t_k)\right)\right]^{-1} \sum_{k=0}^{N-1} e^{j\alpha_y(t_k)} b_x^*(t_k)\mathbf{f}\left(u'_y(t_k)\right)$$

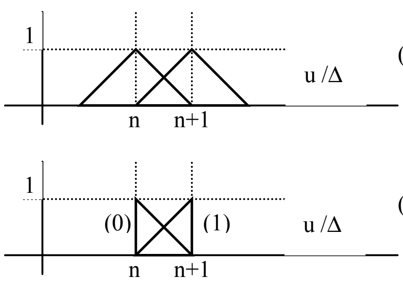

Figure 1: Types of piece-wise linear expansions, (a) triangular and (b) positive/negative slope activation function.

3. Coefficient Update Equations

Two options arise for the estimation of the pre-distortion coefficients: the application of the LMS algorithm or direct application of equation (3) in an open-loop scheme. The former is briefly described here, while the latter, based on diagonalization of ***M***, is exposed in the next section. Application of the LMS coeffcient update equations yields,

$$\mathbf{a}'_{k+1} = \mathbf{a}'_k - \mu\left[\nabla_{\mathbf{a}'^H} J(\mathbf{a}')\right]_{k,N=1}$$
$$= \mathbf{a}'_k + \mu\varepsilon_k^* \mathbf{f}\left(u_y(t_k)\right)$$

where the modeling error ε_k of (2) is defined as,

$$\varepsilon_k = e^{-j\alpha_y(t_k)} b_x(t_k) - \mathbf{a}'^H \mathbf{f}\left(u_y(t_k)\right)$$

It is also possible to consider modulus-dependent step sizes in terms of the modulus activation probability. Therefore,

$$\mathbf{a}'_{k+1} = \mathbf{a}'_k + \mathbf{m} \bullet \left(\varepsilon_k^* \mathbf{f}\left(u_y(t_k)\right)\right)$$

where $\mathbf{m}^T = [\mu_1, \mu_2, ..., \mu_k]$ is the modulus-dependent step size vector and • stands for the Schur-Hadamard (component-wise) vector product. If we set $\mathbf{m} = \mu\mathbf{1}$, with $\mathbf{1}^T = [1, 1, ..., 1]$ the all-ones vector, the rate of convergence of each individual coefficient in $\mathbf{a}'_k$ completely depends on the associated activation probability in $\mathbf{f}\left(u_y(t_k)\right)$. Smaller activation probabilities are related to slower convergence rates. The step-size vector can be made modulus-dependent and trade off rate of convergence with misadjustment.

4. Block Diagonalization of *M*

When positive/negative slope functions are used, we get that ***M*** defined in equation (4) can be expressed in block-diagonal form as,

$$M' = diag\left[M'_i\right]$$

$$M'_i = \begin{bmatrix} m'_{i,0,0} & m'_{i,0,1} \\ m'_{i,1,0} & m'_{i,1,1} \end{bmatrix}$$

$$m'_{i,p,q} = \sum_{k=0}^{N-1} f_i^{(p)}\left(u'_y(t_k)\right) f_i^{(q)}\left(u'_y(t_k)\right)$$

Thus, it is possible to obtain a low-complexity, closed-form expression for the optimum coefficient vector in equation (3). Fast acquisition of the pre-distortion coefficients is then possible via direct application of (3). Let $L=2L_0$ be the number of coefficients using positive/negative slope functions. Then the optimum coefficient vector $\mathbf{a}''_{opt}$ can be directly expressed as a concatenation of 2-component sub-vectors $\mathbf{a}''^T_{opt,i}$.

If we also define $\mathbf{f}(u)$ in terms of 2-component subvectors $\mathbf{f}_i^T(u)$ as,

$$\mathbf{a}''_{opt} = \left[\mathbf{a}''^T_{opt,0}, \mathbf{a}''^T_{opt,1}, \ldots, \mathbf{a}''^T_{opt,L_0-1}\right]^T$$

$$\mathbf{f}(u) = \left[\mathbf{f}_0^T(u), \mathbf{f}_1^T(u), \ldots, \mathbf{f}_{L_0-1}^T(u)\right]^T$$

we finally have the expresión for each $\mathbf{a}''_{opt,i}$ of $\mathbf{a}''_{opt}$,

$$\mathbf{a}''_{opt,i} = \boldsymbol{M}_i'^{-1} \sum_{k=0}^{N-1} e^{j\alpha_y(t_k)} b_x^*(t_k) \mathbf{f}_i\left(u'_y(t_k)\right) \qquad (5)$$

$$\boldsymbol{M}_i'^{-1} = \frac{1}{m'_{i,0,0} m'_{i,1,1} - m'^2_{i,0,1}} \begin{bmatrix} m'_{i,1,1} & -m'_{i,0,1} \\ -m'_{i,0,1} & m'_{i,0,0} \end{bmatrix}$$

Hence, the inversion of $\boldsymbol{M}'$ only requires as many inversions as activation bins.

4.1 Effects of the Modulus Probability Density

The goodness in the estimation of the pre-distortion coefficients relies on the implicit fact that all bins are activated a number of times significant to deliver reliable averages in equation (5). That is, the estimation of $\mathbf{a}''_{opt,i}$ in (5) is driven by the activation functions $\mathbf{f}_i\left(u'_y(t_k)\right)$. Therefore, low activation probabilities of the i-th bin result in noisy estimates of $\mathbf{a}''_{opt,i}$ due to poor averaging. Consequences of this are: (a) Given that the coefficients $\mathbf{a}''_{opt}$ are used to pre-distort the input signal to the HPA, it is desirable that the modulus range of $u'_y(t_k)$ be similar to the modulus range of $u_x(t_k)$, to avoid the situation where the input need be pre-distorted by $\hat{A}'(u)$ for a modulus $u_x(t_k)$ which never occurs in the range of $u'_y(t_k)$. (b) The best estimated coefficients are those pertaining to the most likely bins activated by $u'_y(t_k)$ (during estimation), which do not coincide with the most likely bins activated by $u_x(t_k)$ (during pre-distortion) as it would be desirable.

These apparent drawbacks can be easily circumvented as: (a) it is easy to adjust both modulus ranges to be the same with a very simple algorithm. (b) Further iterations of the pre-distortion procedure result in the coefficient bins being activated by the same modulus density function. That is, the more accurate the estimation of the pre-distorter is, the more similar the HPA output distribution is to the input, as linearization has taken place.

5. Results

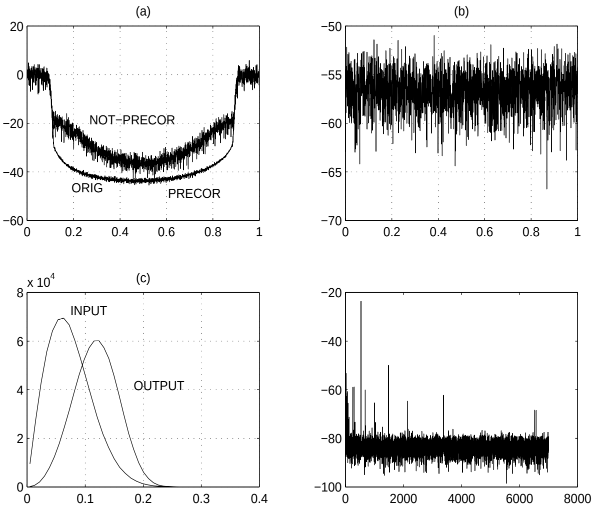

Figure 2: Pre-distortion results for a band-limited Gaussian process as input to the HPA using LMS. Upper-left: NOT-PRECOR, PRECOR and ORIG labels denote the output spectrum of the HPA without pre-distortion, the output spectrum with pre-distortion and the original input spectrum, respectively. Upper-right: normalized spectrum of error of HPA output with respect to perfectly linear HPA (with and without pre-distortion). Lower-left: Histograms of input / output of the HPA (without pre-distortion). Lower right: Time evolution of the adaptation error.

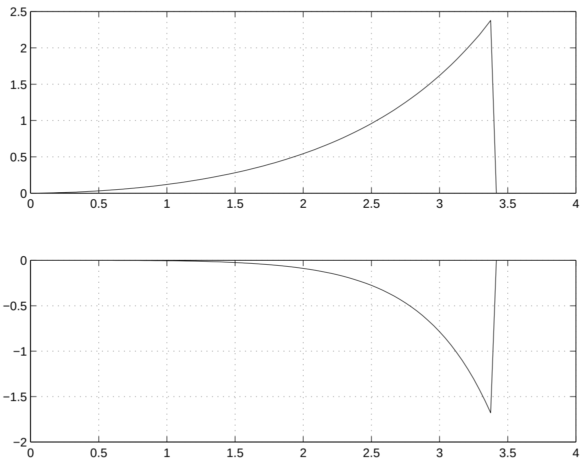

Figure 3: Estimated inverse AM/AM (above) and AM/PM\ (below) distortion curves to be used for pre-distortion. The inverse characteristic is only estimated up to that point where the output histogram used for training is negligible.

Figures (2) and (3) display the results obtained for a non-linear amplifier subject to AM/AM log-distortion and AM/PM quadratic distortion. The estimated pre-distortion characteristic appears in figure (3). The pre-distortion coefficients have been estimated using 33 triangular activation functions. When the block diagonalization approach is used, the minimum frame length is determined by the number of pre-distortion coefficients and the input and output histograms as shown in the lower-left plot in figure (2). If the number of coefficients is sufficiently large, some bins will be activated with very low probability and thus, some of the $\boldsymbol{M}_i$ submatrices defined in equation (5) will have been constructed with poor averages. The condition number of the $\boldsymbol{M}_i$ submatrices is approximately uniform in the likeliest region of the output histogram and tends to oscillate in the low probability density region of the output histogram. Although block diagonalization provides an open-loop way to estimate the pre-distortion coefficients based on positive / negative slope functions, usually, better results in terms of out-of-band residual non-linear distortion are obtained with application of the LMS algorithm, as only half of the coefficients are needed.

6. Conclusions

An unconditionally convergent pre-distortion algorithm has been shown.

Convergence properties are associated with the shape of the modulus probability density functions, there existing a parallelism with the eigenvalue spread of the data covariance matrix that governs the performance of adaptive filtering algorithms. A direct open-loop algorithm based on block-diagonalization of the activation function vector has been presented. An alternative using simpler gradient-based coefficient-update equations has also been presented.

References

[1] Martijn Kuipers, Ramjee Prasad, Pre-distorted Amplifiers for OFDM in Wireless Indoor Multimedia Communications, Vehicular Technology Conference VTC2000, pp. 395-399.

[2] M.Faulkner, M.Yohansson, Adaptive Linearization Using Pre-Distortion: Experimental Results, IEEE Trans. on Vehicular Technology, Vol. VT-43 no.2, pp. 323-332, May 1994.

[3] D.R.Gimlin, C.R.Patisaul, On Minimizing the Peak toAverage Power Ratio for the Sum of N Sinusoids, IEEE Transactions on Communications, Vol. 41 no.4, pp. 631-635, April 1993.

[4] P.Banelli, S.Cacopardi, F.Frescura, G.Reali, Counteraction of Non-Linear Distortion in a Novel MCM-DS-SS Wireless LAN Radio Subsystem, IEEE 0-7803-4198-8/97, pp.320-326, 1997.