DIRECT DIGITAL PREDISTORTION ON A COMPUTER CONTROLLED FPGA

Mark Sterling, Mark Bocko

University of Rochester Electrical and Computer Engineering Rochester NY, 14627 USA

ABSTRACT

We report on the development of direct digital radio frequency predistortion techniques for high power amplifiers and their demonstration in a MF band FPGA-based system. The basic component of our work is a nonlinear cubic voltage characteristic which is realized digitally by means of a look-up table. This method is situated in the context of existing baseband and analog predistortion techniques. In particular, we demonstrate our results experimentally and show a relationship between the bit-depth of our look-up table and the size of intermodulation distortions in the output of a class A amplifier.

Index Terms— Communication System Nonlinearities, Field programmable gate arrays, Microwave Power Amplifiers

1. INTRODUCTION

Modern linear modulation schemes place strict requirements on the linearity of High Power Amplifiers (HPAs). To maintain linearity with class A or class AB operation the input power may be backedoff so that the operating point is not close to saturation. This solution, however, can be very inefficient [1]. Alternatively, the signals involved (either the input or the output) may be altered in some way to compensate for the distortions introduced by the HPA and any other component in the forward link. We explore predistortion [2], which is a procedure for applying corrective distortions to the input signal.

Typical digital predistorters employ a *bandpass* [3] [4] representation of the transmitted signal and amplifier nonlinearity. In such systems, it is assumed that, before up-conversion (I/Q modulation) the input signal is stored as a series of complex numbers to which suitable nonlinear amplitude and phase characteristics are applied [5]. These correct for the two canonical nonlinear functions of a bandpass nonlinearity, respectively the AM/AM and AM/PM characteristics. Further research [6] has augmented this basic bandpass or baseband approach to include cases where frequency dependence and *memory effects* [7] are observed. The *memory polynomial* [8] is one example, noteworthy for the fact that it models the predistorter in a way that is linear in the parameters, which allows for the application of adaptive filtering techniques [9].

Often, wideband signals are the reason cited for the presence of memory effects. However, even in these cases, the quoted signal bandwidth is always still much smaller than the RF operating frequency. For instance, in [8] the signals are shown as having a bandwidth of 15 MHz compared with an RF carrier frequency of \approx 2 GHz. It is unclear how these baseband, or baseband with memory approaches could be feasibly employed when dealing with a very wideband signal, for instance one that stretches across multiple octaves. Of course, we can always write down a complex envelope representation of an arbitrary RF signal (that is, express it as an amplitude and phase modulated sinusoid) but it is unclear exactly what interpretation to attach to this complex envelope except in the narrowband cases. Another way of considering the problem is that baseband predistortion is primarily concerned with in-band distortion products that affect such distortion metrics as adjacent channel power ration (ACPR) [7]. In a standard two-tone test the lowest order in-band intermodulation products are $2f_2 - f_1$ and $-f_2 + 2f_1$. The distortion products at the third and higher harmonics are naturally ignored in baseband approaches or assumed to be attenuated by bandpass channel filters.

The design of certain analog predistorters does not presuppose any assumptions on the bandwidth of the input signal. The analog predistorter described in Roselli [10], for example, is a cubic nonlinearity meant to act instantaneously and directly upon the *RF voltage*. An analog cubic nonlinearity can be achieved by a diode network [11] or by an active circuit as in [10]. To be at all practical, the memoryless nonlinearity must be followed by a tunable filter that adjusts the phase and amplitude of any generated nonlinear components. Perfect *antiphase* between the nonlinear components of the predistorter and the amplifier attains the maximum suppression. Below, we give experimental results for exactly this type of test.

The system we present can be viewed as a combination of the two predistortion schemes mentioned above. On the one hand it is digital, consisting, at its core, of a look-up table routine and thus not dissimilar algorithmically from the baseband LUTs [12]. On the other hand, all of the signals involved and all operations are at RF frequencies. The predistorter and HPA are implicitly modeled by nonlinear block models such as the Wiener, Hammerstein, and Wiener-Hammerstein systems where all signals and coefficients are real-valued.

2. TEST PLATFORM

Our platform for implementing predistortion is an FPGA based DSP kit distributed jointly by Xilinx and Nallatech. The user FPGA on this kit is a Xilinx Virtex2-Pro30. The kit contains two independent on-board ADCs/DACs interfaced to the FPGA. The DACs operate at a sampling frequency of 105 msps with 14 bit resolution. Both the DACs and the ADCs have a voltage range of -1 to +1 volts with 50Ω output/input impedance. The kit provides a header which connects to free pins on the FPGA. This header allows us to send digital control signals (from the parallel port of a PC with Matlab) to the FPGA while it is running. All designs were written in VHDL and will be made available as a VHDL package.

Figure 1 shows the structure of our predistortion program. We

This research was supported by MDA under contract DASG60-02-P-0281 which was a Phase 2 STTR award to HYPRES Inc.



Fig. 1. Block diagram of the predistortion program for the DSP kit.

split the input signal into two branches which are later summed to create the predistorted output signal for the amplifier. In one branch, the input signal passes through unchanged. In the other branch, we use the 14 bit signal as an index into a look up table containing a cubic nonlinearity. In general, the LUT only uses N < 14 MSBs of the 14 bit input signal. That is, the signals are rounded. We later present experimental results showing how the size of the LUT (2^N) affects performance.

The output of LUT is then attenuated and time delayed by amounts set by the external control (PC). This allows us to tune the predistorter for the maximum reduction in distortion. The extra DAC/ADC channel gives a reference signal for the case with no predistortion.

3. NONLINEAR BLOCK MODELS

Here, we attempt to provide some theoretical background for our experiments. By substituting a series of three block models of increasing complexity for the HPA we hope to motivate the given form of our predistortion algorithm. We understand a block model as simply a system consisting of discrete combinations of memoryless nonlinearities and LTI systems. Throughout, $f = x + \gamma x^3$ is the predistorter nonlinearity and $g = x + \beta x^3$ is the nonlinearity occurring in the PA model. We examine a series of three PA models in order of increasing complexity, a memoryless nonlinearity, a Hammerstein System, and a Wiener-Hammerstein System.

The coefficients γ and β and all signals are assumed to be purely real. That is, the AM/PM distortions are considered negligible. We believe that phase distortions could be incorporated in the following discussions by generalizing to complex coefficients and making available the hilbert transform of the input signal. Furthermore, as alluded to above, examples of predistorters which do not account for amplitude-to-phase conversion-and which are, in fact, cuber circuitshave been reported in the literature [10].

3.1. Memoryless Nonlinearity

The first case is that of a memoryless nonlinearity, possibly followed by a gain term K. We include K because we assume that the linear gain of the device has been factored from g. The PA model has the following input/output relationship.

$$y(t) = K(g(x(t))) \tag{1}$$

Assuming a weak nonlinearity the linearization solution is immediate and well known. We are seeking to eliminate all nonlinear terms in the composite function $g \circ f = g(f(x(t)))$. Keeping terms of up to first order in γ or β this gives.

$$\gamma = -\beta \tag{2}$$

From this can be found a relationship between γ and the 1 dB gain compression point for $\gamma \in [0, -\beta]$. The relationship can be derived, for instance, from the results in [13].

$$A_{1\rm dB} = \left(\frac{10^{-\frac{1}{20}} - 1}{\frac{3}{4}(\gamma + \beta)}\right)^{\frac{1}{2}}$$
(3)

3.2. Hammerstein System

The Hammerstein system is a memoryless nonlinearity followed by a linear filter. An extra gain term is not necessary here because the linear gain of the device can be captured by H. The cascade of linearizer and PA model can be written as follows.

$$y(t) = H \cdot g(f(x(t))) \tag{4}$$

Since the nonlinearity g occurs at the input of the model, the linearization problem is essentially the same as in the previous case. The composite function g(f(x(t))) should be linear or close to linear over the range of the input. We suggest the Hammerstein model in cases where only a memoryless predistorter is available but the data—in the form of measured input and output waveforms—is suggestive of frequency dependence or memory.

Hammerstein system identification is well studied and can be performed by driving the DUT with either noise or a broadband communications signal and applying the Narendra-Gallman algorithm [14]. The kernel regression estimate [15] also provides a computationally efficient method to estimate the parameters of a Hammerstein system.

3.3. Wiener-Hammerstein System

The Wiener-Hammerstein system is a three block model consisting of a memoryless nonlinearity between two linear systems, also sometimes called a sandwich system. Again, we give the relationship for the cascade of predistorter and PA model.

$$z(t) = H_2 \cdot g(H_1 \cdot f(x(t))) \tag{5}$$

In this case, there is an extra operator H_1 interposed between f and g. A natural way to approach the problem of linearization might be to attempt to find a class of linear systems which commute with nonlinear functions.

$$[f,H] \equiv fH - Hf = 0 \tag{6}$$

Clearly, if the above relationship holds for H_1 then results from the previous two cases would carry over. However, the only linear systems which, in general, satisfy the condition for commutativity seem to be unity gain time delays.

If, instead of commutativity, H_1 has a linear phase response over some operating bandwidth Ω then we observe that the relative phases of all harmonics generated by f are maintained through H_1 . Thus, the antiphase condition (2) can still be met. That is to say, after H_1 we can always perform a time shift so that all components are, for instance, cosines. To be more precise, let Ω_k be the set of all ω for which the phase response of the system H_1 is $e^{j\omega k}$. Then, for some k, Ω is a subset of Ω_k such that if $\omega \in \Omega$ then $3\omega \in \Omega_k$. The only difficulty is that the potentially nonconstant magnitude response may result in the nonlinearity being undercompensated or overcompensated.

Consider the intermediate signal at the output of W-H nonlinearity,

$$y(t) = g(H_1 \cdot f(x(t))) \tag{7}$$

When $x(t) = A \cos(\omega t)$, y has the following form up to a shifting of the time coordinate,

$$y(t) = \xi_0 \cos(\omega t) + (\gamma \xi_1 + \beta \xi_2) \cos(\omega t) + (\gamma \xi_3 + \beta \gamma \xi_4) \cos(3\omega t)$$
(8)

$$\xi_{0} = A|H(\omega)|$$

$$\xi_{1} = A^{3}\frac{3}{4}|H(\omega)|$$

$$\xi_{2} = A^{3}\frac{3}{4}|H(\omega)|^{3}$$

$$\xi_{3} = A^{3}\frac{1}{4}|H(3\omega)|$$

$$\xi_{4} = A^{3}\frac{1}{4}|H(\omega)|^{3}$$

In the absence of the nonlinearities f and g the only component of y(t) would be $\xi_0 \cos(\omega t) = A|H(\omega)|\cos(\omega t)$. Therefore, for linearization, we attempt to retain this term while minimizing the remaining terms. We construct a functional which measures deviation from linear behavior for sine wave excitations. This is just the magnitude of the strictly nonlinear components averaged over Ω .

$$L(\gamma) = \int_{\Omega} d\omega |\gamma \xi_1 + \beta \xi_2| + |\gamma \xi_3 + \beta \xi_4|$$
(9)

To develop an analytic solution we exchange the magnitude operation with squaring.

$$L(\gamma) = \int_{\Omega} d\omega (\gamma \xi_1 + \beta \xi_2)^2 + (\gamma \xi_3 + \beta \xi_4)^2$$
(10)

The minimum can be found by Leibniz differentiation.

$$\frac{\partial}{\partial\gamma}L = \frac{\partial}{\partial\gamma}\int_{\Omega}d\omega(\gamma\xi_1 + \beta\xi_2)^2 + (\gamma\xi_3 + \beta\xi_4)^2 = 0 \quad (11)$$

The solution is,

$$\gamma = -\beta \frac{\int_{\Omega} d\omega \xi_1 \xi_2 + \xi_3 \xi_4}{\int_{\Omega} d\omega \xi_1^2 + \xi_3^2}$$
(12)

or

$$\gamma = -\beta \frac{\int_{\Omega} d\omega (\frac{3}{4})^2 |H(\omega)| \cdot |H(\omega)|^3 + (\frac{1}{4})^2 |H(3\omega)| \cdot |H(\omega)|^3}{\int_{\Omega} d\omega (\frac{3}{4} |H(\omega)|)^2 + (\frac{1}{4} |H(3\omega)|)^2}$$
(13)

If $|H(\omega)|$ is a constant then we observe the following three cases. When $|H(\omega)| = 1$ then the system is just a time delay, because we have assumed linear phase. When $|H(\omega)| > 1$ then, in 12, $\gamma > (-\beta)$. When $|H(\omega)| < 1$ then, in 12, $\gamma < (-\beta)$. A similar result is obtained in [11].



Fig. 2. A representative set of data from our experimental procedure. The plot shows overlapping plots of the power spectrum of the ZHL-32a output for the cases with (black) and without (light gray) predistortion. The two large spikes in the center are the input signal while all other components are noise or nonlinear distortion components. We note the existence of spurious signals in the output of the FPGA.

4. EXPERIMENT

The experimental procedure was to drive the amplifier with a twotone input (frequencies f_1 and f_2) and vary the phase and amplitude of the cubic correction term by input from a PC. We were also able to perform tests at multiple center frequencies and frequency differences ($f_2 - f_1$). The magnitudes of the IMD were first measured and gave a reference for each individual test. In all cases the frequency difference between the two input components was from 0.1 to 0.3 MHz and the input frequency was varied from 1.1 MHz to 1.7 MHz. The measurements were taken with a LeCroy Waverunner oscilloscope with fast Fourier transform capability. We present data for a frequency difference of 0.3 MHz.

The amplifier we tested was a Mini-Circuits ZHL-32A class-A amplifier terminated with a matched attenuator.

A sample output of an experimental run is shown in Figure 2. A summary of our data for multiple frequencies is shown in Figure 4. And finally, our data showing IMD power versus the number of bits in the LUT is shown in Figure 3. We also have similar data pertaining to the correction of the third harmonic in a single tone test.

5. CONCLUSIONS

As noted above, our FPGA implements a time delay by an integer number of samples in the cubed signal relative to the input signal before the two are added together to create the predistorted signal. At low frequencies or small bandwidths this time delay provides enough control over the phase of the cubic term to obtain adequate predistortion. However, in the general case, (especially when we want to predistort in the vicinity of the fundamental and in the vicinity of the third harmonic *simultaneously*) we need a finer control over the phase response in the cubic branch of our predistorter. The implementation of such a tunable phase shifting program is our next immediate area of research. Another avenue for research would be



Fig. 3. We show the size of the IM3 intermodulation products obtained for the optimal predistorter amplitude and phase settings as a function of the number of bits of resolution in the LUT. As before the $2f_2 - f_1$ components are represented by circles and the $2f_1 - f_2$ components are represented by triangles. Predistortion (black) No Predistortion (light gray).

to implement quintic and higher nonlinear LUTs in addition to the cubic nonlinearity.

A large part of our motivation for this work has been the possibility of realizing this type of predistortion in high speed superconducting electronics. While it may seem impractical to duplicate our results at 1 GHz, such speeds are not outside the projected range of superconductor based electronics [16]. However, it was only our goal to demonstrate the predistorter operating in principle.

6. REFERENCES

- [1] Stephen A. Maas, *Nonlinear Microwave and RF Circuits*, Artech House, second edition, 2003.
- [2] Wolfgang Bosch and Giuliano Gatti, "Measurement and simulation of memory effects in predistortion linearizers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-37, no. 12, pp. 1885–1890, 1989.
- [3] George L. Heiter, "Characterization of nonlinearities in microwave devices and systems," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-21, no. 12, pp. 797–805, 1973.
- [4] Adel A. M. Saleh, "Frequency-independent and frequencydependent nonlinear models of twt amplifiers," *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1715–1720, 1981.
- [5] S. P. Stapleton, "Amplifier linearization using adaptive digital predistortion," *Applied Microwave and Wireless*, vol. 13, pp. 72–77, 2001.
- [6] L. Ding, R. Raich, and G. T. Zhou, "A hammerstein predistortion linearization design based on the indirect learning architecture," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 2002, pp. 2689–2692.
- [7] J. H. K. Vuolevi, T. Rahkonen, and J. P. A. Manninen, "Measurement technique for characterizing memory effects in rf power amplifiers," *IEEE Trans. Microwave Theory Tech.*, vol. 49, no. 8, pp. 1383–1389, 2001.



Fig. 4. This figure provides a summary of the data we collected for two-tone tests at different carrier frequencies. This is a plot of the power of the 3rd order intermodulation products with and without predistortion. The $2f_2 - f_1$ components are represented by circles and the $2f_1 - f_2$ components are represented by triangles. As before the light gray data points are without predistortion and the black data points are with predistortion.

- [8] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417–1418, 2001.
- [9] Lei Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, 2004.
- [10] L. Roselli, V. Borgioni, V. Palazzari, and F. Alimenti, "An active cuber circuit for power amplifier analog predistortion," in *33rd European Microwave Conference*, 2003, pp. 1219–1222.
- [11] T. Nojima and T. Konno, "Cuber predistortion linearizer for relay equipment in 800 mhz band land mobile telephone system," *IEEE Trans. Veh. Technol.*, vol. VT-34, no. 4, pp. 169– 177, Nov. 1985.
- [12] Kathleen J. Muhonen, Moshen Kavehrad, and Rajeev Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: A development and comparison," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1995–2002, 2000.
- [13] C. Cho, W. R. Eisenstadt, B. Stengel, and E. Ferrer, "Ip3 estimation from the gain compression curve," *IEEE Trans. Microwave Theory Tech.*, vol. 53, no. 4, pp. 1197–1202, April 2006.
- [14] E. Eskinat, S. H. Johnson, and W. L. Luyben, "Use of hammerstein models in identification of nonlinear systems," *AIChE Journal*, vol. 37, no. 2, pp. 255–267, 1991.
- [15] E. A. Nadaraya, "On estimating regression," *Theory of Probability and its Applications*, vol. 9, pp. 141–142, 1964.
- [16] D. K. Brock, O. A. Mukhanov, and J. Rosa, "Superconductor digital rf for software radio," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 174–179, 2001.