ABSTRACT

*RF linearization and digital predistortion have been shown to be effective at counteracting the distortion due to the non-linear behavior of high-power amplifiers operating near the power-efficient saturation point. The low cost and simple implementation of digital predistortion has made it a highly desirable alternative to expensive and power hungry RF linearizers. A digital predistorter scales and rotates the input samples based on the samples’ magnitude. The adjustment coefficient depends on the amplifier nonlinearity. These coefficients can be obtained directly by measuring the amplifier characteristics with a network analyzer. The coefficients can also be arrived at by adaptively adjusting an initial estimate such that an error metric is minimized. The error metric is typically the difference between the transmitted and the received samples. The adaptation is accomplished using a standard gradient descent algorithm such as least mean squares (LMS). In this paper we present measurement results comparing the two approaches. The measurements were performed for a space-qualified traveling wave tube amplifier (TWTA). We show that the adaptive approach provides better frequency domain performance (shoulder reduction). However, the improvement in signal to noise ratio and bit error rate is almost identical for the two approaches.

INTRODUCTION

The high-power amplifier that drives the antenna of a wireless transmitter is typically responsible for the bulk of overall power consumption. This is especially true when the transmitter has to cover a wide area (e.g., a cellular base station) or great distances (e.g., a satellite downlink). An ideal amplifier is perfectly linear and does not distort the input signal. Unfortunately, any amplifier exhibits nonlinear behavior when operating close to its maximum output power level (saturation). These nonlinear effects become more severe closer to saturation, distort the transmitted signal, and eventually degrade the performance such that the link can no longer be maintained. Operating close to saturation achieves higher power efficiency by amortizing the fixed power dissipation over larger total power. Thus, a trade off exists between linearity and power efficiency. This trade off is especially severe for power amplifiers used in satellites since the cost of power is very high.

Signals with variable envelope (multilevel) experience more distortion than do constant-envelope signals. Multilevel constellations require a higher signal-to-noise ratio (SNR) to maintain a reliable link and therefore require higher transmitter power. At the same time, they are more sensitive to distortion. Combined, these considerations have prevented widespread use of multilevel modulations when high transmit power is required. Indeed, most satellite and cellular systems in use today still rely on the simple constant magnitude modulations such as QPSK.

Multilevel modulations are highly desirable because they allow more data to be transmitted using the same spectrum, and spectrum is a precious resource. For example, data throughput can be increased fourfold by using a four-level modulation such as 16-QAM instead of a two-level modulation such as BPSK.

Researchers have investigated methods to mitigate the effects of amplifier nonlinearity over the last 30 years [1]. As digital signal processing becomes more powerful and available, new techniques are finally allowing the use of high-order modulations with nonlinear amplifiers. Predistortion of the input signal by a function that approximates the inverse of the nonlinearity is the most popular linearization method. Predistortion can be done in either RF or digital domains. Most linearizers used today operate at RF, but advances in semiconductor technology have opened the door to linearization at baseband. Baseband predistortion is significantly more flexible and adaptable while at the same time reducing the size, weight and power of the linearized amplifier. Results presented in this paper deal only with baseband predistortion.

Baseband predistortion can be either fixed or adaptive. Adaptive predistortion is becoming rather popular [3][7][8], but is not always necessary or desirable. In this paper we discuss the pros and cons of adaptation for different scenarios: satellite, critical ground systems, and non-critical ground systems.

The remainder of the paper is organized as follows. We describe the problem to be solved – nature of the nonlinearity. Baseband predistortion is then introduced, and extended to adaptive predistortion. The measurement results follow, preceded by a short description of the testbed.
HIGH POWER AMPLIFIER

An ideal amplifier provides a linear relationship between the input and the output. A real amplifier starts exhibiting nonlinearity when operated close to its maximum output level, as shown schematically in Figure 1.

![Figure 1. Linear vs. non-linear amplifier](image)

The nonlinearity can be characterized into two types:

- **AM/AM**: An incremental change in input signal power does not correspond to a linear incremental change in the output power.
- **AM/PM**: An incremental change in input signal power results in a *phase* change of the output signal.

The AM/AM and AM/PM responses of the TWTA used in this research were measured using a vector signal analyzer and are presented in Figure 2.

![Figure 2. AM/{AM,PM} characteristics of the TWTA](image)

The effect of these nonlinearities can be observed in both time and frequency domains. The time domain distortion is visible in Figure 3, where an ideal 16-QAM is distorted to the point of undetectability. The frequency domain distortion is visible in Figure 4, where the tight ideal spectrum is broadened by the appearance of ‘shoulders.’ Thus, the performance of a predistorter must be evaluated on its efficacy in reducing the constellation distortion and the ‘shoulders.’

![Figure 3. SNR degradation due to nonlinearity](image)

![Figure 4. Spectrum degradation due to nonlinearity](image)

PREDISTORTION

Consider an amplifier with a nonlinear transfer function $U$. The basic idea of linearization is to apply a function $L$ to the input data $x(t)$ such that $L(a x(t))$ makes the amplifier appear linear. I.e. $U(L(a x(t))) = a x(t)$, where $a$ is a scalar and the relationship holds for all values of $x(t)$ such that $ax(t)$ is less than or equal to the maximum amplifier output. If $U$ is a memoryless\(^1\) function, the linearizer function is given by $L = U^{-1}$.

Of course, no such function $L$ can exist for all values of $x(t)$ since that would require the amplifier with infinite gain for inputs close to saturation. The goal of a practical linearizer is to maintain the linear relationship as close to saturation as possible. If the amplifier function $U$ is measured over the range of interest, the inverse function $L$ can

---

\(^1\) e.g. if the signal bandwidth is much lower than amplifier bandwidth
be obtained by simply exchanging the abscissa and the ordinate.

Linearization is typically performed at the transmitter, before the power amplifier. However, it is theoretically possible to move the linearization to the receiver such that \( L(U(x(t))) = ax(t) \). Linearization at the receiver is typically referred to as nonlinear equalization. It is also possible to split the linearization between transmitter and receiver such that each implements part of the functionality [5][2].

**PREDISTORTER IMPLEMENTATION**

The predistorter is implemented entirely in the digital domain. A programmable lookup table (LUT) provides the sampled version of the inverse nonlinearity function, \( L \). The table values are stored as complex Cartesian numbers. Note that the polar representation adopted by other researchers is equivalent. A block diagram of the predistorter is shown in Figure 5. A value from the LUT is selected based on the magnitude of the incoming sample. Overall hardware complexity of the predistorter is negligible in today’s ASIC processes and uses only 2% of a modern FPGA.

The magnitude of the modulated signal is used for addressing the LUT. The LUT entries can be viewed as correction factors that are multiplied by the original signal to obtain the predistorted version. Entries beyond a certain point in the LUT, defined as the saturation limit, are viewed as ‘beyond saturation’ and result in predistorted signals whose phase and magnitude are the same. This behavior is displayed graphically in Figure 6 where increases in the transmit amplitude beyond the saturation limit of 1.0, result in a received amplitude that is limited to 1.0. Note that the output level of the predistorter at the saturation limit must be calibrated to coincide with the actual saturation power of the amplifier (0 dB backoff).

Sweeping the magnitude of the predistorter input and measuring the output of the power amplifier can verify the basic functionality of the predistorter. The power amplifier output is considered at baseband and is characterized by amplitude and phase. An ideal predistorter generates a perfectly linear amplitude ramp (up to saturation) and a constant phase. The result of this measurement, presented in Figure 7, clearly demonstrates that the predistorter is functional and is well tuned to the amplifier nonlinearity.

**COMPUTING THE PREDISTORTER LUT**

As described in the previous section, the predistorter is a simple circuit\(^2\). The only difficulty lies in determining the values for the LUT. The two options are (shown in Figure 8):

- Measure the nonlinearity using appropriate test equipment and load the data into the LUT
- Use an adaptive algorithm (e.g. LMS) to update the values in the LUT such that the difference between the amplifier input and output is minimized.

\(^2\)The implementation is nontrivial for high bandwidth signals since the predistortion must be done on an oversampled (>4x) signal [9].
Most modern linearizers use feedback for adaptive optimization [7]. Adaptation is desirable since the amplifier transfer function changes with time and temperature. However, the rate of change is typically rather slow under steady-state operation.

The LUT values can be adapted using a simple LMS algorithm, as described in [10] and below. Define the desired transmitted signal, $x(n)$, and the down-converted post-amplifier $y(n)$.

$$x(n) = \rho_{in} e^{j\theta_{in}}$$, input vector

$$y(n) = \rho_{out} e^{j\theta_{out}}$$, output vector

The adaptation equation for the LUT magnitude and phase is then given by:

$$\rho_{pd}(j+1) = \rho_{pd}(j) + \mu_{\rho}\left(\rho_{in} - \rho_{out}\right)$$

$$\theta_{pd}(j+1) = \theta_{pd}(j) + \mu_{\theta}\left(\theta_{in} - \theta_{out}\right)$$

where $K$ is the average gain from $x(n)$ to $y(n)$.

**FIXED OR ADAPTIVE**

Each of the approaches described in the previous section has good arguments both for and against its use. Each argument carriers more or less weight depending on the operating scenario. These scenarios will be considered at the end of the paper.

**Arguments for adaptive predistortion**

- Optimal predistortion as the amplifier characteristics change. Supports wide temperature swings. Supports degradation over time.

- Adapts for non-idealities that may not be observable with standard test equipment (e.g. memory effects).

- Does not require any test or training patterns. Can run continuously while the amplifier is transmitting user data. No disruption of service.

**Arguments against adaptive predistortion**

- Possible catastrophic failure if the adaptation algorithm fails (due to either design error or software upset due to radiation). It is difficult to prove stability in case of anomalous amplifier behavior. Minor oscillations in output power can lead to divergence.

- A power coupler is required at the amplifier output to provide the feedback signal. The coupler may be difficult to manufacture for very high output powers (10s kW). The coupler also introduces some (small) signal attenuation.

- A downconversion chain and a digitizer are required to convert the RF waveform to baseband input to the adaptation algorithm. These components must be of very high quality. The adaptation algorithm will attempt to compensate for any distortion introduced by these components and may not accurately compensate the amplifier. These components also add to cost and weight of the system.

**Arguments for measurement based predistortion**

- Very high precision can be achieved using modern test equipment. The quality of the test equipment front end will likely exceed that of the downconversion chain discussed above.

- The same test equipment can be used to measure many amplifiers, amortizing the cost.

**Arguments against measurement based predistortion**

- The amplifier must be taken off-line to perform any updates. Not practical in many scenarios (e.g. satellite)

**PERFORMANCE MEASUREMENTS**

There’s no substitute for measured data when the processes being studied are nonlinear. Simulation studies must be backed up by measurement, especially when the cost of an error is very high. The testbed, shown schematically in Figure 9 and photographed in Figure 10, uses a space-qualified TWTA with calibrated attenuators and noise source to provide a high fidelity model of a fixed satellite channel [6].
A set of experiments was conducted to determine the bit error rate (BER) performance of a 16-QAM waveform transmitted at different output power levels. The signal spectrum was also evaluated at each output level. Each experiment was repeated three times:

1. Without any linearization. Note that the BER performance is so poor that it will not be shown in the following figures.
2. Using a fixed predistorter.
3. Using an adaptive predistorter. The predistorter was allowed to adapt until the residual error stopped improving.

The BER results are shown in Figure 11 for uncoded 16-QAM. The key observation is that BER performance is very similar for fixed and adaptive. The adaptive predistortion slightly outperforms fixed predistortion for points over 2.5 dB from saturation (dBsat). The difference is appreciable only for low BER values and will likely be negligible if coding is employed.

Note that these BER curves do not represent the best possible results for 16-QAM transmitted through a nonlinear amplifier. Significant improvements can be achieved if additional mitigation techniques are employed at the receiver [2][5].
The next set of experiments deals with the spectral regrowth (shoulders). A spectral mask is typically imposed on the wireless link. Transmitted power may not exceed the mask at a given frequency offset from the carrier. This constraint can be violated by the shoulders and typically limits how close to saturation the amplifier can be operated.

Instead of using any specific frequency mask, we define three different metrics to quantify the spectral regrowth due to nonlinearity:

- Attenuation from the peak to a specific frequency offset from the carrier (see distance A in Figure 12). The frequency offset is specified relative to the symbol rate and is arbitrarily chosen to be $1.7F_b$ for this paper. This data is plotted in Figure 13.

- Ratio of in-band power to out-of-band power. Out of band power is defined as all power at frequencies outside $1.7F_b$, while in-band power is inside $1.7F_b$ (see areas B and C in Figure 12). This data is plotted in Figure 14.

- 99% bandwidth. This standard metric is defined by the bandwidth that contains 99% of all transmitted power. This data is plotted in Figure 15.

As can be seen from these figures, the fixed and adaptive predistortion approaches provide about the same level of improvement when the amplifier is operated close to saturation. When the amplifier is backed-off by more than 3 dB, the adaptive approach is significantly better. Adaptive predistortion lowers the shoulder by over 6 dB relative to fixed predistortion. Likewise, the out of band power is reduced by almost an order of magnitude. Note that the 99% bandwidth utilization is almost the same for both approaches since the power falls off very quickly outside the main lobe.

These results are not unexpected. The predistorter is limited in what it can do close to saturation. The peak-to-average power ratio for a pulse shaped 16-QAM is about 6 dB. Operating closer than 3 dB to saturation makes significant distortion inevitable. However, further from saturation, the adaptive predistorter compensates for subtle nonlinearity variations that are not captured by a continuous wave sweep used for the fixed measurement. It is interesting to note that the improvement in spectral regrowth is a lot more significant than the improvement in the BER, which is about 0.5 dB.
CONCLUSION

The measurement results presented in this paper conclusively show the value of predistortion. The choice between fixed and adaptive predistortion is less clear. Adaptive predistortion does result in reduced spectral regrowth, up to 10 dB better than fixed. However, the effect is significant only when operating more than 3 dB from saturation. The results in [2] show that a coded 16-QAM link can be operated as close as 2 dB from saturation even with a high rate code. An adaptive predistorter offers almost no performance improvement at that point.

Note that all the pro’s and con’s discussed above apply even if the performance of adaptive and fixed are identical. We can now consider a few scenarios where each approach is preferable.

Scenario 1: Satellite transponder. A satellite transponder is expected to provide 100% uptime. It may be required to operate continuously for decades while undergoing large temperature variations between sunlight and darkness. These three considerations make an adaptive predistorter attractive. The adaptation is typically done on the satellite, but may also be done on the ground. The ground-based adaptation suffers from the significantly attenuated received signal and much lower SNR. However, long averaging times and small adaptation constants can overcome the decreased SNR. Note that by putting adaptation on the ground, the satellite transmitter uses an essentially fixed predistorter.

Scenario 2: Ultra high power ground station for satellite uplink. Amplifiers with 10’s of kW of power are not unreasonable for ground-based systems. Such an amplifier may be used only intermittently to send commands to a satellite. The temperature variation on the ground is negligible and aging effects are not as critical since the amplifier can be replaced if it starts degrading. These considerations motivate a fixed predistorter. The amplifier characteristics can be measured at regular maintenance sessions and the LUT table can be updated at the same time. The increased cost and risk associated with putting an adaptive algorithm inline with the amplifier is not justified.

Scenario 3: A cellular basestation. These basestations are typically outdoor units and may experience significant temperature swings. They are expected to operate 99.99% of the time, but the cost of an outage is not dramatic. Unlike the previous scenarios, this one is cost sensitive and therefore uses lower-quality, less linear amplifiers. An adaptive predistorter is an obvious choice for this scenario.

Adaptive predistortion is also called for if the amplifier exhibits memory effects. That is the case if the signal bandwidth is large relative to the bandwidth of the amplifier.

REFERENCES


