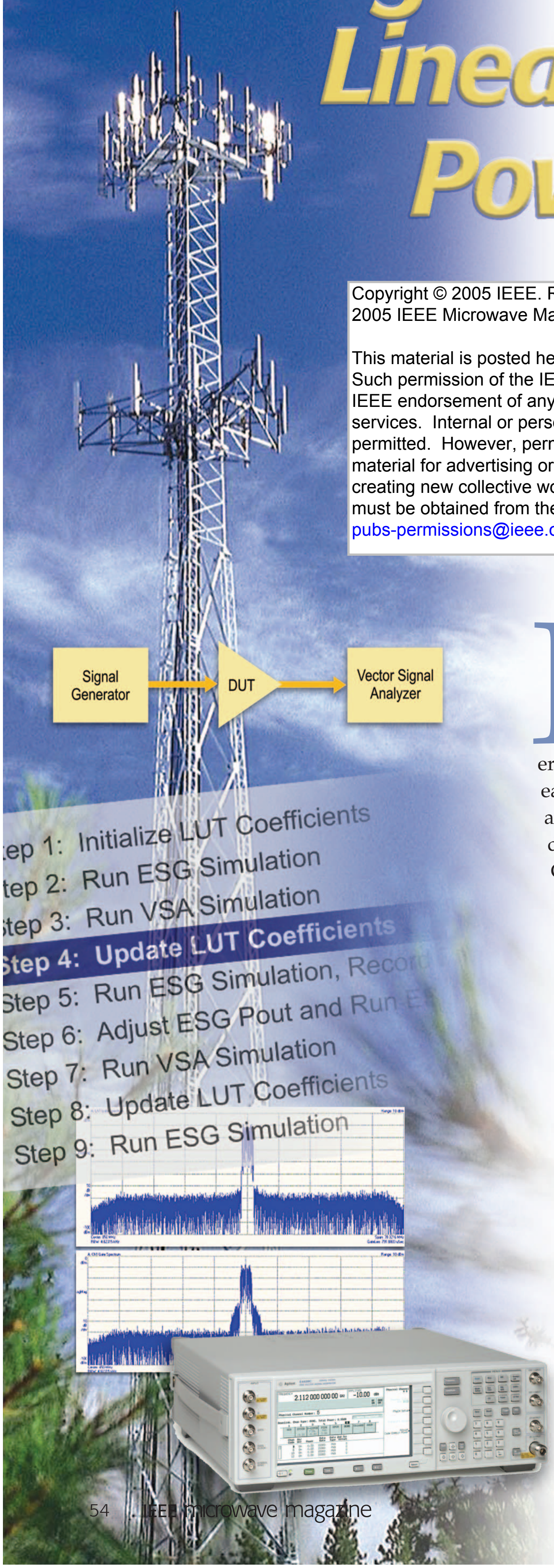
Digital Predistortion Linearizes Wireless Power Amplifiers

Wan-Jong Kim, Shawn P. Stapleton, Jong Heon Kim, and Cory Edelman



Reliable cellular service depends upon clean, consistent transmission from base stations under widely and rapidly changing conditions. The base station's RF power amplifiers (PAs) are key to guaranteeing this reliability. Spectral efficiency has always been important in mobile communications, but modern second- and third-generation digital systems now demand that linearity and efficiency of PAs are crucial performance requirements. Such amplifiers are found in cellular base stations that support any of the code division, multiple access (CDMA) family of wireless standards, such as CDMA2000, the Third Generation Partnership Project (3GPP), wideband CDMA (W-CDMA), etc., as well as improvements to existing standards, such as Enhanced Data Rates for GSM Evolution (EDGE). In many of these applications, due to the use of quadrature modulation and multiple carriers, the signal's power varies or fluctuates significantly over time, as compared with analog frequency modulation (FM) or Gaussian minimum shift keying (GMSK) modulation as used in global system for mobile communications (GSM).

Although all of the above-mentioned systems have good spectral efficiency, the signal's varying envelope generates intermodulation distortion (IMD) when amplified. Most of the IMD power appears as interference between adjacent channels, which requires the use of highly linear amplifiers. However, PAs typically operate at their best linearity and efficiency over only a narrow range of power. It is, therefore, difficult to find a given operating point or "back-off" for a given amplifier and modulation scheme that is both efficient and

Wan-Jong Kim and Shawn P. Stapleton (Shawn_Stapleton@telus.net) are with Simon Fraser University in Burnaby, British Columbia, Canada. Jong Heon Kim is with the Kwangwoon University, in Seoul, Korea. Cory Edelman is with Agilent Technologies, Inc., in Westlake Village, California, USA.

1527-3342/05/$20.00©2005 IEEE

free of distortion products, e.g., linear operation. Also, backing-off the power into an amplifier lowers power efficiency and increases heat dissipation, both of which are highly undesirable effects in the confined space of a base-station shelter. To further complicate things, note that a CDMA-type base station's signal will change depending upon the instantaneous number of users in the cell. In practice, linearity can be improved at the expense of efficiency or through the use of linearization techniques. In this article, we will present a detailed look at one such approach known as digital predistortion (DP). This will be implemented, for our purposes, using electronic design automation (EDA) sofware connected to standard laboratory test equipment. Let us first review some of the available linearization techniques and compare them.

Comparing Linearization Techniques

The feedforward linearization technique (Figure 1) is well known for providing good linearity but with poor efficiency and an undesirable reliance on complex, expensive, and potentially difficult-to-maintain analog hardware. Feedforward works by sampling the amplifier's output and reducing it to the same level as the input signal, then subtracting it from the input, leaving only the distortion generated by the amplifier. The distortion signal is increased with a separate amplifier so that it has the same level as the main output and is then subtracted from the original amplifier output signal. The result is a linearly amplified version of the input signal.

Adding adaptation can further enhance the effectiveness of feedforward. Adaptation may use pilot tones to control the signal cancellation circuit so as to minimize the power of the error signal in the frequency band occupied by the distortion. Other, more complex types of adaptation are based on continuously computing the gradient of the three-dimensional power "surface" for both the signal and error cancellation signals, or by measuring the adjacent-channel power ratio (ACPR) and attempting to minimize it [1]. All of these techniques are complex and depend on hardware that is prone to changing characteristics over time and temperature, making practical implementation and calibration difficult and expensive.

In addition to feedforward, with or without adaptation, several other linearization methods have been used. These include analog predistortion (PD), linear amplification using nonlinear components (LINC), and Cartesian feedback. As with feedforward and its variants, these techniques involve a considerable amount of added analog hardware or require the use of nonlinear components whose characteristics are difficult to control to the degree of precision necessary to achieve the desired improvements.

The most promising linearization technique is adaptive baseband DP [2], [3], in which the adaptation mechanism is based on the difference between the desired modulation and the PA's output. DP involves some complexity in terms of digital signal processing (DSP), but compared to analog-based linearization is relatively simple in terms of how it is interfaced to the PA itself. In a DP linearizer, DSP is often implemented in an application-specific integrated circuit (ASIC), and the only additional analog hardware required over and above the amplifier includes a coupler to sample the output signal and a complex multiplier to apply the

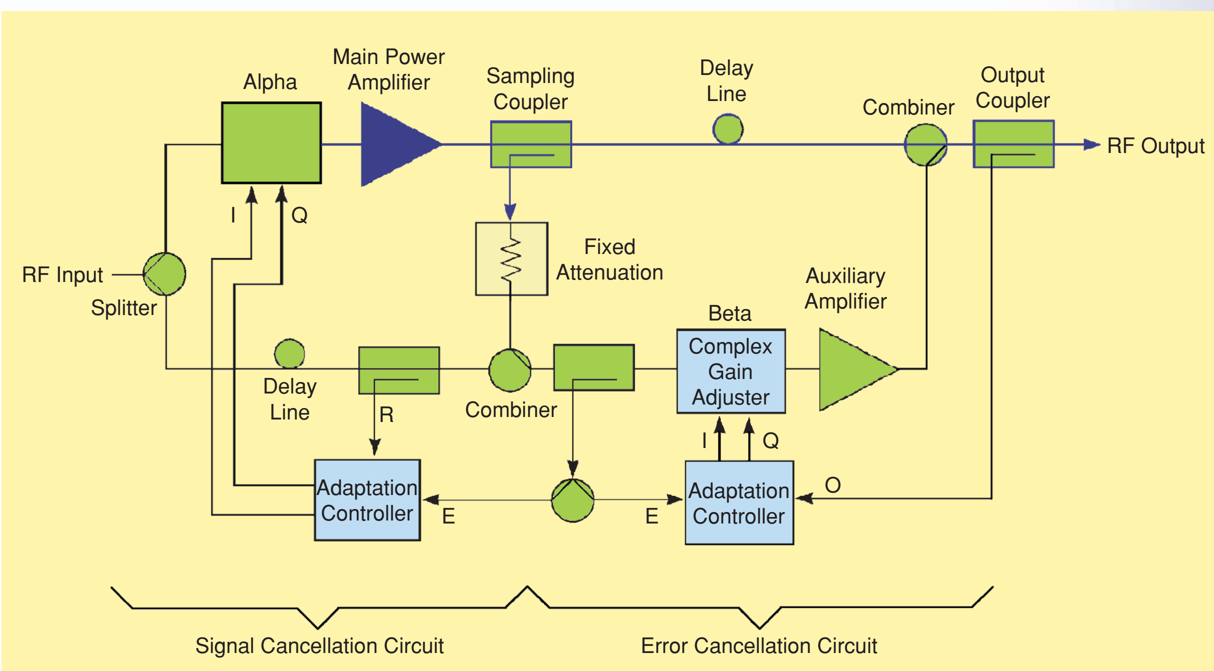


Figure 1. *A detailed block diagram of an adaptive feedforward linearization circuit.*

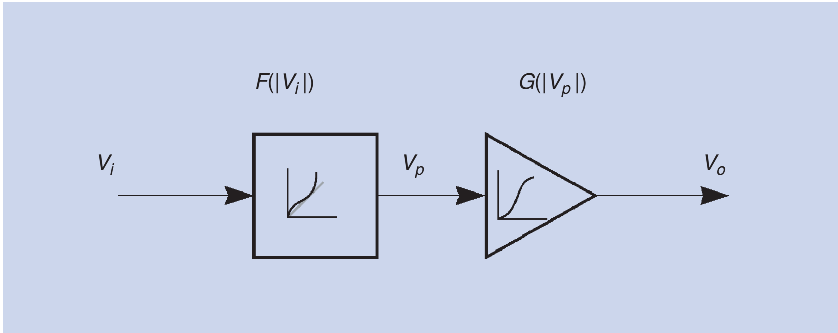


Figure 2. *A digital predistorter followed by a PA.*

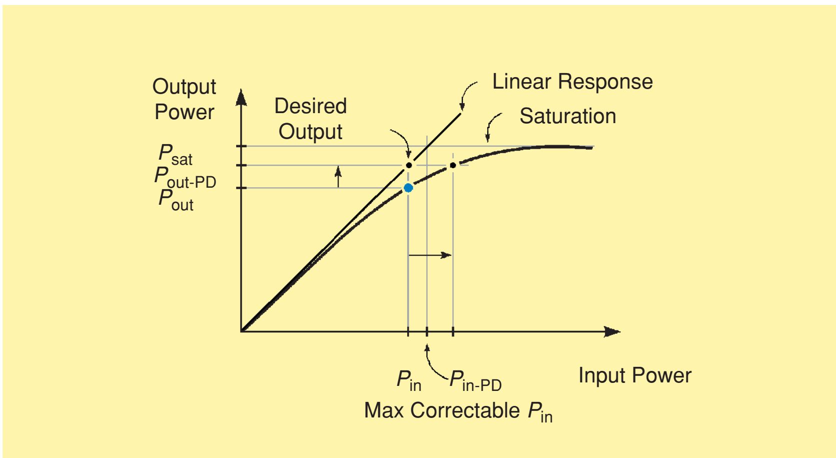


Figure 3. *PA P_{out} versus P_{in} and DP.*

PD. As a result, DP is less costly to implement than feedforward. In addition, the effectiveness of DP does not, unlike the feedforward or analog PD techniques, depend on the operating frequency of the system. It is, therefore, particularly well-suited for software-defined radio applications, in which a single set of RF hardware is used to cover a wide range of frequencies.

This article describes how by using an EDA system, specifically the linearization design-guide under advanced design system (ADS), an adaptive DP system has been developed and implemented to evaluate the performance of DP on a real PA. The simulation system includes the ability to perform the adaptive DP algorithms and to interface with an electronic-signal generator (ESG) and a vector-signal analyzer (VSA) so that the performance of the DP system can be evaluated using an actual high PA, similar to one typically used in a base station.

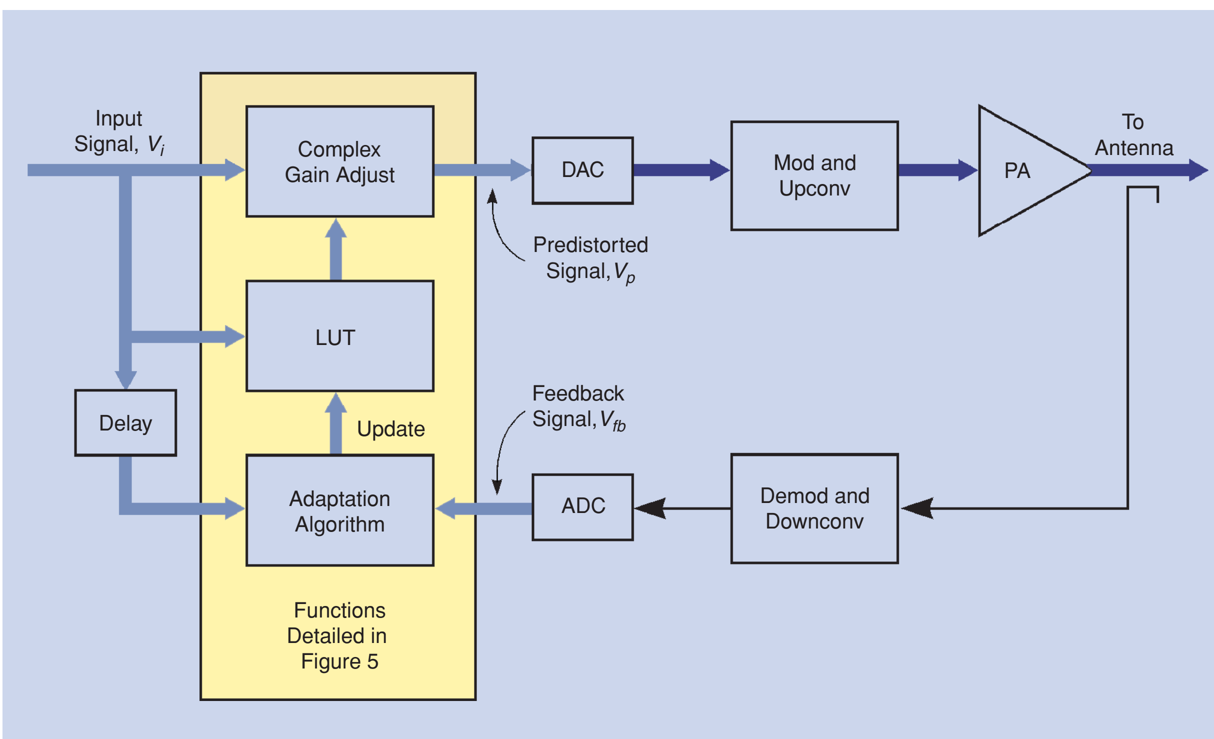


Figure 4. *DP block diagram.*

Implementing the Adaptive DP System

The intent of DP is to linearize the nonlinear response of a PA over an operating region. DP employs DSP techniques to predistort a baseband signal prior to modulation, up-conversion, and amplification by the PA. As a result, the cascade of the DP response and the PA response produces the desired linear response. Figure 2 shows the simplified block diagram. The gain, $\boldsymbol{G}$, of the PA is modeled as a function of the magnitude of the PA input signal, V_{p}. The function $\boldsymbol{G}$ is, in this case, memoryless and nonlinear in amplitude and phase. The use of a memoryless model that is only dependent upon input signal magnitude is a simplification of the actual response of a typical PA. Other variables will impact the PA response, including, most notably, frequency and instantaneous operating temperature.

Similar to $\boldsymbol{G}$, $\boldsymbol{F}$, the transfer function of the PD circuit in Figure 2, is made to be a function of the magnitude of the input signal to the DP, namely, V_{i}. Thus, the cascade of the predistorter and amplifier will result in the desired linear response when $\boldsymbol{F}(|V_{\mathrm{i}}|)\,\boldsymbol{G}(|V_{\mathrm{p}}|) = k$, where k is a constant and $V_{\mathrm{p}} = V_{\mathrm{i}}\,\boldsymbol{F}(|V_{\mathrm{i}}|)$. The manner in which DP works can be understood from Figure 3, which illustrates the typical relationship between the input power and output power of a PA. The heavier dark curve shows that, in the absence of DP, the PA's P_{out} versus P_{in} curve is highly nonlinear. However, through the introduction of DP, the P_{out} versus P_{in} curve is made to have a linear response over a large range of input power levels. The desired linear response of the PA is illustrated by the *linear-output* curve. The slope of the linear-output curve is the desired linear gain of the PA. When the amplifier is operating in compression, the P_{out} versus P_{in} curve falls below the linear-output curve; hence, the actual output power of the PA is not sufficient for linear operation. The inclusion of PD prior to the PA has the effect of introducing expansion; the amplitude of the input signal is increased so that the desired output power (falling on the linear-output curve) is achieved. The expansion effect of DP can be observed in Figure 3 where the input power, P_{in} (resulting in P_{out} before PD), is increased to $P_{\mathrm{in-PD}}$, so that the PA output power is raised to $P_{\mathrm{out-PD}}$, which coincides with the linear-output curve. The region of the P_{out} versus P_{in} curve that can be linearized using DP is limited.

A block diagram of an adaptive DP system is shown in Figure 4. With the inclusion of DP, the digital complex baseband input signal samples are multiplied prior to the digital-to-analog converter by complex coefficients drawn from the look-up table (LUT). The LUT coefficients implement the PD function. The

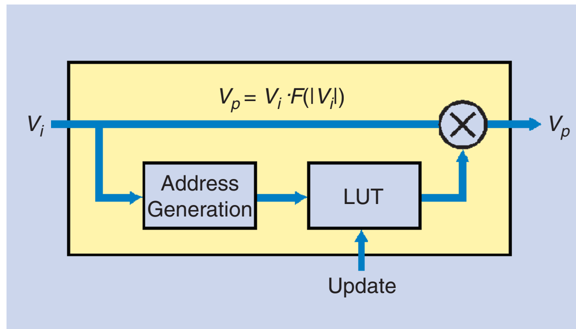


Figure 5. *A complex gain adjust and LUT, corresponding to the blocks highlighted in Figure 4.*

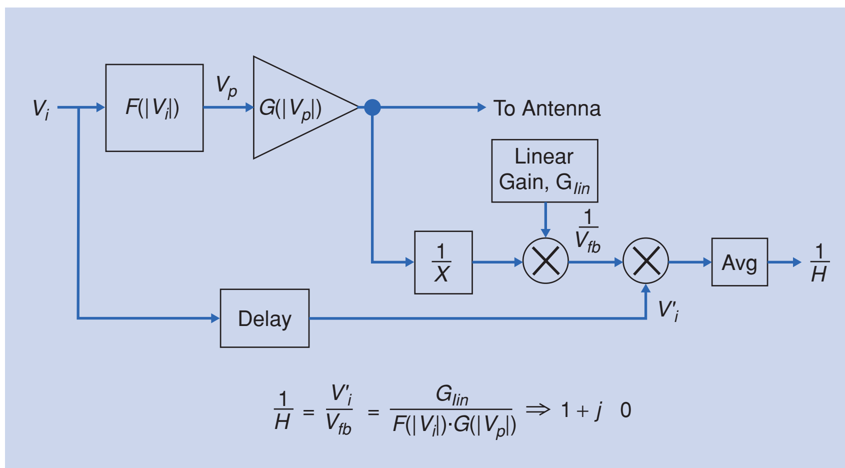


Figure 6. *Implementation of the adaptation algorithm.*

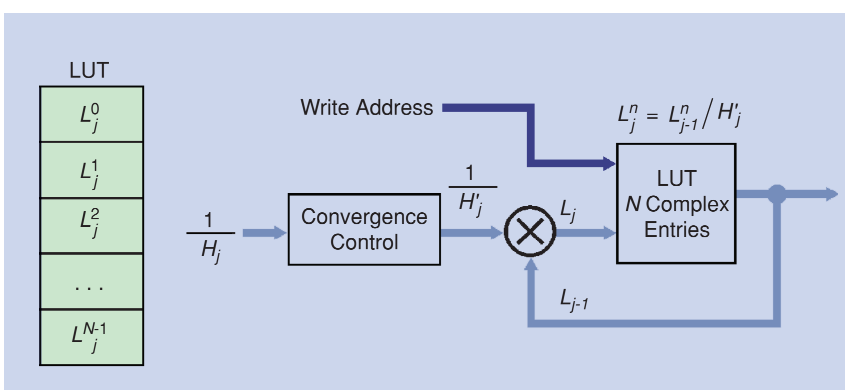


Figure 7. *Calculation of new PD function.*

adaptation algorithm determines the values of the coefficients by comparing the feedback signal and a delayed version of the input signal. The complex gain adjustment and LUT blocks of Figure 4 are shown in greater detail in Figure 5. The PD function is implemented using a complex multiplier, a LUT, and an address-generation block that selects the appropriate coefficient from the LUT, given the magnitude of the input signal.

The size of the LUT employed determines the number of points at which the PD function is calculated. In addition, the distribution of the PD function points need not necessarily be evenly distributed across the range of the input signal magnitude. Instead, it may be desirable to distribute the PD function points across the range of the input signal magnitude using a squared (power) or logarithmic relationship [4], [5].

The function of the adaptation algorithm is to derive the PD function ***F***, i.e., the inverse characteristic of the amplifier response. The PD function may be derived using either a modulated signal input (random signal) or a known training signal. The adaptation algorithm and its implementation are fundamentally different depending upon which type of input signal is used for training: an unmodulated carrier or the modulated signal that corresponds to the intended communications standard. The algorithms that are based upon the use of a modulated signal employ statistical signal processing and typically require some type of curve-fitting algorithm to generate a smooth PD function. The complexity of the adaptation algorithm and its implementation can be significantly simplified by using a known training signal such as an unmodulated carrier that is linearly ramped upward [6].

The adaptation algorithm chosen for the implementation described herein is based upon the use of a training signal. The training signal is a single tone having a frequency equal to the carrier frequency and whose power is ramped upward over the duration of the training period. The power of the tone is set to zero at the start of the training period and will typically peak at, or just below, the maximum correctable input power of the amplifier. The training signal used here is generated by a simulator. In implemented hardware, it would be typically created by a fixed-frequency generator, because the training signal is simply a tone ramped upward in amplitude at the carrier frequency.

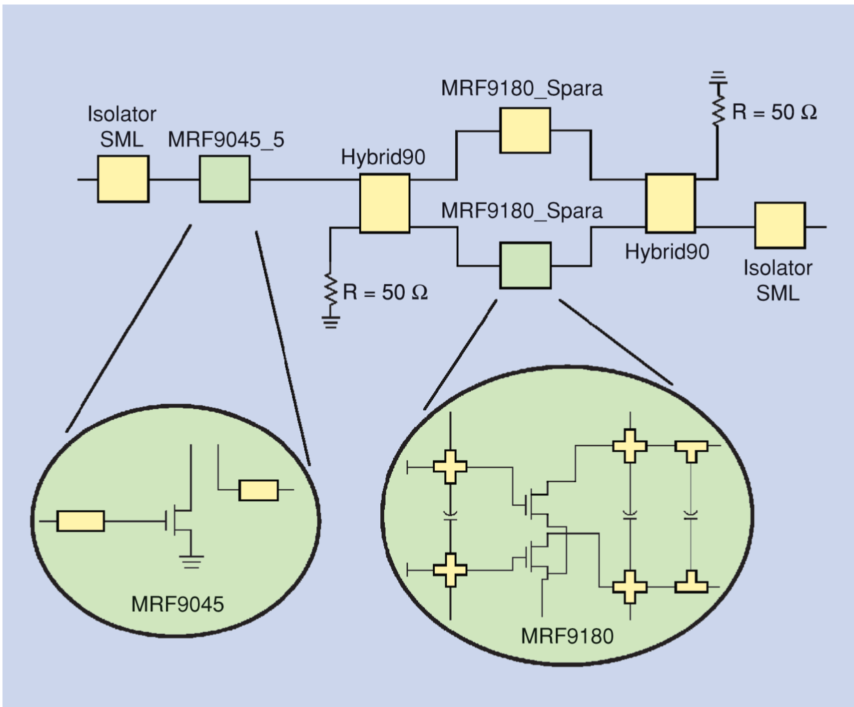


Figure 8. *The designed PA block diagram.*

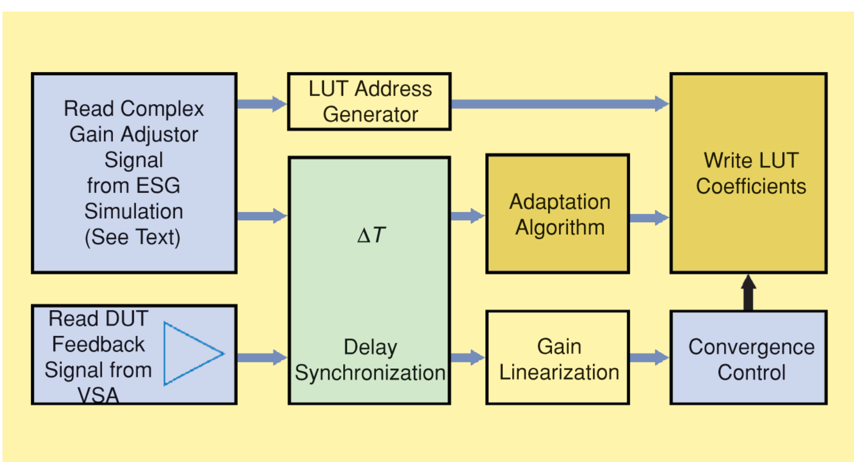


Figure 9. *The schematic of the VSA portion of the digital predistorter. Several significant functional blocks are highlighted, including the blocks on the left-hand side which read-in the reference data from the ESG portion of the PD system, the measured PA output and the required synchronization of the two signals.*

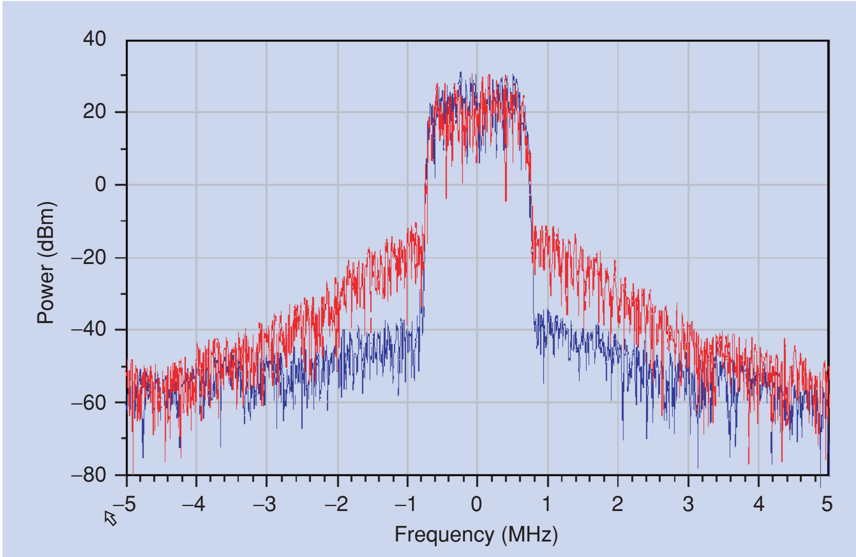


Figure 10. *PA output with and without DP of the PA (red: without PD, blue: with PD). This spectrum is derived using fast-Fourier transform (FFT) methods from time-domain simulation data, thus, the carrier frequency is not shown.*

However, the use of the training signal does require that the modulated signal be interrupted while the training signal is transmitted. In addition, since the training signal is a single tone, the digital predistorter is only correcting for the operation of the amplifier at a single frequency, not across the entire transmission bandwidth. If the amplifier's passband is quite flat, the use of a single-tone training signal in this manner will enable the PD function to be determined accurately. In any event, the use of a single complex coefficient to correct for distortion of the amplifier at a particular power level presupposes that any amplifier memory effects are minimal.

Figure 6 shows a block diagram of the adaptation algorithm that is employed. The algorithm is based upon the determination of the open loop gain H, of the predistorter and amplifier combination, at the power level associated with each LUT entry. Recall that the desired linear response of the predistorter and amplifier cascade requires that $\boldsymbol{F}(|V_{\mathrm{i}}|)\,\boldsymbol{G}(|V_{\mathrm{p}}|) = k$ for all inputs. Hence, if $\boldsymbol{G}_{\mathrm{lin}}$ is set to be equal to k, the desired open-loop gain of the system is unity. $\boldsymbol{G}_{\mathrm{lin}}$ is a constant value after linearization. If the calculated open-loop gain is not equal to unity, the PD function must be adjusted in a manner to drive the open loop towards unity. This can be achieved in the manner illustrated in Figure 7.

The PD function is defined by a set of coefficients stored in the LUT, L^n, where each n corresponds to an input-signal magnitude that is mapped to an LUT address. In order to drive the open-loop gain to unity, the PD function's coefficients are updated by dividing each coefficient by the calculated open-loop gain.

The delay in the feedback path is estimated by calculating the correlation between the magnitude of the input signal and the magnitude of the feedback signal. The use of the magnitude of the signals has the benefit of not requiring phase synchronization in the feedback path. Since the delay in the feedback path will not necessarily be equal to an integer number of DSP sample periods, interpolation is employed to more precisely align the input and feedback signals. The correlation between the input and feedback signal is performed on a modulated signal that precedes the training signal because the gain compression of the amplifier makes the accuracy of the correlation over the training signal suspect. In general, the accuracy of the estimation improves as the block size increases. Unfortunately, a larger block size requires more memory and takes longer to adapt.

The PD function cannot be exactly determined following the transmission of a single training ramp and requires an iterative calculation of the LUT coefficients.

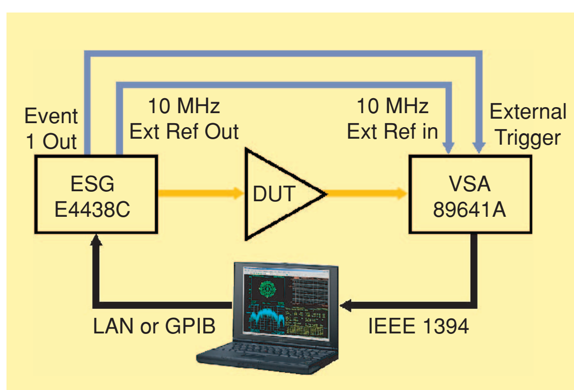


Figure 11. *ESG vector signal generator, VSA vector signal analyzer, DUT and ADS-hosted PC hardware set-up in the connected solution. The DUT is the PA prototype based on LDMOS transistors.*

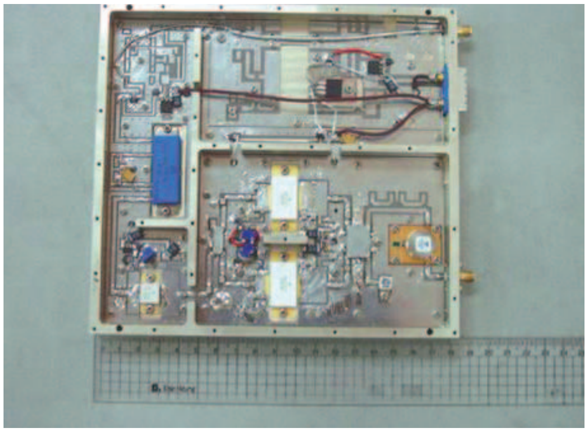


Figure 12. *A photograph of the 30-W PA.*

A series of training ramps will have to be transmitted—although, in practice, significant improvements in the ACPR of the amplifier are often observed even after just two or three training ramps.

Adaptive DP provides sufficient linearization with less complex RF hardware by depending rimarily upon DSP rather than analog manipulation.

The Connected Solution

The use of a mixed software-hardware design method may be referred to as a "connected solution." The traditional design methodology for modern communication systems typically starts with the use of EDA tools to develop an initial design from functional blocks and models. The behavior of the design is then studied by executing a multitude of different simulations. Modifications to the design are made within the EDA environment until acceptable performance is achieved. A prototype based upon the EDA-developed design is fabricated and evaluated within the laboratory setting. Commonly, at this hardware measurement stage, unexpected issues arise due to imperfections in the models and information used during the design phase.

An alternative to the traditional design methodology that does not separate the EDA and laboratory phases of system development can be advantageous for a number of reasons. By combining EDA tools with instrumentation, flexible verification solutions can be created to meet emerging requirements not met by instruments or EDA alone. In addition, the more measurements added to a simulation, the more realistic and predictive are the results. As more realistic, complex, and custom signals and measurements are generated and/or recovered in the EDA environment, the greater the insight into the performance of a device under test (DUT). Moreover, the speed of the design process may be increased as hardware-related issues are uncovered sooner and more easily analyzed in the context of the system.

Simulation and Measurement Results

A 30-W (average) PA was designed using a 170-W laterally diffused metal-oxide semiconductor (LDMOS) and simulated within the adaptive DP system at 880 MHz. It should be noted that due to the use of digital baseband linearization, the operating frequency of the PA does not significantly affect the degree of linearization obtained. For example, this system has also been applied to a W-CDMA PA operating at 2,140 MHz with similar results. The PA consists of three stages: 1) an amplifier based on MRF9045 transistors and driven by an MHL9236 2.5-W ultralinear amplifier module, 2) an MRF9045 amplifier second stage, and 3) a balanced MRF9180 amplifier final stage. Figure 8 represents the second and third stages of the simulated PA block diagram. The designed PA has 49 dB of gain and 54 dBm of saturated power. Figure 9 shows part (described later as the VSA simulation) of the schematic of

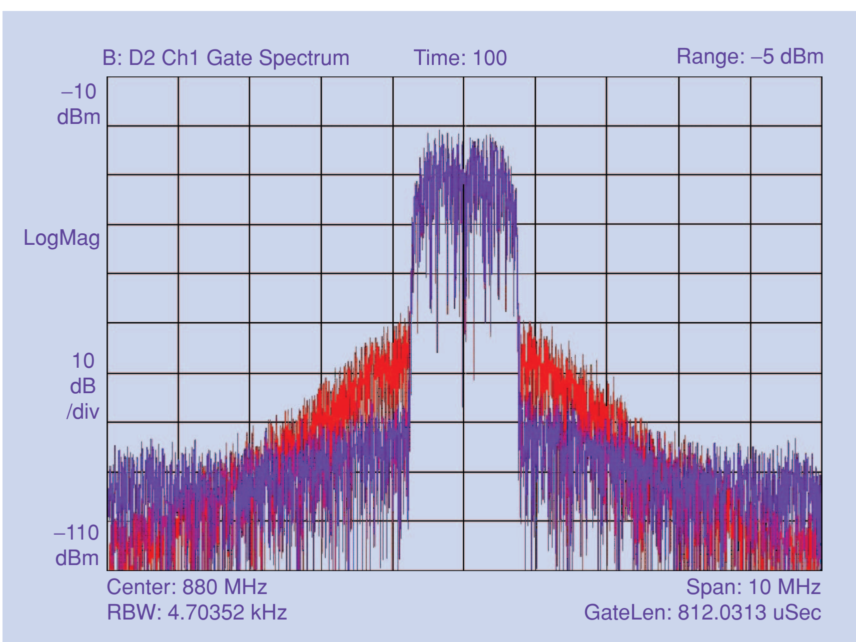


Figure 13. *PA output with and without DP of the actual PA DUT (red: without PD, blue: with PD).*

the designed digital predistorter. All subsystems discussed under the heading "Adaptive DP System" were implemented within the ADS Ptolemy environment. ADS Ptolemy provides a set of DSP and RF model blocks and supports simulation of mixed baseband and modulated RF signals, a requirement for DP. Figure 10 illustrates the simulation results; approximately 20 dB of distortion correction is achieved at the 45 W output power level.

Figure 11 illustrates the set-up for interconnecting the vector signal generator, amplifier (which is the DUT), vector signal analyzer, and the digital predistorter system design. The DUT is the 30-W (average) LDMOS PA (Figure 12) as described above. The process of linearization consists of two separate steps, here referred to as: 1) ESG simulation and 2) VSA simulation. We use the term "simulation" not in the experimental sense but to recognize that a simulation tool is being used to implement the DSP portions of the system. However, the same algorithms may be configured in a pure simulation mode using a model of the PA. In this way, the design may be first predicted using simulation only but then verified and improved using the connected mode with the real PA hardware. The ESG simulation generates and predistorts the modulated signal, passing it on to the ESG, where the signal is modulated onto an RF carrier. The RF output of the ESG is fed to an amplifier, the output of which is connected to the VSA through an attenuator to protect the VSA's input circuitry. The VSA captures the downconverted signal and transfers the signal into the ADS domain, whereby ADS performs signal processing to update the set of LUT coefficients. A CDMA2000 modulated signal with 1.2288-MHz chip rate and 9.6 dB of peak-to-average ratio (PAR) was used for the simulations and measurements. Figure 13 shows the PA output with DP applied after three iterations at 45 W of output power. The out-of-band spectra were reduced by over 15 dB, as seen by comparing the blue trace representing the output signal with PD to the red trace, which is the uncorrected output. Compared to the simulation results, a 5-dB difference was observed. This discrepancy may be due to temperature-based memory effects, the limit on the VSA's bandwidth (36 MHz), and/or inaccuracies of the active device's "Root" model used for simulation.

Although this article focuses on adaptive digital linearization using a simple training signal for correcting only memoryless nonlinearities of a PA, understanding the effects of the system's bandwidth is valuable when implementing practical linearized amplifiers. Bandwidth limitations, for instance, may depend on the wireless system's ACPR specification. For adequate cancellation of distortion products, the digital predistorter should consider the fifth order IMD products of the PA under consideration. For example, if the signal bandwidth is 10 MHz, the DSP bandwidth should be 50 MHz. Moreover, memory effects limit the performance of DP, because memory effects become significant as signal bandwidth increases.

A simple adaptive algorithm may be employed to update the LUT coefficients until an optimum setting is achieved and used to predistort the input to the amplifier.

Conclusions

When compared to other linearization methods, adaptive DP provides sufficient linearization with less complex RF hardware by depending primarily upon DSP rather than analog manipulation. The adaptive DP system may be implemented in EDA software along with interconnected test equipment (arbitrary RF signal source and vector signal analyzer). Through the use of a training signal, a simple adaptive algorithm may be employed to update the LUT coefficients until an optimum setting is achieved and used to predistort the input to the amplifier. This "connected solution" approach can provide key information to design engineers for optimizing the DSP architecture of a PA. For future work, other issues will be taken into account, such as temperature and electrical memory effects.

References

[1] D. Hilborn, S.P. Stapleton, and J.K. Cavers, "An adaptive direct conversion transmitter," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 223–233, May 1994.

[2] O. Ichiyoshi, "High-power linear amplification using periodically updated amplitude and phase correction values," U.S. Patent 5 699 383, Dec. 16, 1997.

[3] F. Zavosh, M. Thomas, C. thron, T. Hall, D. Artusi, D. Anderson, D. Ngo, and D. Runton, "Digital predistortion techniques for RF power amplifiers with CDMA applications," *Microw. J.*, Oct. 1999.

[4] "Adaptive predistortion using the ISL5239," Intersil Application Note AN1028, Sept. 2002.

[5] K.J. Muhonen, M. Kavehrad, and R. Krishnamoorthy, "Look-up table techniques for adaptive digital predistortion: A development and comparison," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1995–2002, Sept. 2000.

[6] J. de Mingo and A. Valdovinos, "Performance of a new digital baseband predistorter using calibration memory," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 1169–1176, July 2001.

5989-9020EN