RF predistortion straightens out your signals

Paul Rako  - May 12, 2011

Modern RF amplifiers need both linearity and high efficiency. The linearity requirement is due to the use of modern modulation schemes, such as QAM (quadrature-amplitude modulation) and OFDM (orthogonal frequency-division modulation, Reference 1). These amps need efficiency so that they consume less power and reject less heat. Developers often mount modern RF-amplifier assemblies on the antenna pole. These “masthead”-amplifier designs’ enclosures can have no fan and are exposed to direct sunlight. Every watt that you can save on power dissipation is a watt that the heat sink need not dissipate. In addition, driving the amplifier too hard will cause it to distort, creating harmonic spurs and ruining your demodulation. These spurs will fall in adjacent frequency bands, perhaps those that cell-phone companies don’t own. The FCC (Federal Communications Commission) places strict limits on this ACLR (adjacent-channel-leakage ratio).

Hence, you have two reasons to achieve good linearity: so that you can accurately demodulate the signal and so that your signals don’t interfere with adjacent signals. It is also important that you have the best achievable power efficiency in the output stage. The problem is that linearity and efficiency are mutually exclusive.

You can view RF-amplifier distortion in both the frequency and the time domains. You can visualize a rounded-off or flattened sine wave in the time domain passing through the RF amplifier, just like an audio signal that you drive too close to the rails (Figure 1). In the frequency domain, amplifier distortion shows up as “skirts” comprising harmonics that lie in adjacent frequency bands (Figure 2). The more power you expect from any amplifier, the more distortion you will get. At RF frequencies, you get not only amplitude distortion but also phase distortion and distortion from thermal transients and electrical-memory effects (Figure 3). Phase distortion occurs when the RF output lags behind the input signal in the areas of fast slew rates, such as when the carrier signal is passing through ground or when a modulation envelope must instantly move to a different level.
Figure 2 In the frequency domain, distortion shows up as skirts that drape from your intended transmission bandwidth (a). Predistortion schemes reduce those skirts, improving adjacent-channel interference and bit-error rate (b) (courtesy Scinterra Networks).
To pack more information into a given bandwidth, modern modulation techniques depend on accurate reception of the envelope of the RF signal. The exact voltage and phase decode a constellation of points that represent a digital code. This code creates a digital data stream that you then further decode into a baseband voice or data signal.

Older modulations are less sensitive to amplifier linearity. AM (amplitude-modulation) radios and analog broadcast TV use AM, which depends on the peaks of the RF signal. Any distortion affects all the peaks equally and has less impact on the quality of the received signal. FM (frequency-modulation) radios and the audio of analog-TV signals use FM, which depends only on the zero crossing of the waveform. Any amplitude nonlinearity has no effect whatsoever. Phase distortions have an effect on the zero crossings, but they tend to be uniform effects and do not interfere with the FM demodulation.

You can use several techniques to improve the linearity of an RF amplifier. First, you can use better transistors. For this reason, manufacturers use GaAs (gallium-arsenide) and other III-V semiconductor processes—chemical compounds with at least one Group III element and at least one Group V element—to make RF transistors. You can also try using SiGe (silicon-germanium) transistors, perhaps in conjunction with a CMOS process (Reference 2). Although slower and noisier than GaAs, SiGe can often get the job done, especially at frequencies lower than 3 GHz. Engineers face constant pressure to use CMOS for RF amplifiers—because of its low cost—but the low operating voltage of CMOS makes it difficult to implement in a power amplifier. CMOS also has a high noise factor, which you can reduce by increasing the size of the transistor structures, but that approach also increases the stray capacitance and lowers the frequencies at which you can use the product. RFMD and other companies offer CMOS on sapphire with a dielectric isolation layer under all the transistors (Reference 3). This approach exploits the cost advantage and reduces stray capacitance.

The market-driven realities are that engineers can make low-power RF amplifiers in CMOS for Wi-Fi hot spots. Cell phones require more exotic processes, such as SOI (silicon on insulator), and GaAs will predominate for the near future in cell-phone base stations.
Once you have a sufficiently linear transistor technology for your power amplifier, you need to look at the amplifier architecture. You can switch from an intermittently driven architecture, such as Class C, to a more continuously driven one, such as Class AB. Class C offers high efficiency because it employs one transistor pulsing a tank circuit to create the RF sine wave you are trying to transmit. Class C amplifiers are woefully inadequate for modern linearity requirements, however, especially in base stations. One way to obtain good linearity is by underdriving the amplifier, so the transistors do not approach saturation and the output-voltage swing is well within the range of the power rails. Unfortunately, this approach is the worst thing you can do for efficiency. To combat the problem, you can try using a Doherty amplifier, a compound device that uses a main path and a subsidiary RF path to allow you to save power at low signal strengths and still be able to accommodate larger signal swings when you need greater transmitted power (Figure 4). Doherty-amplifier architecture works well, but it adds parts and complexity to what ideally is a simple amplifier stage.

Once you push the RF amplifier toward saturation to get efficiency, you can try to linearize it with feedforward techniques. RF designers have for more than a decade successfully used these techniques in cell-phone base stations. The problem now is that the new modulation schemes for 4G (fourth-generation) LTE (long-term-evolution) base stations are even more demanding. To deliver better bandwidth efficiency, measured as bits per hertz, these new modulations put difficult linearity requirements on the best of amplifiers. This situation has led engineers to use predistortion techniques to linearize RF power amplifiers (Reference 4). Because the techniques involve sampling the output of the antenna feed and sending it back to the input, it seems like a familiar feedback technique to all analog engineers. Predistortion provides no feedback signal to an error amplifier, however, because the RF signal moves too fast to send a real-time signal of the carrier frequency back to an error amplifier. Instead, predistortion uses algorithms that accurately predict the effect on various operating conditions of the amplifier to adjust the input signal so that it ends up linear once it passes through the RF power amp. You can imagine the fundamental contribution of the algorithm. All RF amplifiers flatten a sine-wave carrier that is large enough to swing close to the power-supply rails. So your predistortion algorithm would make these larger-amplitude sine waves have sharper peaks. In that way, you get a purer sine wave from the amplifier. It is easy to visualize this scenario in the time domain. In the frequency domain, you can imagine the predistortion as adding harmonic content at phase angles that cancel out the spurs that the nonlinear RF power amplifier creates. When you switch on a predistortion circuit, you see the adjacent-channel spurs fall to much smaller amplitudes.

With a similar thought experiment, you can see how a predistortion algorithm can compensate for phase error in an amplifier. Because the phase error is predictable and repeatable, the algorithm
can modify the input waveform’s timing to null out any amplifier lags. In the time domain, you can imagine the algorithm leading the signal during fast slew rates so that the amplifier ends up outputting a clean sine wave. In the frequency domain, the adjacent-channel spurs fall to acceptable levels.

Modern predistortion algorithms are sophisticated enough to remove distortion even from thermal effects. Hot and cold power transistors distort signals differently. You can develop an algorithm that predicts the power dissipation in the output transistor. From that prediction, you can infer transistor temperature and then adjust the input accordingly to keep the output linear. This algorithm must take into account the thermal time constants of your heat sink and the ambient environment.

Digital or analog predistortion?

Over the last few years, cell-phone-base-station makers have accepted the use of digital predistortion to linearize their amplifiers (Figure 5 and Reference 5). In this scheme, a directional coupler samples the RF output. You use a mixer to downconvert the gigahertz-level signal to a lower frequency. You can then use a fast ADC to sample the waveform. You send those samples into an FPGA, which runs the predistortion algorithm you have developed to modify the input waveforms, which a digital data stream also represents. The FPGA can then output the RF baseband or I (index) and Q (quadrature) signals that you upconvert to the RF-carrier frequency of your cell-phone band.

You can use one of several approaches to build this system (Reference 6). By sourcing separate ADC and downconverter chips, you can optimize the system to your needs and use more standardized parts, which you can obtain from many vendors. For example, Hittite, Analog Devices, Texas Instruments, Linear Technology, and Intersil (Reference 7) all make the silicon chips you need for a discrete digital-predistortion circuit.

Many engineers are familiar with using Altera’s FPGAs in the digital section. The company’s megacore IP (intellectual property) performs the digital part of the predistortion operation (Reference 8). Analog Devices has partnered with Altera to supply a mixed-signal digital-predistortion system board, and Texas Instruments offers parts such as the GC5325 transmitting processor to reduce signal-crest factor and counteract power-amplifier distortion (Figure 6). Xilinx offers a digital-predistortion reference design for its Virtex-4 and Virtex-5 FPGAs. Because cellular base stations are carrying more channels of RF, space is becoming a problem. Companies such as Linear...
Technology have addressed this problem by incorporating the entire digital-predistortion circuit into the LTM9003 micromodule (Figure 7).

Despite cell-phone-base-station manufacturers’ acceptance of digital systems, the vendors are making a fundamentally analog circuit into a sampled-data system, bringing cost, power, and space penalties. An alternative is to use analog techniques to linearize your RF amplifier. Start-up Scintera Networks, for example, targets low-power RF systems in the 5W range and in the signal path of UHF (ultra-high-frequency) TV stations (Figure 8). The scheme samples the RF from the driver stage and keeps that RF signal in the analog domain, but modifies it with coefficients employing a Volterra Series expansion of the waveform. A Volterra Series is a model for nonlinear behavior that is similar to a Taylor Series, except that the Volterra Series can represent memory effects. Scintera’s scheme samples and digitizes the RF output, and that sampling goes into digital circuits in the company’s chip. The design uses the digital section to compute the analog coefficients for the RF-signal chain and then uses another directional coupler to mix the Volterra-coefficient-modified RF signal back into the RF path. The system needs to handle only enough RF in the chip to correct the amplifier distortions. Most of the RF power stays in the main RF path, bypassing the IC. By keeping the RF in the analog domain, Scintera provides a system that consumes much less power than does a digital-predistortion setup (Figure 9).
Be aware that the design and testing of digital-predistortion systems are not trivial tasks. You will need sophisticated RF-design tools, such as AWR’s Microwave Office and Agilent’s ADS (Reference 9). In addition to advanced test equipment to characterize the RF path, you may need to buy and understand specialized test equipment, such as a real-time spectrum analyzer (Reference 10).

No matter if you use analog or digital predistortion, you can reduce interference and use advanced modulation schemes in your RF design. Best of all, the predistortion techniques allow you to drive the RF amplifier closer to saturation, which improves power efficiency. You can roll your own systems from discrete chips or use a micromodule that integrates all of the functions into one package. Achieving the necessary linearity in the ADC and the downconverter ICs is an achievement for the semiconductor companies. These companies all have application experts that can help design an RF-signal path that will meet all of your regulatory requirements, sip power, and deliver the most bits per megahertz.

You can reach Technical Editor Paul Rako at 1-408-745-1994 and paul.rako@ubm.com.

References
Editor's note: The original version of this article contained an error, which has been corrected in the text above. Reference 3 was changed from "Jazz Semiconductor, Fall 2007" to "Advanced Substrate News, October 31, 2007" on May 12, 2011.