Application Note

Addressing Efficiency and Linearity With a Single-Ended Class AB Power Amplifier With Optimized Third-Harmonic Termination

Overview

The single-ended RF power amplifier (PA) is a standard design approach in the RF engineer's tool kit, but high efficiency and good linearity may only be achieved if the harmonic terminations and biasing conditions are carefully examined. This application note describes the design of a single-ended PA using Class AB "sweet spots" and an optimized third-harmonic termination based on a design flow that encompassed analysis of the requirements and initial concepts to the simulation, fabrication, and measurement of the built prototype. The design won first place at the 2016 International Microwave Symposium's High-Efficiency PA Student Design Competition.

Design Challenge

The main challenge of this design was to maximize the overall power-added efficiency (PAE) of the PA while amplifying a time-varying envelope signal, without compromising the linearity performance. The design specs required the achievement of the highest PAE as measured for a two-tone input signal while at the same time ensuring that the carrier-to-intermodulation ratio (measured third-order intermodulation distortion [IMD3] level) was lower than –30 dBc.

A figure of merit (FOM) was calculated from this measurement according to:

$$\mathrm{FOM} = \mathrm{PAE} \times (f_r)^{1/4} = \frac{P_{\mathrm{out}} - P_{\mathrm{in}}}{P_{\mathrm{DC}}} \times (f_r)^{1/4}, \quad \textit{Equation 1.}$$

where the PAE is weighted by the operating frequency f_r in gigahertz to compensate for the greater technical challenge in the higher-frequency design. P_{out}, P_{in}, and P_{DC} refer to the RF output power of the two tones of interest, the RF input, and DC power supplied, respectively.

The PA design operated at 3 GHz and used a packaged Wolfspeed [formerly Cree] gallium nitride (GaN) high-electron mobility transistor (HEMT). The prototype PA, shown in Figure 1, achieved a maximum continuous-wave (CW) output power of 36.2 dBm and a two-tone PAE of 44 percent at –30 dBc IMD3, while delivering 34.04 dBm of output power.

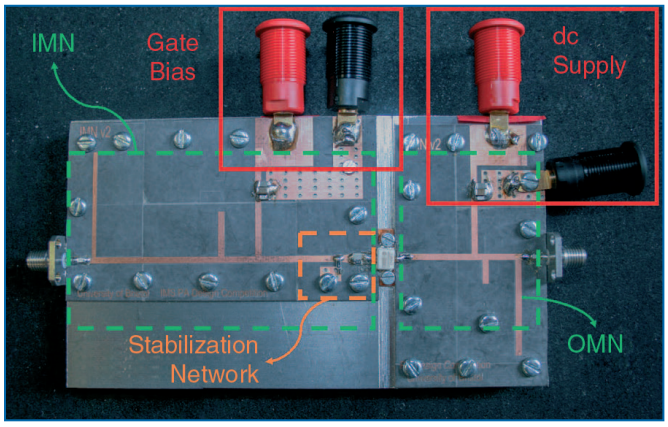

Figure 1: Annotated photograph of completed PA.

NATIONAL INSTRUMENTS

Problem Formulation

At the beginning of a design project, it was necessary to analyze the full range of specifications and survey prior designs that achieved similar results.

Specifications for the design were:

1. Fewer than 24 dBm (250 mW) of input power to reach the point of saturation in the presence of a CW signal and produce an output power between 36 dBm (4 W) and 50 dBm (100 W) at saturation
2. Fewer than 22 dBm (158 mW) per tone to reach the point of saturation in the presence of a two-equal-tone signal with 5-MHz tone spacing
3. Use a maximum of two DC power sources
4. Operate in the 1–10 GHz frequency range

The main constraint placed on the design was the large-signal CW gain, which had to be above 12 dB at saturation to achieve the minimum output power of 36 dBm. For the operating frequency, values between 3 and 5 GHz were considered because, due to the frequency dependence of the FOM in Equation 1, a higher value would result in only a small FOM improvement. GaN transistors were chosen for their performance at multiple gigahertz frequencies and because by operating at the low end of the 36–50 dBm range, their efficiency could be further enhanced.

A two-tone signal with 5-MHz frequency spacing was used to evaluate the design. Closely spaced two-tone signals have a variable time-domain envelope with a 3-dB peak-to-average power ratio, so PAE over a 3-dB backoff range needed to be improved to avoid degrading the average efficiency. Finally, it should be noted that the PAE was measured at -30 dBc IMD3 rather than at saturation. For conventional amplifiers, the PAE was seen to increase monotonically until heavy compression occurred, so, to maximize the measured PAE, it was helpful to maintain a linear power transfer characteristic until saturation. A tradeoff between linearity and efficiency was, therefore, inevitable, and a good balance was crucial to achieve the desired FOM.

Design Solution

The linear DPA architecture has been shown in other designs to adapt well to the design requirements providing remarkable results due to a number of features that complement the specifications. Broadly speaking, the active load pull of the main amplifier by the peaking amplifier ensures improved efficiency in the backoff region. Through the appropriate choice of biasing—Class AB for the main and Class C for the peaking—the compressive and expansive nature of the transfer characteristics of the two branches can be exploited to compensate for one another, thereby suppressing IMD3 and enabling linear operation up until saturation.

Improving on previous design performance was thus no easy task. One option considered was to increase the degree of load modulation compared to the previous linear DPAs, which could boost the efficiency at backoff. However, this would only be obtained if the linearity of the main amplifier's transfer characteristic could also be improved, allowing for the peaking stage to turn on when the main amplifier has maximum (or close to maximum) voltage swing. Ultimately, the designer decided to focus on extending the linearity of a single-ended PA through exploiting specific operating condition values for Class AB gate bias values that are found to improve the linearity of the PA's power transfer characteristic, while also tuning the third harmonic termination.

Despite having drawbacks in terms of performance when compared to a linear DPA solution, a single stage design has several practical advantages, including a single active device, less extensive biasing and matching circuitry, simplicity for implementing and iterating, and low cost and form factor.

Moreover, a strong understanding and characterization of a linear, single-ended Class AB amplifier provides a good foundation for the design of the main amplifier stage in a DPA or other load modulation-based architectures.

In Class AB PAs, the nonlinearities introduced by the output current waveform's truncation (which is, itself, a function of input drive level and bias) can compensate for the weak nonlinearities introduced by the transistor to produce favorable linearity performance at a given drive level. Specifically, for deep Class AB conduction angles, the current waveform truncation causes a compression of its fundamental frequency component with increasing drive level, which can compensate for the expansive characteristic exhibited by real-world transistors in the turn-on region. A full analytical description of this phenomenon can be found in [8]. For two-tone signals, this translates into local minima, or "nulls," in the IMD3 profile at certain points in the amplifier's response, as shown in Figure 2. (In the graphs in the figures that follow, "high" and "low" refer to the IMD3 measured for the intermodulated components above and below the two carrier tones, respectively.)

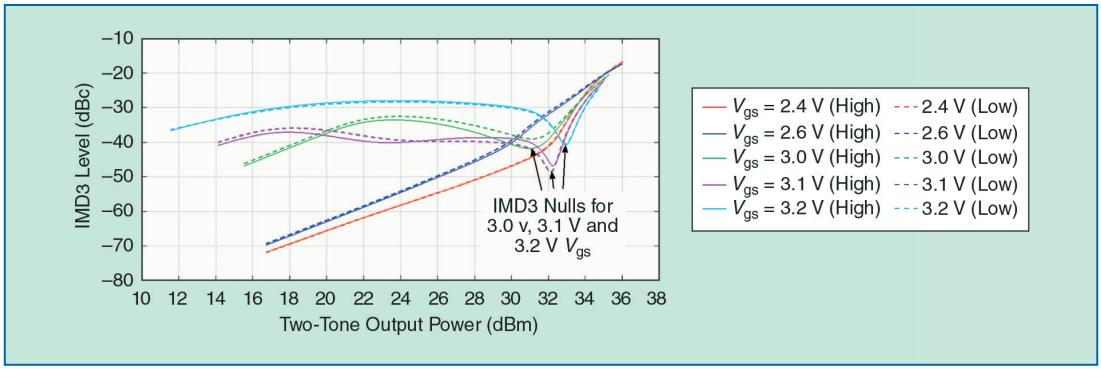

Figure 2: The simulated IMD3 versus output power, showing the varying profile and the nulls in IMD3 for different gate bias voltages.

To further improve efficiency and increase the linear dynamic range of the amplifier, the appropriate impedance terminations must also be considered. Prior work has mostly addressed the impact of the second-harmonic impedance termination on linearity. However, little investigation has been done in considering the effect of third-harmonic termination on linearity, which is particularly significant in Class AB PAs.

Figure 3 shows the result of a harmonic balance (HB) CW simulation sweeping the third-harmonic impedance at a constant input drive level for a Class AB PA, with fundamental and second-harmonic impedances optimized for PAE.

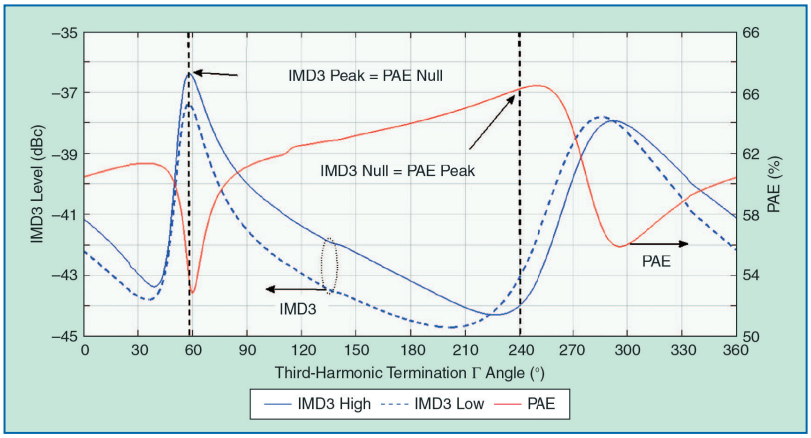

Figure 3: The simulated IMD3 and PAE versus the third-harmonic reflection coefficient angle, illustrating the matching peaks and nulls of the two profiles.

As the figure indicates, the effect on PAE and IMD3 is notable, with the null in IMD3 corresponding to a peak in PAE and vice versa. This is due to an appropriate third-harmonic short, which promotes a correct shaping of the voltage waveform and reduces the overlap in current and voltage at the current generator (CG) plane of the transistor. This is further shown in Figure 4, which illustrates the intrinsic voltage and current waveforms for different cases of third-harmonic termination, which will be further explained in the following section.

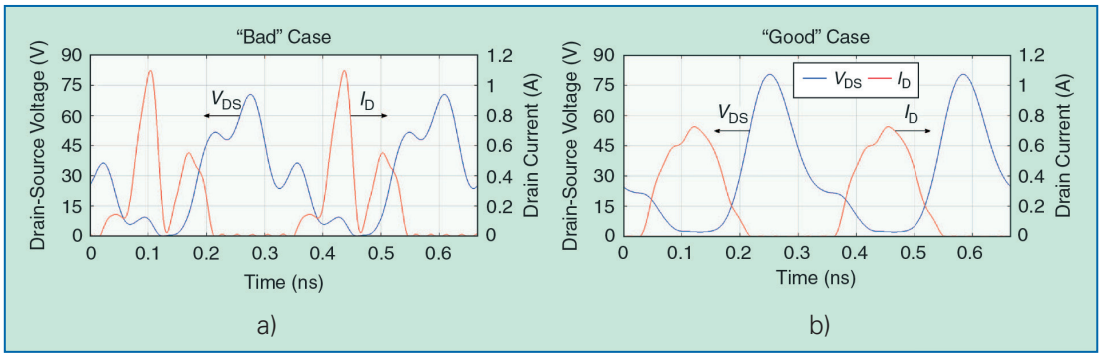

Figure 4: The intrinsic drain-source voltage and drain current waveforms for a) "bad" and b) "good" third-harmonic impedance terminations.

The designer chose NI AWR software, specifically Microwave Office circuit design software [8] for all circuit simulations and focused on finding accurate models for the devices and lumped elements used. Following an extensive survey of packaged GaN HEMT devices, the Cree CGH40006P was selected based on its highly accurate and verified large-signal model and its power and gain-frequency performance [1]. The microstrip circuitry was modeled using AXIEM 3D planar electromagnetic (EM) simulator and Murata high-precision ceramic capacitors were chosen for their accurate parasitic model.

Design and Simulations

A simulation-based design approach was carried out using Microwave Office, resulting in a prototype PA for which an annotated schematic is shown in Figure 5.

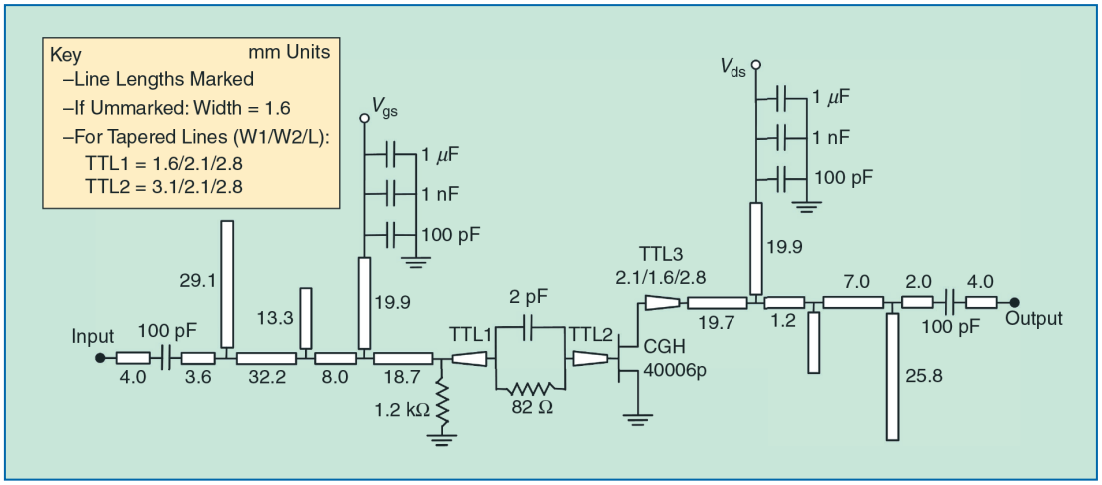

Figure 5: The schematic of the PA showing the tapered transmission line.

Ensuring transistor stability at all frequencies present in the amplified signal is a crucial first step in the design process. A practical way to achieve this is by meeting the k > 1 requirement for unconditional stability from DC to 10 GHz, in which the Rollet factor k is a conventional stability parameter derived for linear two-port networks given passive load conditions. The designer used an input stabilization network consisting of a series resistor-capacitor pair (C=2 pF, R = 82 Ω) and a shunt (to ground) 1.2 kΩ resistor.

Before starting the simulation process, the designer made some a priori qualitative considerations based on the DC-simulated I–V curves of the device model and the absolute maximum ratings provided by the manufacturer. From PA load-line theory, the load-line match is the impedance at the CG plane, which results in a full knee-to-rail voltage swing when maximum drain-source current is reached. The PA was required to saturate at 24-dBm input power, but the device would reach current saturation only at a much higher input drive level. Therefore, the load-line match had to be adjusted to present a higher resistance at the CG plane of the device, allowing for voltage saturation (a full knee-to-rail voltage swing) at lower drain current levels.

To achieve this, a CW harmonic balance (nonlinear) simulation was set up in Microwave Office, where the source and load pull of the large signal model of the device were performed at the fundamental and second-harmonic impedance for maximum PAE, while the input drive was fixed to the required level. Provisional deep Class AB biasing was also maintained throughout the load pull simulations, ensuring strong harmonic content present in the current waveform. Two iterations of source and load pull were performed to account for the bilateral nature of the device. At this point, the third harmonic and bias were tuned with the aim of achieving the highest two-tone PAE with IMD3 levels below -30 dBc while meeting the CW gain requirement. The designer began the optimization process by considering the CW power requirement to narrow down the available design space of usable gate-source voltages.

As expected, it was observed that decreasing the quiescent current from the Class A value to the Class B value led to a decrease in the gain. The smallest quiescent current, which allowed 12-dB gain to meet specifications, was recorded as the minimum tolerable value. Having established a range for the quiescent current, load pull contours were generated for IMD3 and PAE based on third harmonic impedances at a number of bias settings spanning the range. The combination of bias and third-harmonic terminations achieving the highest PAE and two-tone IMD3 at -30 dBc was then selected.

The CGH40006P NI AWR Design Environment model features intrinsic ports that allowed for RF time-domain measurements at the CG plane of the device. Figure 4 shows the resulting intrinsic voltage and current waveforms for both "good" and "bad" third-harmonic terminations. For the "bad" case, as the voltage swings into the knee region, the current is heavily clipped, generating third-order nonlinearities (also evident in the voltage wave), degrading PAE and linearity. In the "good" case, the third harmonic in the current is shorted, and waveform integrity is improved as the voltage swings into the knee region.

To provide the required impedance at the package plane of the device and appropriately bias the gate and drain of the transistor, a multi-harmonic input and output matching network was synthesized as shown in Figure 5, initially using ideal transmission lines and components. The combination of the three open and short stubs at the input and output allowed all three frequencies to be matched at once. The DC bias was provided at both the input and output with quarter-wavelength lines, as high-valued shunt to ground capacitors (shorted the even harmonics while presenting an open at the fundamental. Real Murata component models were included and the planar distributed structure was reoptimized using AXIEM. The designer particularly noted the baseband impedances, as these were critical for bias oscillation (stability) and efficiency when amplifying a modulated signal. Large-value baseband decoupling capacitors were, therefore, included and tuned both at the drain and gate bias lines to ensure a short was present at baseband (megahertz) frequencies. A linear k factor analysis was rerun using the bias insertion points as RF ports with no potential instability being highlighted. Finally, some adjustments were made in the layout to allow for the length of the stubs to be tuned on the bench and the position of the input and output RF decoupling capacitors to be moved up and down the quarter-wavelength bias line.

Manufacturing

To minimize the possibility of error in the manufacturing process, high-quality components were used and careful attention was paid to the mechanical construction of the PA. A Rogers Duroid 5880 glass microfiber polytetrafluroethylene composite with fr = 2.2, tand = 0.0004 was chosen for the substrate due to its low loss and constant, uniform dielectric coefficient. The printed circuit boards (PCBs) were manufactured using a laser-cutter machine and a LPKF ProConduct conductive compound to achieve good electrical contact in the cavity of the via holes. The thickness of the substrate was 0.51 mm, which allowed a 50-Ω track to be made the same width as the transistor package tabs. The input and output matching networks were created on two separate PCBs to facilitate tuning and both boards were attached to a 10-mm–thick aluminum backplane, which provided rigidity and a thermal sink. The transistor was fixed mechanically to the backplane, and good thermal contact was ensured using a thermally conductive paste between the source and the backplane. Finally, subminiature type A (SMA) connectors were mounted mechanically onto the backplane before being soldered to the PCBs.

Measurement Setup

The measurement setup needed to be capable of measuring input, output, and DC powers, as well as the spectral composition of the amplified signal, for both CW and two-tone test signals. A block diagram of the complete apparatus (device under test [DUT]) is shown in Figure 6.

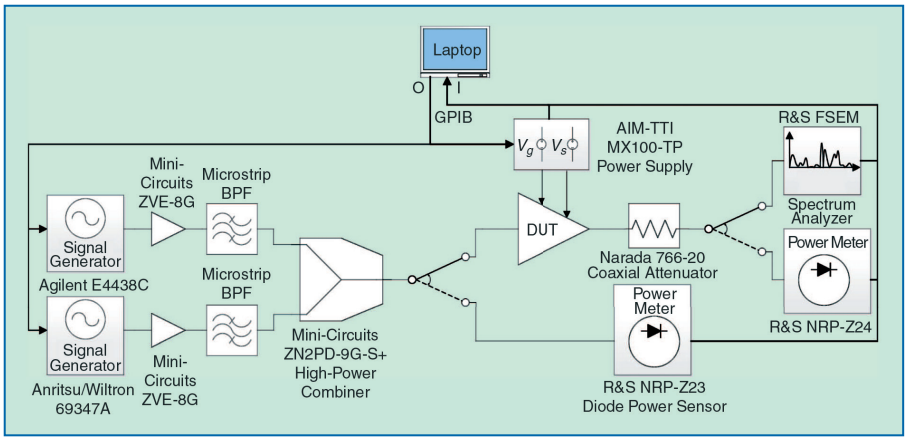

Figure 6: Block diagram of the measurement setup, detailing the instruments used.

A TTi laboratory power supply was used to provide both the gate-bias voltage and the drain-supply rail voltage. This was calibrated using an external current probe, allowing the dc power and bias point to be automatically calculated from the gate and drain currents, together with the associated supply voltages.

For the two-tone tests, two signal generators were used, both operating at a low output power to avoid generating any distortion products. The two signal generator outputs were fed into a linear amplification stage with additional filtering to remove any harmonic tones generated, before being combined into the two-tone signal. The filters were realized on a Rogers 1.575-mm Duroid 5880 using an edge-coupled bandpass topology in combination with a double radial stub low-pass filter to reject re-entrant modes at higher harmonics. The output power from the DUT was reduced with a 20-dB attenuator before it could be fed into a Rohde & Schwarz (R&S) spectrum analyzer to measure the magnitude of the original two tones and all intermodulation products of interest. All instruments were connected with a general purpose interface bus (GPIB), and a MATLAB script was used to set appropriate source power levels and receive data from the DC supply and spectrum analyzer, allowing PAE and IMD3 levels to be calculated. Input power was calibrated replacing the DUT with the NRP-Z23 RF power sensor, while the spectrum analyzer was calibrated across the entire output power range of the PA, referencing to the NRP-Z24 power sensor.

Results

The simulated results of the CW, two-tone, and IMD3 power measurements are shown, along with those measured directly from the built PA, in Figures 7, 8, and 9, respectively.

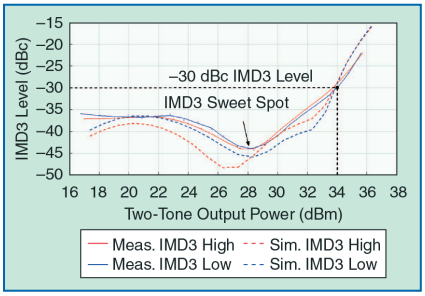

Figure 7: The measured vs. simulated single-carrier gain, drain efficiency (DE), and PAE vs. output power, indicating saturation onset.

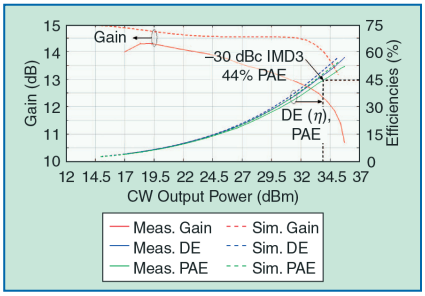

Figure 8: The measured vs. simulated two-tone gain, DE, and PAE vs. output power, with the -30 dBc PAE displayed.

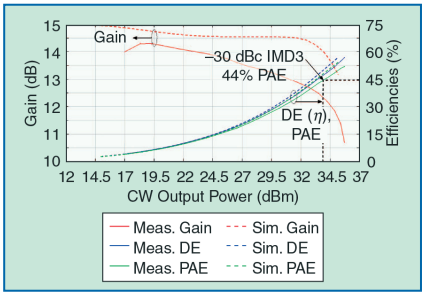

Figure 9: The measured vs. simulated IMD3 levels vs. output power, illustrating the -30 dBc IMD3 level.

Figure 10 expands on the graph in Figure 9, showing a 3D plot of the measured IMD3 versus gate bias and output power.

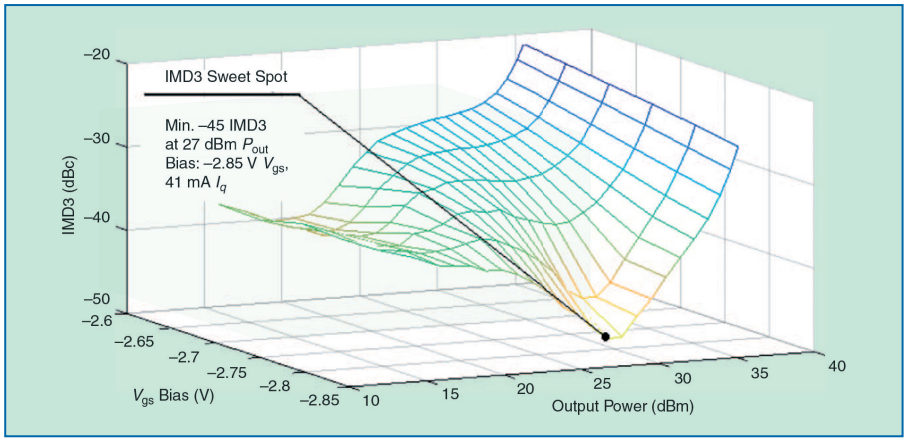

Figure 10: A 3D plot of measured IMD vs. output power vs. gate bias, showing a sweet spot at 41-mA quiescent current.

The different values of gate-bias voltage ^Vgsh displayed in Figure 10 correspond to a range of deep to middle Class AB quiescent currents. As shown, the agreement between the simulations and measurements is strong throughout the results.

During the CW measurements, significant thermal effects were experienced, which made taking reliable and repeatable measurements a challenging process. Nonetheless, a maximum PAE exceeding 50 percent at saturation was recorded, which is indicated as the peak measured value in Figure 7. Good PAE and linearity were also recorded with the two-tone test: 44 percent PAE and 34-dBm output power were measured at the -30 dBc point, giving a FOM score of 58.8, according to Equation 1. The deviations observed in the two-tone data can be explained, in part, by the fact that higher-order nonlinear effects play a greater role in the resulting behavior, meaning that the simulation will be less accurate.

The baseband decoupling capacitors were tuned on the bench, which resulted in the reduced asymmetry between the higher and lower IMD3 curves in the experiment data. The linearity sweet spot highlighted in Figure 9 at around 28-dBm output power is also clearly visible in Figure 10. The fact that the IMD3 profile varies so significantly over the gate-bias range demonstrates the potential of the proposed design methodology. The deepest sweet spot occurs at 41-mA quiescent current, which was predicted by simulation and then targeted in the design. The -30-dBc IMD3 point occurs at 34 dBm, which is only backed off 2 dB from the onset of saturation and, therefore, results in a favorable PAE measurement, as shown in Figure 8.

Conclusion

A linear, single-ended PA operating at 3 GHz was designed by exploiting the sweet spots present in Class AB amplifier responses and optimizing the third-harmonic termination. The improvements to PAE and IMD3 from the third-harmonic tuning are clear from the simulation results, and the Class AB sweet spots were shown in both simulation and measurement. The actual measured performance was 44 percent two-tone PAE at -30 dBc IMD3, giving a FOM of 58.8; this is a good result considering the variable-envelope signal used, which supports the proposed methodology. While the PA was unable to reach the outstanding FOM performance shown in previous years, the single-ended design is simpler and more cost effective than a linear DPA. The good agreement seen between the simulation results and measured data also demonstrates the value of the design method when applied using NI AWR Design Environment.

Special thanks to Paolo Enrico de Falco, James Birchall, and Laurence Smith for their contributions to the application note.