Linearity versus efficiency in mobile handset power amplifiers: a battle without a loser

This paper deals with linearity issues of microwave/RF power amplifiers used in wireless mobile handsets. Beginning with the common-knowledge trade-off between power added efficiency and intermodulation distortion, it is theoretically shown that neither of them have to be sacrificed. Indeed, the use of large-signal IMD sweet-spots - valleys on the distortion vs input power pattern - enable the operation of PAs close to saturation, and thus in a highly efficient regime, still with linearity figures only expected from backed-off class A PAs.

After the end of the cold war, microwave industry changed its focus from defence related business to the commercial telecommunications market, adopting strategies of mass production. Driven first by the mobile telephony, and, more recently, also by wireless data services, issues such as low cost, integration with other technologies and ruggedness have become as important for microwave circuits as they already were for the mass market of low frequency CMOS products. So, special attention was paid to find new active device technologies and design methods, especially for circuits where these conflicting features presented more stringent challenges.

From those, the power amplifier, PA, of mobile handsets is the circuit where power gain, G_T, output power, P_{out}, power added efficiency, PAE, and intermodulation distortion, IMD, (usually assessed in terms of the adjacent - or alternate - channel power ratio, ACPR) requirements are still difficult to be simultaneously fulfilled. Indeed, a recent survey [1] has shown that there is no device technology whose performance can be said to be superior to all the others. If, for example, GaAs HBTs are known for their higher power densities, and so for comparably lower die size, GaAs HFETs based PAs withstand higher output SWRs and offer higher G_T and PAE. Even Si LDMOS or MOSFET technologies (usually associated with the basestation PA segment, or with baseband processing circuits) can play a role due to their low cost, but can lead to larger chip sizes [1]. So, this work will not concentrate in any particular PA technology, but will address generally the linearity/efficiency compromise in mobile handset PAs.

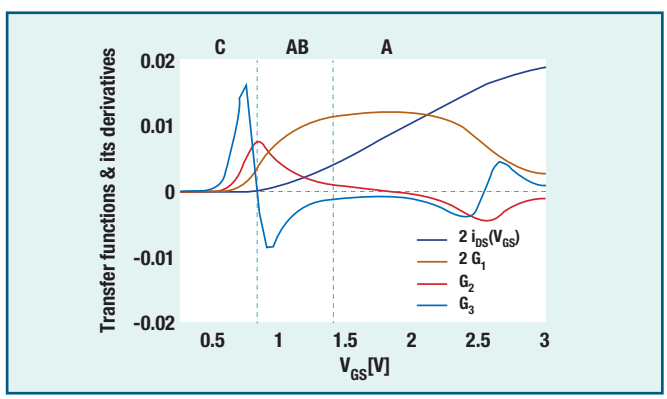


Figure 1: Typical transfer function, $i_O(v_i)$, $[i_{DS}(v_{GS})]$ and its first (G_1), second (G_2) and third (G_3) order derivatives, with different operation classes for each of the quiescent point ranges indicated on top of the figure

The most important characteristics of mobile handset PAs are output power and PAE. The first one defines (with the receiver sensitivity) the wireless cell radius, while the second ultimately determines battery life. As P_{out} is mostly determined by device size, load impedance and supply voltage, the designer simply needs to appropriately select the output termination using Cripps' load-line theory [2]. On the other hand, achieving a certain PAE specification is far from trivial, as it depends on the same parameters as P_{out}, but also demands a careful selection of the quiescent point (PA operation class) beyond output voltage, $v_O(t)$, and current, $i_O(t)$, waveform shaping. This is accomplished by selecting appropriate harmonic terminations [2] and by driving the device into saturation, i.e., strong nonlinear regimes.

Considering that our device is an ideal transfer nonlinearity, $i_O(t) = i_O[v_i(t)]$, which, for explanation purposes, is approximated by a 3rd degree polynomial with memory:

$$i_O(t) = G_1 v_i(t-\tau_1) + G_2 v_i(t-\tau_2)^2 + G_3 v_i(t-\tau_3)^3 \quad (1)$$

(1)

the fundamental zone output to an amplitude, $A(t)$, and phase, $\theta(t)$, modulated excitation $v_i(t) = A(t)\cos[\varpi_c t+\theta(t)]$ will be:

$$i_O(t) = G_1 A(t-\tau_1)\cos[\omega_c t+\theta(t-\tau_1)-\phi_1]+ \frac{3}{4} G_3 A(t-\tau_3)^3 \cos[\omega_c t+\theta(t-\tau_3)-\phi_3] \quad (2)$$

This shows that nonlinearity causes no impact on phase modulation, but can be disastrous for the information carried in the amplitude.

So, PAs for only phase modulation systems, i.e. of constant envelope, such as GSM900 or DCS1800, are designed for P_{out} and PAE specs, while PAs for CDMA and W-CDMA applications, which use non-constant envelopes, must meet those P_{out} and PAE specs but now also guaranteeing certain minimum ACPR masks.

It is for these latter circuits that highly linear PA design methods are a determining issue. Biasing the active device for class

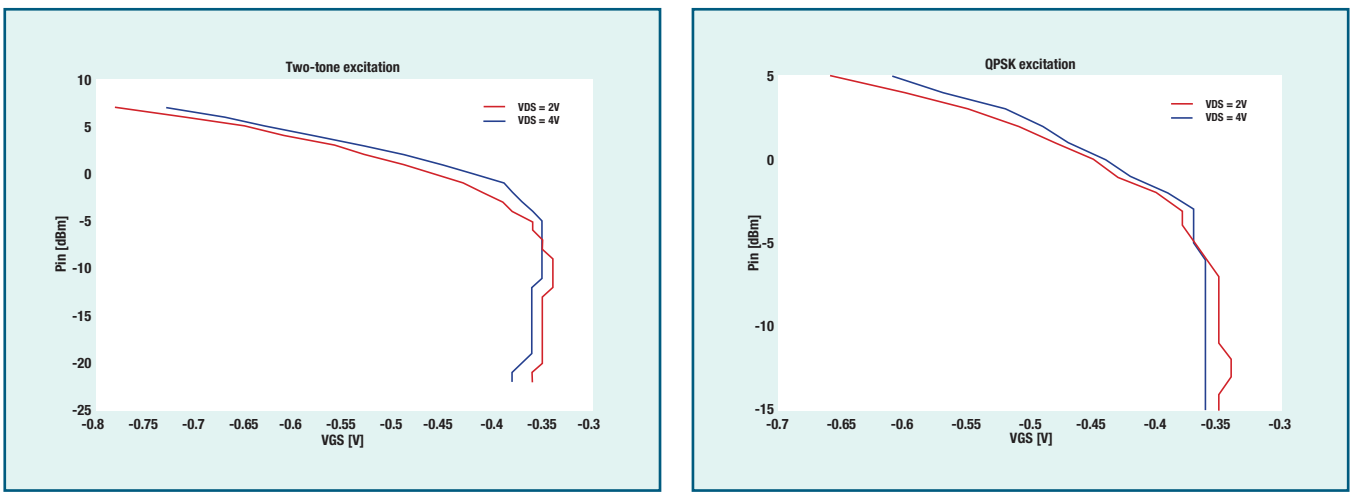


Figure 2: Evolution of the input power level IMD sweet-spot position with VGS and VDS bias, as measured in a microwave HEMT, under two-tone (a) and QPSK (b) excitations

A operation, or backing off a usual class AB or C amplifier, for achieving ACPR specs is not a solution, as it reduces P_{out} and is detrimental for PAE. Neither is driving the device into compression, for guaranteeing PAE, as it jeopardizes linearity. Fortunately, this battle between PAE and IMD does not necessarily have to have a loser. Indeed, there are some peculiar points of the ACPR versus input power characteristic [3], the so-called large-signal IMD sweet-spots, where only a few dB of output-power back-off (and thus a few percent of efficiency degradation) can lead to astonishing high levels of distortion reduction.

The remaining parts of this paper will show how to benefit from these IMD sweet-spots and thus turn them into a practically resourceful linear PA design tool.

Physical origins of IMD sweet-spots

IMD sweet-spots, or valleys in the IMD versus P_{in} characteristic, can be understood as the result of interactions between the PA mild and strong nonlinearities. The former are determined by the PA quiescent point, and govern the circuitís small-signal nonlinear distortion. The latter are determined by the device current turn-on and current saturation (found when signal excursion reaches the triode region of FETs or the saturation region of bipolar devices), and govern large-signal distortion effects.

Since the IMD is intrinsically related to the active device input-output nonlinearity, any attempt to explain its behaviour requires that we examine the transistor transfer characteristic. For that, consider the typical $i_O=f_{NL}(v_i)$ curve of a general nonlinear active device, as depicted in Figure 1. This transfer characteristic describes the dependence of the device output current on the input controlling voltage for a determined output load-line. In a FET based PA (the case shown in Figure 1 was actually obtained from a Si MOSFET) this corresponds to $i_{DS}(v_{GS})$. In a bipolar based PA that would relate i_C with v_{BE}, where v_{BE} does not stand for the usual intrinsic base-emitter voltage, but for the extrinsic one, which accounts for all voltage drops in all input base and emitter series resistances [3].

Also shown in Figure 1 are the first three derivatives of $i_O(v_i)$, which therefore correspond to the coefficients of the polynomial model (in this case a Taylor series) of equation (1).

As seen in Figure 1, the device presents a very rich small-signal nonlinear distortion behaviour, since G_3 can be made either positive (class C operation), null (for ideal class B) or negative (for both class A and AB) [3].

On the contrary, the large-signal distortion is imposed by a fundamental (and thus universal) limit: power saturation. To understand this, we must abandon our small-signal Taylor series model to adopt a large-signal polynomial. Although a formal theory would be blurred by complex mathematics [4]-[6], we can let the model keep its low order polynomial format of equation (1), but now using effective coefficients, G_1, G_2 and G_3, that depend on the excitation amplitude A: $G_1(A)$, $G_2(A)$ and $G_3(A)$. In this way, for example, the coefficients shown in Figure 1 would become $G_1(0)$, $G_2(0)$ and $G_3(0)$.

Returning to the power saturation mechanisms, we know that they impose a large-signal asymptote where gain suffers inevitable compression. So, (2) clearly indicates that the effective large-signal G_3, $G_3(A\rightarrow\infty)$, must oppose the fundamental $G_1(A\rightarrow\infty)$. This means that, whatever the sign of the small-signal $G_3(0)$, the effective large-signal $G_3(A\rightarrow\infty)$ must always be negative [5].

So, depending on the quiescent point, we can have a $G_3(A)$ that begins by being negative or positive, for small-signal - $A\rightarrow 0$, but must always end up as a negative value, for deep large-signal - $A\rightarrow\infty$.

In case $G_3(0)<0$, it is observed that for excitation amplitudes, A, where strong nonlinear operation is reached, intermodulation distortion suffers a sudden rise. For example, the low negative values of $G_3(0)$, shown in Figure 1 for class A, are indicative of very linear operation in highly backed-off regimes (and thus low PAE). But, our large-signal analysis also predicts strong IMD degradation when PAE is increased by reducing output power back-off.

In case $G_3(0)>0$, there must be an input amplitude, A_0, - typically at the onset of saturation - where $G_3(A)$ must change sign. But, as $G_3(A)$ is a continuous function, that can only happen if, for that amplitude, $G_3(A_0)$ vanishes. This implies an extremely low distortion, and so the observation of an IMD sweet-spot. Therefore, referring again to Figure 1, typical class C IMD can be quite high at small-signal, will present a dip at the onset of saturation, and then rise again when finally dominated by the PA strong nonlinear operation. Unfortunately, experience shows that these class C IMD valleys can be very narrow and that they are still dominated by the high small-signal distortion. So, the practical use of these IMD sweet-spots of class C microwave PA circuits tends to be

compromised unless a way is found to avoid these drawbacks.

Finally, let us see what happens if the device is biased for class AB operation. According to Figure 1, we would expect an IMD versus P_{in} characteristic similar to the one shown by class A amplifiers, although eventually starting with an even higher small-signal IMD. Fortunately, that is not necessarily the case. Indeed, if the device is biased in class AB, but close to the $G_3(0)$ null (i.e. close to class B), strong nonlinear operation will be first determined by current turn-on, at a certain A_1 input level, then by current saturation, at the higher A_0. The consequence of this is that the class AB amplifier begins by showing a mild negative $G_3(0)$, then presenting a slightly positive $G_3(A)$, as if it were a class C PA. But, as we have seen before, this change of $G_3(A)$ sign implies an IMD sweet-spot, at A_1, which must then be accompanied by another change of sign when the signal excursion finally enters current saturation, at A_0. In conclusion, a class AB PA is likely to present two IMD sweet-spots in the IMD versus P_{in} characteristic, and so a broad range of signal excursion - not far from the onset of saturation - where the corresponding IMD is very low [6]. Hence, in contrast to both class A and class C PAs, carefully designed class AB PAs can present very good linearity without sacrificing PAE performance. This however requires a rigorous control of the positions (A_1 and A_0) of the IMD sweet-spots, and so a perfect knowledge of which PA design parameters can affect them.

PA design parameters affecting IMD sweet-spots

The first linear handset PA design parameter to be addressed is the quiescent point selection. Given a certain bias voltage, usually set by the battery voltage, quiescent point selection becomes a matter of selecting quiescent output current, or conduction angle. This immediately determines maximum attainable P_{out} and PAE [2].

As the location of the IMD sweet-spots could be related to the input amplitude in which a strong nonlinearity was crossed, it is clear that such conditions can be reached sooner or later depending on the small-signal starting point: the quiescent bias.

In general terms, the sweet-spot typical of class C operation moves higher in input power level as the input bias voltage, VI, is reduced, or when diminishing the conduction angle, in the search for higher collector/drain efficiency. Using the same criteria, the dips in the dual class AB sweet-spot come closer to each other as V_I is increased, up to a point where they collapse into a fairly flat region in the IMD versus P_{in} characteristic, just before disappearing. In Figure 2, the evolution of the input power level IMD sweet-spot position with V_{GS} and V_{DS} bias, as measured in a microwave HEMT device, supports the dependence previously described.

When increasing the drain-to-source voltage, two effects can be perceived. Firstly, the widely known FET pinch-off modulation, usually described by $V_p = V_{po} + \gamma V_{DS}$, produces an expected shift to the left in the $G_3(0)$ null (ideal class B), and, consequently, in the actual PA conduction angle. In this sense, an amplifier working in a soft class C condition (i.e. close to class B) could, for example, become a class AB, due to V_{DS} induced pinch-off modulating action. Secondly, as V_{DS} is enlarged,

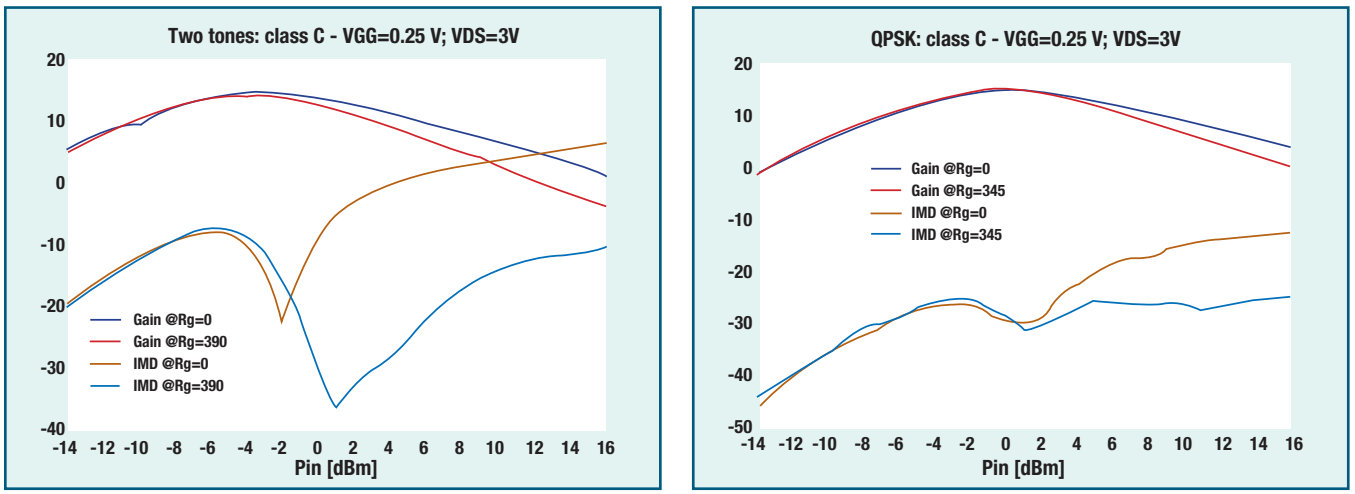


Figure 3: Gain and IMD power level for two-tone and QPSK signals in a class C quiescent point, with and without a resistor in the gate DC path

the load-line is moved to the right in the device output I/V characteristics, which is why a higher input amplitude excursion would be required to reach the current saturation (to enter the triode region). This usually results in a higher sweet-spot position, in terms of input power level, for the same gate-to-source voltage.

When operating on real modulated signals, the deterministic envelope amplitude of the two-tone excitation is substituted by one with a complex statistical distribution, and the deep sweet-spot notches are replaced by smoother IMD valleys. Nevertheless, they keep the same dependence on the PA quiescent point.

As a result of studying this behaviour, a modification of the IMD versus P_{in} profile seems to be feasible. In [7], the possibility of adjusting both biasing voltages with the input power level, to assure the simultaneous existence of a class C type large-signal IMD sweet-spot and constant gain, over a certain excitation amplitude range, was considered. However, as this type of bias adaptation [8], or envelope tracking technique [9], can be quite complex, in particular when a control over the drain voltage is required, other simpler solutions are desirable in mobile handset PAs.

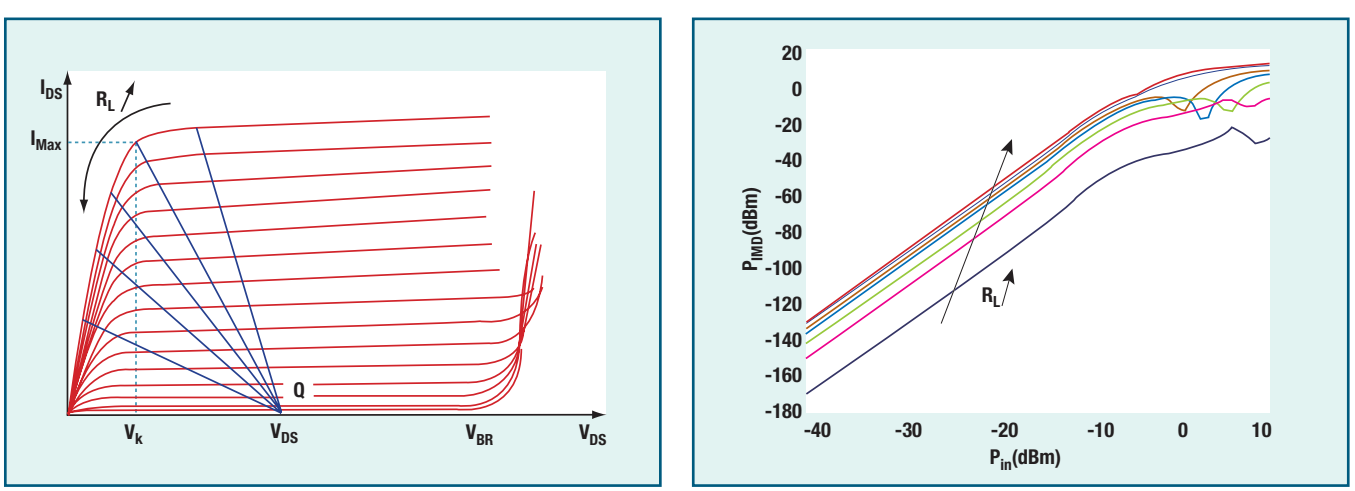
Following this idea, the introduction of a resistor with an appropriate value into the gate DC path has been proposed in [10]. As the input power is increased, the DC current resulting from gate rectification may be used as a sample to adjust "dynamically" V_{GS}, transforming the single class C type sweet-spot in a dual or wider one, see Figure 3.

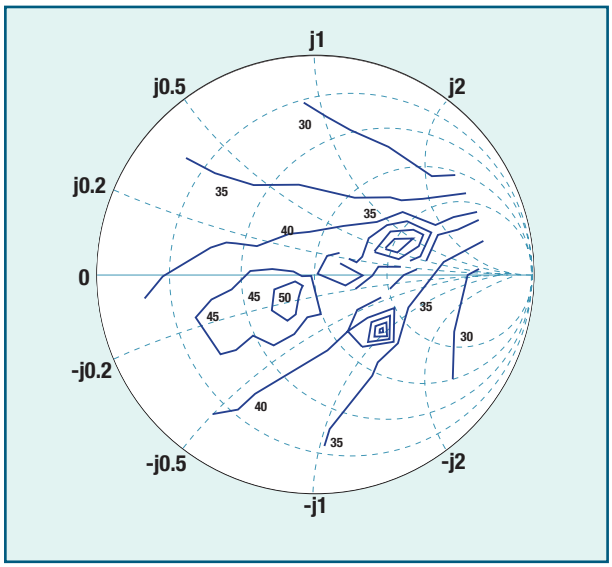
Another important design issue is that concerning the impedance terminations. An interesting conclusion that can be drawn from the observation of Figure 4(a) is that, for simple memory-less active devices (i.e. without any reactive components), if there are conditions for the appearance of an IMD sweet-spot, then every intrinsic resistive load generates its own IMD minimum [3]. As complemented by Figure 4(b), the variation of the load-line slope only affects the value of input power for which the minimum appears.

In case the device reactive parasitics are included, the selection of the optimum load should be the one that, after de-embedding those parasitics, presents a pure resistive load to the current source of the intrinsic device. As seen in the simulated load-pull chart of Figure 5, this corresponds to a slight rotation of the Smith chart real-axis [3]. The impedances obtained in this way define the contour points of highest intermodulation distortion ratios, IMR (ratio of output signal to two-tone IMD power).

Surprisingly enough, out-of band component terminations, mostly at the second harmonic and baseband, have a tremendous impact on IMD sweet-spots [5], [11].

Figure 4: Impact of resistive PA load resistance, R_L, on large-signal IMD sweet spots





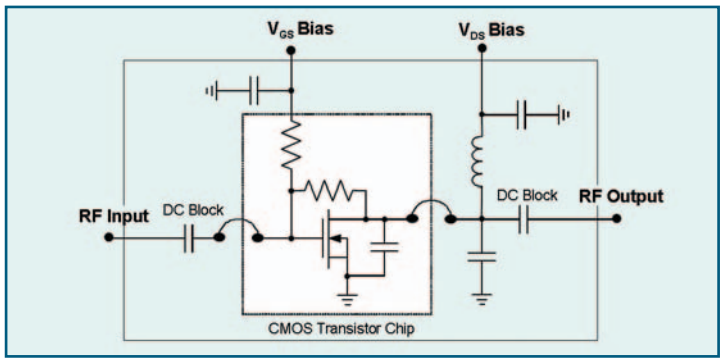


Figure 6 (above): Simplified circuit schematic of the Si MOSFET based PA prototype

Figure 5 (left): Load-pull simulation of IMR values at the IMD sweet-spot

To understand this, we need to realise that either the components at $2\varpi_2$ can be mixed again, in an even order nonlinearity, with the fundamentals at ϖ_1, or the components at ϖ_2-ϖ_1 can be remixed with ϖ_2, to generate new in-band IMD products. So, any residual phase in those indirectly generated IMD products (eventually caused by improper bias circuit design at ϖ_2-ϖ_1 or by uncontrolled second harmonic resonances at $2\varpi_2$), in comparison to the ones that arise directly from the odd order device nonlinearities, will produce quadrature IMD components. These can escape from perfect small-signal/large-signal distortion cancellation, and thus jeopardising the anticipated IMD sweet-spot benefits [5], [11].

Power amplifier example

In order to illustrate the above general considerations, we will now show nonlinear distortion, output power and power added efficiency data, measured from a practical PA prototype. The amplifier, a Si MOSFET based broadband 950MHz PA, was manufactured in a 0.6µm CMOS process from the Austria Mikro Systeme (AMS) foundry. Although built over a previous constant envelope class-F chip design [12], with a device of low total gate periphery and in a low cost standard CMOS process, this PA prototype can already give us a picture of the intermodulation issues above discussed. Its simplified schematic is shown in Figure 6.

The class AB dual IMD sweet-spots, that were predicted and explained using the preceding large-signal analysis, are clearly visible in the two-tone IMD measurements of Figure 7. Solid lines represent circuit simulations during the design phase while dots constitute the laboratory measured data.

Figure 8 shows the output spectrum from a 950MHz CDMA input signal, with a channel bandwidth of 1.2MHz and a chip rate of 1Mchip/s. The adjacent channel power ratio (ACPR) was measured versus output power for different classes of operation, as shown in Figure 9.

Although the statistical nature of the CDMA input signal amplitude tends to wipe out the IMD sweet-spot characteristics, it is obvious from Figure 9 (a) that the mechanisms responsible for the dual sweet-spots seen in the two-tone case are still causing an increased ACPR for class AB operation. The class AB amplifier can in fact deliver higher output power than class A for an ACPR specification of 35dB.

Figure 7: Measured output power, Pout, and in-band intermodulation distortion, IM3

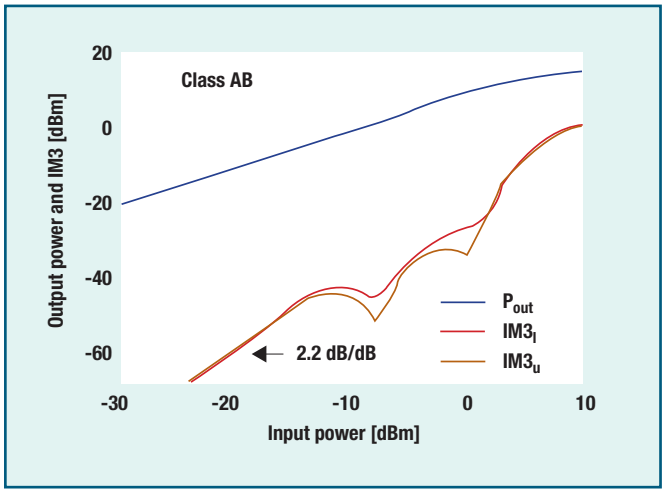
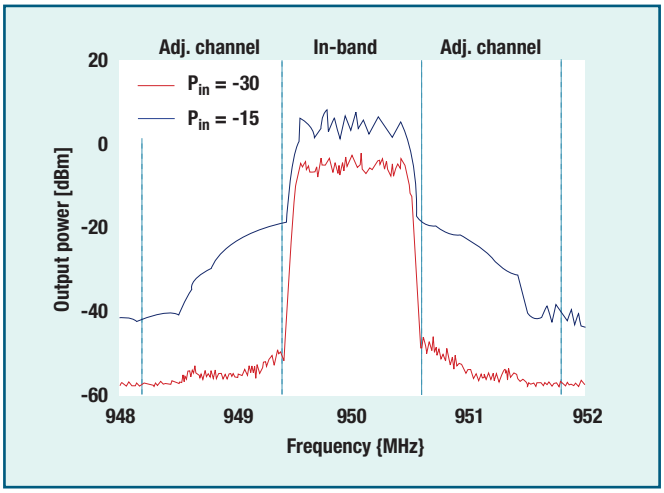


Figure 8: Two output power spectra for different input powers of a 950MHz CDMA excitation



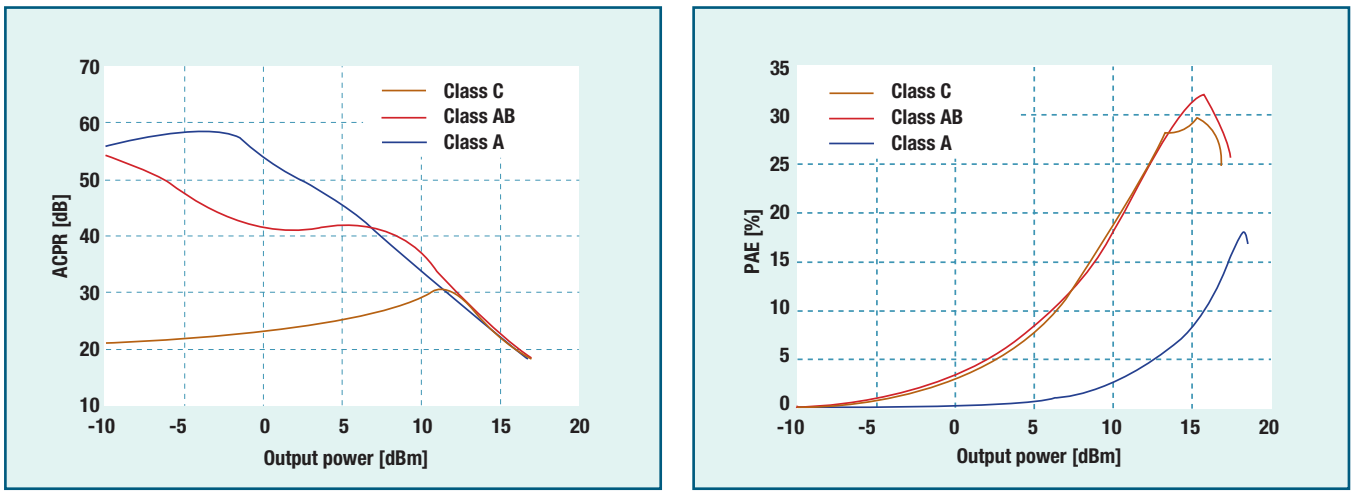


Figure 9: (a) Measured adjacent channel power ratio (ACPR), and (b) measured power added efficiency (PAE), vs. output power for different classes of operation

The advantage offered by the dual IMD sweet-spots in class AB becomes even more evident when comparing measured power added efficiency (PAE) for a given desired ACPR spec as shown in Figure 9 (a) and (b). The maximum PAE under the 35dB linearity requirement is about 20% in class AB, whereas for class A it is below 5%.

Conclusions

This article presented an outline of the physical origins of the IMD sweet-spots, associating them to the typical form of input-output active device transfer characteristics. By explaining on what PA design parameters they depend, practical hints were given to control them. That turned these scientific interesting IMD nonlinear properties into useful tools for practical highly linear and efficient handset PA design.

Acknowledgments

The authors would like to thank Dr F Fortes and Prof M J Rosário from IT-Lisbon for providing the CMOS PA chips.

This work was supported by the EU and carried out under the Network of Excellence Top Amplifier Research Groups in a European Team - TARGET contract IST-1-507893-NOE.

References

[1] *C. E. Weitzel, "RF Power Devices for Wireless Communicationsî, 2002 IEEE Int. Microwave Symp. Dig., pp.285-288, Seattle, Jun. 2002.*

[2] *S. C. Cripps, RF Power Amplifiers for Wireless Communications, Artech House, Inc., Norwood, 1999.*

[3] *J. C. Pedro and N. B. Carvalho, Intermodulation Distortion in Microwave and Wireless Circuits, Artech House, Inc., Norwood, 2003.*

[4] *M. Schetzen, "Nonlinear System Modeling Based on the Wiener Theory", Proc. of the IEEE, vol. 69, pp.1557-1573, Dec. 1981.*

[5] *N. B. Carvalho and J. C. Pedro, "Large and Small Signal IMD Behavior of Microwave Power Amplifiers", IEEE Trans. on Microwave Theory and Tech., vol. MTT-47, pp.2364-2374, Dec. 1999.*

[6] *C. Fager, J. C. Pedro, N. B. Carvalho, H. Zirath, F. Fortes and M. J. Rosário, "A Comprehensive Analysis of IMD Behavior in RF CMOS Power Amplifiers", IEEE Journal of Solid State Circuits, vol. JSSC-39, pp.24-34, Jan. 2004.*

[7] *E. Malaver, J. A. García, A. Tazón and A. Mediavilla, "Characterizing the Linearity Sweet-spot Evolution in FET Devices", European Gallium Arsenide and other Compound Semiconductor Application Symp., Munich, Oct. 2003.*

[8] *G. Hanington, P. F. Chen, P. M. Asbeck and L. E. Larson, "High-efficiency Power Amplifier using Dynamic Power Supply Voltage for CDMA Applications,"IEEE Trans. Microwave Theory and Tech., vol. MTT-47, pp. 1471-1476, Aug. 1999.*

[9] *J. Staudinger, "An Overview of Efficiency Enhancements with Application to Linear Handset Power Amplifiers," IEEE RFIC Symp. Dig., pp.45-48, Seattle, Jun. 2002.*

[10] *C. Gómez, J. A. García and J. C. Pedro, "IMD Sweet-spot Control on Junction FET Devices using a Gate Bias Resistor," accepted for presentation in the 34th European Microwave Conf., Amsterdam, Oct. 2004.*

[11] *N. B. Carvalho and J. C. Pedro, "A Comprehensive Explanation of Distortion Sideband Asymmetries", IEEE Trans. on Microwave Theory and Tech., vol. MTT-50, pp.2090-2101, Sep. 2002.*

[12] *F. Fortes and M. J. Rosário, "A Second Harmonic Class-F Power Amplifier in Standard CMOS Technology," IEEE Trans. Microwave Theory Tech.,vol. MTT-49, pp. 1216ñ1220, Jun. 2001.*

AUTHORS INFORMATION

[1] **Institute of Telecommunications**
University of Aveiro
3810-193 Aveiro
Portugal

[2] **Microwave Electronics Laboratory**
Department of Microtechnology and Nanoscience
Chalmers University of Technology
SE-41296 Göteborg,
Sweden

[3] **Department of Communications Engineering**
University of Cantabria
39005 Santander
Spain