

EDGE modulation – how linearization improves amplifier performance

The design and implementation of a novel polar loop linearizer will improve efficiency and reduce distortion in power amplifiers for EDGE modulation

By **T.J. Fergus**

Historically, the global system for mobile Communications (formerly Groupe Speciale Mobile – GSM) was intended to make power amplifier design simpler. Typical GSM base transceiver system (BTS) power amplifier (PA) efficiencies are in the order of 40 percent. The simplified amplifier

With the introduction of high order modulation schemes for 2.5 and 3G, the design challenges presented to the PA have increased significantly, for both base stations and mobile applications. The 2.5G (enhanced data rates for GSM evolution – EDGE) and 3G (wideband code-division multiple access – WCDMA) standards have significant RF envelope variation due to the high order modulation schemes employed. They require linear amplification to minimize error-vector magnitude (EVM) distortion and limit adjacent channel power (spectral re-growth).

Clearly this requires highly linear power amplifiers operating in class A or AB, depending on the application. Typically, the required linearity is achieved by backing the PA away from compression. This, however results in poor efficiency (the base-station delivers less RF power and dissipates significant energy as heat). Typical efficiency for an EDGE PA is in the order of 15 percent (for 2.5 percent EVM) for laterally diffused metal oxide silicon (LDMOS) transistors.

This article describes the design and implementation of a polar loop (closed loop phase and amplitude feedback system) to improve the efficiency of a gallium-arsenide field effect transistor (GaAs FET) power amplifier operating at 1.91 GHz.

This approach offers the following potential benefits when applied to an RF amplifier.

- Significant improvement in efficiency and RF power handling for low distortions (For 2.5 percent EVM, improved power-added efficiency (PAE) from 13.9 percent to 18.1 percent and output power capability increased by 1.1 dB.
- Minimum complexity due to novel architecture (direct comparison at RF without mix down to IF).
- Gain and phase compensation for amplifier aging, tolerance and temperature.

The techniques presented here are aimed at base-station design where it is more beneficial to apply feedback in this manner. There is no reason why these techniques could not be used on mobile products given sufficient integration. Although applicable to most medium-bandwidth modulation schemes the experiments presented are for GSM EDGE. A GaAs FET PA was chosen over LDMOS to show clearly the benefits of the polar loop in this application.

Feedback and distortion

All amplifiers distort even at low levels. These distortions frequently become problematic at high drive levels. As an amplifier approaches compression, it imparts both phase and amplitude errors onto the signal passing through it. The result is AM to AM and AM to PM distortion, which generally gets worse as we approach compression. It is highly dependent on the amplitude of the signal and is directly related to the amplitude component of the modulation.

The effectiveness of any solution is dependent on the amplifier's characteristics, which should be deter-

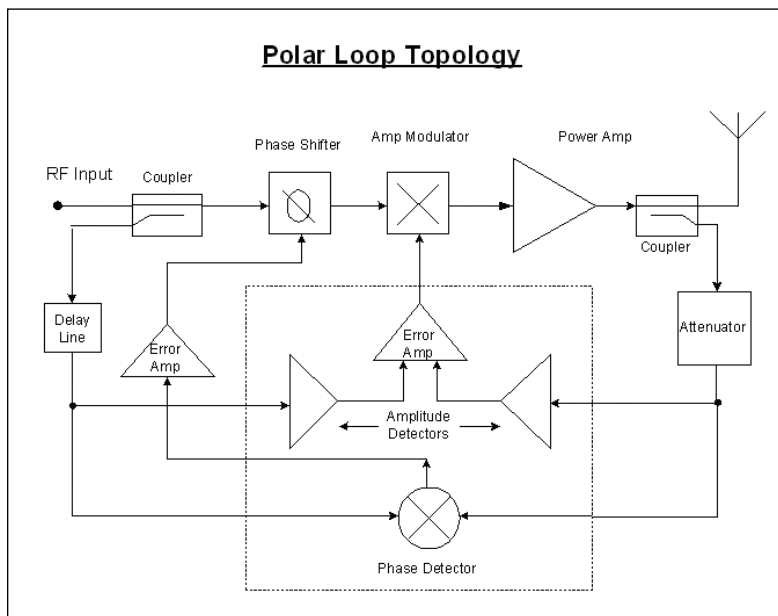


Fig. 1. Block diagram of the polar loop.

design requirement results from the use of Gaussian minimum-shift keying (GMSK). This modulation scheme maintains a constant RF envelope with only the phase of the signal varying with time.

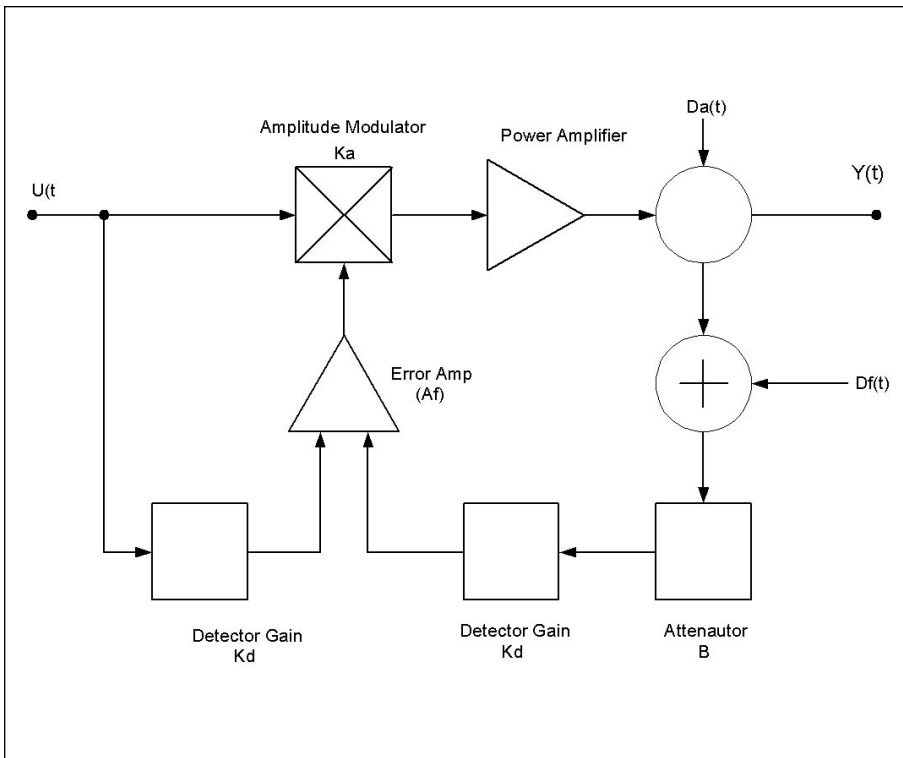


Figure 2. Loop analysis flow diagram.

mined in the first instance. If the amplifier exhibits a gentle compression characteristic then there is a good chance of improving efficiency by linearization.

The technique of using negative feedback to linearize power amplifiers is well known. Although the standard principles still apply, it is not a simple matter to implement negative feedback at RF frequencies when time delays in circuit elements are significant. The dangers of possible instabilities can be minimized using a thorough design process. This limits the bandwidth of such techniques – in effect they correct for the RF envelope distortion rather than the RF cycle itself. In other words it is the intermodulation products being corrected rather than the harmonics. This implies the need to demodulate the RF signal to determine the distortion artefacts. The approach adopted here does not need to demodulate the signal and does a direct comparison at RF.

For the purpose of this article, the application of feedback will be called “linearization” or “correction.”

Linearization techniques

The level of distortion in a typical GSM/EDGE PA is comparatively low – the devices used offer good linearity (normally LDMOS) at the onset. The use of open loop correction is unlikely to deliver the performance required to meet the requirements of EDGE in terms of EVM and adjacent channel

power (ACP). Adaptive predistortion (baseband) is complicated and requires complete system integration. It cannot fully compensate for dynamic amplifier distortions and is approximate.

Closed loop systems can deliver significant performance advantage provided reasonable care is taken over their implementation. There are two main tech-

niques used: feedforward and feedback. Feedforward requires the use of a second amplifier, which is an unwanted complication and requires careful calibration.

The polar loop is an elegant feedback system that applies the correction signals before the amplifier in the forward path using phase and gain modulators. This article discusses such a loop operating at 1.91 GHz. It further presents the mechanisms by which the loop operates with corresponding design compromises. Theoretical analysis is presented to the reader while the practical considerations of loop bandwidth, correction efficiency, stability and adjacent channel power are discussed in detail.

The polar loop

The polar loop addresses both phase and amplitude distortion with two separate feedback loops. This implementation of the polar loop is novel in the sense that the phase and amplitude comparison are done at RF. Generally, polar loops require a mix down to an intermediate frequency or baseband itself (this eases the detection process). Recent advances in analogue ICs have made phase and amplitude detection simpler at RF frequencies.

The benefits of the device are used in the application of this novel polar loop. A simplified block diagram of the polar system is shown in Figure 1.

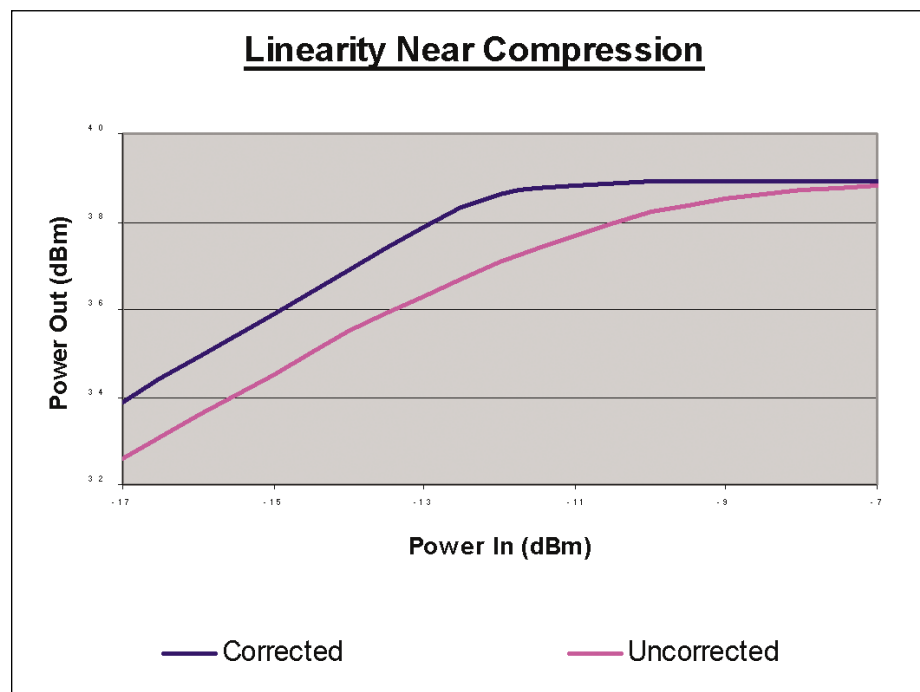


Figure 3. Plot of linear performance near compression.

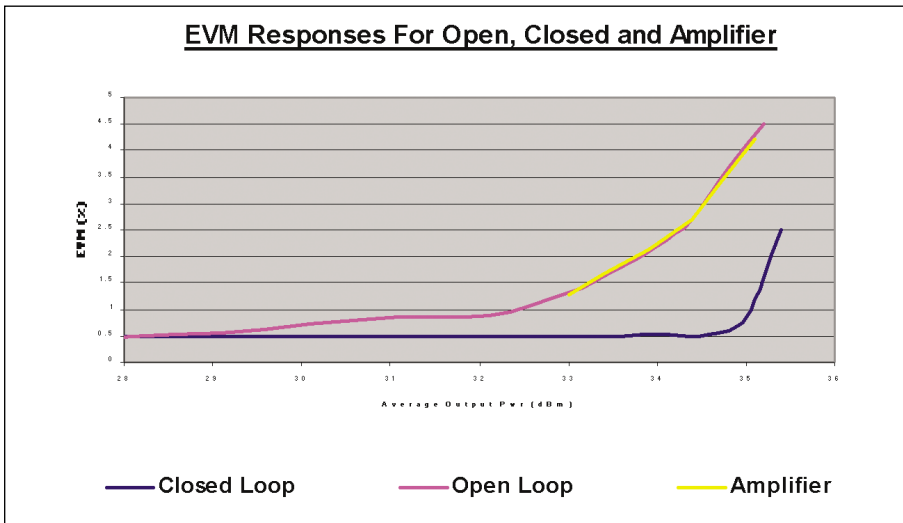


Figure 4. Measured response of the amplifier.

The theory of operation is simple. The phase and amplitude errors are detected by comparison between the reference and output signals. The amplitude error is derived from the difference between the reference and feedback signal envelopes. The detected phase error is the actual phase between RF reference and feedback cycles.

These errors are amplified (or integrated) and used to drive opposing modulators in the forward path. Therefore the loops attempt to minimize the amplifier distortions by cancelling them in the forward path.

A major problem is time delays in the system. In effect, we are comparing a reference (no delay) with a feedback signal (delayed by the forward path and amplifier and feedback circuits). It is possible to equalize this delay using a delay line in the reference path so that the comparison takes place at the right time.

The resultant error is now being applied to the forward path after the event has occurred! This fundamentally limits the correction bandwidth. The delay line can be omitted for narrow band frequency operation when the phase variation is smaller and can be accommodated by the loop. Generally, the modulation bandwidth is much lower than the RF carrier frequency so that the envelope and phase variations over the time delay in the circuits are minimal.

As a general rule of thumb, the correction bandwidth needs to be five times the modulating signal bandwidth. For example, assume we have a

GSM signal with a bandwidth of 200 KHz. The correction loop bandwidth would need to be in excess of 1 MHz to offer a significant benefit.

Amplitude loop mathematics

The following section uses mathematical analysis to predict the behavior of the polar loop.

Where:

$D_a(t)$ are the distortions generated in the amplifier;

$D_f(t)$ are the distortions due to the feedback path;

A is PA gain;

K_a is the amplitude modulator;

A_f is the feedback amplifier;

K_d is the amplitude detector gain;

$U(t)$ is the input signal;

$Y(t)$ is the output signal.

Next, performing the analysis:

$$Y(t) = A \cdot (U(t) + F(t)) + D_a(t) :$$

Where $F(t)$ = amplitude error signal.

$$F(t) = K_d \cdot A_f \cdot K_a \cdot (U(t) - B \cdot Y(t) + B \cdot D_f(t)) :$$

Let $K = K_d \cdot A_f \cdot K_a$, and therefore:

$$Y(t) = A \cdot \left(U(t) + \left(K \cdot \left(\begin{matrix} U(t) - B \cdot Y(t) \\ + B \cdot D_f(t) \end{matrix} \right) \right) + D_a(t) \right) :$$

$$Y(t)(1 + ABK) = A \cdot U(t) + AK \cdot U(t) + ABK \cdot D_f(t) + D_a(t)$$

$$Y(t) = \frac{A(1 + K)U(t)}{1 + ABK} + \frac{ABK \cdot D_f(t)}{1 + ABK} + \frac{D_a(t)}{1 + ABK} :$$

Hence, assuming $ABK \gg 1$ and $K \gg 1$, then:

$$Y(t) = \frac{U(t)}{B} + D_f(t) + \frac{D_a(t)}{ABK}$$

Generally the phase loop can be analyzed in the same manner.

Conclusions drawn from analysis

“B” sets the closed loop gain. Any distortion caused by the PA is reduced by the loop gain ABK . A loop gain of 20 dB should give a reduction of 20 dB in any intermodulation products falling within the loop bandwidth.

The feedback loop cannot distinguish between PA distortions and non-linearities in the reference and feedback paths. It is important to ensure that the correction circuits add minimal distortion — applying correction to any system results in inherent distortion due to the correction loop itself. It can only improve matters if the distortion in the PA is much greater than that inherent in the closed feedback loop itself. This is a low pass effect and can be ignored outside the loop bandwidth.

Any distortion introduced in the forward path (due to modulators) will be significantly reduced by the feedback. It is, however, sensible to ensure that the distortion caused by the modulators is low otherwise we are correcting for the modulators rather than the PA.

The phase margin of the loop sets the maximum dynamic phase variations that can be corrected and still ensure stability. This is affected by variations in PA gain (assumed reasonably linear for EDGE) and the accuracy of the static phase shifter.

As the amplifier starts to distort, the additional phase distortion directly detracts from the phase margin of the loop.

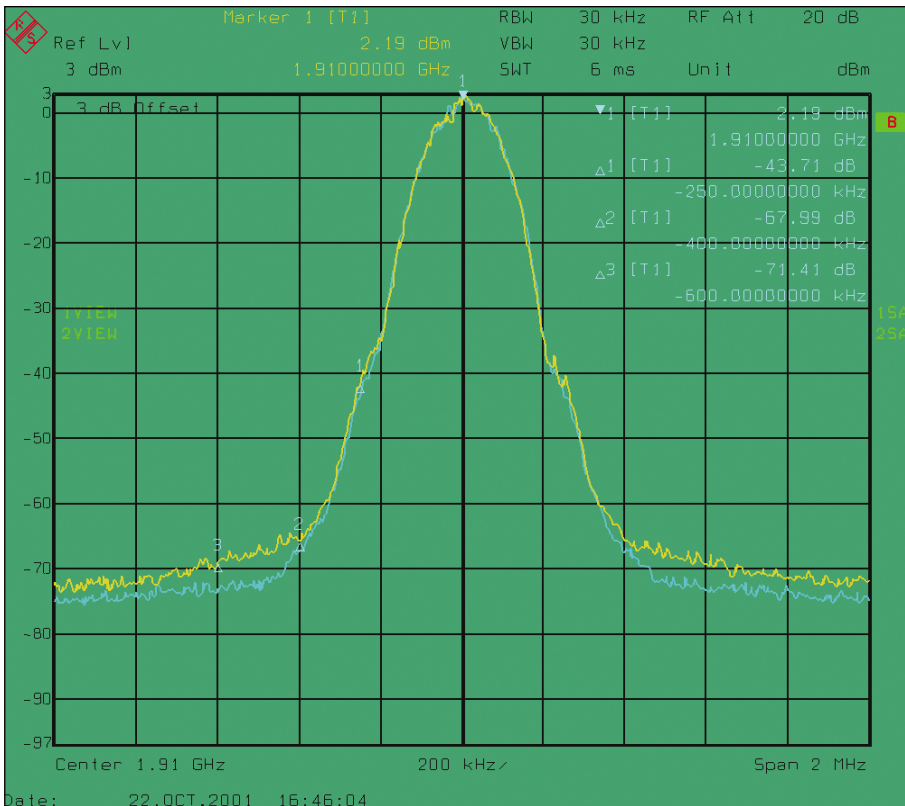


Figure 5. Amplifier plot: Blue = open loop, yellow = closed loop.

Ideally, the phase and amplitude loops should be orthogonal. Any cross-coupling may detract from phase margin.

Practical considerations

The results presented here are for a 3-stage lineup operating at 1.91 GHz. The particular lineup for this discussion consists of the following devices: A NGA386 (Stanford Microdevices); a CMM1321 (Celeritek) and a MGF0909 GaAs FET (Mitubishi).

This is a realistic scenario where the forward path modulators can be applied before the NGA386. The low signal levels mean that the forward path components add little distortion to the complete lineup (good linearity) and little power is wasted (insertion loss).

The DC power consumption of the PA is 12.5 W (10 VDC at 1.25 A controlled by active bias). The main polar loop components are an integrated amplitude modulator based on a 3 dB, 90° hybrid and an AD8302 gain/phase detector.

Each circuit was constructed then tested for performance. Critical issues for these devices were obtaining sufficient modulation bandwidth while maintaining minimal cross-coupling with the other loop. Typical modulation bandwidths achieved were in the order of 15 MHz.

Initially the loops were designed

using traditional theory to ensure stability with adequate phase margin. This set the maximum loop bandwidths that could be safely set without danger of oscillation (these were approximately 2 MHz wide). This represents the maximum performance for error correction of phase and amplitude and results in

the best EVM improvements.

A switch was provided for each loop to select very low loop bandwidth — effectively turning the correction off but centering the loops. The amplitude loop was locked, then the phase using length of coax to center the loop at the midpoint.

Linearity

Once locked and centered, the linearity of the closed loop was measured and is shown in Figure 3.

It is easy to see the effect of the amplitude correction loop on the lineup characteristic (the curves have been offset for clarity). The lineup manages a compression point of 37.6 dBm, which increases this to approximately 38.5 dBm when the linearizer is turned on. Effectively we have traded a soft knee characteristic for a more linear slope with a hard limit. This is a static test and does not verify the loop bandwidths necessary for modulated signals. However, the potential improvement is obvious.

Error vector measurement

To verify the operation of amplitude and phase loops a modulated signal (EDGE) was used. An important parameter of the EDGE signal is the EVM, which gives a measure of the phase and amplitude distortion of the EDGE signal. The measured results are shown in Figure 4. The three traces

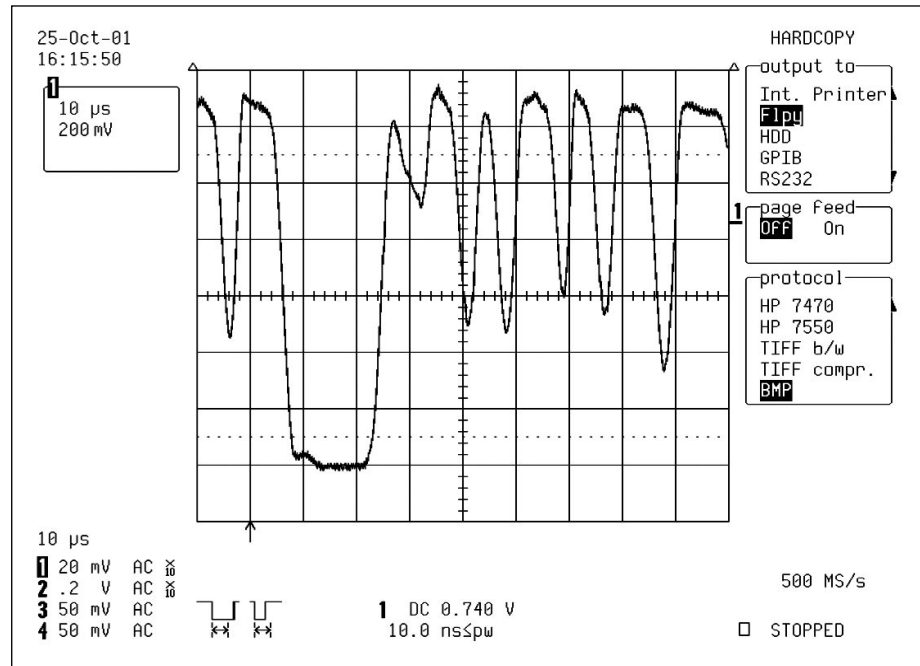


Figure 6. Plot of the amplitude error signal with envelope signal clipping.

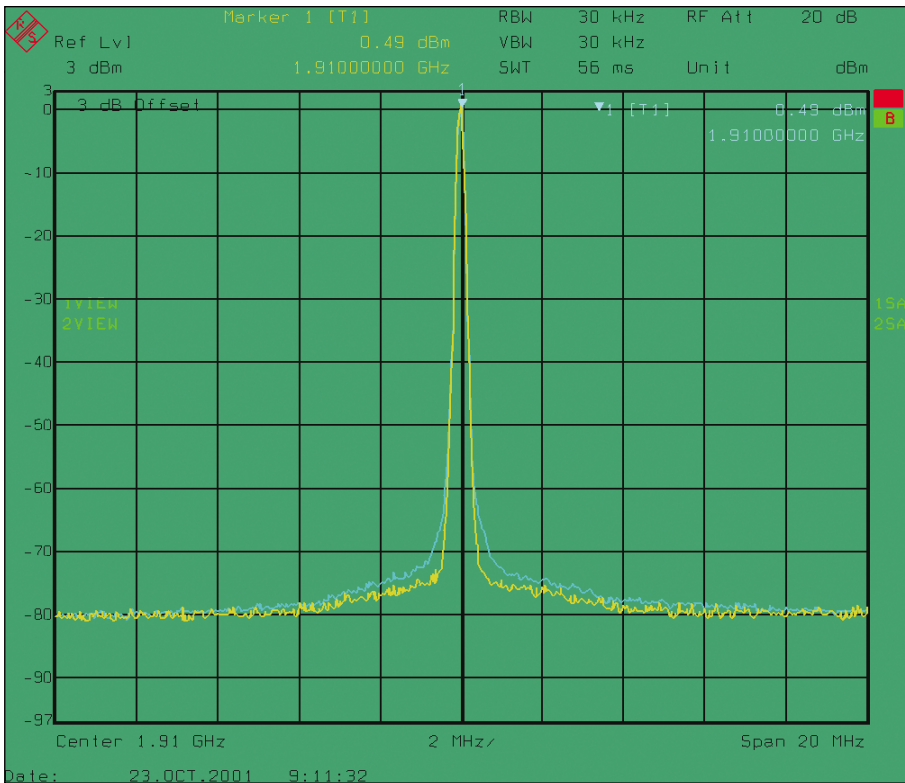


Figure 7. Loop wideband noise performance: Yellow = lineup on its own; Blue = closed loop system.

show EVM for closed loop (polar loop on) open loop (lineup with phase and amp modulators in forward path) and lineup on its own.

The open loop system shows the polar modulators add little distortion to the

lineup in open loop mode. With the polar loop on, the loop holds the EVM very low until it finally increases at approximately 34.8 dBm. The 0.40 percent EVM is the residual present in the test kit.

The potential gains are very good if

the required system EVM is very low (roughly 6.0 dB more output power for 0.5 percent EVM). The main EVM improvement comes from the amplitude loop, however, the phase loop needs to be operational to achieve sub 1 percent EVM. The benefits decrease if higher EVM values are specified. For a target EVM of 2.5 percent we can obtain 1.3 dB more output power from the PA with the linearizer active.

ACP due to modulation

The widest loop bandwidths and highest loop gains give the maximum benefit for EVM. However, this results in increased ACPM and wideband noise. This is an intuitive conclusion as the loop adds noise and distortion that spreads out within its bandwidth and appear at the PA output. In reality, there is a balance to be struck with loop bandwidth and gain to achieve the best compromise and that depends on the modulation, symbol rate and system specification.

The loop bandwidths were optimized for EDGE operation that gave the best EVM performance while still meeting ACPM specifications.

This gives the best EVM while meeting ACPM. The loops cannot be narrowed further, otherwise the loops will be too slow and will not be able to find lock in the ramp up period when frequency hopping. Figure 5 shows the close in ACPM plot for the amplifier producing +33.9 dBm.

It can be seen that the loop raises the modulation skirts however they still meet the requirement in the GSM specification 05.05. ACPM is affected almost entirely by the amplitude loop. In this situation, as the system starts to clip, the growth of ACPM is much greater for the linearized case. This can be easily understood by looking at the amplitude error signal. As EDGE has significant envelope variation, the amplitude loop tries to produce a faithful copy at the output.

As the input drive is increased there comes a point when the amplifier can no longer deliver the peak powers (hard clipping) and the amplitude loop clips.

Figure 6 shows a plot of the amplitude error signal — signal clipping on the envelope peaks can be clearly seen. (Note that a negative voltage gives max gain in this instance).

Clipping is a highly nonlinear action producing spectral spreading and poor ACPM. Clipping should not be allowed

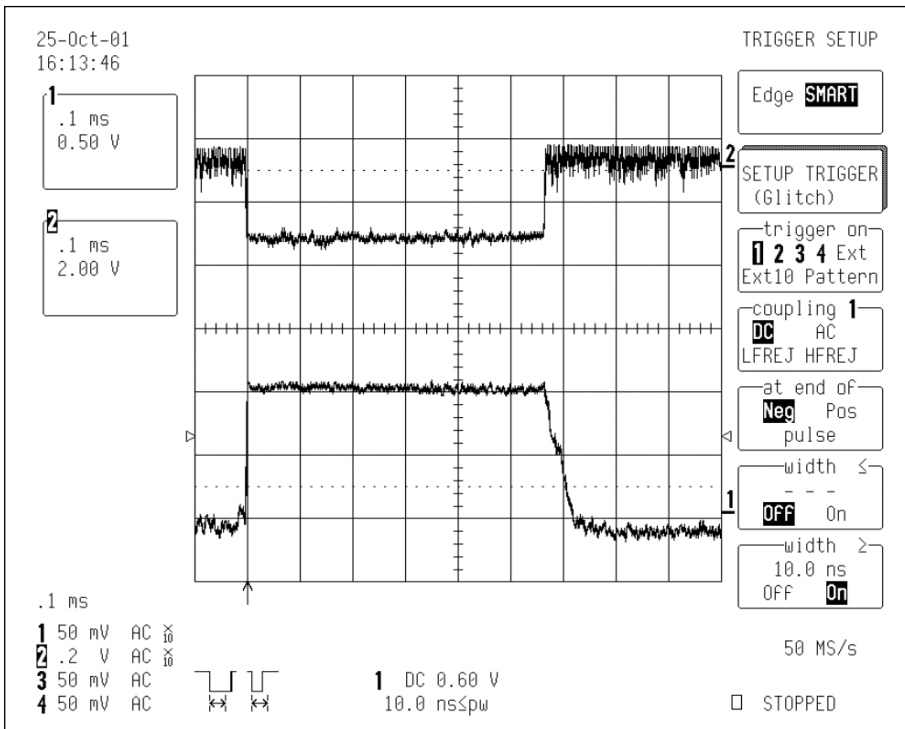


Figure 8. The loop response to framed EDGE data.

to occur in normal conditions.

The phase loop does not exhibit this behavior provided the range is sufficient and correctly centered. After reducing the loop gains and bandwidths, the EVM improvement is not as marked as before, reducing an uncorrected 7.5 percent down to 2.5 percent (the amplifier can now be driven harder and still meet ACPM specs). This represents a 1.1 dB improvement in power over the basic lineup producing 2.5 percent EVM. This is the trade off between EVM and ACPM.

Both loops eventually lose lock at low power levels. During these periods its important that the loops do not saturate, otherwise there will be an extended delay before the RF returns with the next TX burst. A simple diode arrangement can prevent saturation of the error amplifiers.

This system offers 55 dB of closed loop dynamic range (before phase loop fails).

Wideband noise

The loop wideband noise performance is shown in Figure 7 for the amplifier (with linearizer on) delivering +33.9 dBm.

It can be seen that the noise floor increases for the closed loop case within the loop bandwidths. The effect is quite symmetrical, which is a good sign. Often, if the loop stability is marginal or there is loop cross-coupling between phase and amplitude loops, the wideband noise can have asymmetrical humps appearing above the floor. The performance here is good enough for a GSM BTS and does not display any signs of instability.

Hopping/static operation

The speed of operation is a concern. For example, in GSM, the guard period between slots is 10 μ S. Therefore the loop needs to be able to attain lock in this period when frequency hopping between slots. Therefore this impacts on the bandwidth that the correction loop must achieve. The feedback loops must not be allowed to saturate when the RF signal is not present. Figure 8 shows the loop response to framed EDGE data (one slot on, 7 slots off).

In burst mode, the loop lock time is good (less than 10 μ s), with EVM being maintained at 2.5 percent at the rated power. The falling edge is slow — the RF has ramped down and the amplitude drifts off due to offsets in the feedback and reference comparison signal.

Harmonics and filtering

As proven by the earlier analysis, any distorting element in the feedback path will result in output distortion. There are some instances when there are needs for filtering. For example, harmonics of the amplifier could enter the feedback path and confuse the detection circuits without appropriate filtering. Any AM/PM distortion in this filter will distort the PA output.

In this instance a filter is not required due to the high linearity demanded by the application.

Conclusions

This article discussed the design and implementation of a polar loop design using direct comparison at RF. The GaAs FET PA tested in this paper proved a suitable candidate for linearization with a good improvement in efficiency for a given EVM. However, similar results can be expected with any manufacturer's devices with similar parameters.

The measurements conducted showed the trade offs required to satisfy GSM EDGE 05.05 specification. Clearly this technique works well for narrow and medium modulation bandwidths (such as EDGE) and can achieve very low distortion performance acceptable for BTS applications. The implementation of the system minimizes the overheads associated with such linearization techniques while optimizing performance.

The application of closed loop polar feedback has been shown to increase the PAE of the GaAs FET amplifier. The potential benefit is more marked for lower EVM limits.

For 0.6 percent EVM, PAE increased from 5.2 percent to 14.7 percent. For 2.5 percent EVM, PAE increased from 13.9 percent to 18.1 percent with an increase in output power capability of 1.1 dB.

The performance improvement depends on many parameters including the amplifier linearity, amplifier delay, EVM, ACPM, modulation type, symbol rate etc. These need to be considered in addition to the cost overhead, additional power consumption, PCB area and general complexity. Only then is it possible to determine the business case for such a system. Undoubtedly, the obvious gains will be for high power BTS. With suitable integration and the ever-increasing

demand for battery lifetime this technique may prove useful on mobile handsets in the future.

RF

References

[1] GSM 05.05 "Radio Transmission and Reception"

[2] S. C, Cripps "RF Power Amplifiers For Wireless Communications" Artech House 1999 ISBN 0-89006-989-1

[3] P B Kenington "High Linearity RF Amplifier Design" Artech House 2000 ISBN 1-58053-143-1

[4] V Petrovic and W Gosling "Polar Loop Transmitters" Elec. Letters, 10 May 1979, Vol 15, pp 286 – 287

[5] Analog Devices AD8302 Datasheet.

About the author

The author would like to thank all of his colleges in the wireless technology practice for their assistance and helpful comments during this research.

Tim Fergus is a consultant working at the Wireless Technology Practice of PA Consulting Group, Melbourn, Hertfordshire, United Kingdom. He holds a 1st Class B.Eng in Electronic and Electrical Engineering from Brunel University, Uxbridge, Middlesex, in 1993. Tim has worked in consultancy for more than six years specializing in RF and high speed analogue electronics. He has worked on a broad range of communication systems from bespoke formula 1 telemetry systems through to digital cellular systems (handset and base station) such as GSM, DECT and PCN. He also has worked for a wide range of clients providing both system and circuit level design expertise. He can be contacted at *Timothy.Fergus@paconsulting.com*