Increasing the Talk-Time of Mobile Radios with Efficient Linear Transmitter Architectures

Stephen Mann\textsuperscript{1,2}, Paul Warr\textsuperscript{1}, Mark Beach\textsuperscript{1}, Joe McGeehan\textsuperscript{1}

\textsuperscript{1} University of Bristol, http://www.fen.bris.ac.uk/elec, steve.mann@bristol.ac.uk
\textsuperscript{2} PhD Student at the University of Bristol, sponsored by Tait Electronics Ltd, New Zealand

Abstract: In order to meet the linearity specification of standards such as GSM EDGE and TETRA the power amplifier (PA) linearisation methods in common usage today result in low DC to RF efficiency (e.g., Cartesian loop and adaptive predistortion). The transmitter places a heavy demand on the battery and therefore utilising more efficient transmitters can increase talk-time or reduce the battery size. For GSM EDGE and TETRA modulation it is shown that to obtain a transmitter efficiency (of \geq 50\%) the PA needs to be biased as class-C. The effect power amplifier efficiency has on talk-time is demonstrated for a handportable with fixed power overhead owing to the digital processing required to compile and transmit a burst or packet. It is shown that the PA efficiency for a transmitter operating backed-off (under power control), has a marginal impact on the talk-time and that the effect is largest at high output powers. This is put in context by a review of the various linearisation schemes such as Cartesian loop, polar loop, 2nd harmonic injection, envelope elimination & restoration (EER), predistortion methods, and synthesis techniques (e.g. LINC & CALLUM). The review also discusses which methods are not suitable for handportable design, and documents the strengths and weaknesses of the remaining methods. The paper shows that to increase the power efficiency (and thus talk-time), hybrid solutions to power amplifier linearisation are required. Finally, existing hybrid solutions and their shortfalls are discussed.

1. Introduction

Most commonly used linearisation schemes trade efficiency for linearity\textsuperscript{[1]}. The linearisation scheme can only apply a certain amount of linearity-correction after which the power amplifier (PA) must provide the remaining linearity. Thus if high linearity is required, the PA must achieve a certain linearity performance which comes at the expense of power efficiency. The extra linearity is obtained by backing-off the PA so that its output power is well below its peak power rating. Typically, for every dB of power back off an overall improvement of 2 dB is gained in the adjacent channel power ratio (ACPR). The GSM EDGE specification for linearity\textsuperscript{[2]} is easier to meet than that of TETRA\textsuperscript{[3]}, and as such could possibly be met by backing off a class-AB power amplifier, albeit at the expense of efficiency.

The current market demand (or trend) for a small lightweight handportable terminal is, to a large extent, determined by the battery and user interface requirements. To the end user the metric of power efficiency is not as meaningful or as well understood as the metrics of talk-time and battery size. Additionally GSM users have grown accustomed to the current talk-time available for GMSK modulation and the move to EDGE and its linear modulation may be unattractive if talk-time is adversely compromised. The transmitter places a large power drain on the battery and can dominate the length of talk-time for a given battery capacity. In section 3 it is shown that this is only true at high transmit powers where the transmitter’s power requirement dominates.

This paper is organised into 5 sections. Section 2 discusses transmitter efficiency as opposed to PA efficiency and where each is most suitably applied. Section 3 demonstrates how talk-time is related to PA efficiency at higher transmit powers. From sections 2 and 3 a framework for a high-efficiency transmitter is formed and section 4 summarises the linearisation methods in use today against this framework. Section 5 discusses the results of section 4 and suggests using hybrid architectures as a way forward to achieve a highly efficient linear transmitter. Section 7 concludes this paper.

2. Transmitter Efficiency

Two efficiency measurements used in PA design are drain (or collector) efficiency (DC supply efficiency) and Power Added Efficiency (η\textsubscript{PAE}). The DC supply efficiency (η\textsubscript{DC}) is a measure of the ability of the PA to convert DC power to RF power as delivered to the antenna and is defined in (1), where \( P_\text{o} \) is the RF power in watts delivered to the antenna and \( P_\text{DC} \) is the power in watts supplied to the PA. PAE is a measure of the signal power added by the PA and is defined by (2), where \( P_i \) and \( P_\text{DC} \) are as for (1) and \( P_\text{o} \) is the input signal power to the PA. Neither (1) nor (2) account for the power consumed in any of the driver stages or linearisation circuitry. These powers can be included by redefining \( P_\text{DC} \) in equation (2) as \( P'_\text{DC} = P_\text{L} + P_\text{D} + P_\text{DC} \), where \( P_\text{L} \) is the linearisation...
circuit power, and \( P_D \) is the driver stage powers. Using \( P_{DC}' \) in (2) defines transmitter PAE, where \( P_i \) is now the input signal power to the entire transmitter. In (2), assuming \( P_i \) is negligible and \( P_o = 1 \) W, an increase in \( \eta_{PAE} \) from 40% to 50% reduces \( P_{DC} \) from 2.5 W to 2 W, a power saving of 0.5 W. An increase from 50% to 60% reduces \( P_{DC} \) from 2 W to 1.66 W, a saving of 0.33 W. Thus, as transmitter efficiency increases above 50% the power saved diminishes whilst the PA design becomes increasingly difficult. Because of these reasons the metric used throughout this paper to compare linearisation methods shall be \( \eta_{PAE} = 50\% \). Equation (2) can be arranged to determine \( \eta_{DC} \) given \( \eta_{PAE} \), \( P_L + P_D \) and \( P_o \) as in (3).

Equation (3) is plotted in figure 1 for \( P_o = 1 \) Watt and \( P_i \approx 0 \). Meeting the target of \( \eta_{PAE} = 50\% \), requires a PA supply efficiency of 53% to 100% for \((P_L + P_D) = 125 \) mW to 1 W respectively. The case of \((P_L + P_D) = 250 \) mW, requires a 57% efficient PA and this is the favoured compromise between low-power linearisation circuitry and high efficiency PA.

Equation (4) gives the instantaneous efficiency of a class B biased, where \( V_{RF(out)}(t) \) is the amplitude of the RF signal at the PA output, and \( V_{DC} \) is the supply voltage to the PA. The peak efficiency is \( \eta_{DC} = 78.5\% \). For a varying envelope signal the efficiency is not constant, and the mean efficiency is a dependent on the envelope statistics. Figure 2 shows the pdf of EDGE and TETRA signals for randomly chosen symbols.

The time-average efficiency for each signal can be calculated using (5). This gives average efficiencies of 50.5% for EDGE, and 54.5% for TETRA. Practically it will be less as further back-off of the PA (from peak power) is needed to reduce the intermodulation distortion to acceptable levels. So to reach a PA efficiency of \( \eta_{DC} = 57\% \), the use of non-linear class amplifiers such as C, D or E are required. Note to use a class B PA, and achieve \( \eta_{PAE} = 51\% \) & 54% requires the linearisation and drive circuitry to draw only 39 mW and 148 mW respectively.
3. Effect of power control on talk-time.

The above examples have not taken account of power control or the fixed over-head power of the digital and baseband processing necessary to generate the data and control the radio. This includes tasks such as speech codec, error coding, protocol stacks and baseband processing. The relevant standard defines most of the processing requirements of these tasks and there is little opportunity to minimise this overhead with new techniques. It will require further evolution of VLSI technology in order to reduce the power consumption.

The effect that PA supply efficiency and power control has on talk-time is shown in Figure 3, for the following assumptions,

- A maximum of +30 dBm transmitter power, and minimum of 0 dBm power
- 1 Ah of useable battery capacity; the battery voltage is assumed to be 7.2 volts.
- Additionally, it is assumed that the overhead circuitry is supplied from 3 volts via a 90% efficient regulator and that the PA is supplied from the battery.
- Digital (and baseband) subsections and PA drivers that all consume 0.3 – 1.2 W, (i.e. everything but the PA), this will be referred to as the power overhead. An example of a possible power range is given in Table 1, where the lower limit is considered achievable using state of the art technology, and the upper limit is for readily available off-the-shelf parts.
- The last simplification is the assumption that the transmitter is on continuously.

<table>
<thead>
<tr>
<th>Task</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseband processor</td>
<td>25 mW</td>
<td>100 mW</td>
</tr>
<tr>
<td>Linearisation circuitry</td>
<td>75 mW</td>
<td>300 mW</td>
</tr>
<tr>
<td>DSP</td>
<td>75 mW</td>
<td>200 mW</td>
</tr>
<tr>
<td>RISC processor</td>
<td>75 mW</td>
<td>200 mW</td>
</tr>
<tr>
<td>Misc</td>
<td>50 mW</td>
<td>300 mW</td>
</tr>
<tr>
<td>Power regulator (efficiency = 90%)</td>
<td>30 mW</td>
<td>110 mW</td>
</tr>
<tr>
<td><strong>Total with efficiency</strong></td>
<td>330 mW</td>
<td>1210 mW</td>
</tr>
</tbody>
</table>

Table 1. Example power budget for a handportable mobile radio.

For unsaturated PAs the efficiency will decrease as they are backed-off. For the PA efficiency to remain constant, it would need to be operated in a saturated mode. If the PA supply efficiency was constant (or nearly constant), it is immediately obvious that 5 dB of power control increases the talk time dramatically. Therefore, as the mean output power is reduced, the talk time would still increase even though the transmitter efficiency reduces. The dashed lines in Figure 3 indicate the maximum (theoretical) efficiency for a class B amplifier determined using (4) and (5). For low RF output powers e.g. <20 dBm, the talk-time has only a weak dependence on PA supply efficiency as indicated by a flattening of the curves in Figure 3. At +15 dBm, the transmitter power is only 31.6 mW and for reasonable PA efficiencies (>10%) the dissipated power in the PA is small compared to the overhead power. The conclusion drawn from Figure 3 is that PA efficiency at low transmit powers is of less importance than at higher transmit powers.

Figure 3. The effect of a fixed processing power and power back off on talk time for a 1 amp-hour battery.

(A) = 330 mW power overhead, (B) = 1210 mW.
4. Linear PA Technologies

Future techniques for linear RF PA’s must be capable of much higher power efficiency and be easily implemented in compact robust designs, when compared to present technologies. Some technologies in the literature are not capable of meeting either of these requirements. For example, the complexity of adaptive feedforward systems[5] coupled with low efficiency at handportable powers eliminates this scheme from consideration. Second harmonic injection[7] is a recent method; it requires a RF input and only corrects for AM-AM distortion. It is desirable to have linearisation methods that accept baseband input signals allowing direct conversion transmitters. This baseband requirement rules out RF predistortion[8] which gives only modest amounts of PA correction (10-25 dB). CALLUM[9] and LINC[10] use two highly efficient PA’s and, to date, an efficient means of combining their outputs without loss of efficiency has not been found. Assuming 100% efficient amplifiers and a hybrid combiner (with no insertion loss) the overall efficiency is \( \approx 42\% \) for TETRA[11]. For CALLUM and LINC to be feasible power recovery[12] or non-linear combining is necessary to raise the efficiency.

The main focus of the following four sub-sections are on schemes which are likely to meet the high linearity of the TETRA specification and also the bandwidth of GSM EDGE. The ability to combine techniques is also a factor in their selection. The techniques considered are dynamic bias, Cartesian loop, Envelope Elimination & Restoration (EER), and adaptive baseband predistortion methods.

4.1 Dynamic bias techniques

Dynamic bias alters the operating point of the power transistor to either maximise efficiency or linearity[3]. To improve efficiency, the supply voltage tracks the envelope level to minimise the DC bias power of the power transistor. The gate (or base) bias is used to maintain constant gain, which improves or restores the intermodulation performance[4]. High efficiency switching mode amplitude modulators are used to modulate the supply voltage, and linear techniques can be used for the gate (or base) bias due to the lower power. At peak output power, the technique is less efficient than the PA without dynamic bias due to the switching modulator efficiency. Good increases in efficiency (by a factor of two) are gained at the mid power regions of the PA[15].

4.2 Cartesian Loop

Figure 5 shows a block diagram of the Cartesian loop transmitter. A Coupler samples the output of the transmitter where it is quadrature demodulated to Cartesian co-ordinates, and subtracted from the input to form a complimentary distorted error signal.

The Cartesian components are narrowband, easing the baseband implementation. Linearity, stability and bandwidth are traded off in the design flow[16]. However, errors in the feedback path are not corrected and feedback components must be as linear as the desired loop linearity. Noise (particularly wideband) in the Cartesian loop can be increased by insufficient stability and inattention to gain distribution around the loop[17,18]. The phase shifter is needed to counter for RF delays[19,20]. This must be set uniquely for every channel, and generally requires a training sequence to optimise the setting. Lastly, DC offsets in the feedback
loop give rise to carrier feed-through and special circuitry is needed to null these offsets[21]. Whilst it is possible to linearise a class-C PA using Cartesian loop, it is extremely difficult to produce a stable design that meets the TETRA specification [22,23]. This is due to the abrupt phase and amplitude characteristics of class-C biased devices.

4.3 Polar loop and Envelope Elimination & Restoration (EER)
The EER transmitter (figure 6), whilst simple, achieves only modest amounts of correction and requires delay lines to compensate for differences between the envelope and phase path delays [24,25]. The addition of envelope feedback is popular and some implementations of this have achieved intermodulation distortion (IMD) of −30 to −50 dBc (for 1W to 20W) and efficiencies of up to 50% [26,27].

Envelope feedback requires two envelope detectors, the first is at the input to the loop, and the second in the feedback path; both must be linear and require close matching. Additionally as envelope feedback does not correct AM-PM distortion some form of phase feedback is needed. The phase error can be obtained by limiting the output and obtaining a phase-error signal to drive a VCO. The technique is now called polar loop and is shown in figure 7.

For good levels of performance the envelope detector can limit the linearity[28]. Two other problems of polar loop are; 1) the phase feedback forms a PLL, which, for small instantaneous power can lose lock, and 2) the envelope and phase signals can have much larger bandwidth’s than the RF signal bandwidth [24]. The advantage of EER and polar loop is that the bandwidth and linearity of the amplitude modulator determine linearity rather than the PA. The PA amplifies a constant envelope signal, thus high efficiency devices can be used.

4.4 Digital Adaptive Baseband Predistortion
Three types of baseband predistortion that have been widely investigated are data predistortion[29], Cartesian mapping[30] and complex gain mapping[31,32], (see Figure 8).

Complex gain mapping has been the most successful of these due to its small look up table (LUT) size (16 – 100 entries), fast adaption rate (milliseconds); it’s main disadvantage is increased computational load. Complex gain mapping uses the squared amplitude of the signal to index a 1-D LUT that stores the AM-AM and AM-PM
characteristic necessary for predistortion. It does not correct for errors in the quadrature modulator (whereas Cartesian mapped does). Special algorithms (and training) are required to correct for the quadrature errors[33]. The feedback loop for the adaption process is similar to Cartesian loop, requiring quadrature demodulated sample of the PA output. As the adaption does not take place in real time, stability is guaranteed. It is also possible to construct simpler feedback paths for the adaption[34].

5. Comparison of Techniques with High Efficiency Requirements

From section 3, it was seen that a 57% supply-efficient PA was needed to achieve 50% or better transmitter PAE, this implies that the PA will have to be biased as class-C. Most class-C amplifiers have IMPs in a two-tone test of ≈-15 dBc and the performance does not improve with back-off[6], thus to meet the TETRA specification 45 dB of IMD correction is needed. No implementations of adaptive bias, EER or predistortion known to the authors can achieve this requirement. Cartesian Loop can meet this performance under ideal conditions, i.e. handpicked components, but it is unlikely that it would be suitably stable to manufacture in high yields or achieve the bandwidth required for EDGE. Additionally, the integration of Cartesian loop is difficult at these levels of performance. It is unlikely that any single method outlined above will provide significant improvement in efficiency for the next generation handportable. There has been some research done on hybrid solutions in the past, these will be briefly outlined.

Two techniques based on feedforward as the main linearisation scheme use either RF predistortion[6] to increase the efficiency and linearity of the main amplifier, or RF feedback[35] to improve the efficiency. Both require RF inputs resulting in additional complexity. This approach is suited to high power base-station applications. An alternative technique is based on Cartesian loop with either predistortion[36] (added after the error amplifier in the forward path), or adaptive bias[15]. The adaptive bias technique increases the supply efficiency of the PA at mid power levels (18 dBm) by ≈15%. The power of the Cartesian loop and envelope detection cintuity was not given in[15], and would reduce the efficiency for the entire transmitter substantially. By applying adaptive bias in the Cartesian loop, the envelope tracking does not have to be as accurate; the Cartesian loop will correct for deficiencies. The predistortion technique provides greater linearisation than adaptive bias, this can be used to increase PA efficiency as well as increase the Cartesian loop’s stability. Practical results achieving IMD better than ~67 dBc have been obtained, additionally the DACs used in the predistorter can be as little as 4-8 bits, with the Cartesian loop correcting for any spurs that may result. This suggests that the extra complexity of the adaptive predistorter would be easily integrated.

6. Conclusion

The linearisation circuits’ power requirements must be taken into account when comparing the efficiency of linearisation schemes. The PA supply efficiency is important to talk-time only at mid to high transmit powers. To greatly improve talk-time through efficient transmitter design, the effort must be concentrated at higher powers (e.g. > 20 dBm). If the PA drivers and linearisation circuitry can be designed to consume less than 250 mW; then for a 1 Watt PA with a 57% supply efficiency the PAE will reach 50%.

Of the techniques explored in this paper, Cartesian loop has been most successful to date, but to substantially increase efficiency, class-C PAs and hybrid techniques are required. The amount of research on hybrid schemes is small, and further results are needed if high efficiency transmitters for TETRA are to be a reality. Two promising hybrid techniques use Cartesian loop as the basic lineariser. One technique uses predistortion and the other adaptive bias techniques. The amount of research on hybrid schemes is relatively small, and further results are needed.


