Evolution of power amplifiers for mobile phone terminals from the 2nd generation to the 5th generation

Satoshi Tanaka†a), senior member Kenji Mukai†, Shohei Imai†, and Hiroshi Okabe†, members

SUMMARY Mobile phone systems continue to evolve from the 2nd generation, which began in the early 1990s, to the 5th generation, which is now in service. Along with this evolution, the power amplifier (PA) is also evolved. The characteristics required for PA are changing with each generation. In this paper, we will give an overview of the evolution of PAs from the 2nd generation mobile phones such as GSM (global system for mobile communications) to the 5th generation mobile phones that is often called NR (new radio), in particular, the circuit system. Specifically, the following five items will be described. (1) Ramp-up and ramp-down power control circuit corresponding to GSM, (2) Self-bias circuit technology for improving linearity that becomes important after W-CDMA (wideband code division multiple access), (3) Power mode switching methods for improving efficiency at low output power, (4) Power combining methods that have become important since LTE (long term evolution), and (5) Backoff efficiency improvement methods represented by ET (envelope tracking) and Doherty PA.

key words: power amplifier, mobile phone, power mode switching, envelope tracking, Doherty.

1. Introduction

Since power amplifiers for mobile phones start with saturated amplifiers corresponding to the 2nd generation GSM (global system for mobile communications). After the 3rd generation W-CDMA (wideband code division multiple access), the application of linear amplifiers corresponding to modulated signals with high PAPR (peak to average power ratio) is progressing. In this paper, the transition of requirements in each generation of mobile phone is summarized in Section 2. The first of the technical problems of PA is high efficiency operation. However, there are technical problems other than simple high efficiency in order to meet the unique requirements of each generation of mobile phones. Section 3 introduces control technology that suppresses out-of-band surpius by smoothly launching and lowering for intermittent operation, which is important in GSM. Section 4 introduces W-CDMA, LTE (long term evolution) critical technology for reducing supply current and gain at low outputs with the highest probability of the presence of average output power. In addition, HPUE (high power user equipment), which further increases 3dB output, is applied in LTE. Therefore power combining technology becomes important. This power combining technology is also introduced. In order to cope with the increase in PAPR of modulated signals after LTE, efficiency improvement in the backoff region is a major problem in PA. Section 5 introduces the ET (envelope tracking) system as a representative technology for efficiency improvement in the backoff region. Furthermore, when the modulation band is more than 100MHz, there is a concern that the efficiency of the envelope tracker used by the ET method will be degraded. Doherty PA is introduced as an alternative technology candidate for ET.

2. Requirements for mobile phone PAs

Table 1 summarizes the changes in basic specifications from the 2nd generation (2G) to the 5th generation (5G) mobile phones and the accompanying changes in power amplifiers.

Table 1: Specifications for Mobile Phone PAs

<table>
<thead>
<tr>
<th>Generation</th>
<th>Modulation Method</th>
<th>Subband</th>
<th>Transmit Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G</td>
<td>FDD-CDMA</td>
<td>850MHz</td>
<td>33 dBm</td>
</tr>
<tr>
<td>3G</td>
<td>W-CDMA</td>
<td>1800MHz</td>
<td>30 dBm</td>
</tr>
<tr>
<td>4G LTE</td>
<td>FDD</td>
<td>850MHz</td>
<td>24 dBm</td>
</tr>
<tr>
<td>5G NR</td>
<td>NR</td>
<td>2800MHz</td>
<td>30 dBm</td>
</tr>
</tbody>
</table>

Copyright © 20XX The Institute of Electronics, Information and Communication Engineers
Table 1 Summary of generations of mobile phones and power amplifier technologies.

<table>
<thead>
<tr>
<th>Duplex Method</th>
<th>GSM (EDGE)</th>
<th>W-CDMA</th>
<th>LTE</th>
<th>5G (FR1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Access</td>
<td>FDD</td>
<td>FDD</td>
<td>FDD/TDD</td>
<td>FDD/TDD</td>
</tr>
<tr>
<td>Modulation Scheme</td>
<td>TDMA</td>
<td>CDMA</td>
<td>SC-FDMA</td>
<td>CP-OFDM</td>
</tr>
<tr>
<td>PAPR (dB) /BW (MHz)</td>
<td>0, (3.2)/0.27</td>
<td>3.0-3.5/1.5</td>
<td>4.0-6.5/ 5-40</td>
<td>&lt;8.5/~200</td>
</tr>
<tr>
<td>Maximum Transmit Power (dBm)</td>
<td>33 (27) @ 850/900 MHz</td>
<td>24</td>
<td>23/26</td>
<td>23/26/29</td>
</tr>
<tr>
<td>PA Architecture</td>
<td>Saturated Amplifier</td>
<td>Linear Amplifier</td>
<td>Linear Amplifier with Mode Switching</td>
<td>Linear Differential Amplifier with Mode Switching</td>
</tr>
<tr>
<td>Transistor Type</td>
<td>LD MOS</td>
<td>HBT</td>
<td>HBT</td>
<td>HBT</td>
</tr>
<tr>
<td>PA Operation Method</td>
<td>APC/LDO</td>
<td></td>
<td>Envelope Tracking</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Polar Loop</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Due to the multi-value modulation, the PAPR of the modulated signal increases to 4.0-6.5 dB, which is larger than that of W-CDMA [3]. The modulation bandwidth is 5-40 MHz. The maximum transmit power is basically 23 dBm, which is 1 dB lower than W-CDMA.

Fig. 1 shows the relationship between the maximum instantaneous transmit power and the probability density of W-CDMA and LTE modulation signals [2] calculated by “Ptolemy system simulator”. The average transmit power for each modulated signal is indicated by a circle. In the case of W-CDMA, the maximum transmit power is 24 dBm, and the probability density is maximized at a maximum instantaneous transmit power that is about 0.6 dB higher, and the probability density decreases rapidly at powers higher than that. In the case of LTE, the maximum instantaneous transmit power differs depending on the modulation method and modulation bandwidth. In the case of 8RB (resource block) with QPSK modulation, the modulation bandwidth is 1.6 MHz for the channel bandwidth of 5 MHz. The average transmit power in this case is 23 dBm. When the modulation method and RB change, the PAPR of the modulated signal changes. Reduce the average transmit power according to the MPR (maximum power reduction) defined for each condition. Therefore, the peak power of maximum instantaneous transmit power of LTE is about 1.5 dB higher than that of W-CDMA.

NR (new radio), which is 5G, is applied to the millimeter wave band (FR2) in addition to the frequency (FR1) of conventional mobile phones, but FR1 will be described here. NR is an extension system of LTE, and like LTE, there are two types of duplex systems, FDD and TDD.

There are two types of multiple access methods on the transmitting side: CP (cyclic prefix)-OFDM (orthogonal frequency division multiplexing) and DFTS (discrete Fourier transform spreading) -OFDM. 256QAM is added as the modulation signal, and the maximum PAPR is about 8.5 dB [3].

The modulation bandwidth has been expanded to 200 MHz according to the standard, but the current service is up to 100 MHz.
In order to correspond to the characteristics of the communication standards of each generation, the characteristics required for PA have changed as follows. In GSM, GMSK which is the main modulation method, is a modulation signal with a constant amplitude, and therefore high efficiency as a saturation amplifier was required. In addition, it is necessary to perform intermittent operation in order to support the TDMA. It is also necessary to smoothly switch between the transmission state and the off state. APC (auto power control) method and LDO (low dropout regulator) method were used for this switching control and transmit power control. In addition, the polar loop method that can support EDGE (enhanced data rates for GSM evolution) was also used. These control methods will be described in detail in Section 3.

LDMOS (laterally diffused metal-oxide semiconductor) [4-6], HBT (Hetero-junction bipolar transistor) [7,8], and GaAs MESFET (gallium arsenide metal-semiconductor field effect transistor) [9] were considered as the device of GSM. LDMOS and HBT do not require a negative bias voltage, therefore these devices were widely applied for GSM.

The modulation schemes in W-CDMA era include amplitude modulation, and it is required to improve the performance as a linear amplifier. Therefore, control of the self-bias effect [10] was investigated.

In the voice communications such as a WCDMA, the probability distributed function (PDF) of output power from a power amplifier (PA) is highest around 0 dBm (See below Fig. 7). Therefore, it is important to reduce the power consumption at low transmit power. In addition, in order to reduce the dynamic range of the output power of RFIC (radio frequency integrated circuit), a method to reduce the gain at low output power is also important. Therefore, power mode switching method is important.

To expand the coverage area per base station of 2.4 GHz band or higher in LTE and NR, HPUE, which applies an antenna output power of 26 dBm, and which is 3 dB higher than the FDD method, was applied to the TDD method. In the case of GSM, PA was able to support the maximum transmit power of 33 dBm at frequencies below 1 GHz, but it is difficult to generate large power in the band above 2.4 GHz. The power combining methods are important to deal with this.

PA circuit technologies will be described in detail in Section 4. In GSM, multiple devices have been applied, but since W-CDMA, HBT devices that have excellent frequency characteristics and operate with a single power supply have been widely applied to PA [11,12].

After LTE, the PAPR of the modulated signal increases and it is important to improve the backoff efficiency of the PA. To improve the backoff efficiency, the ET (envelope tracking) method, which dynamically changes the power supply of the PA according to the amplitude of the modulated signal, is effective. In NR, the PAPR is further increased and the modulation bandwidth is further expanded.

In the case of the ET method, there is a concern that the power efficiency of the ET power supply will decrease as the bandwidth increases. The application of Doherty PA has begun to be considered as a method for improving backoff efficiency without using an ET power supply. Section 5 introduces these two technologies for improving backoff efficiency.

3. Power control method for GSM

As described in Section 2, the GSM-compatible PA needs to perform intermittent operation in order to support the TDMA, and it is necessary to smoothly switch between the transmission state and the off state. For this reason, GSM stipulates in the standard the allowable power leakage amount to the neighboring frequency band due to the ramp-up and ramp-down operations at the start and end of the transmission operation as shown in Table 2 [13]. In order to satisfy this, accurate ramp-up and ramp-down envelope control is required.

<table>
<thead>
<tr>
<th>Separation Frequency</th>
<th>Leakage Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 kHz</td>
<td>23 dBm</td>
</tr>
<tr>
<td>600 kHz</td>
<td>26 dBm</td>
</tr>
<tr>
<td>1200 kHz</td>
<td>32 dBm</td>
</tr>
<tr>
<td>1800 kHz</td>
<td>36 dBm</td>
</tr>
</tbody>
</table>

Table 2 Allowed maximum leakage power at neighboring frequency band

Measurement Conditions: Filter BW: 30 kHz, Video BW: 100 kHz

Fig. 2 shows the relationship between the rising and falling waveforms and the frequency spectrum. When a steep change such as a square wave occurs, the spectrum spreads widely. On the other hand, if a smooth change is made, the spectrum regrowth can be small and narrowed. Therefore, it can be seen that smooth envelope control is important.

The followings are typical power control methods. Fig. 3 summarizes typical methods for performing ramp-up and ramp-down control. The comparisons of each method are also summarized on Table 3. Fig. 3 (a) is an operating principle of the APC method using negative feedback. In this method, the input power of the PA is always constant. The output power level of PA is detected by the coupler, and this is detected by the detection circuit. The detected signal is compared with the control signal for power control, and the error between the two is obtained. The PA bias circuit is controlled according to the error. Accurate power control can be realized because of control by negative feedback. In order to make it smaller, a method such as installing a sense transistor in the final stage transistor without using a coupler, detecting the current of the sense transistor, and using it for
negative feedback was also considered [5].

Fig. 3 Typical methods for performing ramp-up and ramp-down control.

Table 3 The comparisons of each method

<table>
<thead>
<tr>
<th>Power Control Method</th>
<th>APC Method</th>
<th>LDO Method</th>
<th>PIN Control Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Scheme</td>
<td>Closed Loop</td>
<td>Open Loop</td>
<td>Open (Closed) Loop</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Good</td>
<td>Fair</td>
<td>Good</td>
</tr>
<tr>
<td>Output Power Margin</td>
<td>Not Required</td>
<td>Not Required</td>
<td>Required</td>
</tr>
</tbody>
</table>

Fig. 3 (b) shows the LDO method to which the LDO power supply [14] is applied. The output power is controlled by controlling the power supply voltage of the PA with the LDO power supply. Even with this method, the input power of the PA is always constant. The square of the power supply voltage and the output power are in proportional relationship, and good control is possible even with open-loop control. The loss of LDO is added to total performance. For this reason, efficiency is lower than the other two methods.

Fig. 3 (c) shows the PIN control method. In this method, ramp-up and ramp-down envelope control are performed by the RFIC that gives the input signal of the PA. In this method, if the maximum power is completely saturated at the PA, the output power cannot be controlled by the input power.

Therefore, some power margin at the maximum output power level is required for the PA. It can have a simple structure. In this method, open loop control is applied in most cases. However, it is also possible to realize closed loop control like the APC method by detecting the coupler output in the RFIC.

Fig. 4 shows an example of the polar loop method that can also support 8-PSK modulation [15, 16]. The LDO method is applied to the PA part. The output power of PA is monitored by a coupler and converted into an IF (intermediate frequency) signal. The IF signal is divided into an amplitude signal and a phase signal. The I (in phase) and Q (quadrature phase) baseband signals are converted into an IF signal by an orthogonal modulator, and this IF signal is further divided into an amplitude signal and a phase signal, and the error is obtained by comparing with the monitored signal, and the phase is obtained. The error is feedback to the VCO (voltage-controlled oscillator) via the loop filter, and the amplitude is also fed back to the LDO via the loop filter and integrator. With this negative feedback, it is possible to support ramp-up and ramp-down envelope control as well as 8-PSK modulation. Although the negative feedback is introduced here, there is also an example of open-loop control regarding amplitude control [17].

4. PA circuit technologies for W-CDMA and LTE

4.1 Linearity improvement technology

The main modulation method after W-CDMA is that the amplitude changes, and it is required to improve the performance as a linear amplifier. Therefore, control of the self-bias effect [10] and harmonic control [18-20] were investigated. Fig. 5 shows an example of a typical PA for W-CDMA and LTE [2]. It has a two-stage configuration of a driver stage and a final stage.
The bias circuit is composed of an emitter follower. In order to improve efficiency, the final stage is often biased toward class AB. In class AB bias, the quiescent current, which is the collector current at the time of no input signal, is set low to some extent. Therefore, when the signal is small, the collector DC current becomes low. However, as the input signal becomes larger, the collector DC current becomes larger due to the self-bias effect, and it corresponds to a large output power. This can achieve high efficiency operation, but often trades off with linearity.

The self-bias effect is shown in Fig. 6. The final stage of the HBT amplifier in Fig. 5 is taken as an example. As shown in Fig. 6 (a), let us first briefly consider the case where a bias is supplied from the voltage source \( V_{bias} \) to the base of the HBT \( Q_2 \) via the base ballast resistor \( R_{bias2} \). Fig. 6 (b) shows the bias line of the base of \( Q_2 \). The horizontal axis is the base-emitter voltage \( V_{BE} \), and the vertical axis is the base-emitter current \( I_{BE} \). The curve shown by the solid line is an exponential function which is a diode characteristic between the base and the emitter. The diagonal line is the load line when the current is supplied from \( V_{bias} \) via \( R_{bias2} \), and the intersection with the diode curve is the bias point. When an RF input voltage is applied to the base in this state, an RF current flows, but the average current increases due to the characteristics of the exponential function of the diode. This phenomenon can also be understood as follows. The exponential function can be expressed as a power series, and as the trigonometric input increases, the DC current component of each term of the output increases. As a result, the DC current increases. Therefore, when an RF input voltage is added to DC bias voltage, the DC diode characteristics change so that a large amount of current flows even under the same \( V_{BE} \) conditions as shown by the dotted line. Along with this, the \( V_{BE} \) of the bias point changes in the low direction and the \( I_{BE} \) changes in the high direction.

In the actual bias circuit, an emitter follower (\( Q_4 \)) is used instead of the power supply \( V_{bias2} \) as shown in Fig. 6 (c). When an RF signal is input to \( Q_4 \), a large bias current can be supplied with a lower \( V_{BE} \) value as in the case of \( Q_2 \). When the base voltage of \( Q_4 \) is constant, the bias voltage \( V_{bias} \) rises with lower \( V_{BE} \). Therefore, the bias current at the bias point further increases. A method has been proposed in which this effect is positively utilized and the RF input signal is positively supplied to the emitter follower via a capacitance as shown in Fig. 6 (c) [21]. It has been also reported that the self-bias effect can be adjusted by supplying bias from multiple emitter followers [22].

In addition, since linearity largely depends on the signal source impedance with respect to the base of \( Q_2 \), it is necessary to design the base input impedance as well [23].

4.2 Reduced power consumption and gain at low power

Fig. 7 shows the transmit power distribution of voice communication and data communication of mobile phones [24, 25]. The horizontal axis is the transmit power at antenna of mobile phone, and the vertical axis is the probability density. Here, the transmit power is not the instantaneous transmit power shown in Fig. 2, but the average output power of the modulated signal. From Fig. 7, the highest probability density in voice communication is -3 dBm. The highest probability density in data communication is 6 dBm. In both cases, the probability density at low power is high compared to that at the maximum power.

There is an APT (average power tracking) method as a method of improving efficiency when the output power is low. The outline of the APT method is shown in Fig. 8 [3].
In the APT method, the power supply voltage and the bias currents are adjusted according to the average output power in order to improve the efficiency at each average output power. The power supply voltage is controlled by the DCDC converter. The typical efficiency of the buck converter used here is about 95%. The loss of the buck converter is small.

Figs. 8 (b) and 8 (c) show graphs of efficiency and ACLR (adjacent leakage power ratio) when the power supply voltage \( V_{CC} \) is changed 1.0-5.0 V in 0.5 V steps. The efficiencies with above \( V_{CC} \) conditions, when ACLR is -35 dBc are plotted with red dots in Fig. 8 (b). When the power supply voltage is constant, the efficiency decreases significantly when the average output power decreases, but it can be seen that the decrease in efficiency is suppressed to a low average output power by lowering the \( V_{CC} \) according to the average output power.

When reducing the output power of the PA, it is desired to reduce the gain as well as the power consumption at low output. For example, in the case of W-CDMA, the maximum transmit power and the minimum transmit power at antenna of mobile terminal are 24 dBm and -50 dBm, respectively, and the dynamic range is as large as 74 dB. In RFIC, there is a problem that it becomes difficult to achieve SNR (signal to noise ratio) when the output level is small. In order to reduce the dynamic range of the RFIC, it is desirable to reduce the gain when the transmit output power is small.

Many mode switching methods are being studied to reduce power consumption and gain at low output. Table 4 summarizes the proposed mode switching technologies.

The stage bypass method stops the operation of the final stage when the output power is low, and provides a switch to bypass the final stage [26-28]. Since it operates with a small number of stages when the output power is low, both current consumption reduction and gain reduction can be realized. Since a switch is used, it is difficult to realize low loss switch with HBT alone, and it is desirable to use FET together. For this reason, there were many examples where BiFET was applied. A method that omits the switch has also been proposed [29].

In the multi pass method, independent PAs for high output power and low output power are provided, and these are switched and used [30-32]. Performance improvement can be expected because it can be optimized for the operation mode, such as setting different load impedances for each PA for high output power and PA for low output power. However, since this method also requires a changeover switch, it is necessary to sufficiently reduce the switch loss. A method without a switch has also been proposed [33].

The attenuator method is mainly aimed at reducing the gain. An attenuator is connected in series to the input of the PA to reduce the gain [34]. A method of connecting an attenuator between the first stage and the final stage has also been proposed [35,36]. As for the attenuator, it is possible to use a diode as well as a FET, and it is possible to realize it only with HBT. Another advantage of this method is that there is no switch for the output matching circuit, and there is no efficiency degradation at high gain mode comparing with stage bypass and multi pass method.

<table>
<thead>
<tr>
<th>Table 4 Summary of the mode switching methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Configurations</td>
</tr>
<tr>
<td>Switch</td>
</tr>
<tr>
<td>Additional Pass</td>
</tr>
</tbody>
</table>
By reducing the bias current flowing through the transistors in the first and final stages to reduce the transconductance of the transistors, the variable gain method is reducing the gain and at the same time reducing the current consumption. When the current density of a transistor decreases, the linearity deteriorates, so the transistor may be switched [37].

The four methods (Table 4) introduced here may be used alone or in combination.

4.3 Output power combining technology

When outputting 3 dB higher power for HPUE, simply doubling the size of the transistor with 23 dBm output and halving the load impedance will cause an increase in the loss of the matching circuit. Power combining methods is important for loss reduction. Table 5 summarizes the power combining methods.

The single method is setting the load impedance according to the output power described above. In this case, the power supply voltage is $V_{CC}$, the output current of the final stage is $I_{OUP}$, the load impedance is $R_L$, and the voltage amplitude at the collector end at the maximum output of the final stage is $V_{OUT_{peak}}$. $V_{OUT_{peak}}$ is smaller than $V_{CC}-V_{knee}$. Here, $V_{knee}$ represents the upper limit voltage of the saturation region of the bipolar transistor.

In the in-phase combining method, an in-phase signal is added to a transistor divided into two to set the load impedance of $2R_L$, and output power is combined using a Wilkinson coupler or the like [4]. The output current of one transistor is half that of the single method.

The stack method is often used in PAs that apply a CMOSFET (complementary metal-oxide-semiconductor field-effect transistor) [38]. By stacking multiple transistors in a stack and performing the same operation, the CMOSFET with a low breakdown voltage operates at a high voltage supply and obtains high output power. As shown in followings, this method is used for power combining of HPUE-compatible output stages. The power supply voltage is $2V_{CC}$ and the output current is $0.5I_{OUT}$. The load impedance is $4R_L$, which is obtained by connecting the load impedance $2R_L$ of each transistor in series.

Therefore, the impedance conversion ratio in the output matching circuit can be reduced, and the loss in the output matching circuit can be reduced. The maximum $V_{OUT_{peak}}$ of the upper collector is $2(V_{CC}-V_{knee})$. This method requires twice the power supply voltage, but there is also a method of operating at $V_{CC}$ by inserting a capacitor between the collector of the lower transistor and the emitter of the upper stage and separating the power supply [39, 40].

In the differential method, a reverse phase signal is added to the transistor divided into two to set the load impedance of $2R_L$. At this time, the load impedance between the collector ends of the two transistors is $4R_L$. Therefore, the impedance conversion ratio in the output matching circuit can be reduced as in the stack method, and the loss in the output matching circuit can be reduced. Transformers are often used in output matching circuits that require differential/single conversion functions [41].

Comparing the four methods (Table 5), the stack method and the differential method have merits from the viewpoint of reducing the loss of the output matching circuit. Furthermore, in the case of the differential method, there are additional merits such as the reduction in the influence of the emitter parasitic inductor, cancelation of the 2nd harmonic, and the reduction in the power supply bypass capacitance value. We believe that the differential method is desirable as a power combining technology for HPUE-compatible PAs [42].

<table>
<thead>
<tr>
<th>Topology</th>
<th>Single</th>
<th>In-Phase Combining</th>
<th>Stack</th>
<th>Differential</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{CC}$</td>
<td>$V_{CC}$</td>
<td>$V_{CC}$</td>
<td>$2V_{CC}$</td>
<td>$V_{CC}$</td>
</tr>
<tr>
<td>$I_{OUT}$ (RMS)</td>
<td>$I_{OUT}$</td>
<td>$0.5I_{OUT} \times 2$</td>
<td>$0.5I_{OUT}$</td>
<td>$0.5I_{OUT} \times 2$</td>
</tr>
<tr>
<td>$R_L$</td>
<td>$R_L$</td>
<td>$2R_L$</td>
<td>$4R_L$</td>
<td>$4R_L$</td>
</tr>
<tr>
<td>Transistor Size</td>
<td>1</td>
<td>$0.5 \times 2$</td>
<td>$0.5 \times 2$</td>
<td>$0.5 \times 2$</td>
</tr>
<tr>
<td>$P_{OUT}$</td>
<td>$I_{OUT}^2R_L$</td>
<td>$0.5I_{OUT}^2R_L \times 2$</td>
<td>$I_{OUT}^2R_L$</td>
<td>$I_{OUT}^2R_L$</td>
</tr>
<tr>
<td>$V_{OUT_{peak}}$</td>
<td>$&lt;V_{CC}-V_{knee}$</td>
<td>$&lt;2(V_{CC}-V_{knee})$</td>
<td>$&lt;2(V_{CC}-V_{knee})$</td>
<td>$&lt;2(V_{CC}-V_{knee})$</td>
</tr>
<tr>
<td>$L_p$ Cancel</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>2nd Harmonics Cancel</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Bypass Cap. At Power Line</td>
<td>Large</td>
<td>Large</td>
<td>Large</td>
<td>Small</td>
</tr>
</tbody>
</table>

Table 5 Summary of the power combining methods for PA.
5. Backoff efficiency improvement technologies for LTE and NR

As mentioned in Section 2, after W-CDMA, the PAPR of the applied modulated signal tends to increase, and the difference between the peak power of maximum instantaneous transmit output power and the maximum transmit power becomes large. For this reason, the average operation of PA operates when a large backoff is taken from transmit power becomes large. For this reason, the average instantaneous transmit output power and the maximum applied modulated signal tends to increase, and the efficiency of PA decreases as $P_{OUT}$ decreases, as shown in Fig. 8 (b). Therefore, when a large PAPR modulated signal is output, a deterioration in backoff efficiency becomes apparent.

In the APT method introduced in Section 4, the efficiency drops when the average output power is lowered. In order to improve efficiency when transmitting a large PAPR modulated signal, it is necessary to control the power supply voltage with respect to the instantaneous power. The ET method can achieve this operation.

In addition, in the multi pass method introduced in Section 4, when the average output power is high, a low $R_L$ is applied, and when the average output power is low, a high $R_L$ is set to increase the efficiency of each. In order to improve efficiency when transmitting a large PAPR modulated signal, it is necessary to control $R_L$ with respect to instantaneous power. Doherty PA can achieve this operation.

5.1 ET method

Fig. 9 shows the operating principle of the ET PA. Fig. 9 (a) shows the configuration diagram. The envelope signal of the modulated signal is obtained from the amplitude of the I and Q baseband signals in the RFIC. Then, the envelope signal is scaled according to the output level of the RFIC. In order to match the scaled signal with the voltage actually applied to the PA, the minimum voltage $V_{min}$ and the maximum voltage $V_{max}$ are defined on a table that defines a curve connecting the both. The ET tracker changes the power supply voltage of the PA according to the table signal. As for the input/output characteristics when $V_{CC}$ is changed in PA, the higher the $V_{CC}$, the higher the saturation power and the slightly higher gain. The increase in gain when $V_{CC}$ increases here is due to the decrease in the base-collector capacitance and a decrease in the effect of negative feedback as $V_{CC}$ increases. When $V_{CC}$ is changed according to the table, PA operates under the condition of $V_{CC} = V_{min}$ when the output signal is small. As the output signal increases, the gain decreases under the $V_{CC} = V_{min}$ condition and distortion occurs. In this case, by increasing $V_{CC}$ according to the table, the gain is constant even if the output increases.

The effect of this operation will be explained using Figs. 9 (b) and 9 (c). Fig. 9 (b) shows the collector output voltage of the PA in the APT mode. In this case, the $V_{CC}$ is constant. A modulated signal voltage $V_C$ is generated on the collector around $V_{CC}$. The case where the voltage reaches $V_{knee}$ when the voltage drops under the maximum condition of the amplitude of the modulated signal corresponds to the maximum output. In this case, focusing on the modulated signal, the amplitude may be small and a difference may occur between the lowest voltage and $V_{knee}$. This case corresponds to the operation in the backoff state.

This difference corresponds to the power loss of the PA, which causes a decrease in efficiency. Fig. 9 (c) shows the collector output voltage of the PA when ET is operated. Since $V_{CC}$ is changed according to the envelope of the modulated signal, the minimum operating voltage is in the vicinity of $V_{knee}$ regardless of the magnitude of the amplitude of the modulated signal, and power loss can be reduced.

Fig. 10 shows an example of the characteristics of PA for ET [3]. Fig. 10 (a) shows examples of the table used in Fig. 9 (a). Figs. 10 (b), (c), and (d) show the efficiency, gain, and phase curves of PA when $V_{CC}$ is changed. What is shown by the red dotted line here corresponds to the control called "iso-gain" described in Fig. 9 (a). The green dotted line corresponds to a control called "de-troughing" that further lowers the $V_{CC}$. The yellow curve is a series of peak efficiency conditions and has a complicated shape and is not suitable for control. Since "iso-gain" controls to keep the gain constant, there is no AM (amplitude modulation) - AM characteristic distortion. Therefore, when using DPD (digital pre-distortion), only AM-PM (phase modulation) characteristic distortion needs to be considered. The "iso-gain" method is widely applied because of the simplification of DPD.

The ET method introduced here is widely applied to mobile phones, and its operation has been confirmed up to 100 MHz. Therefore, it can also be used for NR. The details of the ET method are described in the paper [3], and please...
refer to that as well.

If the bandwidth is further increased, the efficiency of the ET tracker may decrease and it may be necessary to improve the backoff efficiency by another method. As another method, the Doherty amplifier will be described in the next subsection.

Fig. 10 An example of the characteristics of PA for ET.

5.2 Doherty PA

The operating principle of Doherty PA will be described with reference to Fig. 11. Fig. 11 (a) outlines the configuration of the Doherty PA. Two types of PA are used: a carrier amplifier with a bias of class AB or B and a peak amplifier with a class C bias. The carrier amplifier is connected to the output of the peak amplifier via a λ/4 line whose characteristic impedance is Ropt, which functions as an impedance converter. Connect the load impedance ZL= Ropt/2 of the entire Doherty PA to the output of the peak amplifier.

In the actual circuit, Zl is the impedance of output matching circuit that converts to an impedance of 50 Ω. The carrier amplifier is always operating, but the peak amplifier operates only when the input is large enough due to the class C bias. When the peak amplifier is in the off state, the load impedance ZLC of the carrier amplifier is 2Ropt. This is the load impedance ZL=Ropt/2 of the entire Doherty PA converted to 2Ropt by the λ/4 line. Due to this large load impedance, the carrier amplifier saturates early.

With proper bias setting of the peak amplifier, the operation of the peak amplifier is started at the timing of near saturation of the carrier amplifier. A 90 deg phase difference is provided between the output of the peak amplifier and the output of the carrier amplifier. Therefore, the output signal of the carrier amplifier and the output signal of the peak amplifier, which are 90 deg out of phase via the λ/4 line, are in phase. When IC (the output current of peak amplifier)=Ip, the load impedance of the carrier amplifier is halved, and the saturation power of the carrier amplifier is doubled. Furthermore, when combined with the output of the peak amplifier that outputs the same output as the carrier amplifier, the Doherty PA as a whole has twice the saturation power. In other words, the Doherty PA has a saturation power that is 6 dB higher than the saturation power of the carrier amplifier alone. In Fig. 11 (b), the horizontal axis shows the amount of backoff based on the output power at which the carrier amplifier and peak amplifier are saturated, and the vertical axis shows the efficiency. The efficiency increases as the output power increases and the carrier amplifier saturates. After that, the peak amplifier starts operating, and along with that, the efficiency is slightly reduced, but the high efficiency is maintained. When both the carrier amplifier and the peak amplifier are saturated, the efficiency is maximized again.

\[ Z_{LC} = 2R_{opt} \left( \frac{I_c}{I_c + I_p} \right). \]

Under the condition of \( I_c \) (the output current of carrier amplifier)=Ip, the load impedance of the carrier amplifier is halved, and the saturation power of the carrier amplifier is doubled. Furthermore, when combined with the output of the peak amplifier that outputs the same output as the carrier amplifier, the Doherty PA has a saturation power that is 6 dB higher than the saturation power of the carrier amplifier alone. In Fig. 11 (b), the horizontal axis shows the amount of backoff based on the output power at which the carrier amplifier and peak amplifier are saturated, and the vertical axis shows the efficiency. The efficiency increases as the output power increases and the carrier amplifier saturates. After that, the peak amplifier starts operating, and along with that, the efficiency is slightly reduced, but the high efficiency is maintained. When both the carrier amplifier and the peak amplifier are saturated, the efficiency is maximized again.

It can be seen that the above operation can realize high efficiency even at the output level backed off from the maximum output power, and it is suitable for improving the efficiency of a modulated signal having a large PAPR.

It is desirable to set the load impedance high from the viewpoint of increasing load impedance, a differential load impedance, a series type Doherty PA has been proposed instead of the parallel as introduced here [43]. Also, from the viewpoint of increasing load impedance, a differential, series-added Doherty PA for HPUE compatible with n40 and n41, which is a band for NR [47].

Fig. 12 shows the outline of the prototype Doherty PA module. Fig. 12 (a) shows the circuit configuration. A 90 deg phase shift circuit was provided at the input, and the first stage was single-ended and the final stage was differential. Impedance was converted using a transformer and serial addition was performed. Fig. 12 (b) shows the appearance of
the prototype Doherty PA module. The size of the GaAs HBT chip is 1.0x1.15 mm^2, and the module size created with the PCB 6 layers is 2.6x2.2 mm^2, which is a compact size suitable for terminal application.

![GaAs HBT chip and PCB](image)

**Fig. 12** The outline of the prototype Doherty PA module.

Fig. 13 shows the characteristics of the prototype Doherty PA module. A 100 MHz bandwidth (270 RB) DFTS-OFDM QPSK (PAPR is about 7 dB @ CCDF (Complementary Cumulative Distribution Function) =0.003%) was used for the modulated signal. In the evaluation, we used DPD of the polynomial that does not correspond to the memory effect. Fig. 13 (a) shows the efficiency and ACLR of each output at 2.5 GHz. For ACLR, both the LSB (lower side band) side and the USB (upper side band) side are shown to confirm the memory effect. It can be seen that the difference between the both is small and the memory effect is suppressed. A maximum output power of 30.5 dBm and an efficiency of 36% were achieved under the condition of ACLR=-35 dBc. Fig. 13 (b) shows the efficiency and frequency characteristics of ACLR. A frequency band sufficient to cover the n40 and n41 bands has been achieved. As described above, it was demonstrated that Doherty PA is promising as a backoff improvement technology in NR. The details of Doherty PA introduced here are described in the paper [47], and please refer to that as well.

![Efficiency and ACLR graphs](image)

**Fig. 13** The characteristics of the prototype Doherty PA module.

6. Conclusions

The PA circuit of GSM, W-CDMA, LTE and NR are summarized. Of course, the first issue of PA is to improve efficiency, but there are technical issues other than simple improvement of efficiency in order to deal with the control peculiar to the generation of mobile phones.

Regarding GSM, we have summarized the control technology (that smoothly ramps-up and ramps-down) for intermittent operation and suppresses out-of-band spurious.

In W-CDMA and LTE, the probability of existence of transmit power is highest at low output power, and it is important to reduce current consumption and gain at low output. This paper summarizes these methods. In addition to the FDD technology used in W-CDMA from LTE, TDD technology has come to be applied, and the application of HPUE, which increases the output by 3 dB compared to the past, has progressed, and along with this, power combining method has become in demand. In this paper, we compared the power combining method and showed that the differential method is superior. The possibility of further increase in transmit power is also being discussed, and the power combining method will continue to be an important issue in the future.

As the PAPR of the modulated signal increases with each generation of LTE and NR, improving efficiency in the backoff region has become a major issue. First, we introduced the ET method currently applied. As the generation progresses with LTE and NR, the modulation bandwidth also expands. When the modulation bandwidth is 100 MHz or wider, there is a concern that the efficiency of the envelope tracker used in the ET method will decrease. Introduced Doherty PA as an alternative method candidate. It was shown that the differential and series addition methods are applied to secure the band, and the characteristics that can withstand practical use can be realized by covering n40 and n41 bands.

In the future, it is expected that frequencies above 7 GHz band (FR3) will be assigned to mobile phones, and it will be important to apply them to high frequencies in addition to high efficiency, wideband, and high output, and further technological innovation is desired.

Acknowledgments

We would like to express our deep gratitude to H. Sato, Y. Honda, K. Takenaka, M. Ito, S. Arayashiki, and T. Tsutsui for their useful discussions in compiling this paper.

References

[25] B. Tomas and J Johnson, “New RF Metrics for the smartphone-
[24] GSM Association “Battery life measurement technique 5.1” GSM
[48x260]interface including a loop-bandwidth calibration system,” IEEE ISSCC,
[48x298]B. Kim, “Broadband class-F power amplifiers for handset applications,”
[48x307]45/46/34% PAE linear polar transmitter for EDGE/WCDMA/Mobile-
[40] J. Enomoto, T. Terashima, M. Itou, K. Watanabe, Y . Tanaka, S. Tanaka,
[38] M. Hassan, C. Olson, D. Kovace, J. J. Yan, D. Nobbe, D. Kelly, P. M.
[35] M. Miyashita, T. Okuda, H. Kusuda, S. Shimamura, S. Konishi, J. Udono,
[34] G. Hau, T. B. Nishimura, and N. Iwata, “High efficiency, wide dynamic
[33] T. Tanoue, M. Ohnishi, and H. Matsumoto, “Switch-off-lessempedance-
[31] 3GPP, “3rd generation partnership project; technical specification
Yamabe, K. Maeda, F. Kitabayashi, Y . Sasaki, M. Hirobe, S. Shinjo, K.
Horiguchi, T. Sumino, S. Suzuki, H. Katayama, T. Shimura, and H. Seki,
“A two-power-mode Si-CMOS GaAs-HBT hybrid power amplifier module
for 0.9-GHz-band WCDMA handsets applications,” APMC, pp.668-670, Nov.
2014.
Y. Kikuchi, “A quadrature hybrid amplifier using active bias circuits
and its optimization for class-AB InGaP-GaAs HBT power amplifiers,” IEEE
Trans. Microwave theory and techniques, vol.52, no.5, pp.1455-1465, May
2004.
power amplifier with an improved low-power efficiency using stage-bypass,” IEEE RFIC
Symposium, pp.163-166, June 2010.
[25] B. Tomas and J Johnson, “New RF Metrics for the smartphone-
[24] GSM Association “Battery life measurement technique 5.1” GSM
[23] B. Tomas and J Johnson, “New RF Metrics for the smartphone-
chain GaAs MESFET-TXMUIC dual band power amplifier for GSM,” IEEE

Satoshi Tanaka received the B.S. and M.S. degrees from Waseda University, Tokyo, Japan, in 1983 and 1985, respectively, and the Ph.D. degree in engineering from Tohoku University, Miyagi, Japan, in 2019. In 1985, he joined the Central Research Laboratory, Hitachi, Ltd., Tokyo, where he was engaged in the research and development of mixed analog and digital bipolar CMOS, GaAs ICs, CMOS, and Bi-CMOS RFICs for GSM and W-CDMA applications. In 2006, he moved to Renesas Technology Co., Tokyo, and developed a power amplifier for mobile phones using LDMOS and HBT. In 2012, he joined Murata Manufacturing Co., Ltd., Kyoto, Japan. Currently, he develops RF front-end modules, especially multiband multimode PA modules for GSM/W-CDMA/LTE/5G. Dr. Tanaka was a Technical Committee Member of the IEEE International Conference on Solid-State Circuits (ISSCC) from 2005 to 2009. From 2009 to 2018, he was an RF program committee chair of the IEEE Asian Conference on Solid-State Circuits (A-SSCC). In 2015, he was Chair of the Technical Committee on Circuits and Systems (CAS) of the IEICE. He is a senior member of IEEE.

Kenji Mukai received the B.S. degree from Osaka City University, Osaka, Japan, in 2004, and the M.S. degree from Osaka University, Osaka, in 2006, both in physics, respectively. In 2006, he joined Mitsubishi Electric Co., Tokyo, Japan, where he was engaged in the research and development of GaAs HBTs and GaN HEMT power amplifiers for mobile communications. From 2012 to 2013, he was a visiting scholar with the University of California, San Diego, CA, USA, where he researched GaN HEMT switching-mode power amplifiers for envelope tracking and digital transmitters. In 2017, he joined the RF Device Division at Murata Manufacturing Co., Ltd., Kyoto, Japan, and was instrumental in developing PAMiD and PAMiF for 5G NR. His current research interests include GaAs and Si-based RF MMIC front-end modules and their efficiency enhancement techniques.

Shohei Imai received the B.S. and M.S. degrees in electronics engineering from Okayama University, Okayama, Japan, in 2008 and 2011, respectively. He worked as an Engineer of microwave high-power amplifiers for GaN HEMTs with Mitsubishi Electric Co., Tokyo, Japan, from 2011 to 2016. He worked with Honda R&D Co., Ltd., Saitama, Japan, from 2016 to 2018 as an Engineer of vehicle-to-vehicle communication systems and applications. Since 2019, he has been working with Murata Manufacturing Co., Ltd., Kyoto, Japan. His research interests include efficient RF power amplifiers and mobile communication systems. He is a member of IEEE.

Hiroshi Okabe received the B.S. and M.S. degrees from the University of Electro Communications, Tokyo, Japan, in 1990 and 1992, respectively. From 1992 to 2012, he was with the Central Research Laboratory, Hitachi, Ltd., Tokyo, where he was engaged in the research and development of cellular handsets, built-in antennas, and RF front end modules. From 2001 to 2002, he was a visiting scholar with the Electrical Engineering Department, University of California, Los Angeles, USA, where he researched the circuit applications of metamaterials. In 2012, he joined Murata Manufacturing Co., Ltd., Kyoto, Japan. Currently, he leads the research and development of power amplifier ICs and front-end solutions for the latest cellular handsets. He is a member of IEEE.