CMOS high linearity PA driver with an on-chip transformer for W-CDMA application*

Fu Jian(付健), Mei Niansong(梅年松), Huang Yumei(黄煜梅)†, and Hong Zhiliang(洪志良)

ASIC & System State Key Laboratory, Fudan University, Shanghai 201203, China

Abstract: A fully integrated high linearity differential power amplifier driver with an on-chip transformer in a standard 0.13-μm CMOS process for W-CDMA application is presented. The transformer not only accomplishes output impedance matching, but also acts as a balun for converting differential signals to single-ended ones. Under a supply voltage of 3.3 V, the measured maximum power is larger than 17 dBm with a peak power efficiency of 21%. The output power at the 1-dB compression point and the power gain are 12.7 dBm and 13.2 dB, respectively. The die size is 0.91×1.12 mm^2.

Key words: CMOS; PA driver; on-chip transformer; impedance transformation; linearity; efficiency
DOI: 10.1088/1674-4926/32/9/095006   **EEACC:** 1220

1. Introduction

The last decade has witnessed the phenomenal growth of the wireless industry, and low-cost, reliable and power-efficient wireless systems are in demand. High levels of integration, such as software-defined radio (SDR), are sought after in order to reduce cost. CMOS technology is the prime contender for this level of integration. Although most of the RF building blocks have been successfully integrated into a CMOS process, the power amplifier (PA) with a PA driver (PAD) is mostly implemented in a different process, such as GaAs HBT (hetero-junction bipolar transistors) in conjunction with external passives[1]. The primary reasons for this are the high output power and efficiency required by the PA, which are difficult to achieve in CMOS technology. In addition, on-chip passives suffer from a low quality factor due to the thin metals used, as well as the conductive nature of the silicon substrate. These losses degrade both the output power and efficiency of the PA and the PAD. The design of a CMOS PA and a fully integrated PAD thus remains a complex challenge.

In the design of PADs, a lossy output matching network is detrimental to achieving high output power, efficiency and linearity. To implement output networks, high-Q off-chip components, such as inductors and capacitors, are needed. However, this will increase the size as well as the cost of the PAD module. In order to reduce the size and the cost of the PAD, on-chip output matching network is required.

There have been several efforts to reduce the loss of the on chip output matching network in recent studies of CMOS PAs and PADs, such as using a distributed active transformer[2] and a power combining transformer[3]. A fully integrated CMOS power amplifier driver with a spiral type transformer as the output matching network is presented in this work. The analysis of on-chip transformers is described and a weak class A amplifier is shown. A PAD with an on-chip transformer delivering a maximum output power of 17 dBm and a peak power added efficiency (PAE) of 21% is also described.

2. Transformer as an output matching network for a PAD

2.1. Simplified model of a transformer

On-chip transformers are often implemented as coupled inductors for output impedance matching. Therefore, non-idealities such as leakage flux, feed-through capacitance and various losses cause real transformers to deviate from ideal behaviors. It is necessary to develop a mathematical equation based on a simplified model, as shown in Fig. 1[4]. In a coupled-inductor transformer, the magnetic field created by the port-1 current I_1 through the primary inductor L_1 generates a voltage in the secondary inductor L_2. At the same time, the current through the secondary I_2 will magnetically induce a voltage in the primary circuit. The Z-parameter two-port equations for the transformer model can be expressed as that in Ref. [4] (note that the flow direction of I_2):

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} R_1 + j\omega L_1 & -j\omega M \\ j\omega M & -R_2 - j\omega L_2 \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}, \quad (1)$$

$$M = k\sqrt{L_1 L_2}, \quad (2)$$

where M is the mutual inductance and k is the coupling factor. The losses of the primary and the secondary inductors are
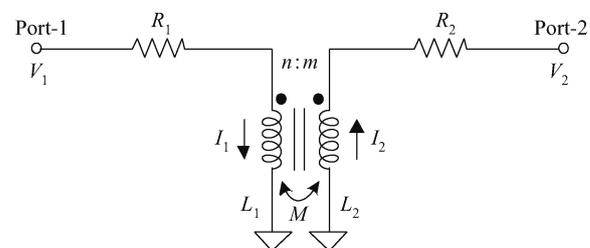
Fig. 1. A simplified model of transformer.
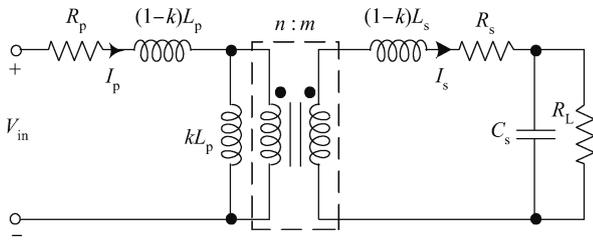
Fig. 2. Transformer equivalent T-model.

usually modelled with the resistors, R_1 and R_2, as shown in Fig. 1, where

$$R_1 = \frac{\omega L_1}{Q_1}, \quad R_2 = \frac{\omega L_2}{Q_2}. \quad (3)$$

The quality factors of the coupled inductors are slightly different from those of the individual inductors due to the current redistribution. The T-model (as shown in Fig. 2) can be modified so it is expressed in terms of leakage inductance and magnetizing inductance. As the transformer is used to achieve output matching in the PAD, it will be necessary to resonate some of the transformer's inductance to minimize the insertion loss.

From Fig. 2, we can calculate the efficiency η of the transformer[4]:

$$
\eta = \frac{P_{\text{load}}}{P_{\text{load}} + P_{\text{diss}}}
$$
$$
= \frac{\dfrac{R_0}{1 + (\omega R_0 C_s)^2}}{\dfrac{R_0}{1 + (\omega R_0 C_s)^2} + R_s + \left|\dfrac{Z_s + j\omega k L_p}{j\omega k L_p}\right|^2 R_p}, \quad (4)
$$

and

$$
\eta_{\max} = \frac{1}{1 + \dfrac{2}{Q_p Q_s k^2} + 2\sqrt{\dfrac{1}{Q_p Q_s k^2}\left(1 + \dfrac{1}{Q_p Q_s k^2}\right)}}, \quad (5)
$$

where P_{load} and P_{diss} are the power delivered to the load and power dissipated in the parasitic resistance of the transformer. Z_s is the secondary side impedance transformed to the primary side. Q_p and Q_s represent the quality factors of primary and secondary windings. This result shows that efficiency can be increased by either increasing the quality factor of the windings or by increasing the coupling between the primary and secondary windings.

2.2. Impedance transformation

The transformer also performs an impedance transformation network. Figure 3 shows the principle of the voltage transformation and the layout of the 3 : 2 transformer. We can calculate transformed impedance as follows from Fig. 3:

$$
\frac{R_{\text{eq}}}{R_{\text{L}}} = \frac{\dfrac{9}{2}V/I}{6V\Big/\left(\dfrac{3}{2}I\right)} = \frac{9}{8}. \quad (6)
$$

(a)
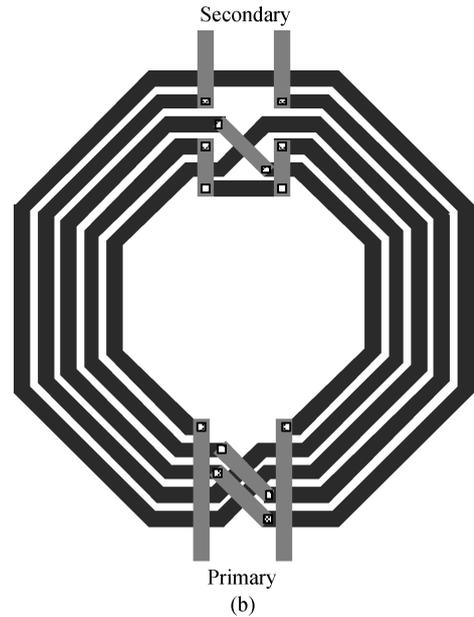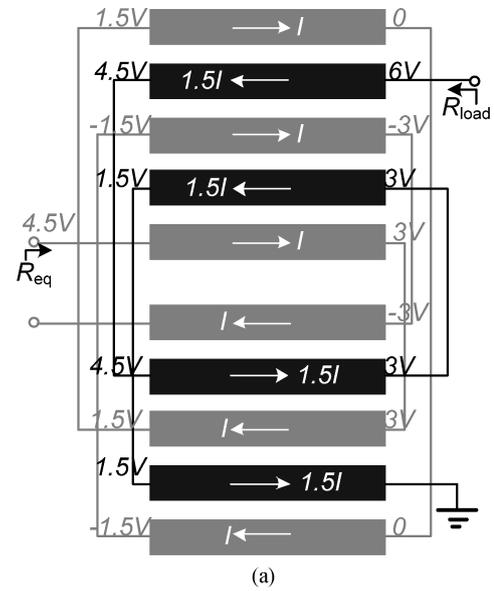
Fig. 3. (a) Schematic diagram and (b) layout of the 3 : 2 transformer.

If $R_{\text{L}} = 50\,\Omega$, $R_{\text{eq}} = 56.25\,\Omega$.

Generally, the impedance transformation ratio is defined as:

$$
r = \frac{R_{\text{L}}}{R_{\text{m}}} = \frac{V_2/I_2}{V_1/I_1} = \frac{V_2/V_1}{I_2/I_1} = \frac{m/n}{n/m} = \frac{m^2}{n^2}, \quad (7)
$$

where R_{m} is the differential load impedance seen at the transistor drain. $1 : N$ transformers in PA design for high output power are designed in most cases. However, the PAD in this work does not need such a high power as a PA. Taking this into consideration, a 3 : 2 transformer for moderate output power is defined. From Eq. (6), the differential load impedance seen at the transistor drain is $112.5\,\Omega$ ($56.25 \times 2\,\Omega$) when the load impedance is $50\,\Omega$ in a 3 : 2 transformer. The sizes of transistors and currents in PADs are much smaller than those in PAs, thus if a differential input and single end output core is used, the differential load impedance is generally about several hundred Ohms. Load-line matching theory[5] is widely used in output matching for PA. In the PAD, the output power is not
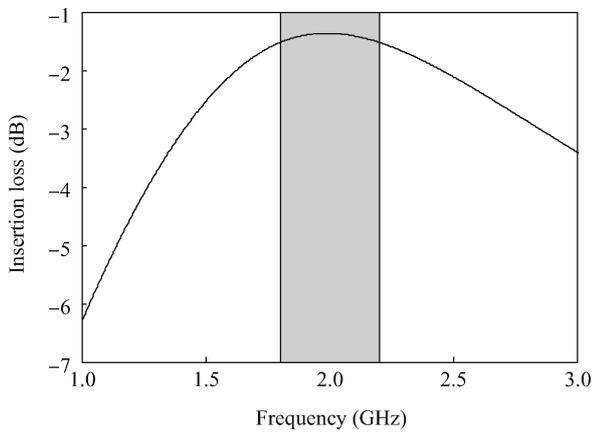
Fig. 4. Simulated insertion loss of the transformer in the band of interest.

Fig. 5. Typical PA transconductor and its derivatives versus input bias v_{gs} (simulation results).

so large that we can still use small signal impedance matching. As mentioned above, the differential load impedance is about 112 Ω in this work as the load impedance is 50 Ω which is the input impedance of the off-chip PA. If the capacitors and inductors (in the transformer) are tuned out, by carefully choosing the sizes of transistors (especially output transistors), the output impedance at the drains of the output transistors can be well transformed to load impedance by the transformer. So there is no need for another matching network for output impedance matching. We call this phenomenon "impedance self-matching".

2.3. 3:2 transformer design

The layout of the symmetric 3 : 2 transformer is shown in Fig. 3(b). A notable feature of this transformer is that only two metal layers (Metal 7 and Metal 8) were used for the windings, which simplifies the transformer design. Only the top metal layer M8 was presented for the main windings, while the underpass elements were realized with only one lower metal layer M7. Electromagnetic simulation has been performed using an Agilent Momentum EM simulator. Custom finger metal–insulator–metal ("MIM") and off-chip surface mount capacitors have been employed on both the primary and secondary sides of the transformer to tune the transformer to the desired center frequency of about 2 GHz, respectively. The simulated insertion loss of the presented 3 : 2 transformer is 1.37 dB at 2 GHz, as shown in Fig. 4, and varies only 0.15 dB over the band of interest from 1.8–2.2 GHz. The size of the transformer is 0.36×0.34 mm^2.

3. Weak class A amplifier

In a general common source amplifier, the voltage-to-current transfer characteristic is given by Eq. (8), where I_{dc} and v_{gs} are dc current and gate-to-source voltage, respectively.

$$i_{DS}\left(v_{gs}(t)\right) = I_{dc} + G_1 v_{gs}(t) + G_2 v_{gs}^2(t) + G_3 v_{gs}^3(t) + \cdots,$$

(8)

G_1 is the amplifier's transconductor while G_2 and G_3 are the transconductor's derivatives versus input bias v_{gs}. It is well known that the coefficient of $v_{gs}^3 G_3$ in Eq. (8) plays an im-

portant role in the third order inter-modulation (IMD3) distortion of a RF PA[6]. A typical current's derivative (G_1, G_2, G_3) characteristics of a NMOS-FET (in fact it is a typical PA with a single transistor) are shown in Fig. 5 (to enhance plot visibility, G_3 is compressed with a factor 4). It is shown that G_3 goes to the positive peak at the sub-threshold region, then crosses zero near the threshold voltage and then reaches the negative peak at the point where the gate voltage is slightly larger than the threshold voltage. Different bias voltages mean different working regions for PAs. G_3 is larger than zero at class C and class B regions, while less than zero at class A regions. As a result, G_3 gets close to zero by carefully biasing the PA works between class A and class B region. The bias current increases with the input signal level[7] in a CMOS class B amplifier. So the dc power consumption dynamically varies with the input signal. Therefore, The PAE of a CMOS class B amplifier is much higher than that of a class A amplifier when the input power level is small.
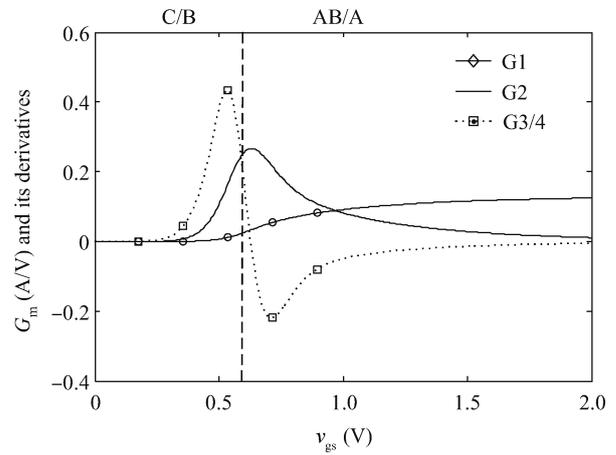
In the design of the PAD, the same trade-off between linearity and efficiency is also made as that in a PA. As discussed above, bias in a PAD operating between the class A and class B regions is preferred. The input transistors are biased slightly larger than the threshold voltage, which is called "weak class A biasing". Weak class A biasing (generally called class AB biasing) for a PAD can achieve both high linearity and moderate efficiency.

4. Weak class A PAD with transformer design

A differential weak class A PAD with an on-chip transformer for W-CDMA application is designed. The fully integrated PAD, as shown in Fig. 6, was fabricated in a standard 0.13-μm CMOS technology. The PAD was implemented in pseudo-differential topology, which creates an ac virtual ground at the common source node and VDD node to reject even harmonic distortion and also reduce the impact of the ground and VDD bond wires. The PAD, designed for a W-CDMA transmitter, needs to drive an off-chip PA to satisfy the system specifications. Generally, off-chip PAs are single-ended, which demands that the output of the PAD is single-ended. Fortunately, the PAD of this work includes an on-chip
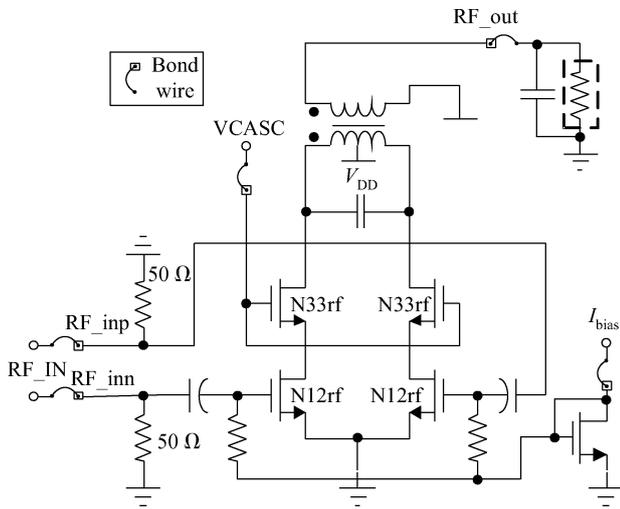
Fig. 6. Schematic of the PAD with an on-chip transformer.
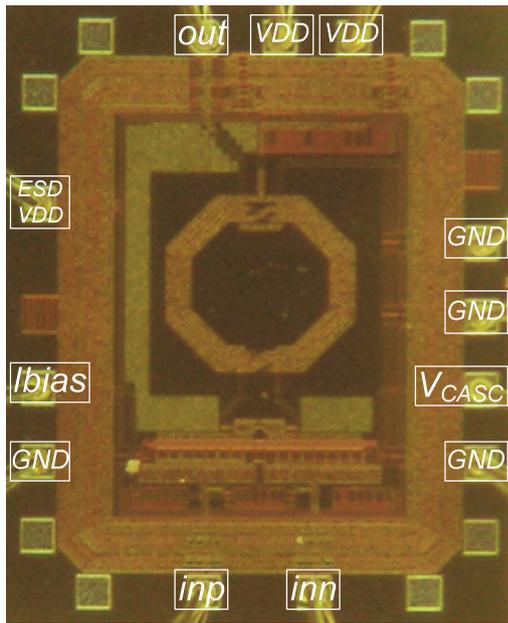
Fig. 7. Micro-photograph of the PAD (0.91×0.112 mm^2).

Fig. 8. Measured small-signal (S-parameter) performance of the PAD.

Fig. 9. Measured power gain and large-signal transfer characteristics of the PAD.

5. Experimental results

The micro-photograph of the PA driver is expressed in Fig. 7. The chip area including the pads is 0.91×0.112 mm^2. A printed circuit board (PCB) was used for chip measurements. It is worth mentioning that the losses caused by the PCB, such as the input balun and signal line, have been carefully de-embedded. To minimize the off-chip parasitic effects, all the bond-wires except for bias are double bonded.

At a supply voltage of 3.3 V, small-signal measurements were performed on the PAD, as shown in Fig. 8. The PAD has a small signal gain of 13 dB at 1.95 GHz. The input and output match are better than –10 and –11 dB over the interesting bands, respectively. The large-signal measurements with a continuous-wave signal at the frequency of 1.95 GHz were performed and the measurement results are presented in Fig. 9. The measured saturated power is larger than 17 dBm with a power gain of 13.2 dB and the output power at the 1-dB compression point is 12.7 dBm. The $P_{\text{1-dB}}$ can be varied by about ± 0.5 dB around 12.5 dBm by adjusting the gate bias of the input devices. The efficiency curve is plotted in Fig. 10. The PAE is 13.6% at the 1-dB compression point with the maximum value larger than 21%.

For linearity measurement, a two-tone test was performed

transformer, which not only accomplishes output impedance matching but also converts the differential signals to single-ended signals. Otherwise it will need an off-chip balun, which increases the design complexity and cost. The cascode configuration is applied to the core to improve reliability. The common source transistors are standard thin-gate-oxide devices while the cascode transistors are high voltage thick-gate-oxide to enhance reliability. As the output power is not too high in the PAD, there is enough margin to keep the devices work safely. To make a trade-off between linearity and efficiency, we bias the PAD in a weak class A region. As mentioned in Ref. [8], a PAD in a weak class A region can achieve high linearity with moderate efficiency. To reduce the power and ground bounce, large on-chip bypass capacitors are connected to the center-tap of the transformer (the supply node) and the ground nodes. Considering the size of the whole chip, only NMOS capacitors are used.

Table 1. Summary of measured performance and comparisons.

Parameter	RFIC06[9]	RFIC07[10]		MWCL09[11]	This work
CMOS technology (μm)	0.18	0.18		0.18	0.13
Frequency (GHz)	2.4	2.4	5.2	2.5	1.95
Supply voltage (V)	1.8	1.8	1.8	1.8	3.3
Power gain (dB)	9.2	10.4	5.1	11.8	13.2
P_{sat} (dBm)	—	13	8.7	—	17
OP @ 1 dB (dBm)	7.8	11.6	7.6	5.5	12.7
OIP3 (dBm)	20 @ –5 dBm	22	17	16.1	23 @ 9 dB back off
PAE (%) Saturated	—	—	—	—	21%
1dBCP	—	16.2	10.8	21%	13.6%
On-chip balun	No	No	No	No	Yes
Sizes (mm^2)	0.348*	0.48*	0.48*	—	1.02
FOM	—	113	113	85	87

* The sizes of the references above are much smaller (Ref. [10] excludes pads) than in this work as they all exclude the balun. So the specifications of the references may degrade a little if applied with balun, especially with an on-chip balun.
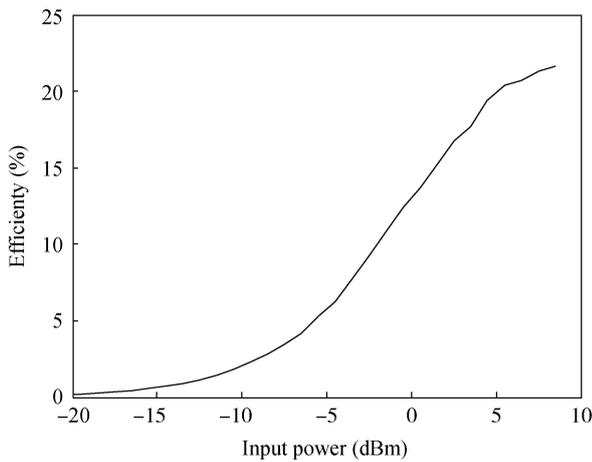
Fig. 10. Power efficiency versus input power of the PAD.
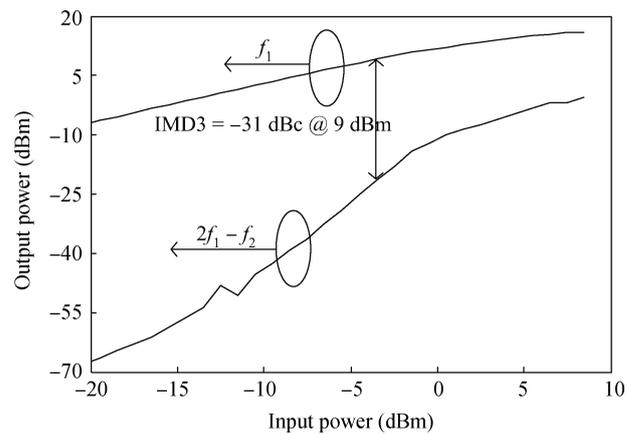
Fig. 12. Measured output power and IM3 for a two-tone test.
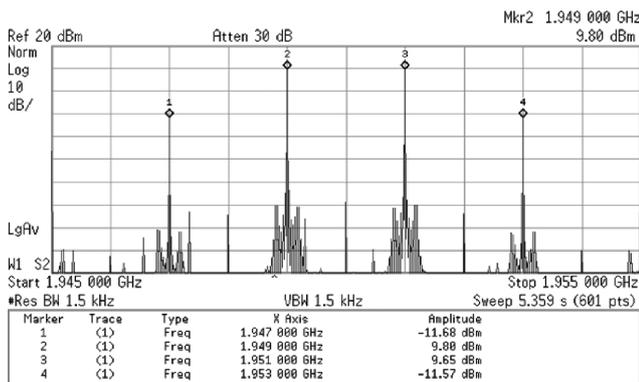
Fig. 11. Measured spectrum for a two-tone test with total output power 12.7 dBm.

with tone-spacing of 2 MHz. Two tones generating a total output power of 12.7 dBm were applied to the input of the PA driver, as shown in the spectrum of Fig. 11. Little asymmetry was observed during testing. The measured IMD3 at 1-dB compression point is less than –21 dBc. The IM3 curve is shown in Fig. 12. The IMD3 is –46 dBc at around 0.5 dBm output power (one tone each, 9 dB back-off). The calculated power at output third order intercept point (OIP3) is better than 23 dBm.

Table 1 presents a summary of measured performance of the PAD and comparisons between this work and other reported designs. It shows that a largest output power at 1 dB compression point and OIP3 with a moderate PAE are achieved in the PAD of this work. However, it should be noted here that the supply voltage in this work is 3.3 V while the others are 1.8 V. To make a fair comparison, a figure of merit (FOM) is defined by the ITRS[12], as shown in Eq. (9) (where we made a minor alteration for $P_{1\text{-dB}}$ instead of saturated power), which links the output power at 1-dB compression point ($P_{1\text{-dB}}$) with the power gain (Gain) and PAE, while the f^2 term reflects the degradation in gain and output power with increasing frequency:

$$\text{FOM} = f^2 \times \text{Gain} \times P_{1\text{-dB}} \times \text{PAE}_{1\text{-dB}}. \qquad (9)$$

As mentioned above, the losses caused by on-chip components are much larger than for off-chip devices. From the comparison of the FOMs, the PAD with an on-chip transformer in this work performs not much worse than those with off-chip impedance matching and a balun. In the transmitter design, in order to meet the high output power and linearity specifications of the transmitter, the power gain and linearity of a PA driver tends to be the key performance parameters. It can be seen from the table that the PAD in this work achieves a high power and gain linearity for W-CDMA application.

6. Conclusion

A CMOS PAD, fabricated in a standard 0.13-μm CMOS process, has been presented. A 3 : 2 transformer for the PAD has been designed to serve as an output matching network and a differential-to-balance converter. Under a 3.3 V power supply, a saturated power of larger than 17 dBm with a peak PAE of 21% is achieved. It can deliver 12.7 dBm power at the 1-dB compression point. The PA driver has a high power gain and linearity with a moderate PAE, which makes it suitable for W-CDMA application.

References

[1] Tseng P, Zhang L, Gao G, et al. A 3-V monolithic SiGe HBT power amplifier for dual-mode (CDMA/AMPS) cellular handset applications. IEEE J Solid-State Circuits, 2000, 35(9): 1338

[2] Aoki I, Kee S, Rutledge D, et al. Fully integrated CMOS power amplifier design using the distributed active-transformer architecture. IEEE J Solid-State Circuits, 2002, 37(3): 371

[3] An K, Lee O, Kim H, et al. Power-combining transformer techniques for fully-integrated CMOS power amplifiers. IEEE J Solid-State Circuits, 2008, 43(5): 1064

[4] Aoki I, Kee S, Rutledge D, et al. Distributed active transformer—a new power-combining and impedance-transformation technique. IEEE Trans Microw Theory Tech, 2002, 50(1): 316

[5] Cripps S. RF power amplifiers for wireless communications. 2nd ed. Norwood: Artech, 2006

[6] Tanaka S, Behbahani F, Abidi A. A linearization technique for CMOS RF power amplifiers. IEEE Symposium VLSI Circuits, 1997: 93

[7] Ding Y, Harjani R. A high-efficiency CMOS +22-dBm linear power amplifier. IEEE J Solid-State Circuits, 2005, 40(9): 1895

[8] Fager C, Pedro J, Carvalho N. et al. A comprehensive analysis of IMD behavior in RF CMOS power amplifiers. IEEE J Solid-State Circuits, 2004, 39(1): 24

[9] Kim J, Kim T, Jeong M, et al. A 2.4-GHz CMOS driver amplifier based on multiple-gated transistor and resistive source degeneration for mobile WiMAX. IEEE Asian Solid-State Circuits Conference, 2006: 255

[10] Magnusson H, Olsson H. A compact dual-band power amplifier driver for 2.4 GHz and 5.2 GHz WLAN transmitters. IEEE Radio Frequency Integrated Circuits Symposium, 2007: 83

[11] Kim D, Hong N, Choi Y. A novel linearization method of CMOS drive amplifier using IMD canceller. IEEE Microw Wireless Compon Lett, 2009, 19(10): 671

[12] International Technology Roadmap for Semiconductors, 2007 edition. [Online], Available: http://public.itrs.net