

# High-Power High-Efficiency Class-E-Like Stacked mmWave PAs in SOI and Bulk CMOS: Theory and Implementation

Anandaroop Chakrabarti and Harish Krishnaswamy

**Abstract**—Series stacking of multiple devices is a promising technique that can help overcome some of the fundamental limitations of CMOS technology in order to improve the output power and efficiency of CMOS power amplifiers (PAs), particularly at millimeter-wave (mmWave) frequencies. This paper investigates the concept of device stacking in the context of the Class-E family of nonlinear switching PAs at mmWave frequencies. Fundamental limits on achievable performance of a stacked configuration are presented along with design guidelines for a practical implementation. In order to demonstrate the utility of stacking, three prototypes have been implemented: two fully integrated 45-GHz single-ended Class-E-like PAs with two- and four-stacked devices in IBM's 45-nm silicon-on-insulator (SOI) CMOS technology, and a 45-GHz differential Class-E-like PA with two devices stacked in IBM's 65-nm low-power CMOS process. Measurement results yield a peak power-added efficiency (PAE) of 34.6% for the two-stacked 45-nm SOI CMOS PA with a saturated output power of 17.6 dBm. The measurement results also indicate true Class-E-like switching PA behavior. A peak PAE of 19.4% is measured for the four-stacked PA with a saturated output power of 20.3 dBm. The two-stacked PA exhibits the highest PAE reported for CMOS mmWave PAs, and the four-stacked PA achieves the highest output power from a fully integrated CMOS mmWave PA including those that employ power combining. The 65-nm CMOS differential two-stacked PA exhibits a peak PAE of 28.3% with a saturated differential output power of 18.2 dBm, despite the poor ON-resistance of the 65-nm low-power nMOS devices. This paper also describes the modeling of active devices for mmWave CMOS PAs for good model-hardware correlation.

**Index Terms**—Class-E, CMOS, high efficiency, millimeter wave (mmWave), power-added efficiency (PAE), power amplifier (PA), power device modeling, stacking.

## I. INTRODUCTION

### A. Millimeter-Wave (mmWave) CMOS Power Generation Challenges

THE advent of scaled CMOS technologies with transistor  $f_{\max} \approx 250$  GHz has generated significant interest in using the mmWave bands above 30 GHz in applications such

Manuscript received November 09, 2013; revised February 09, 2014 and March 13, 2014; accepted May 11, 2014. Date of publication June 12, 2014; date of current version August 04, 2014.

The authors are with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA (e-mail: ac3215@columbia.edu; harish@ee.columbia.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMTT.2014.2327919

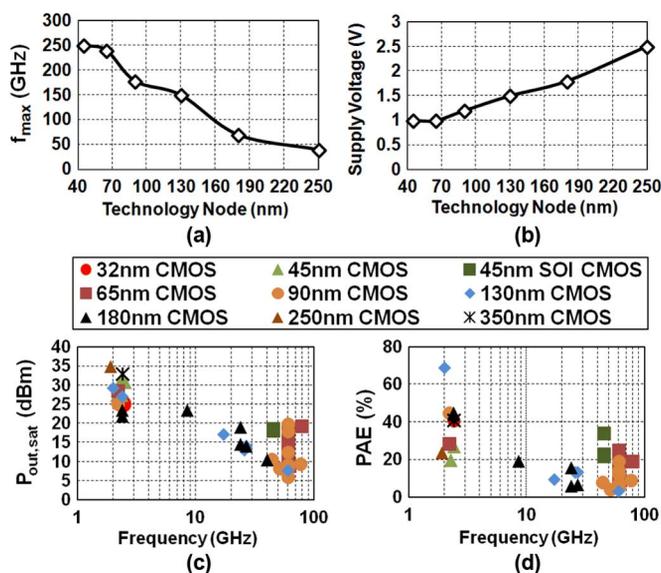


Fig. 1. (a)  $f_{\max}$  of deep-submicrometer CMOS technology nodes from the literature. (b) Supply voltage of deep-submicrometer CMOS technology nodes. (c) Survey of the saturated output power achieved by reported RF and mmWave CMOS PAs. (d) PAE achieved by reported RF and mmWave CMOS PAs.

as wideband commercial wireless communication, satellite communication, automotive radar, and biomedical imaging. This, in turn, has driven research focus on the development of efficient mmWave power amplifiers (PAs). The frequency band around 45 GHz (upper end of the  $Q$ -band, which extends from 33 to 50 GHz) is well suited for satellite communication owing to a low atmospheric attenuation of 0.2 dB/Km. The requirement of high output power for such long range communication necessitates a high-power PA, in addition to energy-efficiency. Traditionally, III–V compound semiconductor technologies were the preferred choice for implementing such amplifiers. This is because implementing high-power high-efficiency PAs in CMOS at mmWave frequencies has proven to be a challenging task owing to the limited breakdown voltage of highly scaled CMOS technologies, low available gain of devices, and poor quality of on-chip passives. The low breakdown voltage limits the output swing, and consequently, the output power that can be delivered to a  $50\text{-}\Omega$  load. The load may be transformed to a lower impedance to enable higher output power, but the poor quality of on-chip passive components limits the efficiency of the transformation. The low available gain results in large input power requirements for mmWave PAs, limiting power-added

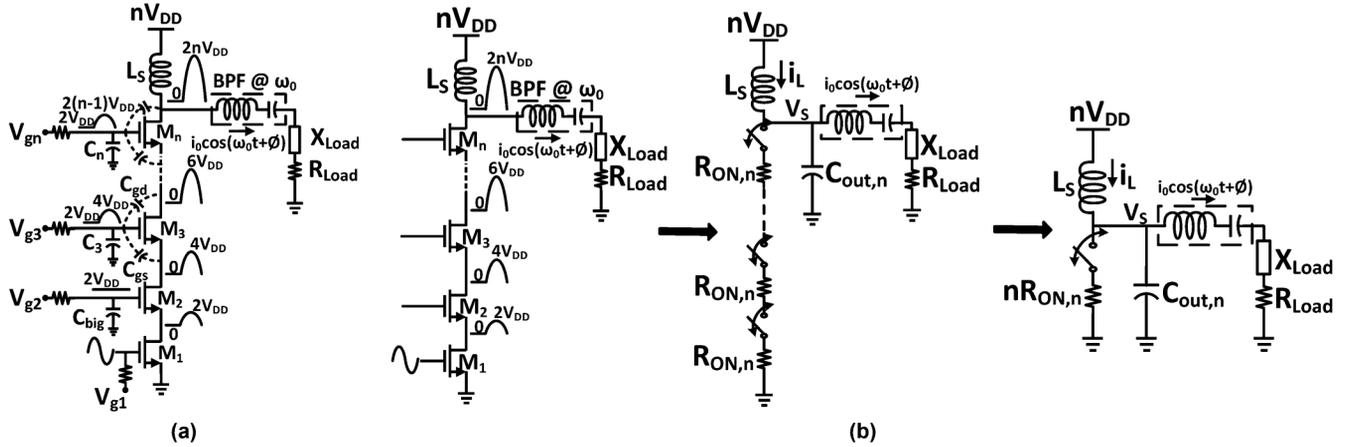


Fig. 2. (a) Stacked CMOS Class-E-like PA concept with voltage swings annotated in volts and (b) loss-aware Class-E design methodology for stacked CMOS PAs.

efficiency (PAE). A possible solution is to power combine the outputs of several PAs. However, the efficiency and scalability of conventional on-chip power-combining schemes such as transformer-based series combining [1]–[5], current combining [6], and Wilkinson combining are limited by the characteristics of the back-end-of-the-line, steepness of inherent impedance transformation, and combiner asymmetry.

These design tradeoffs can be clearly appreciated from a survey of RF and mmWave PAs reported in the literature (Fig. 1). Fig. 1(a) depicts the scaling of  $f_{max}$  of CMOS technology nodes based on prior reports [7]–[10], while Fig. 1(b) depicts the scaling of supply voltage  $V_{DD}$ . Experimentally, it would seem that as technology scales,  $f_{max} V_{DD}^2$  remains approximately constant and equal to 250 GHz-V<sup>2</sup>, although explicitly deriving such a scaling law for constant-field scaling remains challenging due to the complex and layout-dependent nature of loss mechanisms within nanoscale CMOS devices. This observation is, however, consistent with the known fundamental tradeoff (referred to as the Johnson limit [11]) between speed of operation and the breakdown voltage of a technology. If one assumes that a PA designer chooses a technology with an  $f_{max}$  that is three times the operating frequency  $f$  to ensure approximately 10-dB gain, that the PA is designed to directly drive a 50- $\Omega$  load with no impedance transformation to maintain efficiency, and that the output node sustains a peak-to-peak swing that is twice the  $V_{DD}$  (i.e., no harmonic shaping), the output power of such a PA would be  $(250 \text{ GHz-V}^2/3f) \times (1/2 \times 50 \Omega) = (830 \text{ mW} * \text{GHz})/f$ . This scaling law indicates that achieving watt-class output power at frequencies around 1 GHz is feasible, but typical output powers at, say, 60 GHz, would be in the range of 10–15 mW if impedance transformation or power combining are not exploited. Fig. 1(c) largely bears out this trend, with some efforts achieving higher output powers through either impedance transformation, power combining, or a combination of the two. Fig. 1(d) indicates that efficiency also degrades significantly as frequency increases, although an explicit scaling trend for efficiency is more complicated because of the complex nature of loss mechanisms in active and passive devices. State-of-the-art

PAEs at mmWave frequencies are generally below 20%, except for a few outliers.

### B. Device Stacking in CMOS PAs

Series stacking of multiple devices [e.g., Fig. 2(a)] is a potential technique that breaks these tradeoffs associated with CMOS PA design. Stacking of multiple devices increases the voltage swing at the load, as the increased voltage stress can be shared by the various devices in the stack. Thus, for a stack of  $n$  devices, the output voltage swing can be  $n$  times higher than that of a single device (provided that design techniques are incorporated to ensure that each individual device sees  $V_{gs}$ ,  $V_{gd}$ , and  $V_{ds}$  swings that lie within permissible breakdown limits). Stacking, however, does not alleviate the drain–bulk and source–bulk stress of the individual stacked devices. In particular, the topmost device of the stack sees a drain–bulk swing that is equal to the  $n$ -times increased output swing of the stacked PA. Consequently, in bulk CMOS, the junction breakdown voltage limits the maximum number of devices that can be stacked to three or four devices. However, in silicon-on-insulator (SOI) CMOS, the presence of an isolated floating body for each device eliminates this limitation. The number of devices that can be stacked in SOI CMOS is only limited by the breakdown of the buried oxide (BOX) below each device. This voltage is, however, higher than 10 V in IBM’s 45-nm SOI CMOS process [12], enabling five or more devices to be stacked (assuming 2-V peak RF voltage swing per device for long-term reliability [13]).

Prior studies on stacked PAs at RF frequencies have explored the cases where input power is provided only to the bottommost device in the stacked configuration [14], [15], as well as to all the devices in the stack through transformer coupling [16]. These works have demonstrated the characteristics of stacking in a variety of technologies such as GaAs MESFET [14], [16] and SOI [15]. More recently, stacking for mmWave PAs has been investigated in nanoscale SOI CMOS [12], [17]–[21]. However, while stacking has been predominantly explored in the context of linear/quasi-linear PAs, nonlinear switching-type stacked PAs remain an interesting topic of research.

Switching PAs are extensively utilized at RF frequencies owing to their (ideally) lossless operation. The Class-E PA [22] has been of particular interest because of its relatively simple output network. The design of switching PAs in CMOS at mmWave frequencies [12], [17], [19], [20] is challenging due to the lack of ideal square-wave drives (resulting in soft switching), impracticality of harmonic shaping of voltages and currents, low PAE due to the high input drive levels required to switch the devices, and high loss levels in the device/switch. Thus, at mmWave frequencies, one can practically implement a “switch-like” PA at best. In this paper, we explore stacked “Class-E-like” PAs in SOI CMOS and low-power bulk CMOS technologies.

In order to determine if device stacking overcomes the speed-breakdown voltage tradeoff of CMOS technology scaling (quantified earlier using the  $f_{\max}V_{DD}^2$  product), it is important to determine if stacked Class-E-like PAs are able to increase output power without substantial degradation in efficiency (the PA metric that is significantly impacted by transistor speed). A loss-aware Class-E design methodology that accounts for high device loss [23] is applied to stacked Class-E-like mmWave CMOS PAs to understand the design tradeoffs and determine the fundamental limits of performance. These theoretical results are further explained by means of an analysis that identifies technology-dependent metrics that govern the performance of stacked Class-E-like CMOS PAs. Specifically, the analysis introduces a technology constant, referred to as the switch time constant, which is an important technology metric for switching PAs in addition to  $f_{\max}$ . It is shown theoretically that mmWave stacked Class-E-like PAs implemented based on the loss-aware design methodology achieve better efficiency than PAs using conventional power enhancement techniques, such as impedance transformation and power combining, for the same output power level. Two single-ended 45-GHz Class-E-like prototypes with two- and four-stacked devices have been fabricated in IBM’s 45-nm SOI CMOS technology based on this design methodology. Another 45-GHz differential Class-E-like PA with two devices stacked has been implemented in IBM’s 65-nm low-power CMOS technology. These PAs provide experimental validation for the theoretical results.

## II. STACKED CMOS CLASS-E-LIKE PAs

### A. Concepts

Fig. 2(a) depicts the concept of a stacked CMOS Class-E-like PA. The stacked configuration consists of multiple series devices, which might be of equal or different size. In order to preserve input power and improve PAE, only the bottom device is driven by the input signal. The devices higher up in the stack turn on and off due to the swing of the intermediary nodes. The topmost drain is loaded with an output network that is designed based on Class-E principles, and consequently sustains a Class-E-like voltage waveform. The intermediary drain nodes must also sustain Class-E-like voltage swings with appropriately scaled amplitudes so that the voltage stress is shared equally among all devices. In the 45-nm SOI and

65-nm CMOS technologies employed, the nominal  $V_{DD}$  of the high-speed thin-oxide devices is  $\approx 1$  V and for long-term reliability, the maximum swing across any two transistor junctions under large-signal operation is limited to  $2 \times V_{DD} = 2$  V [13]. Consequently, for a PA with  $n$  stacked devices, the peak output swing is  $2n$  V, as marked on Fig. 2(a), and the appropriate intermediary node swings are also noted. Appropriate voltage swing may be induced at the intermediary nodes through techniques such as inductive tuning [13], capacitive charging acceleration [24], and placement of Class-E load networks at intermediary nodes [19]. In this paper, we employ inductive tuning at select intermediary nodes, which, for simplicity, is not shown in Fig. 2(a). The tradeoffs associated with inductive tuning and capacitive charging acceleration for Class-E-like PAs at mmWave frequencies will be discussed later in this paper. In order to conform to the peak ac swing limit across the gate–source junction in the on half-cycle and the gate–drain junction in the off half-cycle, the gates of the devices in the stack must swing, as shown in Fig. 2(a). The swing at each gate is induced through capacitive coupling from the corresponding source and drain node via  $C_{gs}$  and  $C_{gd}$ , respectively, and is controlled through the gate capacitor  $C_n$ . The dc biases of all gates are applied through large resistors.

From Fig. 2(a), it can be seen that for a two-stacked switching PA, the gate of the top device is connected to signal ground via a large capacitor and experiences no signal swing. However, this does not reduce a two-stacked switching PA to a regular cascode configuration because of the following reasons. The main objective of stacking is to allow operation off a higher supply voltage by distributing the overall voltage stress equally amongst the transistors. This is accomplished by engineering the drain and gate nodes voltage profiles [as depicted in Fig. 2(a)] to ensure that all devices have the same  $V_{gs}$ ,  $V_{gd}$ , and  $V_{ds}$  swings, which results in a linear increase in the supply voltage with the number of devices stacked. In stacked switching PAs, the nature of the voltage swings requires a constant gate bias *only* for the second device. Conventional cascode PAs can operate off a higher supply voltage, as well, but a linear scaling in supply voltage cannot be achieved. This is because the gate of the top device is usually connected to the supply voltage to maximize small-signal power gain. The ensuing unequal voltage stress across the devices compromises long-term reliability and can even enforce operation off a single-device supply voltage [25]–[27]. The simulated drain–source and gate–source waveforms for the two-stacked and four-stacked Class-E-like PAs (Fig. 9(c) and (d) respectively, presented in Section IV) demonstrate that voltage swings are indeed equal for the prototypes implemented in this work and serve to distinguish a two-stacked switch-like PA from a conventional cascode PA. This claim is further validated by comparing the measured performance of the two-stacked PA implemented in this work with a prior mmWave cascode PA [27] in Section V-B.

### B. Theoretical Analysis and Fundamental Limits

To facilitate a theoretical analysis, the improved loss-aware Class-E design methodology described in [23] is employed. The basic Class-E design methodology of ensuring zero-voltage switching (ZVS) and zero-derivative-of-voltage

switching (ZdVS) [22] is no longer optimal for achieving high-efficiency operation in the presence of high loss levels in the switching device and/or passive components. The improved loss-aware Class-E design methodology formally takes switch loss and passive loss into account. The methodology also incorporates the input power required to drive the switch and enables optimization of the PAE rather than drain efficiency. In essence, the methodology is an analytical load-pull for optimizing PAE in the presence of high loss levels and input power requirements.

The devices (taken to be equal in size) in a stacked switching PA are assumed to behave as a single switch with linearly increased breakdown voltage and ON-resistance. As far as theoretical results are concerned, only the total ON-resistance of the stacked configuration and output capacitance of the top device are pertinent. The output capacitance of a stacked configuration should ideally scale down linearly with number of devices stacked. However, wiring parasitics are significant at mmWave frequencies and there will be parasitic capacitance to ground from the intermediate drain/source and gate nodes, which will prevent linear scaling of output capacitance with stacking. As a worst case estimate, the overall output capacitance is taken to be the same as that of a single device ( $= C_{gd} + C_{ds}$ , where  $C_{ds} = (C_{db} \times C_{sb}) / (C_{db} + C_{sb})$  since the high body resistance in SOI technology causes  $C_{db}$  and  $C_{sb}$  to appear in series). It is indeed this mechanism that prevents efficiency from remaining constant as we stack more devices, as will be shown later in this paper [graphically in Fig. 3(a) and theoretically in (9)]. The devices can be of different sizes and there could potentially be some benefit in tapering device sizes as well [17], [28] since the gate capacitors conduct a portion of the device current. Thus, progressive device size reduction up the stack would reduce parasitic capacitances and prevent capacitive discharge loss at intermediate nodes. However, device size tapering has not been pursued in this work.

The time-domain equations and corresponding design procedure for a stacked Class-E PA are described in the Appendix. For various levels of stacking ( $n$ ), the design methodology is used to analytically vary device-size and dc-feed inductance to find the design point(s) with optimal PAE under the constraint of a  $50\text{-}\Omega$  load impedance to avoid impedance transformation losses. As an example, for a stack of four devices ( $n = 4$ ), we start with an initial device size of  $100\text{ }\mu\text{m}$  and set the tuning parameter  $\omega_s = 0.8 \times \omega_0$ . The design methodology then determines the optimal load impedance for highest PAE and the corresponding output power. The load impedance is then scaled (along with device size, input, and output powers) to have a real part of  $50\text{ }\Omega$ . The procedure is repeated by changing the tuning parameter  $\omega_s$ . Finally, amongst all these design points for a stack of four devices driving a  $50\text{-}\Omega$  load, the one with the highest PAE is chosen. This yields a device size of  $204\text{ }\mu\text{m}$  for the four-stacked PA, with theoretical output power and PAE of  $145\text{ mW}$  and  $48\%$ , respectively (as shown in Fig. 3). The procedure can similarly be used to determine the corresponding metrics for other levels of stacking. Device ON-resistance, output capacitance, and input-drive-power as functions of device size are determined from post-layout device simulations and are validated through device measurements (discussed in Section III).

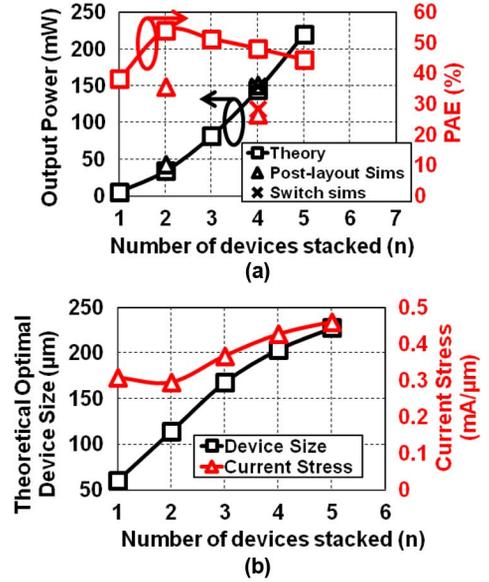


Fig. 3. (a) Theoretical and simulated (post-layout) output power and PAE and (b) device size and theoretical device stress for the optimal design as a function of number of devices stacked based on the loss-aware Class-E design methodology at 45 GHz in 45-nm SOI CMOS. Loss in dc-feed inductance is included for theoretical results. Output power and PAE for a switch+capacitor-based model for the four-stacked configuration are also annotated.

The values for these parameters are shown in Table I, where  $f_0$  is the operating frequency,  $\bar{R}_{ON} = R_{ON} \times W$ ,  $\bar{C}_{out} = C_{out} / W$ ,  $\bar{C}_{in} = C_{in} / W$ , and  $\bar{P}_{in} = P_{in} / W$  are technology parameters normalized to the device width ( $R_{ON}$ ,  $C_{out}$ ,  $C_{in}$ , and  $P_{in}$  being, respectively, the ON-resistance, output capacitance, input capacitance, and input power corresponding to a device of width  $W$ ). At mmWave frequencies, the constants of proportionality in the input power functions take into account the power lost in the gate resistance, and are consequently frequency dependent. The values of those constants reported in Table I are based on 45-GHz simulations. In order to incorporate the loss of the dc-feed inductance, a quality factor of 15 is assumed at 45 GHz based on measurements [20].

Fig. 3(a) depicts the optimal output power and PAE for different levels of stacking in 45-nm SOI CMOS at 45 GHz. The optimal size of each stacked device and the associated device stress (defined as the ratio of the average current drawn from the power supply to the device width) are shown in Fig. 3(b). It is clear that due to the increasing achievable output voltage swing, stacking in Class-E-like CMOS PAs enables dramatic increases in output power (near-quadratic due to linear increase in output swing). The PAE reduces with increased stacking due to increasing total switch loss. However, the methodology ensures that the PAE degradation is gradual. In order to do this, the design methodology requires the size of each stacked device to increase with  $n$  to reduce the individual (and hence, overall) ON-resistance. Consequently, careful device layout is required for high levels of stacking as it is challenging to layout large devices while maintaining a high  $f_{max}$ . Another important consideration for device stacking is the current stress for the stacked devices, which increases with the level of stacking. Current stress (or large-signal current density) is the ratio of the

TABLE I  
NORMALIZED DEVICE PARAMETERS USED IN LOSS-AWARE CLASS-E ANALYSIS

Tech.	$\overline{R_{ON}} = \overline{R_{ON} \times W}^{\S}$	$\overline{C_{out}} = \overline{C_{gd} + C_{ds}} = \frac{C_{gd}}{W} + \frac{C_{ds}}{W}^{\#}$	$\overline{R_{ON} \times C_{out}}$ (Switch Time Constant)	$\overline{C_{in}} = \overline{C_{gs} + C_{gd}} = \frac{C_{gs}}{W} + \frac{C_{gd}}{W}^{\dagger}$	$\overline{P_{in}} = \frac{P_{in}}{W}$
45 nm SOI CMOS	$275 \Omega - \mu m$ ( $V_{gs}=1V$ )	$0.769 \text{ fF}/\mu m$	$211.5 \Omega\text{-fF}$	$0.31 \text{ fF}/\mu m$	$4.191 \times f_0 \times \overline{C_{in}} \times (V_{high} - V_{low})^2$
65 nm low-power bulk CMOS	$820 \Omega - \mu m$ ( $V_{gs}=1V$ )	$0.48 \text{ fF}/\mu m$	$393.6 \Omega\text{-fF}$	$0.28 \text{ fF}/\mu m$	$4.341 \times f_0 \times \overline{C_{in}} \times (V_{high} - V_{low})^2$

Note:  $R_{on}$ ,  $C_{gs}$ ,  $C_{gd}$ ,  $C_{ds} = (C_{db} \times C_{sb})/(C_{db} + C_{sb})$ , and  $P_{in}$  correspond to ON-resistance, gate-source, gate-drain, and drain-source capacitances, and input power (to switch a device between the triode and cutoff regions), respectively, for a device of width  $W$  (including layout parasitics).  $f_0$  is the operating frequency.  $V_{high}$  and  $V_{low}$  refer, respectively, to the high and low amplitude levels of a 45-GHz square-wave input signal.

<sup>§</sup> Estimated in triode region of operation.

<sup>#</sup> Estimated in cutoff region of operation.

<sup>†</sup> Estimated as average of capacitance values in cutoff and triode regions of operation.

average current drawn from the supply under large-signal operation ( $I_{DC}$ ) to the device width. Note that  $I_{DC}$  is different from the supply current  $I_{bias}$  drawn with no input power (i.e., under small-signal operation), the latter being used to determine the current density for operating at highest  $f_{max}$  in linear PAs. The current drawn under large-signal operation is typically 1.5–2 times higher than the small-signal bias current in our implementations. This implies that in a practical implementation, the metallization of the source and drain fingers of the MOS devices must be augmented with additional metal layers, if required, so that they can support the required currents while satisfying electromigration rules for the technology. While Fig. 3(a) shows an increasing trend for output power until five devices, at much higher levels of stacking, the assumption in the theoretical analysis that the switch ON resistance is much smaller than the impedance of its output capacitance [23] would be violated. Furthermore, there would be diminishing returns in output power owing to increased losses with stacking. In practice, the maximum practical device size, the maximum current stress that can be tolerated as per electromigration requirements, and drain-bulk/buried-oxide breakdown mechanisms would determine the maximum number of devices that can be stacked. The post-layout simulated results for output power and PAE for a two-stacked and a four-stacked Class-E-like PA have been annotated on Fig. 3(a), as well and show excellent agreement with the theoretical output power. The post-layout simulated efficiency is lower by  $\approx 20\%$  owing to various implementation losses and soft-switching at mmWave, as well as power loss at intermediate nodes, which are not accounted for in theory. However, the theoretical and simulated trends in PAE are in agreement. Later in this paper, a switch+capacitor-based model for the device is constructed for simulation-based investigation of power loss at intermediate nodes for a four-stacked configuration. The resulting output power and PAE [see Fig. 3(a)] show excellent agreement with post-layout device-based simulations and reaffirm the utility of a simplified theoretical analysis.

### C. Interpretation Using Waveform Figures of Merit

An analysis using the unique properties of switching PAs facilitates a better understanding of the underlying phenomena associated with device stacking and an interpretation of the results

of the loss-aware Class-E design methodology. An excellent description of the characteristics of switching PAs can be found in [29]. We have

$$PAE = 1 - \frac{P_{loss}}{P_{DC}} - \frac{P_{in}}{P_{DC}} \quad (1)$$

where  $P_{in}$  and  $P_{DC}$  are input power to the PA and the dc power consumption, respectively. The loss in the PA ( $P_{loss}$ ) is given by

$$P_{loss} = P_{loss,switch} + P_{loss,cap} \quad (2)$$

$$= I_{RMS,n}^2 \times n \times R_{ON,n} + P_{loss,cap} \quad (3)$$

$$= I_{RMS,n}^2 \times n \times \frac{R_{ON,n}}{W_n} + P_{loss,cap} \quad (4)$$

where  $n$  is the number of devices stacked in series,  $R_{ON,n}$  is the ON-resistance of each device in the  $n$ -stacked PA,  $W_n$  is the width of each device/switch, and  $I_{RMS,n}$  is the root mean square (rms) value of the current flowing through the stack of  $n$  switching devices (excluding the output capacitance).  $P_{loss,cap}$  is the switching loss associated with the output capacitance of the PA and is dependent on the topmost drain voltage value at the switching instant. In general, at mmWave frequencies, the capacitive discharge loss is negligible compared to the conduction loss in the switching device(s). This is evident from Table II, where the conduction loss in the switch and the capacitive discharge loss have been tabulated for the optimal designs at different levels of stacking described in Fig. 3. Indeed, this reinforces our earlier assertion that the conventional ZVS/ZdVS-based Class-E design methodology is not applicable at mmWave frequencies. Therefore, for the purpose of simplifying our analysis, we shall ignore the contribution of the term  $P_{loss,cap}$  to the overall loss. In a switching PA, the average current drawn from the supply is always proportional to  $I_{RMS,n}$ . The proportionality constant depends on the tuning of the load network [29]. Tuning of a Class-E load network is determined by the dc-feed inductance ( $L_s$ ) and the load impedance ( $Z_{load}$ ) in relation to the device output capacitance. Since we are in a regime where conduction loss is significant, the proportionality constant will also depend on the value of the total switch ON-resistance ( $n \times R_{ON,n}$ ) relative to the output capacitance  $C_{out,n}$ . Since  $R_{ON,n} \times C_{out,n}$  is a technology constant, specifying  $n$ ,  $C_{out,n}$ ,  $L_s$ , and  $Z_{load}$  completely characterizes the tuning of the stacked Class-E PA.

TABLE II  
CONDUCTION LOSS AND CAPACITIVE DISCHARGE LOSS FOR THE OPTIMAL DESIGNS AT DIFFERENT LEVELS OF STACKING DESCRIBED IN FIG. 3.  
VALUES FOR THE WAVEFORM FIGURES OF MERIT  $F_I^2$  AND  $F_C$  FOR LOSS-AWARE AND ZVS-BASED DESIGNS ARE ALSO TABULATED

Devices Stacked (n)	Device size ( $\mu\text{m}$ )	ON Resist. ( $\Omega$ )	Output Cap. ( $C_{out,n}$ ) (fF)	Device Conduction Loss (mW) ( $P_{loss,switch}$ )	Output Cap. Switching Loss (mW) ( $P_{loss,cap}$ )	$F_I^2$ (loss-aware)	$F_C$ (loss-aware)	$F_I^2$ (ZVS)	$F_C$ (ZVS)
1	60	4.58	35.46	1.5	0	2.06	2.06	2.36	3.61
2	114	2.41	67.37	10.6	1.3	2.56	0.94	2.17	3.02
3	168	1.64	99.29	32.8	5.7	2.73	0.74	2.1	2.61
4	204	1.35	120.56	80.2	17.8	2.62	0.68	2.07	2.29
5	228	1.21	134.75	176.3	45.4	2.54	0.63	2.05	2.04

It is also therefore clear that the optimal tuning is likely to vary for different levels of stacking due to the increasing total switch loss. Ignoring capacitive discharge loss, (4) becomes

$$\begin{aligned}
 P_{loss} &\approx I_{RMS,n}^2 \times n \times \frac{\overline{R_{ON,n}}}{W_n} \\
 &= \frac{I_{RMS,n}^2}{I_{DC,n}^2} \times I_{DC,n}^2 \times n \times \frac{\overline{R_{ON,n}}}{W_n} \\
 &= F_{I,n}^2 \times I_{DC,n}^2 \times n \times \frac{\overline{R_{ON,n}}}{W_n} \quad (5)
 \end{aligned}$$

where  $F_{I,n} = I_{RMS}/I_{DC}$  is a waveform figure of merit (FOM) defined in [29] and  $I_{DC,n}$  is the average supply current with  $n$  devices stacked.

For a stack of  $n$  devices, the supply voltage scales linearly with  $n$ . On the other hand, in a switching PA, average supply current is proportional to the product of output capacitance, supply voltage, and the operating frequency ( $\omega_0$ ), the constant of proportionality being dependent on the tuning of the circuit. The linear dependence on output capacitance is simply an artifact of circuit scaling properties, while the linear scaling with supply voltage arises from the fact that switching PAs are linear with respect to excitations at the drain node (e.g., supply voltage) [29]. Denoting the impedance of the device output capacitance at the fundamental frequency by

$$Z_C = \frac{1}{\omega_0 \times (\overline{C_{out}} \times W_n)}$$

the waveform FOM  $F_C$  is defined in [29] as

$$F_C = \frac{P_{DC}}{\frac{V_{DC}^2}{Z_C}} \quad (6)$$

For an  $n$ -stacked PA,  $V_{DC} = (n \times V_{DD})$ , where  $V_{DD}$  is the supply voltage for a single device PA. Substituting

$$P_{DC} = V_{DC} \times I_{DC,n} = (n \times V_{DD}) \times I_{DC,n}$$

in (6), we get

$$I_{DC,n} = F_{C,n} \times \omega_0 \times (\overline{C_{out}} \times W_n) \times (n \times V_{DD}). \quad (7)$$

Consequently,

$$\begin{aligned}
 \text{PAE} &= 1 - \frac{F_{I,n}^2 \times I_{DC,n}^2 \times n \times \frac{\overline{R_{ON,n}}}{W_n}}{(n \times V_{DD}) \times I_{DC,n}} \\
 &\quad - \frac{W_n \times \overline{P_{in}}}{(n \times V_{DD}) \times I_{DC,n}} \quad (8)
 \end{aligned}$$

$$\begin{aligned}
 &= 1 - n \times F_{I,n}^2 \times F_{C,n} \times \omega_0 \times \overline{C_{out}} \times \overline{R_{ON}} \\
 &\quad - \frac{k \times \overline{C_{in}} \times (V_{high} - V_{low})^2}{n^2 \times V_{DD}^2 \times F_{C,n} \times \overline{C_{out}}} \quad (9)
 \end{aligned}$$

where  $k$  is a technology- and frequency-dependent constant of proportionality that results from the input power functions shown in Table I.

Table II lists the values of the waveform figures of merit for designs based on the loss-aware and ZVS methodologies for different levels of stacking. While the waveform metric  $F_I$  is comparable for both, the loss-aware methodology shapes the waveforms to minimize  $F_C$ , thereby yielding optimal designs with highest possible PAE.

The foregoing expression captures the variation in PAE in terms of technology constants and number of devices stacked. The only design-related variables in this expression are the waveform figures of merit. The third term captures the PAE benefit of stacking since output power is increased quadratically, but input power is only provided to the bottom device in the stack. This explains why PAE improves when one goes from a single-device Class-E-like PA to a two-stacked Class-E-like PA in Fig. 3(a). However, as stacking is increased beyond two devices, the benefits from the third term wear off and the second term causes a reduction in PAE due to a reduction in drain efficiency. It is well known that  $\overline{R_{ON}} \times \overline{C_{out}}$  is the technology constant (which we shall refer to as the switch time constant) that determines the drain efficiency of a Class-E PA [29] for a given operating frequency and this constant degrades linearly for an  $n$ -stacked device since the ON-resistances add in series while the output capacitance remains that of a single device. However, the loss-aware Class-E design methodology optimizes the output network tuning to ensure that the PAE degradation is gradual as stacking is increased by minimizing the  $F_{I,n}^2 \times F_{C,n}$  product. The benefits of the loss-aware Class-E tuning methodology over the ZVS/ZdVS-based tuning methodology can be appreciated in Fig. 4, where the  $F_{I,n}^2 \times F_{C,n}$  product for the loss-aware design technique can be observed to be lower than that corresponding to the ZVS design methodology by a factor of 2–3 depending on the number of devices stacked.

The preceding analysis also highlights the importance of the switch time constant as a technology metric that determines the efficiency of switching PAs. For linear-type PAs,  $f_{max}$  is a sufficient metric to gauge the PA efficiency. For switching-type PAs,  $f_{max}$  determines the input power requirements (via the technology- and frequency-dependent constant  $k$ ) while the switch time constant determines the drain efficiency.

As the levels of stacking are increased, the switch time constant becomes more significant than  $f_{max}$ . As can be seen in

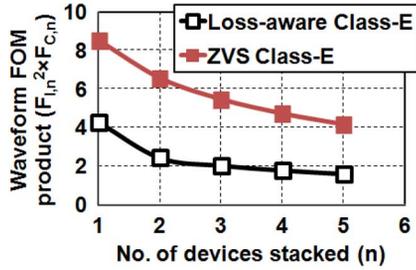


Fig. 4. Product of waveform figures of merit  $F_{1,n}^2$  and  $F_{C,n}$  for stacked Class-E-like PAs in 45-nm SOI CMOS at 45 GHz based on the loss-aware and ZVS based design methodologies.

Table I, and later in this paper, 65-nm low-power bulk CMOS and 45-nm SOI CMOS have similar  $f_{max}$ , but significantly different switch time constants. Consequently, it can be expected that switching-type PAs in 45-nm SOI CMOS will achieve higher efficiencies than those in 65-nm low-power bulk CMOS. This is validated by our experimental results in Section V.

The loss-aware Class-E analysis takes into account several mmWave nonidealities, and therefore, the results in Fig. 3 represent the fundamental limits on achievable performance in stacked CMOS Class-E-like PAs. The main nonideality that causes deviation from these limits in practice is soft switching of the stacked devices due to the lack of square-wave drives. Nevertheless, the optimal design points predicted by the analysis are excellent starting points for simulation-based optimization.

#### D. Stacking Versus Power Combining

In order to appreciate the benefits of device stacking using the loss-aware Class-E design methodology, it is imperative to contrast this approach to conventional impedance transformation and power-combining techniques. To evaluate the performance of power combining, a 45-nm SOI two-stacked Class-E-like PA (resulting from the loss-aware design methodology) with a theoretical output power of 34 mW and a corresponding theoretical PAE of 54% at 45 GHz is chosen since it has reasonable output power as well as the highest efficiency [see Fig. 3(a)]. Since the two-stacked PA is designed for an optimal load impedance of 50  $\Omega$ , a cascaded-tree of two-way 50- $\Omega$  Wilkinson power combiners is chosen. A 2-way 50  $\Omega$  Wilkinson power combiner with 70.7- $\Omega$   $\lambda/4$  transmission lines in 45-nm SOI CMOS technology has an electromagnetic (EM)-simulated efficiency  $\eta = 0.87$  at 45 GHz. An  $N$ -way cascaded-Wilkinson-tree power combiner (where  $N$  is an even multiple of two) will therefore have an overall efficiency of  $\eta^{\log_2 N}$ . Fig. 5(a) compares the theoretical PAE, as a function of output power, for different levels of device stacking with that of two, four-, and eight-way Wilkinson-tree power combining. For a given output power, stacked Class-E-like PAs implemented using the loss-aware Class E design methodology offer  $\approx 10\%$ – $20\%$  higher efficiency compared to Wilkinson power combining (using two-stacked PAs). Power combining using transformers is a better alternative at mmWave frequencies since ideally transformer-based series power combining has a constant efficiency with the number of elements combined.

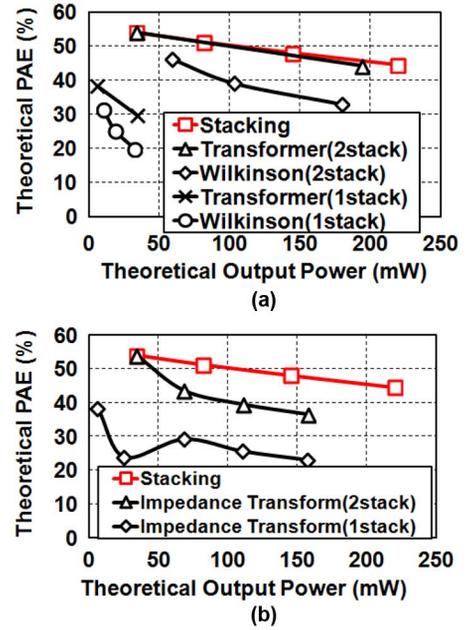


Fig. 5. Comparison of device stacking in Class-E-like PAs (based on loss-aware Class-E design methodology) at 45 GHz in 45-nm SOI CMOS with: (a) two-, four-, and eight-way Wilkinson-tree-based power combining and transformer-based series power combining (the two-stacked and one-stacked Class-E-like PAs obtained from the loss-aware Class-E design methodology are used with both the  $N$ -way Wilkinson-tree-based and transformer power combiners) and (b) impedance transformation at 45 GHz in 45-nm SOI CMOS (the two-stacked and one-stacked Class-E-like PAs obtained from the loss-aware Class-E design methodology are scaled to increase output power and a two-element  $L$ - $C$  network is used to transform the 50- $\Omega$  load to the optimal load impedance for the scaled PAs. Quality factors for the inductor and capacitor are assumed to be 15 and 10, respectively, at 45 GHz).

However, interwinding and self-resonant capacitances introduce asymmetry in transformer power combiners, degrade efficiency, and cause stability problems, usually permitting a maximum of two transformer sections to be combined in series [4]. Ignoring the effect of parasitic capacitances, a two-section series transformer combiner is used to power-combine two two-stacked PAs. The secondary inductance is chosen for maximum efficiency subject to a 50- $\Omega$  load, and the PAs are appropriately scaled to drive the load impedance presented by the primary of the transformer. As shown in Fig. 5(a), transformer power combining utilizing two-stacked PAs can yield results similar to stacking only under ideal conditions and is fundamentally limited to two-way combining. The corresponding results for Wilkinson and transformer-based power combining using one-stacked (single device) Class-E-like PAs obtained from the loss-aware design methodology are also included to emphasize the inefficacy of the traditional design technique of using single-device PAs for high-power amplification.

#### E. Stacking Versus Impedance Transformation

The efficiency of the alternative technique of impedance transformation is dependent on the steepness of transformation as well as the topology of the impedance transformation network. The two- and one-stacked Class-E-like PAs in 45-nm SOI obtained from the loss-aware Class-E design methodology at 45 GHz are again employed for the purpose of comparison. In order to achieve output power comparable to those obtained

from device stacking, the Class-E-like PAs are scaled appropriately while an impedance transformation network is used to transform the  $50\text{-}\Omega$  load to the corresponding lower load impedance for the scaled PAs. A two-element  $L\text{-}C$  impedance transformation network is designed and used in each case. The quality factors of the inductor and capacitor are assumed to be 15 and 10, respectively, at 45 GHz, based on measured characterizations of inductors, capacitors, and transmission lines in the 45-nm SOI CMOS technology [20]. A comparison of the PAEs of impedance transformation and device stacking is summarized in Fig. 5(b). Device stacking results in designs with  $\approx 10\text{--}30\%$  higher efficiency for the same output power compared to impedance transformation.

Once device stacking is exploited to the limit as dictated by secondary breakdown mechanisms (e.g., that of the BOX in the SOI), it is interesting to consider the combination of device stacking with impedance transformation and/or power combining to achieve watt-class output power levels at mmWave frequencies.

### III. IBM 45-nm SOI AND 65-nm CMOS MODELING

In the 45-nm SOI CMOS, an accurate high-frequency model for the device, which accounts for intrinsic input resistance (IIR), as well as layout-related wiring resistances, capacitances, and inductances of the gate, drain, and source fingers and vias is nonexistent. The model provided in the design kit is augmented to incorporate the impact of IIR, which models the distributed characteristics of the channel in a MOSFET [30]. While IIR is controlled by two parameters (XRCRG1 and XRCRG2) in BSIMSOI modeling [31] (which default to 0 in the PDK), we have found that transient simulations in SPECTRE fail to converge when these parameters are assigned values based on our device measurements. Consequently, a bias-independent resistance is added in series with the gate to account for IIR. The bias independence of this resistor, along with its location (outside the PDK device model), is a source of inaccuracy in our transient simulations. Wiring resistances and capacitances are extracted using Calibre PEX. High-frequency models for the gate and drain vias are simulated in the IE3D field solver [32].

The layout of a fabricated floating-body power device test structure in 45-nm SOI technology employs a continuous array of gate fingers (40–70) with a finger width of  $2.793\ \mu\text{m}$ . We make use of a doubly contacted gate with a symmetric gate via on both sides to reduce gate resistance [30]. Wiring resistances and capacitances are extracted for the entire stacked-device layout configuration and high-frequency models for the vias are added to this overall  $R\text{-}C$  extracted model. The layout for the four-stacked configuration is shown in Fig. 6, while the corresponding high-frequency model with parasitics is illustrated in Fig. 7. The source and drain fingers of the devices consist of metal layers  $M_1 - M_3$  strapped so as to conform to electromigration requirements. The connection from the source of the bottom device to the ground node supports a large current under large-signal operation. Consequently, this connection is augmented with thick metal strips in metal layers  $M_2$  and  $M_3$  strapped together, which also helps minimize the source inductance.

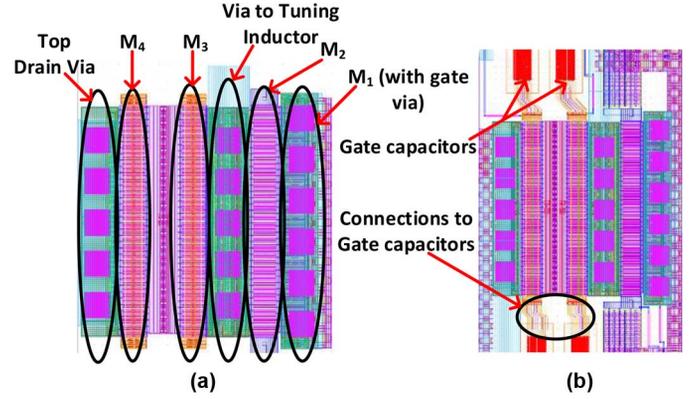


Fig. 6. Close-up of: (a) devices and (b) devices with connections to gate capacitors in four-stacked PA layout implemented in 45-nm SOI CMOS.

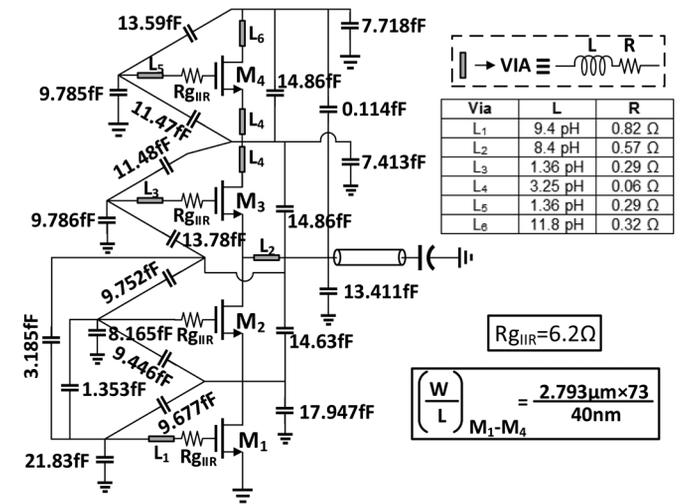


Fig. 7. Augmented schematic of four-stacked power device in 45-nm SOI CMOS with capacitive and inductive layout parasitics.

The measured  $f_{\text{max}}$  and  $f_T$  of the test structures of the individual power devices used in the implemented 45-nm SOI two-stacked and four-stacked PAs (distinct from the custom stacked layout as discussed earlier) are shown in Fig. 8(a) and (b), respectively. Device measurements were conducted up to 65 GHz using a pair of coaxial 1.85-mm (dc–65 GHz) ground–signal–ground (GSG) probes, calibrated at the probe tip planes. The industry-standard open-short de-embedding was performed to a reference plane at the top of the gate and drain vias. The measured  $f_{\text{max}}$  and  $f_T$  were obtained by extrapolating the measured Mason’s unilateral power gain ( $U$ ) and  $h_{21}$  at 20 dB/decade. The measured  $U$  is observed to have 20-dB/decade slope up to 65 GHz and the modeled  $U$  exhibits the same slope up to  $f_{\text{max}}$ . Peak  $f_{\text{max}}$  of  $\approx 180$  GHz and  $\approx 190$  GHz are achieved for these power devices.

It is difficult to achieve  $f_{\text{max}}$  for power devices that is similar to that of smaller devices due to layout challenges [33]. For reference, our measurements reveal that a  $(1\ \mu\text{m} \times 10)/40\ \text{nm}$  device achieves an  $f_{\text{max}}$  of  $\approx 250$  GHz in this technology [30]. The use of a compact device layout with a continuous array of large number of gate fingers with large finger width reduces the parasitic capacitance and causes the 204- $\mu\text{m}$  device to have a higher  $f_T$  compared to the 115- $\mu\text{m}$  device. However, the layout

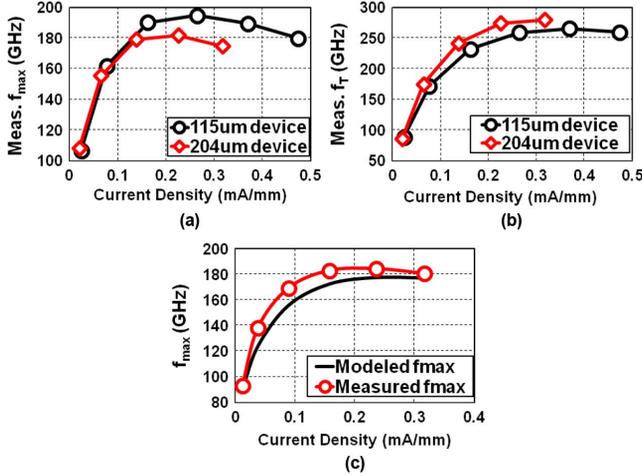


Fig. 8. Measured (extrapolated): (a)  $f_{max}$  and (b)  $f_T$  of  $(2.793 \mu\text{m} \times 41)/40 \text{ nm}$  and  $(2.793 \mu\text{m} \times 73)/40 \text{ nm}$  power devices in IBM 45-nm SOI CMOS across current density. These devices are used in designing the two-stacked and four-stacked PAs, respectively. (c) Measured and simulated  $f_{max}$  for a  $(3 \mu\text{m} \times 50)/60 \text{ nm}$  65-nm low-power bulk-CMOS power device across current density.

also results in an increased gate resistance and lower  $f_{max}$  for the larger device. This prevents the devices of the four-stacked PA from being driven into a hard-switching condition, as will be discussed later in this paper. Splitting the overall device into several smaller devices (each with reduced finger width and small number of gate fingers) wired appropriately in parallel should further improve the  $f_{max}$  to approach 250 GHz and available gain [33]. It should be noted that such a multiplicity-based layout approach might compromise  $f_T$  due to increased wiring capacitance. In a switch-like PA, a good balance between  $f_T$  and  $f_{max}$  must be maintained.

A similar device layout approach and modeling strategy is employed for power devices in IBM's 65-nm low-power bulk-CMOS technology. A key point of difference, however, is the presence of IIR modeling within the PDK, eliminating the need for an external IIR resistance. A  $(3 \mu\text{m} \times 50)/60 \text{ nm}$  power device test structure is measured using the same approach as mentioned earlier [see Fig. 8(c)]. A peak  $f_{max}$  of approximately 180 GHz is observed in measurement. It should, however, be noted that while power devices in 65-nm low-power bulk CMOS are able to achieve similar  $f_{max}$  to power devices in 45-nm SOI CMOS, the width-normalized ON-resistance (quantified as  $\overline{R}_{ON}$  in Table I) is almost three times higher for the same gate drive level due to the high threshold voltage of the low-power process ( $V_{th} = 560 \text{ mV}$  at the PA bias point). As will be demonstrated experimentally, this leads to inferior performance for mmWave Class-E-like PAs in 65-nm low-power bulk CMOS.

#### IV. IMPLEMENTATION DETAILS

The schematics in Fig. 9(a) and (b) depict the Class-E-like PAs implemented by stacking two and four floating-body devices in 45-nm SOI CMOS technology. Device sizes and dc-feed inductance values are chosen based on the theoretical analysis, while supply and gate bias voltages and gate capacitor values are selected based on the considerations described

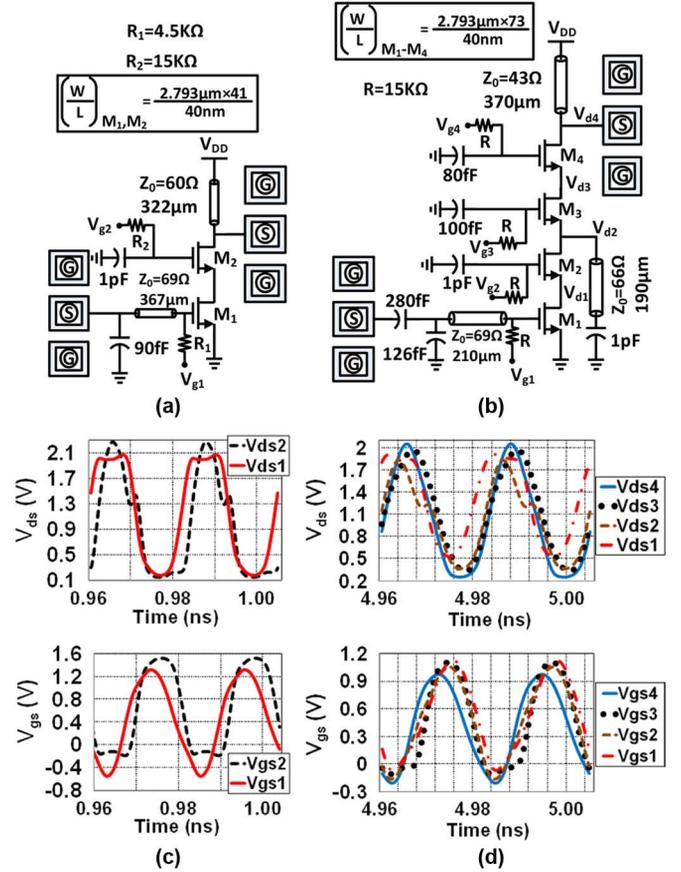


Fig. 9. Schematics of 45-nm SOI CMOS  $Q$ -band Class-E-like PAs with: (a) two devices stacked and (b) four devices stacked. Simulated drain-source and gate-source voltage waveforms of the  $Q$ -band (c) two-stacked Class-E-like PA ( $V_{g1} = 0.4 \text{ V}$ ,  $V_{g2} = 1.7 \text{ V}$ ,  $V_{DD} = 2.4 \text{ V}$ ), and (d) four-stacked Class-E-like PA in 45-nm SOI CMOS ( $V_{g1} = 0.4 \text{ V}$ ,  $V_{g2} = 1.8 \text{ V}$ ,  $V_{g3} = 2.8 \text{ V}$ ,  $V_{g4} = 4 \text{ V}$ , and  $V_{DD} = 4.8 \text{ V}$ ).

earlier. For the first stacked device ( $M_2$  in both designs), the gate voltage must be held to a constant bias as discussed previously. This can be accomplished through a large bypass capacitor placed as close as possible to the gate to mitigate stray inductance that can result in oscillations. DGNCAPs (which are device capacitors) are suitable for this purpose since their wiring is in the lowest metal layer and they provide higher capacitance density than VNCAPs (interdigitated finger capacitors). All other capacitors, including gate capacitors for the higher stacked devices, which are not large in value, are implemented using VNCAPs. For both the designs, the output harmonic filter is eliminated to avoid passive loss with minimal impact on performance.

As was mentioned earlier, a tuning inductor may be placed at intermediary nodes to improve their voltage swing and make them more Class-E-like. Simulation results indicate that the improvement in swing for the two-stacked PA is offset by an increase in the conduction loss of the top device. This can be explained as follows. The voltage swing at the intermediate node controls the turn-on and turn-off of the top device. As shown in Fig. 10(a), in the absence of the tuning inductor, the intermediate node voltage gets clipped to  $V_{g2} - V_{th2}$  once the top device turns off during the OFF half-cycle [13]. The voltage remains



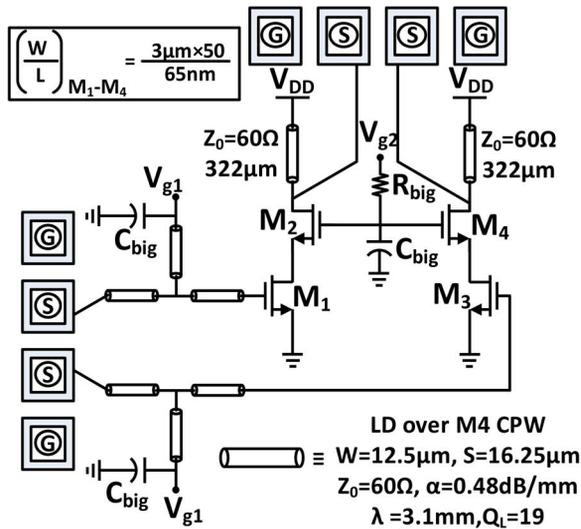


Fig. 12. Schematic of differential two-stacked Class-E-like PA implemented in 65-nm low-power bulk CMOS.

flowing through the “switch.” As a first-order approximation, the currents through the external wiring parasitic capacitances  $C_{gs}$ ,  $C_{gd}$ ,  $C_{ds}$ , and  $C_{d0}$  are scaled in proportion to the ratio of the intrinsic to external wiring parasitic capacitance, and their sum is subtracted from the total device current to arrive at the switch-current in simulation. Fig. 13 shows the  $V_{DS}$  and the corresponding  $I_{switch}$  for the various devices in the two- and four-stacked PAs implemented in 45-nm SOI CMOS from which the nonoverlapping characteristic of voltage and current waveforms is clearly evident. Figs. 14 and 15 compare the switch voltage and switch current waveforms for devices  $M_2$  and  $M_4$  of the two- and four-stacked PAs, respectively, with theory. Aside from the sharp current spikes in the theoretical waveforms at switch turn-on, there is excellent correspondence between theory and simulation. The current spikes arise from the assumption of hard switching, which is not possible at mmWave. However, the soft switching in simulation does not compromise the shaping of voltages and currents and their harmonic content for the rest of the switching cycle. These results clearly indicate the feasibility of switching operation at mmWave frequencies.

As mentioned before, the theoretical loss-aware Class-E design methodology approximates the stacked configuration as a series connection of switches and assumes that appropriate voltage swings are somehow ensured at the intermediate nodes. A more realistic model for the circuit is a stack of switches, each accompanied by the corresponding intrinsic device capacitances ( $C_{gs}$ ,  $C_{gd}$ ,  $C_{ds}$ , and  $C_{d,ground}$ ) and gate capacitors (except for the bottom switch). An “Elmore network” of RC delays is encountered in stacked linear PAs (owing to the simultaneous presence of capacitances and finite device output resistances), which can cause phase shift in the voltages and currents as one moves up the stack. It is unclear that there would be a similar “Elmore delay” in stacked switching PAs since during the OFF cycle the switch devices have very high OFF resistance. However, the capacitive discharge loss at intermediate nodes might be nonnegligible and can have a considerable influence on overall efficient

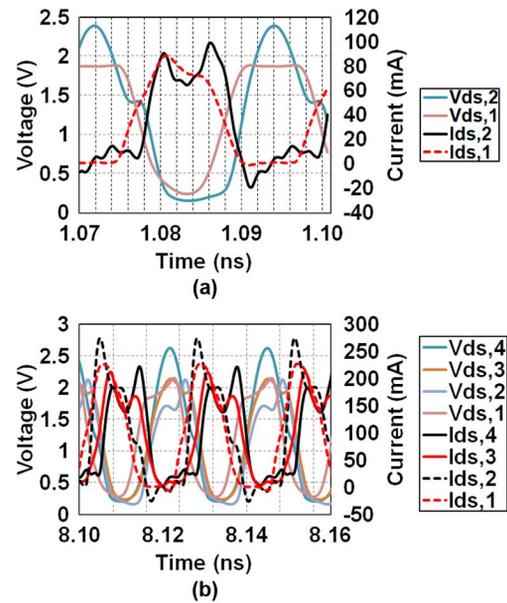


Fig. 13. Post-layout simulated drain–source voltages and corresponding switch currents for: (a) two-stacked PA and (b) four-stacked PA in 45-nm SOI CMOS.

operation of the PA. Ignoring these losses in the theoretical analysis results in PAE higher than what is obtained from actual device-based simulations.

For linear stacked PAs, the impact of “Elmore delay” can be accounted (and compensated) for theoretically [28] since a small-signal model for the devices is used for the preliminary analysis and design procedure. A similar endeavor for stacked switching PAs is challenging owing to nonlinear operation of the circuit. Consequently, we adopt a simulation-based approach to investigate this effect for a four-stacked configuration (without the tuning inductor) using a switch+capacitor-based model for the devices. The resulting circuit resembles that in Fig. 9(b) (sans the tuning inductor, the input matching network, and with a square-wave drive instead of a sinusoidal input) with each device modeled as a switch augmented with intrinsic device capacitances. Layout parasitics (capacitors, resistors, and inductances) based on Fig. 7 are also incorporated in the circuit to facilitate better correlation with device-based results. The delay in the switch-voltage waveforms exhibit close correspondence with those obtained from device-based simulations as well, as shown in Fig. 16. Since the voltage profiles confirm no significant delay, the phenomenon of “Elmore delay” is not a concern in stacked switching PAs. The resulting output power and PAE are reported in Fig. 3(a). The reduction in efficiency ( $\approx 20\%$ ) for the switch+capacitor-based model indicates that capacitive discharge loss at intermediate nodes is a more important practical consideration. These results, in conjunction with the comparison presented in Fig. 15 indicate that ignoring “Elmore delay” and the additional loss mechanisms in the theoretical analysis is not a crippling limitation since it does not significantly alter the waveform characteristics, and hence, the impact on switching behavior of the stacked configuration at mmWave frequencies. One would therefore obtain output power similar to that predicted by theory, but at a lower PAE, which follows the theoretical trend [see Fig. 3(a)]. This also demonstrates the efficacy of

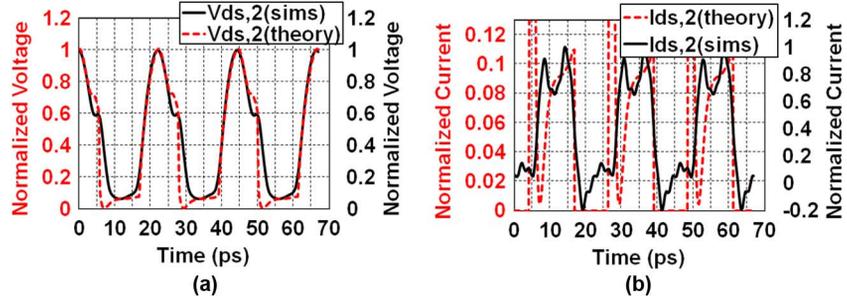


Fig. 14. Comparison of post-layout simulated waveforms for device  $M_2$  of the two-stacked PA in 45-nm SOI CMOS with theory.

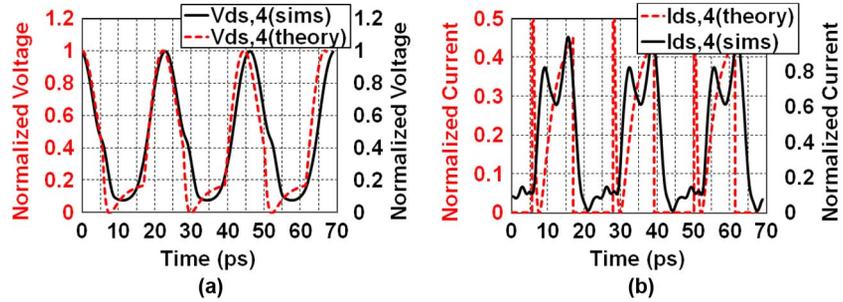


Fig. 15. Comparison of post-layout simulated waveforms for device  $M_4$  of the four-stacked PA in 45-nm SOI CMOS with theory.

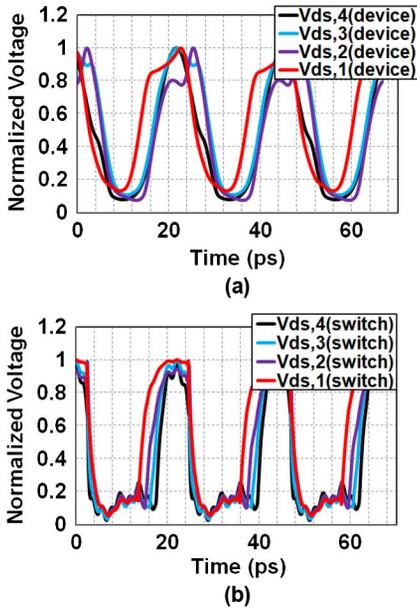


Fig. 16. (a) Post-layout simulated device voltages for the four-stacked PA prototype without tuning inductor in 45-nm SOI CMOS. (b) Simulated switch voltages for the same four-stacked configuration without tuning inductor, using a switch+capacitor-based model for the devices and layout parasitics from Fig. 7.

a simple switch+capacitor-based model to predict the performance of a practical implementation.

## V. EXPERIMENTAL RESULTS

The chip microphotographs of the PAs are shown in Fig. 17. The PAs are tested in chip-on-board configuration through on-chip probing using two coaxial 1.85 mm (dc–65 GHz) GSG probes.

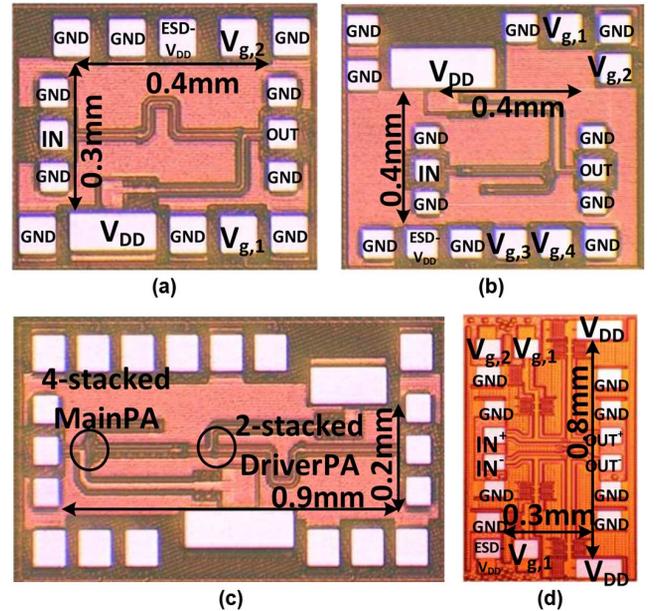


Fig. 17. Chip microphotographs of the mmWave stacked Class-E-like PAs with: (a) two devices stacked in 45-nm SOI CMOS, (b) four devices stacked in 45-nm SOI CMOS, (c) a two-stage cascade of a main PA with four devices stacked with a two-stacked driver stage in 45-nm SOI CMOS, and (d) two devices stacked in 65-nm low-power bulk CMOS.

### A. Small-Signal Measurements

The small-signal measurement setup is calibrated at the probe tip planes. The small-signal measurements are conducted up to 65 GHz using an Anritsu 37397E Lightning vector network analyzer (VNA). Figs. 18 and 19 illustrate the simulated and measured small-signal  $S$ -parameters of the two-stacked PA and the four-stacked PA implemented in 45-nm SOI CMOS. The measured peak gain of the two-stacked PA is 13.5 dB at 46 GHz

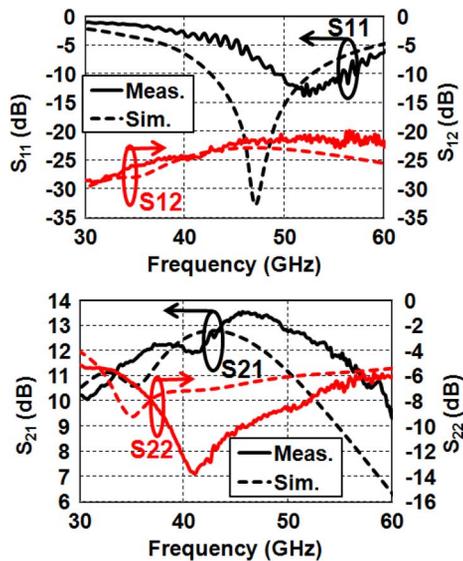


Fig. 18. Small-signal  $S$ -parameters of 45-nm SOI two-stacked Class-E-like PA ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.7$  V,  $V_{DD} = 2.4$  V). Power consumption = 49 mW.

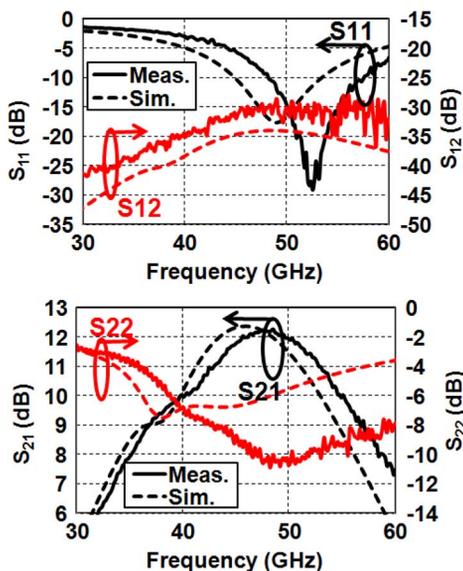


Fig. 19. Small-signal  $S$ -parameters of 45-nm SOI four-stacked Class-E-like PA ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.8$  V,  $V_{g3} = 2.8$  V,  $V_{g4} = 4$  V,  $V_{DD} = 4.8$  V). Power consumption = 206 mW.

with a  $-3$ -dB bandwidth extending from 32 to 59 GHz. The  $-1$ -dB bandwidth extends from 42 to 52 GHz, making it suitable for wideband applications. The measured peak gain of the four-stacked PA is 12.3 dB at 48.5 GHz with a  $-3$ -dB bandwidth extending from 37 to 56 GHz. The measured  $-1$ -dB bandwidth spans a wide frequency range from 43.5 to 52.5 GHz. A frequency shift of  $\approx 3$ –5 GHz is observed between measured and simulated curves for both PAs in both  $S_{11}$  and  $S_{21}$ . This can probably be attributed to overestimation of capacitive parasitics at design time. The fact that the PAs have a significant small-signal gain goes against the concept of conventional Class-E PA design, but is simply an outcome of the “Class-E-like” design methodology described in this paper. The PA is designed for optimum performance at a Class-E input drive level, at which point the devices can be regarded as hard switching. However,

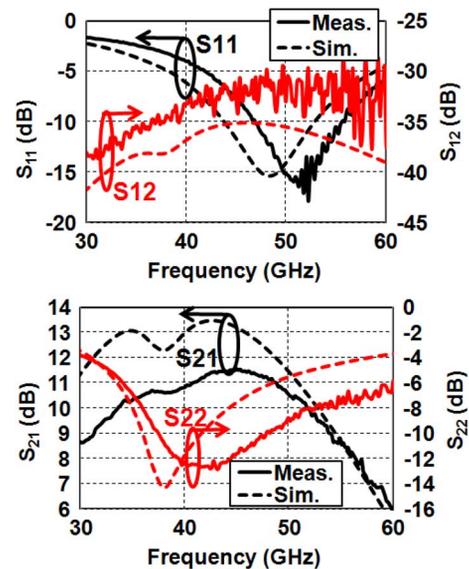


Fig. 20. Small-signal  $S$ -parameters of 45-nm SOI four-stacked Class-E-like PA with the tuning inductor eliminated using laser trimming ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.8$  V,  $V_{g3} = 2.8$  V,  $V_{g4} = 4$  V,  $V_{DD} = 4.8$  V). Power consumption = 206 mW.

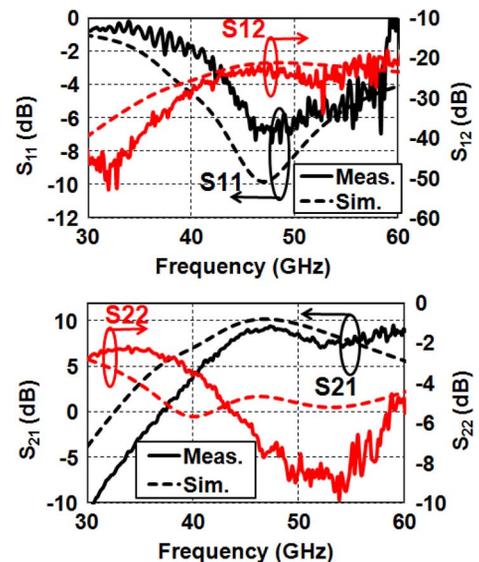


Fig. 21. Small-signal  $S$ -parameters (single ended) of the 65-nm differential two-stacked Class-E-like PA ( $V_{g1} = 0.8$  V,  $V_{g2} = 2$  V,  $V_{DD} = 2$  V). Power consumption = 89 mW under small-signal operation.

at the dc bias point, the devices are biased somewhat above the threshold voltage, imparting the circuit with small-signal gain. Of course, this gain is less than the maximum gain available from the stacked devices as the output load is designed based on Class-E principles. A modified version of the 45-nm SOI four-stacked PA, obtained by laser-trimming the tuning inductor, was also characterized, and its simulated and measured small-signal  $S$ -parameters are reported in Fig. 20. The measured peak gain is 11.6 dB at 45 GHz with a  $-3$ -dB bandwidth extending from 30 to 55 GHz. The  $-1$ -dB bandwidth extends from 36 to 50 GHz.

The measured and simulated small-signal  $S$ -parameters of the two-stacked differential PA implemented in 65-nm bulk CMOS are illustrated in Fig. 21. The measurement setup and

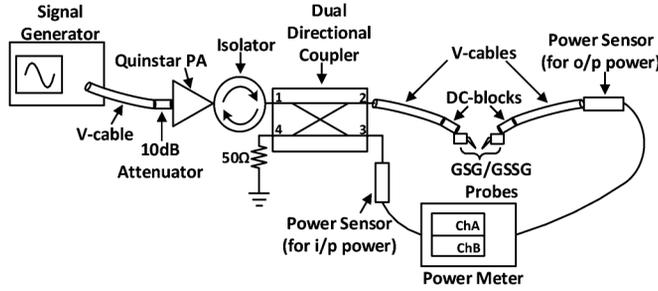


Fig. 22. Large-signal  $Q$ -band measurement setup for the fabricated PAs.

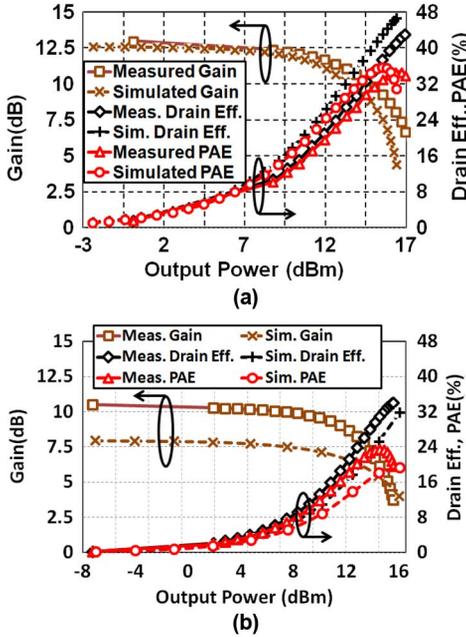


Fig. 23. Measured gain, drain efficiency, and PAE as a function of output power for: (a) the 45-nm SOI two-stacked Class-E-like PA at 47 GHz ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.7$  V,  $V_{DD} = 2.4$  V) and (b) the 65-nm differential two-stacked Class-E-like PA at 47.5 GHz ( $V_{g1} = 0.8$  V,  $V_{g2} = 2.1$  V,  $V_{DD} = 2.8$  V).

calibration procedure are the same as discussed before with the exception that coaxial 1.85-mm coplanar wave GSSG probes are used for the measurements. However, one probe of each differential pair is terminated with  $50 \Omega$  so, in essence, single-ended measurements are being performed. This stems from the practical challenges in creating a differential mmWave measurement setup. The measured peak gain is 9.5 dB at 47 GHz with a  $-1$ -dB bandwidth extending from 44 to 50 GHz.

### B. Large-Signal Measurements

The large-signal measurement setup for the aforementioned PAs is shown in Fig. 22. The large-signal characteristics of the 45-nm SOI PAs are shown in Figs. 23(a) and 24. Measurement results yield a peak PAE of 34.6% for the two-stacked PA with a saturated output power of 17.6 dBm at 47 GHz. Compared to the cascode PA in [27] operating at a similar frequency at supply voltages close to the nominal single-device  $V_{DD}$  of the technology, the two-stacked prototype achieves 3–7-dB higher output power along with 10%–12% higher PAE. The four-stacked PA has measured saturated output power of

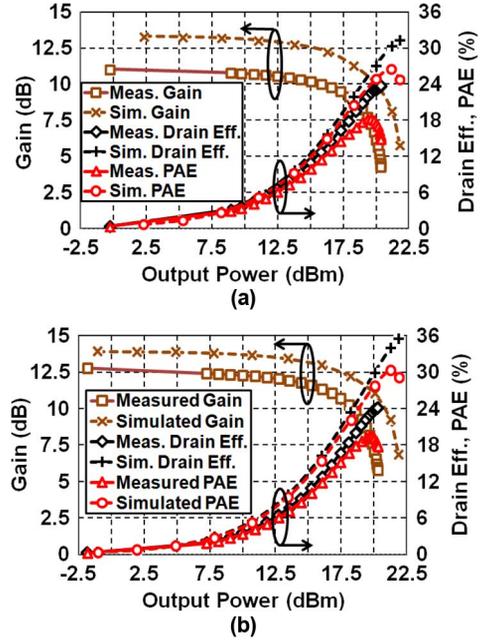


Fig. 24. Measured gain, drain efficiency and PAE as a function of output power for: (a) the 45-nm SOI four-stacked Class-E-like PA with the tuning inductor eliminated through laser trimming at 42.5 GHz and (b) the 45-nm SOI four-stacked Class-E-like PA at 47.5 GHz ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.8$  V,  $V_{g3} = 2.8$  V,  $V_{g4} = 4$  V, and  $V_{DD} = 4.8$  V for both designs).

20.3 dBm at 47.5 GHz at a peak PAE of 19.4%. For the trimmed version of the four-stacked PA without the tuning inductor, a peak PAE of 18.3% was achieved at 42.5 GHz along with a saturated output power of 20.3 dBm. Unlike the two-stacked PA, the measured performance metrics of the four-stacked PAs (particularly efficiency) are somewhat lower than those predicted by simulations. This is an indication of unmodeled active losses, as there is good correspondence between the measured and simulated characteristics of the various passive components [20]. The loss in the active components depends on a proper choice of device layout, as well as accurate modeling, as discussed in Section III.

Large-signal measurements were also conducted for the two- and four-stacked PAs across frequency (at the optimal bias point) and for different supply voltages (at a fixed frequency, keeping gate biases constant). The results are depicted in Figs. 25 and 26. Large-signal measurement beyond 48 GHz was limited by the characteristics of the measurement equipment (specifically, the Quinstar PA used to drive the PAs under test, as well as the isolator, dual-directional coupler, and the power sensors used in the measurement setup). Unlike the two-stacked PA, the output power does not increase with increasing supply voltage for the four-stacked prototype. Once again, this can probably be attributed to the device layout discussed previously that results in lower  $f_{max}$ .

This hypothesis is tested in measurement. As discussed previously, an important characteristic of switching PAs is linearity with respect to supply voltage, which causes the average supply current and the output power to scale linearly and quadratically with supply voltage, respectively. This unique feature distinguishes switching PAs from the class of linear PAs. At mmWave

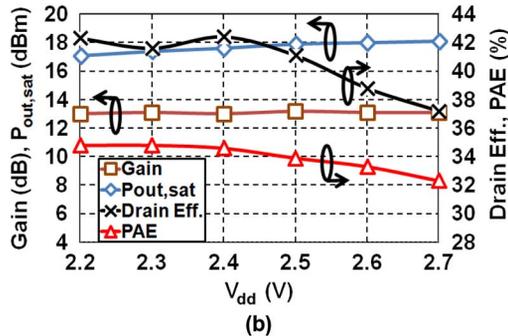
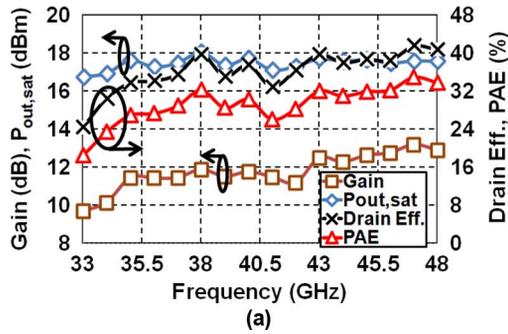


Fig. 25. Measured gain, saturated output power, drain efficiency, and PAE: (a) across frequency ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.7$  V,  $V_{DD} = 2.4$  V) and (b) across supply voltage at 47 GHz of the 45-nm SOI two-stacked Class-E-like PA ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.7$  V).

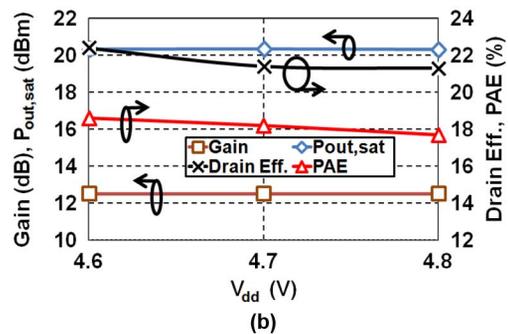
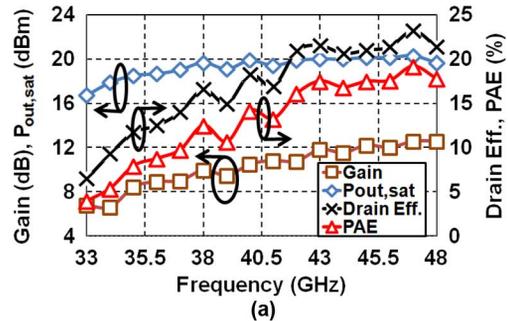


Fig. 26. Measured gain, saturated output power, drain efficiency, and PAE: (a) across frequency ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.8$  V,  $V_{g3} = 2.8$  V,  $V_{g4} = 4$  V) and (b) across supply voltage at 47 GHz of the 45-nm SOI four-stacked Class-E-like PA.

frequencies, the various sources of nonidealities result in deviation from ideal Class-E characteristics. Thus, the scaling of supply current and output power with supply voltage can be utilized as a useful metric to determine the extent of switching characteristics of a PA in the mmWave regime. Fig. 27 illustrates the measured average supply current and saturated output power of the two-stacked PA in 45-nm SOI CMOS as a function of  $V_{DD}$  and  $V_{DD}^2$ , respectively. The respective linear trends

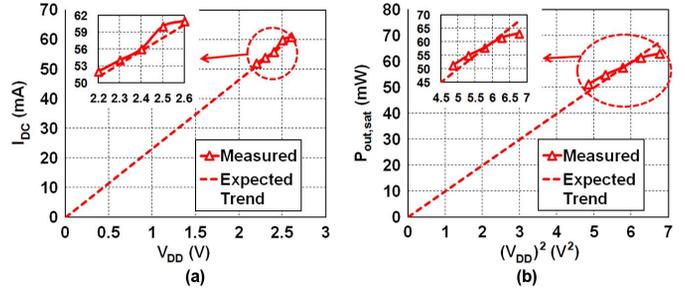


Fig. 27. Measured and expected: (a) average supply current versus  $V_{DD}$  and (b) saturated output power versus  $V_{DD}^2$  for two-stacked Class-E-like PA in 45-nm SOI CMOS. The profiles display the linearity with respect to supply voltage associated with switching Class-E PAs, thereby establishing the Class-E-like characteristics of the PA even at mmWave frequencies.

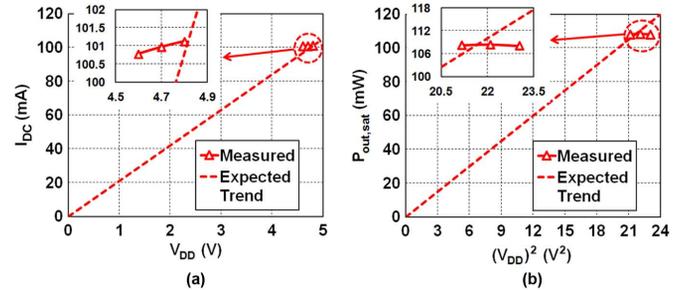


Fig. 28. Measured and expected: (a) average supply current versus  $V_{DD}$  and (b) saturated output power versus  $V_{DD}^2$  for four-stacked Class-E-like PA in 45-nm SOI CMOS. The profiles do not display the linearity with respect to supply voltage characteristic of switching Class-E PAs, owing to layout-induced increased gate resistance, which prevents hard switching at mmWave frequencies.

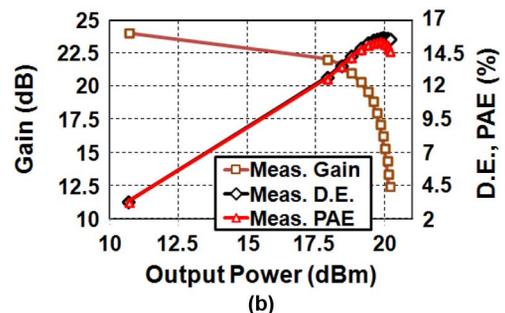
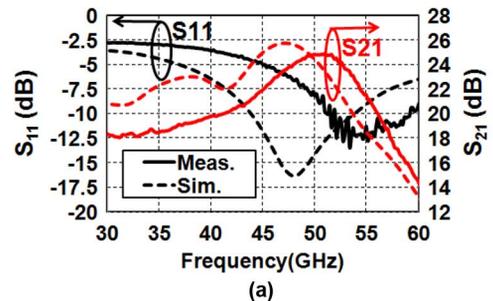


Fig. 29. (a) Measured small-signal  $S$ -parameters and (b) measured gain, drain efficiency, and PAE as a function of output power for the two-stage 45-nm SOI PA comprising a four-stacked main PA and a two-stacked driver stage at 47 GHz ( $V_{g1} = 0.4$  V,  $V_{g2} = 1.7$  V,  $V_{g3} = 2.8$  V,  $V_{g4} = 4$  V,  $V_{DD,1} = 4.8$  V,  $V_{g5} = 0.4$  V,  $V_{g6} = 1.6$  V, and  $V_{DD,2} = 2.4$  V). Power consumption = 255 mW under small-signal operation.

can be clearly observed, thereby corroborating the Class-E-like nature of the design. This also proves that switch-like PAs can indeed be implemented at mmWave frequencies with appropriate design methodology. The corresponding results for the

TABLE III  
COMPARISON OF FABRICATED CLASS-E-LIKE PAs WITH STATE-OF-THE-ART CMOS AND SiGe mmWAVE PAs  
(REFERENCES ARE ORGANIZED IN ORDER OF DECREASING PAE, FOR CMOS AND SiGe PAs SEPARATELY)

Reference	Technology	Freq. (GHz)	$V_{DD}$ (V)	$P_{sat}$ (dBm)	$\eta$ (%)	Peak PAE (%)	Gain (dB)	ITRS FOM **	Class of Operation	Power Combining	Fully Integrated?
This work	45 nm SOI	47	2.4	17.6	42.4	34.6	13	59.43	Class E, 2-stacked	None	Yes
This work	45 nm SOI	47.5	4.8	20.3	23	19.4	12.8	59.51	Class E, 4-stacked	None	Yes
This work	45 nm SOI	47	2.4 (Driver), 4.8 (PA)	20.1	15.6	15.4	24.9	70.32	Class E, Two-stage cascade	None	Yes
This work	65nm	47.5	2.8	18.2	35.8	28.3	11.2	57.45	Class E, 2-stacked	Diff., with diff. output <sup>§</sup>	Yes
[21]	45 nm SOI	45	2.7	18.6	N/A	34	9.5	56.48	Class AB, 2-stacked	None	Yes
[34]	32 nm SOI	60	0.9	12.5	N/A	30	10	52.83	Class-E	Diff. with diff. output <sup>§</sup>	Yes
[5]	40 nm	60	1	17.4	35.9	29.3	21.2	68.83	Class AB	2-way diff. transformer combined	Yes
[19]	45 nm SOI	47.5	2.8	17.9	33.8	25.5	9.8	55.3	Dual-Output Class E, 2-stacked	None	Yes
[35]	40 nm	60	1	15.6	N/A	25	N/A	N/A	N/A	2-way diff. transformer-combined (outphasing)	Yes
[27]	65 nm SOI	60	1.8	14.5	N/A	25	16	60.04	Class AB, cascode	None	Yes
[18]	45 nm SOI	45	4	18.2	N/A	23	8	52.88	Class AB, 3-stacked	None	Yes
[36]	65 nm	79	1	19.3	N/A	19.2	24.2	74.29	N/A	8-way transformer and t-line combiner	Yes
[19]	45 nm SOI	47.5	2.9	19.1	24.5	16	8.2	52.88	Dual-Output Class E, 2-stacked	2-way current combined	Yes
[37]	65 nm	60	1	18.6	N/A	15.1	20.3	45.95	N/A	4-way transformer combined	Yes
[17]	45 nm SOI	45	5.1	24.3	21.3	14.6	>18	67	Class B/AB, 4-stacked	Diff. with diff. output <sup>§</sup>	No <sup>+</sup>
[38]	90 nm	60	1.2	19.9	N/A	14.2	20.6	67.59	N/A	4-way Wilkinson-tree combiner	Yes
[39]	0.13 $\mu$ m SiGe	41	4	23.6	N/A	31	12.5	63.3	Class E, 2-stacked	None	Yes
[40]	0.13 $\mu$ m SiGe	45	2.5	21.7	25	22	9.3	57.5	Class E	2-way Wilkinson combiner	Yes
[41]	0.13 $\mu$ m SiGe	42	2.4	19.4	N/A	14.4	6	49.4	Class-E	2-way Wilkinson combiner	Yes

\*\* Defined as  $P_{sat}$  (dBm) + Gain (dB) +  $20\log_{10}$ (Freq. (GHz)) +  $10\log_{10}$ (PAE).

§ Ideal external lossless output balun assumed.

+ Uses off-chip bias-T for providing power supply.

four-stacked PA with the tuning inductor are shown in Fig. 28. The four-stacked PA's measured characteristics deviate from expected trends. This indicates that the devices are not being driven to a hard-switching condition, likely due to reduced device  $f_{max}$ . It should be noted that [18] and [17] have realized large power devices at mmWave with high  $f_{max}$ , and hence, the foregoing results for the four-stacked PA should not be taken to mean that Class-E operation is not possible for high levels of stacking.

The small-signal  $S$ -parameters and large-signal performance metrics of the two-stage PA implemented in 45-nm SOI CMOS are summarized in Fig. 29. The measured peak gain is 24.9 dB at 51 GHz while a peak PAE of 15.4% was achieved at 47 GHz along with a saturated output power of 20.1 dBm.

Large-signal measurements of the two-stacked differential PA implemented in 65-nm bulk CMOS yield a peak PAE of 28.3% with a saturated output power of 15.2 dBm at 47.5 GHz [see Fig. 23(b)], implying a saturated differential output power of 18.2 dBm. The lower efficiency of this PA compared to the two-stacked 45-nm SOI PA stems from the higher ON-resistance of the 65-nm devices.

### C. Comparison With State-of-the-Art

Table III depicts a comparison of these PAs to state-of-the-art mmWave CMOS and SiGe PAs. The references are arranged in order of decreasing PAE. The 65-nm PA is comparable to state-of-the-art implementations in efficiency, despite the poor ON-resistance characteristics of the technology. This is a direct consequence of the loss-aware Class-E design methodology. On the other hand, the two-stacked PA in 45-nm SOI CMOS exhibits the highest PAE reported for a CMOS mmWave PA. The

PA reported in [21] exhibits similar PAE and output power, and also employs device stacking in 45-nm SOI CMOS, albeit in the context of Class-AB operation. The four-stacked PA in 45-nm SOI CMOS exhibits the highest output power achieved from a fully integrated CMOS mmWave PA. The work in [17] uses an off-chip bias-T to provide the supply voltage, and consequently does not integrate the mmWave dc-feed inductor. Furthermore, it is a differential implementation with a differential output, and assumes ideal 3-dB external differential-to-single-ended conversion. It is important to study the output power delivered to a single-ended output pad when comparing works. An on-chip dc-feed inductor is seen in simulation to introduce approximately 1-dB output-side loss based on the quality-factor achievable in this technology. When these are factored in, the work in [17] achieves comparable output power to our four-stacked PA, with an associated PAE that is lower than our four-stack and comparable to our cascade PA. Other prior fully integrated CMOS mmWave PAs with comparable output power [19], [36], [38] rely on power combining. Since most of the works reported in Table III operate at higher frequencies, it is important to use a FOM to ensure fair comparison. The ITRS FOM, defined as

$$\text{ITRS FOM} = P_{sat}(\text{dBm}) + \text{Gain}(\text{dB}) + 10\log_{10}\text{PAE} + 20\log_{10}f_0 \quad (10)$$

where  $f_0$  is the operating frequency in gigahertz, takes into account four important performance metrics of a PA. The implemented single-stage prototypes achieve ITRS FOM comparable to current state-of-the-art mmWave CMOS PAs and the highest

amongst fully integrated PAs, which do not employ power combining. In particular, the two-stage cascade PA in 45-nm SOI CMOS achieves the highest ITRS FOM amongst PAs, which do not employ power combining, and second highest overall.

## VI. CONCLUSION

This work indicates that stacked switching CMOS PAs potentially take us one step closer to implementing efficient PAs in CMOS with watt-level output power at mmWave frequencies for the first time. Topics for future research include large-scale low-loss power-combining techniques so that multiple such PAs may be combined to approach watt-class output power, and linearizing architectures that enable such mmWave switching-type PAs to be used with complex modulation formats with high average efficiency.

## APPENDIX

Referring to Fig. 2(b) and utilizing the loss-aware Class-E design methodology described in [23], the analytical equations describing the switch voltage  $V_{S,ON}$  and  $V_{S,OFF}$  during the ON ( $(T/2) < t < T$ ) and OFF ( $0 < t < (T/2)$ ) half-cycles, respectively, for a stacked configuration can be derived to be

$$V_{S,ON}(t) = nV_{DD} + a_1 e^{\beta t} + a_2 \cos(\omega_0 t + \phi) + a_3 \sin(\omega_0 t + \phi) \quad (11)$$

and

$$V_{S,OFF}(t) = nV_{DD} [1 - \cos(\omega_s t) + V_{S,OFF}(0) \cos(\omega_s t) + \frac{V'_{S,OFF}(0)}{\omega_s} \sin(\omega_s t) + \frac{i_0 \omega_0 \sin(\phi)}{C_{out,n}(\omega_s^2 - \omega_0^2)} [\cos(\omega_0 t) - \cos(\omega_s t)] + \frac{i_0 \omega_0^2 \cos(\phi)}{C_{out,n}(\omega_s^2 - \omega_0^2)} \left[ \frac{\sin(\omega_0 t)}{\omega_0} - \frac{\sin(\omega_s t)}{\omega_s} \right]] \quad (12)$$

where  $\omega_0$  is the switching frequency,  $\omega_s = 1/\sqrt{L_s C_{out,n}}$ ,

$$\begin{aligned} a_1 &= V_{S,ON} \left( \frac{T}{2} \right) - nV_{DD} \\ &\quad - \frac{nR_{on,n} i_0 e^{-\frac{\beta T}{2}} + \frac{\beta}{\omega_0} nR_{on,n} i_0 e^{-\frac{\beta T}{2}} \sin(\phi)}{\left( 1 + \frac{\beta^2}{\omega_0^2} \right)} \\ a_2 &= \frac{-nR_{on,n} i_0}{1 + \frac{\beta^2}{\omega_0^2}} \\ a_3 &= \frac{-nR_{on,n} i_0 \beta}{\omega_0 \left( 1 + \frac{\beta^2}{\omega_0^2} \right)} \\ \beta &= \frac{-nR_{on,n}}{L_s} \end{aligned} \quad (13)$$

while  $V_{S,ON}(T/2)$ ,  $V_{S,OFF}(0)$  and  $V'_{S,OFF}(0)$  are constants determined by imposing continuity of dc-feed inductor current  $i_{L_s}$

at the switching instant ( $t = T/2$ ) and the periodicity of waveforms as follows:

$$i_{L_s,OFF}(0^+) = i_{L_s,ON}(T^-), V_{S,OFF}(0^+) = V_{S,ON}(T^-) \quad (14)$$

$$i_{L_s,OFF} \left( \frac{T^+}{2} \right) = i_{L_s,ON} \left( \frac{T^-}{2} \right). \quad (15)$$

The capacitive discharge loss at the switching instant can be estimated as

$$P_{loss,cap} = 0.5 f_0 C_{out,n} \left[ V_{S,OFF}^2 \left( \frac{T^-}{2} \right) - V_{S,ON}^2 \left( \frac{T^+}{2} \right) \right]. \quad (16)$$

The expressions for  $i_{L_s}$  can be derived using

$$i_{L_s,ON} = \frac{V_{S,ON}}{nR_{on}} + i_0 \cos(\omega_0 t + \phi) \quad (17)$$

and

$$i_{L_s,OFF} = C_{out,n} \frac{dV_{S,OFF}}{dt} + i_0 \cos(\omega_0 t + \phi). \quad (18)$$

The loss in the switch and the dc-feed inductance are given by

$$P_{loss,switch} = nR_{ON} * \frac{1}{T} \int_{\frac{T}{2}}^T \left( \frac{V_{S,ON}}{nR_{ON}} \right)^2 dt \quad (19)$$

$$P_{loss,choke} = R_{choke} * \frac{1}{T} \left( \int_0^{\frac{T}{2}} i_{L_s,OFF}^2 dt + \int_{\frac{T}{2}}^T i_{L_s,ON}^2 dt \right) \quad (20)$$

respectively, while the input power ( $P_{in}$ ) required to switch a device between the triode and cutoff regions is approximated as

$$P_{in} = k f_0 C_{in,n} V_{on}^2 \quad (21)$$

where  $C_{in,n} = C_{gs,n} + C_{gd,n}$  is the input capacitance of the bottom device in the triode region,  $V_{on}$  is the input drive level in the ‘‘ON’’ half-cycle and  $k$  is a fitting parameter determined from schematic simulations. The foregoing lead to a complete expression for PAE,

$$PAE = 1 - \frac{P_{loss}}{P_{DC}} - \frac{P_{in}}{P_{DC}} \quad (22)$$

where

$$P_{loss} = P_{loss,switch} + P_{loss,choke} + P_{loss,cap} \quad (23)$$

and

$$P_{DC} = nV_{DD} \times I_{DC,n} = nV_{DD} \times \frac{1}{T} \left( \int_0^{\frac{T}{2}} i_{L_s,OFF} dt + \int_{\frac{T}{2}}^T i_{L_s,ON} dt \right). \quad (24)$$

A MATLAB code subsequently sweeps the magnitude  $i_0$  and phase  $\phi$  of the load current to arrive at a design point with optimal PAE for a given device size, input drive level  $V_{on}$ , the tuning parameter  $\omega_s$ , and the number of devices stacked  $n$ .

## REFERENCES

- [1] P. Haldi, G. Liu, and A. Niknejad, "CMOS compatible transformer power combiner," *Electron. Lett.*, vol. 42, no. 19, pp. 1091–1092, Sep. 2006.
- [2] Y. Zhao, J. Long, and M. Spirito, "Compact transformer power combiners for millimeter-wave wireless applications," in *IEEE Radio Freq. Integr. Circuits Symp.*, May 2010, pp. 223–226.
- [3] D. Chowdhury, C. Hull, O. Degani, Y. Wang, and A. Niknejad, "A fully integrated dual-mode highly linear 2.4 GHz CMOS power amplifier for 4G WiMax applications," *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3393–3402, Dec. 2009.
- [4] J. W. Lai and A. Valdes-Garcia, "A 1 V 17.9 dBm 60 GHz power amplifier in standard 65 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Tech. Dig.*, Feb. 2010, pp. 424–425.
- [5] D. Zhao, S. Kulkarni, and P. Reynaert, "A 60 GHz dual-mode power amplifier with 17.4 dBm output power and 29.3% PAE in 40-nm CMOS," in *Proc. ESSCIRC*, Sep. 2012, pp. 337–340.
- [6] M. Bohsali and A. Niknejad, "Current combining 60 GHz CMOS power amplifiers," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2009, pp. 31–34.
- [7] T. Dickson, K. H. K. Yau, T. Chalvatzis, A. Mangan, E. Laskin, R. Beerkens, and P. Westergaard *et al.*, "The invariance of characteristic current densities in nanoscale MOSFETs and its impact on algorithmic design methodologies and design porting of Si(Ge) (Bi)CMOS high-speed building blocks," *IEEE J. Solid-State Circuits*, vol. 41, no. 8, pp. 1830–1845, Aug. 2006.
- [8] A. Niknejad, S. Emami, B. Heydari, M. Bohsali, and E. Adabi, "Nanoscale CMOS for mm-wave applications," in *IEEE Compound Semicond. Integr. Circuits Symp.*, Oct. 2007, pp. 1–4.
- [9] B. Heydari, M. Bohsali, E. Adabi, and A. Niknejad, "Millimeter-wave devices and circuit blocks up to 104 GHz in 90 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 42, no. 12, pp. 2893–2903, Dec. 2007.
- [10] S. Nicolson, A. Tomkins, K. Tang, A. Cathelin, D. Belot, and S. Voinigescu, "A 1.2 V, 140 GHz receiver with on-die antenna in 65 nm CMOS," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2008, pp. 229–232.
- [11] E. Johnson, "Physical limitations on frequency and power parameters of transistors," in *IRE Int. Convention Rec.*, Mar. 1965, vol. 13, pp. 27–34.
- [12] I. Sarkas, A. Balteanu, E. Daquay, A. Tomkins, and S. Voinigescu, "A 45 nm SOI CMOS class-D mm-wave PA with >10 Vpp differential swing," in *IEEE Int. Solid-State Circuits Conf. Tech. Dig.*, Feb. 2012, pp. 88–90.
- [13] A. Mazzanti, L. Larcher, R. Brama, and F. Svelto, "Analysis of reliability and power efficiency in cascode class-E PAs," *IEEE J. Solid-State Circuits*, vol. 41, no. 5, pp. 1222–1229, May 2006.
- [14] A. Ezzeddine and H. Huang, "The high voltage/high power FET (HiVP)," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2003, pp. 215–218.
- [15] S. Pornpromlikit, J. Jeong, C. Presti, A. Scuderi, and P. Asbeck, "A 33-dBm 1.9-GHz silicon-on-insulator CMOS stacked-FET power amplifier," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2009, pp. 533–536.
- [16] J. McRory, G. Rabjohn, and R. Johnston, "Transformer coupled stacked FET power amplifiers," *IEEE J. Solid-State Circuits*, vol. 34, no. 2, pp. 157–161, Feb. 1999.
- [17] A. Balteanu, I. Sarkas, E. Daquay, A. Tomkins, and S. Voinigescu, "A 45-GHz, 2-bit power DAC with 24.3 dBm output power, >14 Vpp differential swing, and 22% peak PAE in 45-nm SOI CMOS," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2012, pp. 319–322.
- [18] S. Pornpromlikit, H.-T. Dabag, B. Hanafi, J. Kim, L. Larson, J. Buckwalter, and P. Asbeck, "A Q-band amplifier implemented with stacked 45-nm CMOS FETs," in *IEEE Compound Semicond. Integr. Circuit Symp.*, Oct. 2011, pp. 1–4.
- [19] A. Chakrabarti, J. Sharma, and H. Krishnaswamy, "Dual-output stacked class-EE power amplifiers in 45 nm SOI CMOS for Q-band applications," in *IEEE Compound Semicond. Integr. Circuit Symp.*, Oct. 2012, pp. 1–4.
- [20] A. Chakrabarti and H. Krishnaswamy, "High power, high efficiency stacked mmWave class-E-like power amplifiers in 45 nm SOI CMOS," in *IEEE Custom Integr. Circuits Conf.*, Sep. 2012, pp. 1–4.
- [21] A. Agah, H. Dabag, B. Hanafi, P. Asbeck, L. Larson, and J. Buckwalter, "A 34% PAE, 18.6 dBm 42–45 GHz stacked power amplifier in 45 nm SOI CMOS," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2012, pp. 57–60.
- [22] N. Sokal and A. Sokal, "Class E-A new class of high-efficiency tuned single-ended switching power amplifiers," *IEEE J. Solid-State Circuits*, vol. 10, no. 3, pp. 168–176, Jun. 1975.
- [23] A. Chakrabarti and H. Krishnaswamy, "An improved analysis and design methodology for RF class-E power amplifiers with finite DC-feed inductance and switch on-resistance," in *IEEE Int. Circuits Syst. Symp.*, May 2012, pp. 1763–1766.
- [24] O. Lee, J. Han, K. H. An, D. H. Lee, K.-S. Lee, S. Hong, and C.-H. Lee, "A charging acceleration technique for highly efficient cascode class-E CMOS power amplifiers," *IEEE J. Solid-State Circuits*, vol. 45, no. 10, pp. 2184–2197, Oct. 2010.
- [25] D. Sandstrom, B. Martineau, M. Varonen, M. Karkkainen, A. Cathelin, and K. A. I. Halonen, "94 GHz power-combining power amplifier with +13 dBm saturated output power in 65 nm CMOS," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2011, pp. 1–4.
- [26] S. Ko and J. Lin, "A linearized cascode cmos power amplifier," in *IEEE Annu. Wireless Microw. Technol. Conf.*, 2006, pp. 1–4.
- [27] A. Siligaris *et al.*, "A 60 GHz power amplifier with 14.5 dBm saturation power and 25% peak PAE in CMOS 65 nm SOI," *IEEE J. Solid-State Circuits*, vol. 45, no. 7, pp. 1286–1294, Jul. 2010.
- [28] H. Dabag, B. Hanafi, F. Golcuk, A. Agah, J. Buckwalter, and P. Asbeck, "Analysis and design of stacked-FET millimeter-wave power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 4, pp. 1543–1556, Apr. 2013.
- [29] S. Kee, "The class E/F family of harmonic-tuned switching power amplifiers" Ph.D. dissertation, Dept. Elect. Eng., California Inst. Technol., Pasadena, CA, USA, 2001. [Online]. Available: <http://resolver.caltech.edu/CaltechETD:etd-04262005-152703>
- [30] J. Sharma and H. Krishnaswamy, "216- and 316-GHz 45-nm SOI CMOS signal sources based on a maximum-gain ring oscillator topology," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 1, pp. 492–504, Jan. 2013.
- [31] "BSIM SOI Manual," BSIM Group, Univ. California at Berkeley, Berkeley, CA, USA.
- [32] "IE3D User Manual," Mentor Graphics Corporation, Wilsonville, OR, USA.
- [33] U. Gogineni, J. del Alamo, and C. Putnam, "RF power potential of 45 nm CMOS technology," in *Silicon Monolithic Integr. Circuits RF Syst. Top. Meeting*, Jan. 2010, pp. 204–207.
- [34] O. Ogunnika and A. Valdes-Garcia, "A 60 GHz class-E tuned power amplifier with PAE >25% in 32 nm SOI CMOS," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2012, pp. 65–68.
- [35] D. Zhao, S. Kulkarni, and P. Reynaert, "A 60 GHz outphasing transmitter in 40 nm CMOS with 15.6 dBm output power," in *IEEE Int. Solid-State Circuits Conf. Tech. Dig.*, Feb. 2012, pp. 170–172.
- [36] K.-Y. Wang, T.-Y. Chang, and C.-K. Wang, "A 1 V 19.3 dBm 79 GHz power amplifier in 65 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Tech. Dig.*, Feb. 2012, pp. 260–262.
- [37] J. Chen and A. Niknejad, "A compact 1 V 18.6 dBm 60 GHz power amplifier in 65 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Tech. Dig.*, Feb. 2011, pp. 432–433.
- [38] C. Law and A.-V. Pham, "A high-gain 60 GHz power amplifier with 20 dBm output power in 90 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Tech. Dig.*, Feb. 2010, pp. 426–427.
- [39] K. Datta, J. Roderick, and H. Hashemi, "Analysis, design and implementation of mm-wave SiGe stacked class-E power amplifiers," in *IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2013, pp. 275–278.
- [40] K. Datta, J. Roderick, and H. Hashemi, "A 22.4 dBm two-way Wilkinson power-combined Q-band SiGe class-E power amplifier with 23% peak PAE," in *IEEE Compound Semicond. Integr. Circuit Symp.*, Oct. 2012, pp. 1–4.
- [41] N. Kalantari and J. Buckwalter, "A 19.4 dBm, Q-band class-E power amplifier in a 0.12  $\mu\text{m}$  SiGe BiCMOS process," *IEEE Microw. Wireless Compon. Lett.*, vol. 20, no. 5, pp. 283–285, May 2010.



**Anandaroop Chakrabarti** received the B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 2010, the M.S. degree in electrical engineering from Columbia University, New York, NY, USA, in 2011, and is currently working toward the Ph.D. degree at Columbia University, New York, NY, USA.

In Summer 2013, he was with the IBM T. J. Watson Research Center, Yorktown Heights, NY, USA, on a three-month internship. His research interests include mmWave and RF circuits and systems in silicon, massive mmWave multi-input-multi-output (MIMO) systems and related applications.



**Harish Krishnaswamy** received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 2001, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California (USC), Los Angeles, CA, USA, in 2003 and 2009, respectively.

In 2009, he joined the Electrical Engineering Department, Columbia University, New York, NY, USA, as an Assistant Professor. His research interests broadly span integrated devices, circuits, and systems for a variety of RF and mmWave applications. His current research efforts are focused on silicon-based mmWave PAs, sub-mmWave circuits and systems, and reconfigurable broadband RF transceivers for cognitive and software-defined radio.

Dr. Krishnaswamy serves as a member of the Technical Program Committee (TPC) of several conferences, including the IEEE RFIC Symposium and IEEE VLSI-D. He was the recipient of the IEEE International Solid-State Circuits Conference (ISSCC) Lewis Winner Award for Outstanding Paper in 2007, the Best Thesis in Experimental Research Award from the USC Viterbi School of Engineering in 2009, and the Defense Advanced Research Projects Agency (DARPA) Young Faculty Award in 2011.