Analysis and Design of a Stacked Power Amplifier With Very High Bandwidth

David Fritsche, Robert Wolf, and Frank Ellinger, *Senior Member, IEEE*

***Abstract*—In order to simplify and optimize the design process of stacked amplifiers, this paper presents a novel analytical method to dimension the input network for ideal output behavior. To verify this new structural design process, a fully integrated stacked power amplifier (PA) in 0.25-µm SiGe BiCMOS technology is proposed. The stacked architecture enables broadband matching networks, therefore the designed PA reaches a very high bandwidth of 800 MHz around 2 GHz. At 2 GHz, the small-signal gain is 23.8 dB. The output power in the 1-dB compression point and the saturated output power are 26.2 and 27.3 dBm, leading to a power-added efficiency (PAE) of 34% and 40%, respectively. Using a long-term evolution (LTE) modulated input signal without any predistortion, the amplifier reaches an average output power of 21 dBm and a PAE of 12%, fulfilling the LTE specifications in terms of adjacent channel leakage ratio and error vector magnitude.**

***Index Terms*—HBT, high voltage/high power (HiVP), long-term evolution (LTE), SiGe BiCMOS, stacked power amplifier (PA), voltage doubler.**

I. Introduction

THE CONTINUOUS scaling in modern semiconductor technologies leads to a continuous decrease of the breakdown voltages of the devices. Therefore, the achievable output power of conventional power amplifier (PA) architectures, where a single device provides the output voltage swing, also decreases more and more. To overcome the low breakdown voltage, it is possible to stack several devices. In case of optimal tuning at the input of the stacked devices, it is theoretically possible to obtain a multiple higher output voltage swing. For an unchanged output current swing, the output power and the optimal load impedance are multiple times higher, allowing a high output power, as well as a broadband behavior.

The idea of transistor-stacking is not new. Already in 1985, Ezzeddine *et al.* presented an amplifier with stacked transistors [1]. In this circuit, the transistors are stacked only for dc and not for ac allowing an operation under supply voltages higher than the breakdown voltage of a single device. However, the ac-voltage swing of each stage remains the same. Thus, their inputs and outputs have to be connected to a power divider and power combiner, respectively. Further investigations followed [2], [3]. In 1992, the first amplifier stacked transistors for dc and ac was presented by Shifrin *et al.* [4]. By directly connecting the stacked devices, the output power of the stacked devices can be combined without needing additional elements. The required voltage swing at the inputs of the stacked devices is enabled by capacitors. In 2003, Ezzeddine *et al.* proposed a circuit also based on this structure [5].

In the following years, many research groups presented circuits with stacked transistors based on this structure. A different approach for the implementation of the input network was presented in [6] and [7], where the tuning of the stacked devices is done with transformers. Two other appreciable publications are [8] and [9], where the compensation of output capacitances is investigated. In [9], this is done with parallel inductors, and in [8], with feedback capacitors between the source and drain of the stacked transistors, using the positive voltage gain and the Miller effect of the stacked transistors.

The amplifier presented in this paper consists of a cascode stage with an additional stacked transistor. By doubling the output voltage for the same output power, the output current can be halved and a very high bandwidth of 800 MHz at 2 GHz is reached, allowing an operation in multiple frequency bands. To meet the stringent requirements of the long-term evolution (LTE) standard, the circuit is designed for high linearity. In contrast to most other reported stacked amplifiers, the proposed amplifier is fully integrated, and thus usable without extra off-chip components. For the most important step in the design process, the dimensioning of the optimal input network that drives the stacked transistors, a novel analytical method is presented. Based on small-signal simulations, the optimal values of all input elements can be calculated independently of the specific implementation.

This paper is organized as follows. In Section II, the concept of transistor stacking is introduced. The design process with the novel method to calculate the input network is explained in Section III. In Section IV, the designed circuit is presented. Section V is about the measurement results and the comparison with the state-of-the-art. A short conclusion is given in Section VI.

II. Concept of Transistor Stacking

The performance and the efficiency of a PA have a huge influence on the performance and the efficiency of the whole wireless communication system. Therefore, designing a PA is a challenging task. The most important design goals are preferably high output power and efficiency. To obtain this goal, the system

Manuscript received January 23, 2012; revised June 19, 2012; accepted July 02, 2012. Date of publication August 16, 2012; date of current version September 27, 2012. This work was supported by the CoolBaseStations Project (part of the Leading-Edge Cluster "Cool Silicon"), sponsored by the Federal Ministry of Education and Research (BMBF) within the scope of its Leading-Edge Cluster Competition.

The authors are with the Chair for Circuit Design and Network Theory, Dresden University of Technology, 01062 Dresden, Germany (e-mail: david.fritsche@tu-dresden.de).

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.2012.2209439

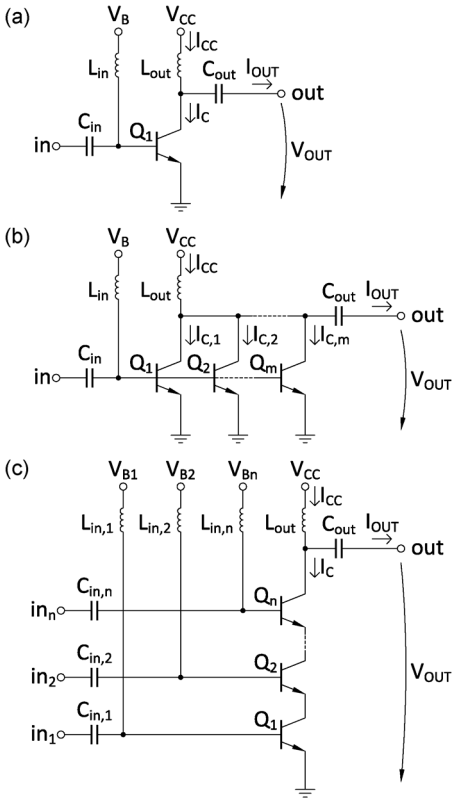

Fig. 1. Amplifier with: (a) one device, (b) m parallel devices, and (c) n stacked devices.

impedance of 50 Ω at the output is transformed to the optimal load impedance that enables the maximum output power. This optimum load impedance allows the maximum voltage swing, as well as the maximum current swing at the output of the PA, and thus maximum output power and maximum efficiency.

For a better understanding, the following investigations assume that voltage and current at the output of the transistor field are equal to the voltage and current at the output of the circuit, respectively. Thus, the optimal load is directly connected to the output of the amplifier. In conventional architectures, the output voltage swing is enabled by a single transistor, as shown in Fig. 1(a). Thus, the maximum possible voltage swing across this transistor is exactly the maximum possible output voltage swing. For this reason, for a given operating-point current, the low breakdown voltage limits the maximum output power.

To obtain higher output power, a typical approach in conventional architectures is to increase the number m of parallel devices. This is shown in Fig. 1(b). By doing that, the output current swing and thus the output power are increased by a factor of m. However, the increase of output power by placing many parallel devices is limited since the optimal load impedance has to be reduced accordingly. For high output power, the optimal load impedance gets very low, which leads to significant losses in the matching network and to low bandwidth.

One possibility to exceed this limit is the stacking of n devices, as shown in Fig. 1(c). By tuning the inputs of the stacked devices in a suitable way, it is theoretically possible to obtain n collector–emitter voltages that are equal in amplitude, as well as in phase. Compared to a single transistor, this leads to an n times higher output voltage, and thus for a given operating-point current to an n times higher output power. As a positive side effect, the optimum load impedance is increased by factor n, providing a basis for a broadband behavior and low matching losses. Table I summarizes the comparison of the three circuits, whereby $\hat{V}_{\mathrm{MAX}}$ and $\hat{I}_{\mathrm{MAX}}$ stand for the amplitudes of voltage and current at maximum excitation, respectively.

TABLE I
COMPARISON OF SINGLE, PARALLEL, AND STACKED CONFIGURATION

Parameter	single device	m parallel devices	n stacked devices
V_{CC}	$V_{CE,OP}$	$V_{CE,OP}$	$n\,V_{CE,OP}$
I_{CC}	$I_{C,OP}$	$m\,I_{C,OP}$	$I_{C,OP}$
$\hat{V}_{OUT,MAX}$	$\hat{V}_{CE,MAX}$	$\hat{V}_{CE,MAX}$	$n\,\hat{V}_{CE,MAX}$
$\hat{I}_{OUT,MAX}$	$\hat{I}_{C,MAX}$	$m\,\hat{I}_{C,MAX}$	$\hat{I}_{C,MAX}$
$R_{OUT,OPT}$	$\frac{\hat{V}_{CE,MAX}}{\hat{I}_{C,MAX}}$	$\frac{1}{m}\cdot\frac{\hat{V}_{CE,MAX}}{\hat{I}_{C,MAX}}$	$n\cdot\frac{\hat{V}_{CE,MAX}}{\hat{I}_{C,MAX}}$
$P_{OUT,MAX}$	$\frac{\hat{V}_{CE,MAX}\,\hat{I}_{C,MAX}}{2}$	$m\cdot\frac{\hat{V}_{CE,MAX}\,\hat{I}_{C,MAX}}{2}$	$n\cdot\frac{\hat{V}_{CE,MAX}\,\hat{I}_{C,MAX}}{2}$

Ideally, the efficiencies of all circuits are equal. Nevertheless, for a given output power, the stacked topology needs only $1/n$ times the current compared to a conventional circuit, leading to less losses in the dc supply path. Furthermore, the increase of the optimal load impedance by factor n^2 for a given output power leads typically to much lower matching losses.

By stacking n devices of m parallel transistors, it is possible to combine the advantages of both approaches. The maximum output power $P_{\mathrm{OUT,MAX}}$ and the optimum load impedance $R_{\mathrm{OUT,OPT}}$ are then determined by

$$P_{\mathrm{OUT,MAX}} = m \cdot n \cdot \frac{\hat{V}_{\mathrm{CE,MAX}}\hat{I}_{C,\mathrm{MAX}}}{2}$$

$$R_{\mathrm{OUT,OPT}} = \frac{n}{m} \cdot \frac{\hat{V}_{\mathrm{CE,MAX}}}{\hat{I}_{C,\mathrm{MAX}}}.$$

The product $m \cdot n$ is proportional to $P_{\mathrm{OUT,MAX}}$. The ratio n/m can be chosen in such a way that $R_{\mathrm{OUT,OPT}}$ is close to 50 Ω. By a proper choice of m and n, it is therefore possible to achieve a high output power, as well as a very broadband output matching network.

As shown in this section, transistor stacking is a promising way to increase the output power and bandwidth of PAs in modern semiconductor technologies since it is possible to obtain a high output voltage swing in spite of low breakdown voltages. Furthermore, the efficiency can be increased by means of lower current and thus lower losses.

III. Design Process

A. Selection of Architecture

The first step of the design process is the selection of the amplifier architecture. The simplest way of stacking transistors to obtain a higher output voltage swing is shown in Fig. 2(a), where $n-1$ transistors are stacked on top of a common emitter stage. In case of optimal tuning of the n input nodes, all n collector–emitter voltages are equal in amplitude and phase, and thus contribute to an n times higher output voltage.

A second possibility is to stack n transistors on top of the cascode stage. This is shown in Fig. 2(b). This architecture

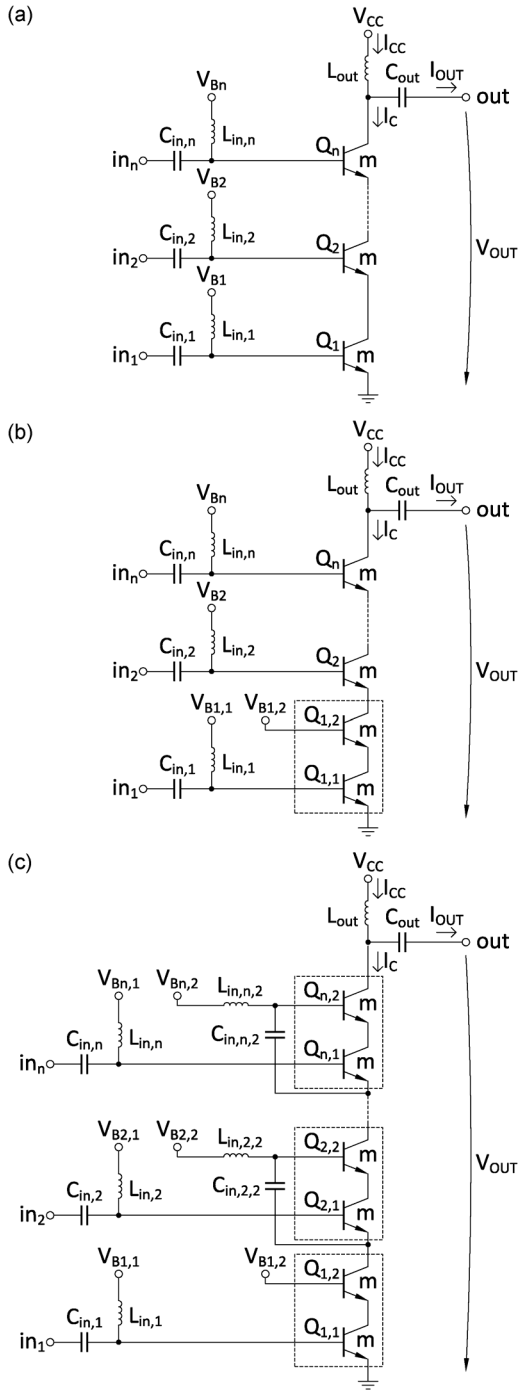

Fig. 2. Amplifier with: (a) stacked single transistors (b) stacked transistors on top of a cascode, and (c) stacked cascode stages.

consists of overall $n + 1$ devices. Nevertheless, only n collector–emitter voltages contribute to the output voltage swing. This is because of the properties of the cascode stage, where the collector–emitter voltage of $Q_{1,1}$ is nearly constant. The usage of a cascode has the benefit of a better decoupling of the bottom input from the output and thus easier input matching, improved stability and higher power gain. The drawback is a lower efficiency, especially for a low number of stacked devices, but since $Q_{1,1}$ does not contribute to the output voltage swing, its collector–emitter voltage can be chosen close to the saturation voltage, reducing the influence on the efficiency.

The third possibility is to stack n cascode stages, as shown in Fig. 2(c). In this case, for adding n collector–emitter voltages, $2n$ stacked devices are needed. In this case, every stacked stage adds a dc-voltage drop, which leads to a significant decrease of the efficiency.

Altogether, there are three possible architectures of stacking transistors to get a multiple output voltage. The highest efficiency can be reached with the architecture in Fig. 2(a) because all collector–emitter voltages contribute to the output voltage swing. Nevertheless, replacing the bottom transistor with a cascode stage has especially the advantage of an improved decoupling between input and output and higher stability. Which of these possibilities is the best choice has to be decided case-by-case.

B. Choice of the DC Operating Point

To meet the stringent linearity requirements of the LTE standard, an ideally distortionless amplification of the input signal is desired. Most suitable for this is class-A or class-AB operation mode of the amplifier. Therefore, the collector–emitter voltage and the collector current of every device have to be adjusted for highest possible symmetrical voltage and current swing.

To obtain high output power, the collector–emitter voltage of devices that contribute to the output voltage swing should be as high as possible. This enables, on the one hand, the highest possible output voltage swing, and thus output power, and on the other hand, the highest possible efficiency because the ratio of saturation voltage and operating point voltage gets as low as possible. The collector–emitter voltage of devices that do not contribute to the output voltage swing should be as low as possible in order not to decrease the efficiency more than necessary. Typically, a voltage of several hundred millivolts above the saturation voltage is enough.

Selecting the collector current per device is a tradeoff between efficiency and bandwidth. A low collector current per transistor leads to low saturation voltage, and thus to a higher efficiency. Since the total collector current is given by the output power, a low collector current per device results in a high number of parallel devices. Simulations showed that a higher number of parallel devices reduces the input impedance, and thus the bandwidth of the matching networks. In the used technology, a collector current between 0.25 and 0.5 mA per device is a good tradeoff.

Independent of the selected architecture, the voltage at the base of the bottom transistor mainly determines the collector current through all stacked devices. Due to the high sensitivity of the collector current to the baser–emitter voltage, the collector current should be adjusted by a current mirror. By tuning the other bias voltages, it is possible to adjust each collector–emitter voltage in the operating point.

C. Determining the Optimal Input Network

The most important step in the design process of a PA with stacked transistors is the design of the input network. The input network has the task to tune the stacked transistors in such a way that their collector–emitter voltages are ideally equal in amplitude and phase.

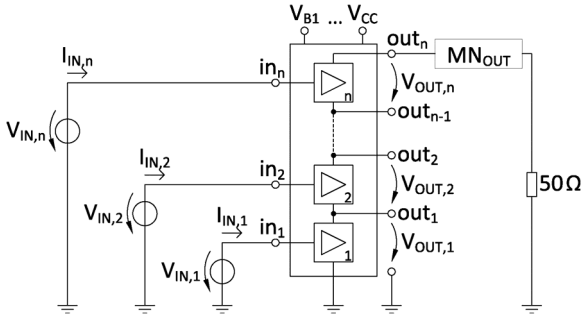

Fig. 3. Circuit with stacked amplifiers with ideal voltage sources at the inputs and matched output.

To find a suited input network and to dimension it, it is necessary to further investigate the circuit. As shown in Fig. 3, the circuit can be represented by a multiport network. The n inputs of the multiport network are the inputs of the n stacked amplifiers, where a stacked amplifier is either a single transistor or a cascode stage. The n outputs of the multiport network are connected to the outputs of the n stacked amplifiers. The output out_n is connected to an output matching network, which transforms the 50-Ω system impedance to $\underline{Z}_{\mathrm{OUT,OPT}}$. To investigate the circuit in the desired operating point, the correct bias voltages have to be applied.

The first step is to find the dependencies between the n output voltages and the n input voltages with small signal simulations. For this purpose, only one input voltage source is activated and there influence on the n output voltages is determined. After n simulations, it is possible to arrange the n transfer functions as a complex matrix, shown at the bottom of this page. With this matrix it is easily possible to calculate the output voltages $\underline{V}_{\mathrm{OUT},1}$ to $\underline{V}_{\mathrm{OUT},n}$ for any input voltages $\underline{V}_{\mathrm{IN},1}$ to $\underline{V}_{\mathrm{IN},n}$

$$\begin{pmatrix} \underline{V}_{\mathrm{OUT},n} \\ \vdots \\ \underline{V}_{\mathrm{OUT},2} \\ \underline{V}_{\mathrm{OUT},1} \end{pmatrix} = \underline{M} \cdot \begin{pmatrix} \underline{V}_{\mathrm{IN},n} \\ \vdots \\ \underline{V}_{\mathrm{IN},2} \\ \underline{V}_{\mathrm{IN},1} \end{pmatrix}.$$

Inverting the matrix for desired output voltages the required input voltages can be calculated

$$\begin{pmatrix} \underline{V}_{\mathrm{IN},n} \\ \vdots \\ \underline{V}_{\mathrm{IN},2} \\ \underline{V}_{\mathrm{IN},1} \end{pmatrix} = \underline{M}^{-1} \cdot \begin{pmatrix} \underline{V}_{\mathrm{OUT},n} \\ \vdots \\ \underline{V}_{\mathrm{OUT},2} \\ \underline{V}_{\mathrm{OUT},1} \end{pmatrix}.$$

Since the goal of an amplifier with stacked transistors are output voltages that are equal in phase, the complex output voltages can be replaced by their real amplitudes

$$\begin{pmatrix} \underline{V}_{\mathrm{IN},n,\mathrm{OPT}} \\ \vdots \\ \underline{V}_{\mathrm{IN},2,\mathrm{OPT}} \\ \underline{V}_{\mathrm{IN},1,\mathrm{OPT}} \end{pmatrix} = \frac{1}{\sqrt{2}} \cdot \underline{M}^{-1} \cdot \begin{pmatrix} \hat{V}_{\mathrm{CE,MAX}} \\ \vdots \\ \hat{V}_{\mathrm{CE,MAX}} \\ \hat{V}_{\mathrm{CE,MAX}} \end{pmatrix}.$$

By applying the optimal input voltages $\underline{V}_{\mathrm{IN},1,\mathrm{OPT}}$ to $\underline{V}_{\mathrm{IN},n,\mathrm{OPT}}$ by the ideal voltage sources to the circuit, the desired in-phase output voltages with the maximum amplitude $\hat{V}_{\mathrm{CE,MAX}}$ can be obtained.

Equivalent to $\underline{M}$, another matrix $\underline{N}$ can be determined that contains the dependencies between the input voltages and the input currents, shown in the equation at the bottom of this page. With the optimal input voltages, it is possible to calculate the flowing input currents under the optimal output conditions

$$\begin{aligned} \begin{pmatrix} \underline{I}_{\mathrm{IN},n,\mathrm{OPT}} \\ \vdots \\ \underline{I}_{\mathrm{IN},2,\mathrm{OPT}} \\ \underline{I}_{\mathrm{IN},1,\mathrm{OPT}} \end{pmatrix} &= \underline{N} \cdot \begin{pmatrix} \underline{V}_{\mathrm{IN},n,\mathrm{OPT}} \\ \vdots \\ \underline{V}_{\mathrm{IN},2,\mathrm{OPT}} \\ \underline{V}_{\mathrm{IN},1,\mathrm{OPT}} \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \cdot \underline{N} \cdot \underline{M}^{-1} \cdot \begin{pmatrix} \hat{V}_{\mathrm{CE,MAX}} \\ \vdots \\ \hat{V}_{\mathrm{CE,MAX}} \\ \hat{V}_{\mathrm{CE,MAX}} \end{pmatrix}. \end{aligned}$$

$$\underline{M} = \begin{pmatrix} \left.\frac{\underline{V}_{\mathrm{OUT},n}}{\underline{V}_{\mathrm{IN},n}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq n)}} & \cdots & \left.\frac{\underline{V}_{\mathrm{OUT},n}}{\underline{V}_{\mathrm{IN},2}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 2)}} & \left.\frac{\underline{V}_{\mathrm{OUT},n}}{\underline{V}_{\mathrm{IN},1}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 1)}} \\ \vdots & \ddots & \vdots & \vdots \\ \left.\frac{\underline{V}_{\mathrm{OUT},2}}{\underline{V}_{\mathrm{IN},n}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq n)}} & \cdots & \left.\frac{\underline{V}_{\mathrm{OUT},2}}{\underline{V}_{\mathrm{IN},2}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 2)}} & \left.\frac{\underline{V}_{\mathrm{OUT},2}}{\underline{V}_{\mathrm{IN},1}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 1)}} \\ \left.\frac{\underline{V}_{\mathrm{OUT},1}}{\underline{V}_{\mathrm{IN},n}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq n)}} & \cdots & \left.\frac{\underline{V}_{\mathrm{OUT},1}}{\underline{V}_{\mathrm{IN},2}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 2)}} & \left.\frac{\underline{V}_{\mathrm{OUT},1}}{\underline{V}_{\mathrm{IN},1}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 1)}} \end{pmatrix}$$

$$\underline{N} = \begin{pmatrix} \left.\frac{\underline{I}_{\mathrm{IN},n}}{\underline{V}_{\mathrm{IN},n}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq n)}} & \cdots & \left.\frac{\underline{I}_{\mathrm{IN},n}}{\underline{V}_{\mathrm{IN},2}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 2)}} & \left.\frac{\underline{I}_{\mathrm{IN},n}}{\underline{V}_{\mathrm{IN},1}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 1)}} \\ \vdots & \ddots & \vdots & \vdots \\ \left.\frac{\underline{I}_{\mathrm{IN},2}}{\underline{V}_{\mathrm{IN},n}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq n)}} & \cdots & \left.\frac{\underline{I}_{\mathrm{IN},2}}{\underline{V}_{\mathrm{IN},2}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 2)}} & \left.\frac{\underline{I}_{\mathrm{IN},2}}{\underline{V}_{\mathrm{IN},1}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 1)}} \\ \left.\frac{\underline{I}_{\mathrm{IN},1}}{\underline{V}_{\mathrm{IN},n}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq n)}} & \cdots & \left.\frac{\underline{I}_{\mathrm{IN},1}}{\underline{V}_{\mathrm{IN},2}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 2)}} & \left.\frac{\underline{I}_{\mathrm{IN},1}}{\underline{V}_{\mathrm{IN},1}}\right|_{\underline{V}_{\mathrm{IN},i=0(\forall i\neq 1)}} \end{pmatrix}, \qquad i \in [1,n]$$

Knowing the input currents, it is possible to analyze the input behavior of the circuit. For that purpose, the input impedances are

$$\underline{Z}_{\mathrm{IN},i,\mathrm{OPT}} = \frac{\underline{V}_{\mathrm{IN},i,\mathrm{OPT}}}{\underline{I}_{\mathrm{IN},i,\mathrm{OPT}}} = R_{\mathrm{IN},i,\mathrm{OPT}} + j \cdot X_{\mathrm{IN},i,\mathrm{OPT}}.$$

Typically, these input impedances under optimal input and output conditions are very different between the input in_1 and the inputs in_2 to in_n.

- The real part of $\underline{Z}_{\mathrm{IN},1,\mathrm{OPT}}$ is positive and the imaginary part negative. The magnitudes of the real and imaginary parts are in the same order. This means that the circuit consumes real power and that the reactive power is in the same order.
- The input impedances $\underline{Z}_{\mathrm{IN},2,\mathrm{OPT}}$ to $\underline{Z}_{\mathrm{IN},n,\mathrm{OPT}}$ are, in contrast to $\underline{Z}_{\mathrm{IN},1,\mathrm{OPT}}$, very inductive. This means, that the reactive power is multiple times higher than the consumed or delivered real power. Whether real power is consumed or delivered depends on the concrete circuit.

The next step is to find a suitable input network that replaces the ideal voltage sources by implementable elements without affecting the input and output behavior of the circuit. To provide real power to the input in_1, $\underline{Z}_{\mathrm{IN},1,\mathrm{OPT}}$ has to be matched by an input matching network to a 50-Ω source. The dimensioning of a matching network is a common problem, and thus not further described. However, at the inputs in_2 to in_n, the reactive power is much higher than the consumed or delivered real power. Therefore, it is sufficient to compensate the reactive power by a capacitor and not to connect these inputs to a 50-Ω source. The size of the capacitor can be calculated with the real and imaginary part of the input impedances of $\underline{Z}_{\mathrm{IN},i,\mathrm{OPT}} (i \in [2,n])$

$$C_{\mathrm{B},i} = \frac{1}{\omega} \frac{X_{\mathrm{IN},i,\mathrm{OPT}}}{R_{\mathrm{IN},i,\mathrm{OPT}} + X^2_{\mathrm{IN},i,\mathrm{OPT}}}, \qquad i \in [2,n].$$

If the circuit delivers real power at these inputs in the optimal case $(R_{\mathrm{IN},i,\mathrm{OPT}}) < 0)$, it is possible to consume the delivered real power by a parallel resistor

$$R_{B,i} = -\frac{R^2_{\mathrm{IN},i,\mathrm{OPT}} + X^2_{\mathrm{IN},i,\mathrm{OPT}}}{R_{\mathrm{IN},i,\mathrm{OPT}}}, \qquad i \in [2,n].$$

In this case, the ideal behavior can be reached. The resulting circuit is shown in Fig. 4.

If the circuit consumes real power in the ideal case $(R_{\mathrm{IN},i,\mathrm{OPT}} > 0)$ that cannot be provided with this practical network, there is a difference to the ideal behavior. Typically, this difference is very low because of the much higher reactive power compared to the real power. Finally, the inputs in_2 to in_n can be implemented by passive networks so that only in_1 has to be driven by a source.

It is also possible to choose the inputs in Fig. 3 in a different way, which leads to different network implementations with equivalent behavior. For example, it is possible to connect the ideal voltage sources not between the input and ground, but between the inputs and an additional common node, which leads to a network as in [10].

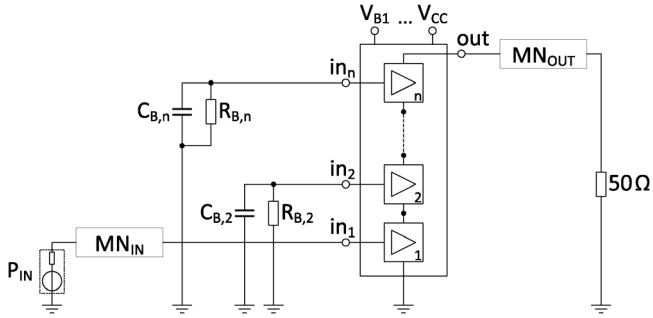

Fig. 4. Circuit with stacked amplifiers with practical input network.

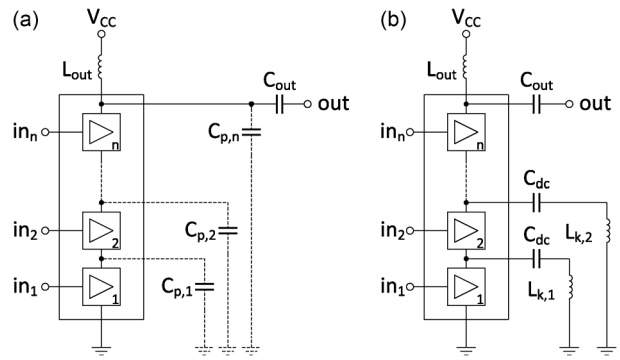

Fig. 5. Circuit with: (a) parasitic capacitances at the output and (b) compensation of parasitic capacitances by additional inductors.

D. Compensating Parasitic Output Capacitances

At high frequencies, the parasitic capacitances at the output have to be considered. Simplified, the parasitic capacitances can be combined as shown in Fig. 5(a) to an equivalent capacitance between the outputs of the stacked amplifiers and ground. These parasitic capacitances lead to the circumstance that every stacked amplifier sees a complex load impedance instead of a purely resistive impedance. Due to that, voltage and current at the intrinsic transistor get out of phase, leading to a reduced output power and efficiency. To avoid this, the parasitic capacitances should be compensated in a suited way.

The capacitance $C_{\mathrm{p},n}$ can be compensated by the inductor L_{OUT} that is needed anyway for biasing and perhaps for output matching. However, in conventional architectures, it is not possible to compensate the remaining capacitances $C_{\mathrm{p},1}$ to $C_{\mathrm{p},n-1}$. Nevertheless, a compensation of these capacitances can be very advantageous. For these reasons, $n-1$ dc-decoupled inductors $L_{\mathrm{k},1}$ to $L_{\mathrm{k},n-1}$ are added. This is shown in Fig. 5(b) and also proposed in [9]. In differential circuits, the dc-decoupling capacitors can be circumvented. The additional inductors can offer a much higher output power and efficiency, provided that they have a high quality factor. Otherwise, the resistive losses dominate, leading to an evanescent effect of the compensation of the parasitic capacitances.

An alternative possibility is described in [8], where between the inputs and outputs of the stacked amplifiers additional capacitors are applied. Due to the positive voltage gain of the stacked amplifiers and the Miller effect, the previous stage sees a negative capacitance. In an optimal case, this negative load capacitance compensates the output capacitance of the previous stage. The advantage of this concept is that capacitors are much smaller than inductors. On the other hand, the disadvantage is

that the added capacitance has an inductive behavior only for the previous stage. With other words, the added capacitance increases the output capacitance of the current stage. Thus, the next stage has to compensate a higher capacitance. This circumstance leads to a decrease of the bandwidth. Due to the desired high bandwidth, the approach with true inductors was chosen.

IV. Designed Circuit

The circuit in Fig. 6 was designed based on the described design process and fabricated in 0.25-µm BiCMOS IHP SGB25V technology. The used transistors provide a maximum transit frequency of 50 GHz and a maximum frequency of oscillation of 90 GHz. A pseudo-differential structure was chosen to reduce the effect of parasitic elements. For stability reasons, the circuit of Fig. 2(b) with a cascode stage as a bottom stage is used, having the advantage of a better isolation between input and output compared to a common emitter stage. On top of a cascode stage, one additional transistor is stacked, leading to overall three stacked transistors. Two transistors contribute to the output voltage swing. Thus, the circuit can be called an output voltage doubler.

The operating points and the tuning ranges of the stacked transistors are given in Table II. The collector–emitter voltage of $Q_{1,1\mathrm{p/n}}$ is 0.7 V, and therefore slightly above the saturation voltage. The collector–emitter voltages of $Q_{1,2\mathrm{p/n}}$ and $Q_{2\mathrm{p/n}}$ are 3.5 V to allow a high voltage swing. To decrease their saturation voltage, the number of parallel devices of the transistors $Q_{1,2\mathrm{p/n}}$ and $Q_{2\mathrm{p/n}}$ is doubled compared to $Q_{1,1\mathrm{p/n}}$. The total collector current $I_{CC+/-}$ is 60 mA in the operating point and 80 mA in a 1-dB compression point. This behavior can be achieved by proper choice of the two resistors (10 and 200 Ω) of the current mirror. A deeper operation in class-AB would be disadvantageous regarding the stringent linearity requirements of the LTE standard.

Assuming an ideal voltage doubler behavior, the optimal load impedance is determined by

$$R_{\mathrm{OUT,OPT}} = \frac{n}{m}\frac{\hat{V}_{\mathrm{CE,MAX}}}{\hat{I}_{C,\mathrm{MAX}}} = \frac{2}{160}\frac{3.2\ \mathrm{V}}{0.5\ \mathrm{mA}} = 80\ \Omega.$$

The output matching network transforms the 50-Ω system impedance to 80 Ω and compensates parasitic capacitances at the collector nodes of the transistors $Q_{2\mathrm{p/n}}$. Taking the losses in the inductors into account, it consists of a differential inductor $L_{\mathrm{OUTp/n}} = 3$ nH and two capacitors $C_{\mathrm{OUTp/n}} = 1.7$ pF. To compensate the parasitic capacitance between the cascode stage and the additional stacked transistor, an inductor $L_{K1} = 7.6$ nH is added.

For this configuration, the optimal input network can be determined as described in Section III. The bases of the bottom transistors $Q_{1,1\mathrm{p/n}}$ are matched to 50 Ω by an input matching network of $L_{11} = 6.6$ nH and $C_{21\mathrm{p/n}} = 3.3$ pF. To increase the bandwidth, the series 15-Ω resistors are added. The 10-Ω resistor, which is needed for biasing, contributes to resistive matching and thus also higher bandwidth. The bases of the top transistor $Q_{2\mathrm{p/n}}$ are connected to a 260-fF capacitor and a 2-kΩ resistor, respectively. The capacitor compensates the reactive power. The resistor consumes the delivered real power and is used for biasing.

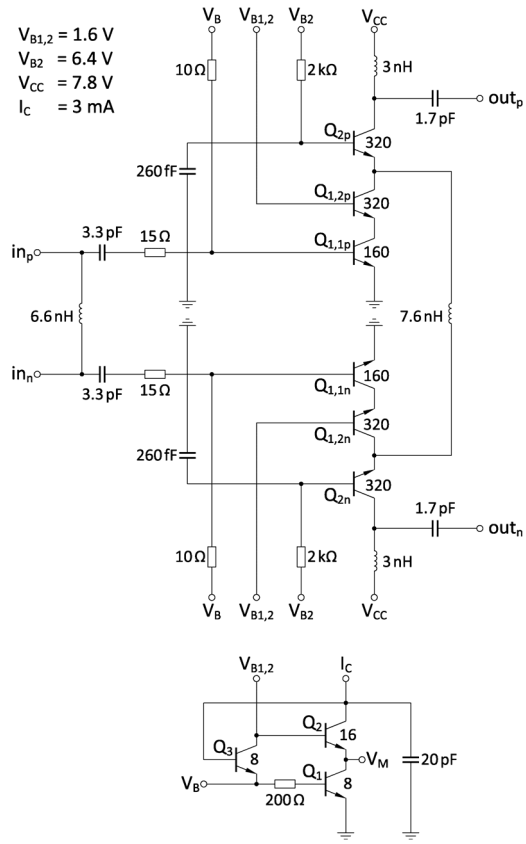

Fig. 6. Final circuit schematic.

TABLE II
Operating Point and Tuning Range of the Stacked Transistors

Parameter	$Q_{1,1\mathrm{p}}$ and $Q_{1,1\mathrm{n}}$	$Q_{1,2\mathrm{p}}$ and $Q_{1,2\mathrm{n}}$	$Q_{2\mathrm{p}}$ and $Q_{2\mathrm{n}}$
$V_{\mathrm{CE,OP}}$	0.7 V	3.5 V	3.5 V
$I_{\mathrm{C,OP}}$	0.38 mA	0.19 mA	0.19 mA
$I_{\mathrm{C,1dB}}$	0.5 mA	0.25 mA	0.25 mA
$V_{\mathrm{CE,SAT}}$	0.5 V	0.3 V	0.3 V
$\hat{V}_{\mathrm{CE,MAX}}$	-	3.2 V	3.2 V
$\hat{I}_{\mathrm{C,MAX}}$	0.5 mA	0.25 mA	0.25 mA
Number	160	320	320

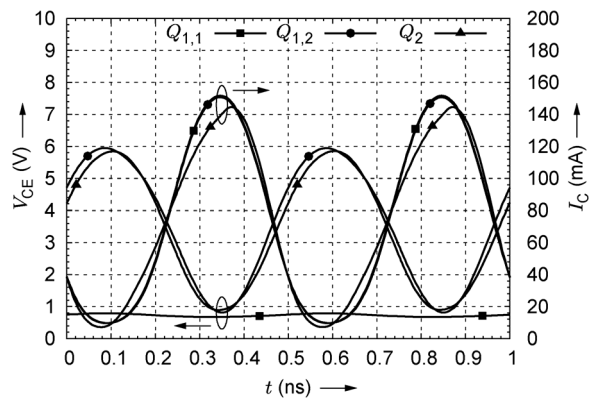

Fig. 7. Transient simulation results.

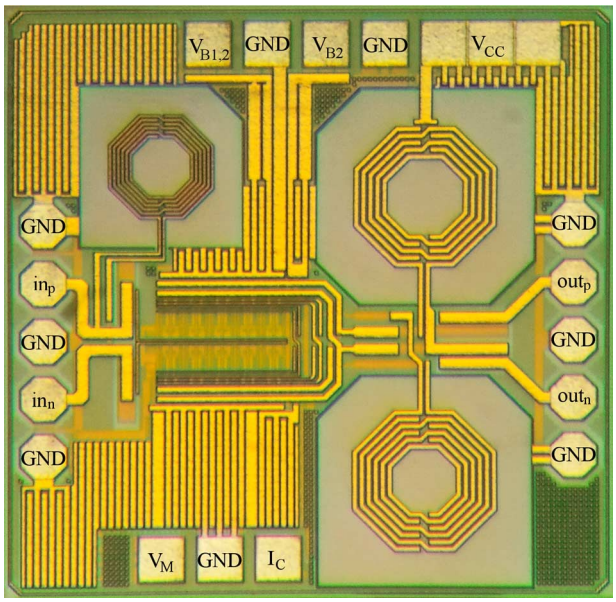

Fig. 8. Photograph of the fabricated chip.

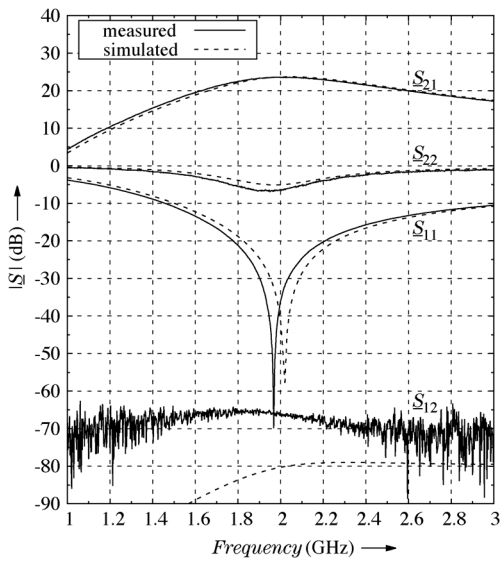

Fig. 9. Small-signal measurement and simulation results.

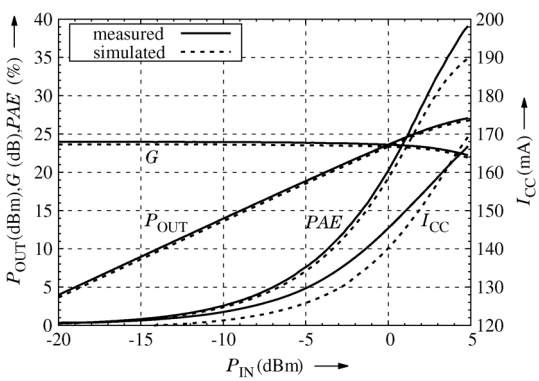

Fig. 10. Large-signal measurement and simulation results.

To prove the validity of the small-signal simulation-based design process for large-signal excitation, Fig. 7 depicts the simulated collector–emitter voltages and intrinsic collector currents of the three stacked transistors for a power level of 2 dB

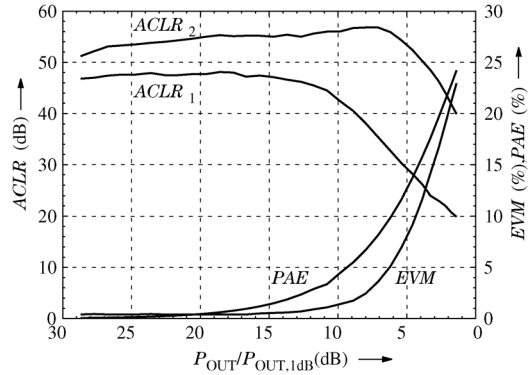

Fig. 11. Linearity measurement results.

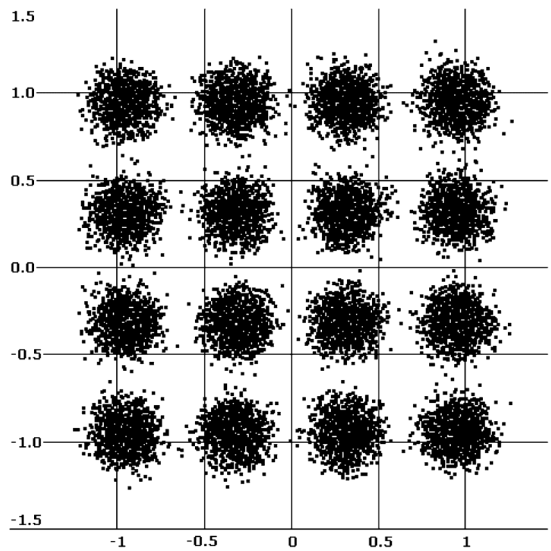

Fig. 12. QAM-16 constellation diagram for 3.5-dB back-off.

below the 1-dB compression point. As intended, the deviations of amplitude and phase between the collector–emitter voltages of $Q_{1,2}$ and Q_2 and between the intrinsic collector currents of $Q_{1,1}$, $Q_{1,2}$, and Q_2 are very low.

A chip photograph is shown in Fig. 8. The chip has a total size of 1 mm $\times$ 1 mm.

V. Measurement Results

A. Continuous-Wave Measurements

The manufactured circuit was measured on-wafer with the network analyzer Rohde & Schwarz ZVA67. The small-signal behavior is shown in Fig. 9. $|\underline{S}_{11}|$ is below -15 dB from 1.7 to 2.5 GHz. The small-signal gain $|\underline{S}_{21}|$ has a maximum value of 23.8 dB at 2.0 GHz. In the range from 1.7 to 2.5 GHz, $|\underline{S}_{21}|$ drops off less than 3 dB from its maximum value. Referring to the behavior of $|\underline{S}_{11}|$ and $|\underline{S}_{21}|$, the bandwidth of the circuit is about 800 MHz. In the whole measured range from 1 to 3 GHz $|\underline{S}_{22}|$ is less than 0 dB and $|\underline{S}_{12}|$ is less than -60 dB. The circuit is unconditionally stable for differential- and common-mode excitation.

Output power P_{OUT}, power gain G, and power-added efficiency (PAE) at 2.0 GHz are plotted versus input power P_{IN} in Fig. 10. The 1-dB compression point is at $P_{\mathrm{IN,1\,dB}} = 3.4$ dBm and $P_{\mathrm{OUT,1\,dB}} = 26.2$ dBm. The PAE in this point is 34%. The

TABLE III
COMPARISON WITH STATE-OF-THE-ART

	Design	Technology	Description	G_0 (dB)	$P_{OUT,SAT}$ (dBm)	PAE_{SAT} (%)	$P_{OUT,1dB}$ (dBm)	PAE_{1dB} (%)	V_{DD} (V)	f (GHz)	BW (GHz)	chip size (mm^2)
external devices needed	[9]	65 nm CMOS	cascode with stacked transistor	11.6	30.1	21.2	n. s.	n. s.	4.6	6.5	n. s.	2.9
	[12]	0.25 µm SOS CMOS	three stacked transistors	17.1	21	44	18.3	29.2	3.9	1.9	0.3	0.6
	[13]	65 nm CMOS	two stacked cascode stages	27	31	60.5	28.0	39	3.6	0.9	n. s.	0.6
	[14]	0.13 µm CMOS	three stacked transistors	11.5	29.5	34.5	27.0	20	3.6	0.9	n. s.	1.0
	[15]	2 µm GaAs HBT	two stacked transistors	15	25	42	24.2	36.4	8	5	0.6	0.4
	[16]	0.13 µm SOI CMOS	four stacked transistors	14.6	32.4	47	30.8	46.1	6.5	1.9	n. s.	0.7
	[17]	65 nm CMOS	two stacked cascode stages with cascode driver	34	29.4	51	25.4	37.9	3.4	1.8	n. s.	n. s.
integrated	[7]	2 µm HBT	four stacked transistors	13.1	31	37	26.0	23	15	4.8	0.8	2.0
	[11]	0.5 µm GaAs PHEMT	three stacked transistors	15	22.6	19.5	22.2	17.7	12	5.2	3.5*	1.5
	This	0.25 µm SiGe HBT	cascode with stacked transistor	23.8	27.3	40	26.2	34	7.8	2.0	0.8	1.0

fully integrated: [7], [11]; output matching off-chip: [12], [15], [17]; input and output matching off-chip: [9], [13], [14], [16], * input matching not specified

output power under saturation conditions is about 27.3 dBm. The PAE is in that case about 40%.

B. LTE Measurements

With the vector signal generator Rohde & Schwarz SMBV100A and the vector signal analyzer Rohde & Schwarz FSV7, it was possible to measure the behavior of the circuit with an LTE input signal. No predistortion has been applied. For a center frequency of 2 GHz and channel bandwidth of 20 MHz, the adjacent channel leakage ratio (ACLR), error vector magnitude (EVM), and PAE are plotted versus average back-off output power $P_{\mathrm{OUT}}/P_{\mathrm{OUT,1\,dB}}$ in Fig. 11. The limit of the ACLR in the LTE specification for user equipment is 30 dB. ACLR_1 is below 30 dB for the output power range more than 5.2 dB in back-off and ACLR_2 is below 30 dB for the complete back-off range. The limits for EVM are 17.5% for quadrature phase-shift keying (QPSK), 12.5% for quadrature amplitude modulation 16 (QAM-16), and 8% for quadrature amplitude modulation 64 (QAM-64). These limits are met for more than 2.5-, 3.5-, and 5-dB back-off, respectively. Fig. 12 shows the constellation diagram for a QAM-16 modulated signal for 3.5-dB back-off. The maximum average output power to meet ACLR, as well as EVM specification of LTE is 20.8 dBm, leading to a PAE of 12%.

Table III compares the presented measurement results with other presented designs of stacked amplifiers. Except for [7], where transformers are used to tune the stacked amplifiers, the voltage swing at the inputs of the upper stacked devices is enabled by capacitors. Only [7] and [11] are fully integrated designs. In the other designs, at least the output matching networks are off-chip. Since off-chip elements have typically a much higher quality factor, a fair comparison of the efficiency is only possible with [7] and [11].

Compared to other designs, the small-signal gain of 23.8 dB is very high. The linearity of the circuit is also very high, as one can see in the low difference between the saturated output power (27.3 dBm) and the output power in the 1-dB compression point (26.2 dBm). The output power in the 1-dB compression point is comparable to most of the other designs. The PAE in the 1-dB compression point is 34%. This is much higher than the efficiency of the other fully integrated circuits. The total chip size is very small, taking into account that no additional off-chip elements are needed. For bandwidth comparison, the 3-dB bandwidth of $|\underline{S}_{21}|$ and the range of $|\underline{S}_{11}| < -15$ dB was considered. Normalized to the operation frequency, the presented circuit has a very high bandwidth of 40%, which is only exceeded by [11]. This is mainly caused by the optimal load impedance close to 50 Ω and the reduced operating point current and transistor widths, showing impressively the potential of stacked PAs.

VI. CONCLUSION

In this paper, the design process of stacked amplifiers is described and optimized. A novel way to analytically dimension the input network is proposed that allows calculating the values of all needed elements. Based on this, a fully integrated circuit is presented showing a very high bandwidth, high gain, high efficiency, and small chip size, especially compared to the other fully integrated circuits.

REFERENCES

[1] A. Ezzeddine, H. L. A. Hung, and H. C. Huang, "High-voltage FET amplifiers for satellite and phased-array applications," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1985, pp. 336–339.

[2] K. E. Peterson, H. L. Hung, F. R. Phelleps, T. F. Noble, and H. C. Huang, "Monolithic high-voltage FET power amplifiers," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 1985, vol. 943, pp. 945–948.

[3] K. E. Peterson, H. L. A. Hung, F. R. Phelleps, E. Y. Chang, J. L. Singer, H. E. Carlson, and A. B. Cornfield, "30-V MMIC power amplifier with novel bias circuitry," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Jul. 1991, vol. 822, pp. 823–826.

[4] M. Shifrin, Y. Ayasli, and P. Katzin, "A new power amplifier topology with series biasing and power combining of transistors," in *Proc. IEEE Microw. Millim.-Wave Monolithic Circuits Symp.*, Jun. 1992, pp. 39–41.

[5] A. K. Ezzeddine and H. C. Huang, "The high voltage/high power FET (HiVP)," in *Proc. IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2003, pp. 215–218.

[6] J. G. McRory, G. G. Rabjohn, and R. H. Johnston, "Transformer coupled stacked FET power amplifiers," *IEEE J. Solid-State Circuits*, vol. 34, no. 2, pp. 157–161, Feb. 1999.

[7] M.-F. Lei, Z.-M. Tsai, K.-Y. Lin, and H. Wang, "Design and analysis of stacked power amplifier in series-input and series-output configuration," *IEEE Trans. Microw. Theory Tech.*, vol. 55, no. 12, pp. 2802–2812, Dec. 2007.

[8] A. K. Ezzeddine, H. C. Huang, and J. L. Singer, "UHiFET—A new high-frequency high-voltage device," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2011, pp. 1–4.
[9] M. Fathi, D. K. Su, and B. A. Wooley, "A stacked 6.5-GHz 29.6-dBm power amplifier in standard 65-nm CMOS," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2010, pp. 1–4.
[10] Y. Luque, N. Deltimple, E. Kerherve, and D. Belot, "A 65 nm CMOS fully integrated 31.5 dBm triple SFDS power amplifier dedicated to W-CDMA application," in *Proc. IEEE Int. Electron., Circuits, Syst. Conf.*, Dec. 2010, pp. 595–598.
[11] C.-C. Shen, H.-Y. Chang, and G. D. Vendelin, "Comparison of enhancement- and depletion-mode triple stacked power amplifiers in 0.5 μm AlGaAs/GaAs PHEMT technology," in *Proc. Eur. Microw. Integr. Circuits Conf.*, Sep. 2009, pp. 222–225.
[12] J. Jeong, S. Pornpromlikit, P. M. Asbeck, and D. Kelly, "A 20 dBm linear RF power amplifier using stacked silicon-on-sapphire MOSFETs," *IEEE Microw. Wireless Compon. Lett.*, vol. 16, no. 12, pp. 684–686, Dec. 2006.
[13] S. Leuschner, S. Pinarello, U. Hodel, J. E. Mueller, and H. Klar, "A 31-dBm, high ruggedness power amplifier in 65-nm standard CMOS with high-efficiency stacked-cascode stages," in *Proc. IEEE Radio Freq. Integr. Circuits Symp.*, May 2010, pp. 395–398.
[14] L. Wu, I. Dettmann, and M. Berroth, "A 900-MHz 29.5-dBm 0.13-μm CMOS HiVP power amplifier," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 9, pp. 2040–2045, Sep. 2008.
[15] C.-C. Shen, F.-H. Huang, C.-K. Lin, H.-Y. Chang, Y.-J. Chan, and Y.-C. Wang, "A broadband stacked power amplifier using 2-μm GaAs HBT process for *C*-band applications," in *Proc. Asia–Pacific Microw. Conf.*, Dec. 2008, pp. 1–4.
[16] S. Pornpromlikit, J. Jeong, C. D. Presti, A. Scuderi, and P. M. Asbeck, "A watt-level stacked-FET linear power amplifier in silicon-on-insulator CMOS," *IEEE Trans. Microw. Theory Tech.*, vol. 58, no. 1, pp. 57–64, Jan. 2010.
[17] S. Leuschner, J. E. Mueller, and H. Klar, "A 1.8 GHz wideband stacked-cascode CMOS power amplifier for WCDMA applications in 65 nm standard CMOS," in *Proc. IEEE Radio Freq. Integr. Circuits Symp.*, Jun. 2011, pp. 1–4.

David Fritsche was born in Bautzen, Germany, in 1986. He received the Masters degree in electrical engineering from the Technische Universität Dresden (TUD), Dresden, Germany, in 2011, and is currently working toward the Ph.D. degree at TUD.

His main research interest is the design of PAs and circuits for operation at sub-terahertz frequencies.

Robert Wolf was born in Karl-Marx-Stadt (now Chemnitz), Germany, in 1984. He received the Masters degree in electrical engineering from the Technische Universität Dresden (TUD), Dresden, Germany, in 2009, and is currently working toward the Ph.D. degree at TUD.

His main research interests include system analysis and the design of integrated control systems for efficiency enhancement of RF PAs.

Frank Ellinger (S'97–M'01–SM'06) was born in Friedrichshafen, Germany, in 1972. He received the Diploma degree in electrical engineering from the University of Ulm, Ulm, Germany, in 1996, and the MBA and Ph.D. degree in electrical engineering and Habilitation degree in high-frequency circuit design from ETH Zürich (ETHZ), Zürich, Switzerland, in 2001 and 2004, respectively.

Since August 2006, he has been a Full Professor and Head of the Chairfor Circuit Design and Network Theory, Technische Universität Dresden (TUD), Dresden, Germany. From 2001 to 2006, he was Head of the RF Integrated Circuit (RFIC) Design Group, Electronics Laboratory, ETHZ, and a Project Leader of the IBM/ETHZ Competence Center for Advanced Silicon Electronics hosted by IBM Research, Rüschlikon, Switzerland. He has been the Coordinator of the RESOLUTION, MIMAX, and FLEXIBILITY projects funded by the European Union. He has authored or coauthored over 170 refereed scientific papers. He authored the lecture book *Radio Frequency Integrated Circuits and Technologies* (Springer, 2008).

Prof. Ellinger is an elected IEEE Microwave Theory and Techniques Society (MTT-S) Distinguished Microwave Lecturer (2009–2011). He was the recipient of several awards including the IEEE Outstanding Young Engineer Award, the ETH Medal, the Denzler Award, and a Young Ph.D. Award of ETHZ. He was a two-time recipient of the Rohde & Schwarz/Agilent/Gerotron EEEf-COM Innovation Award.