A 21–26-GHz SiGe Bipolar Power Amplifier MMIC

Tak Shun Dickson Cheung, Member, IEEE, and John R. Long, Member, IEEE

Abstract—A three-stage 21–26-GHz medium-power amplifier fabricated in $f_T = 120$ GHz 0.2 $\mu$m SiGe HBT technology has 19 dB small-signal gain and 15 dB gain at maximum output power. It delivers 23 dBm, 19.75% PAE at 22 GHz, and 21 dBm, 13% PAE at 24 GHz. The differential common-base topology extends the supply to $B_{V_{CEO}}$ of the transistors (1.8 V). New on-chip components, such as on-chip interconnects with floating differential shields, and self-shielding four-way power combining/dividing baluns provide inter-stage coupling and single-ended I/O interfaces at the input and output. The $2.45 \times 2.45$ mm$^2$ MMIC was mounted as a flipchip and tested without a heatsink.

Index Terms—Balun, microwave power amplifiers, millimeter-wave power amplifiers, MMIC power amplifiers, monolithic transformer, SiGe, SiGe power amplifiers, silicon-germanium.

I. INTRODUCTION

Wireless communication beyond 20 GHz offers several advantages when compared to the 1–6-GHz regime. The reduction in wavelength results in at least 3× improvement in angular resolution (which is proportional to wavelength), thereby allowing electromagnetic waves reflected from objects to be used for remote sensing applications. New applications for wireless consumer electronics, such as automotive radars that operate in the spectrum from 22 to 29 GHz (assigned by the FCC [1]) or the 21.65–26.65 GHz1 band (allocated by the ETSI in Europe [2]), are feasible with the resolution afforded by short (almost millimeter) wavelengths, and will be migrated to the 76–77-GHz band in North America [3] and to 77–81 GHz in Europe [4] when suitable technologies become available. Wireless local-area networks (WLANs) such as 2.4–2.48-GHz IEEE 802.11b/g could also benefit from up-bandaging to the 24–24.25-GHz ISM band. The 3× increase in channel bandwidth at 24 GHz implies a threefold improvement in data capacity. A 24-GHz antenna is approximately 1/10 the size of a 2.4-GHz radiator, making compact, multiple antenna arrays practical. A directional antenna at 24 GHz can transmit 20 dB more signal compared to omni-directional radiators [5]–[7], giving 24-GHz point-to-point WLANs greater immunity to interference. The electronic beam steering or beam forming capabilities of short wavelength array antennas will simplify their deployment in practical WLAN applications [8].

Commercial viability of consumer products such as WLAN or automotive radar hinge upon their implementation in silicon CMOS or BiCMOS technologies, because they offer the potential for low manufacturing cost due to their inherent economies of scale. Device transit frequencies (i.e., $f_T$) in the 100–200-GHz range now extend the reach of these technologies to the 24-GHz band for exploratory circuit development. Although current CMOS and SiGe bipolar devices have excelled in small-signal (e.g., output power $< -10$ dBm) radio frequency (RF) circuits [9]–[14], silicon-based power amplifiers in this frequency range have rarely delivered power above 10–20 dBm with a power-added efficiency (PAE) above 15% at frequencies beyond 10 GHz.

In both Europe and North America, automotive radars are permitted to transmit an average EIRP (effective isotropic radiated power) at 24 GHz up to $-41.3$ dBm/MHz of spectral width, and a peak EIRP of 0 dBm per 50 MHz of spectral width [1], [2]. In practice, about $+12$ dBm of output power is required for a vehicle tracking radar transmitter [15]. For 24-GHz WLANs, the signal limit imposed in Europe is $+20$ dBm peak EIRP with a duty cycle of 10% (or less) for average output powers above $-10$ dBm EIRP [2]. Similar regulations are imposed by the FCC in North America [5], [16], [17]. Therefore, a gain of 15–20 dB is required from the power amplifier (PA) to amplify the modulated WLAN signal from approximately 0 dBm at the upconverting mixer output to 20 dBm at the antenna.

This paper describes a 21–26-GHz three-stage SiGe power amplifier which produces 10%–20% PAE, with 15 dB gain at $+21$ dBm output power [18], making it suitable for either WLAN or automotive radar applications. It is implemented in a production SiGe BiCMOS technology ($f_T = 120$ GHz, $B_{V_{CEO}} = 1.8$ V) on medium-resistivity (10 $\Omega$-cm) silicon substrates [19]. The topology selected for the PA is described in Section II. It uses magnetically coupled Class-AB stages to realize relatively high output power over a wide bandwidth from a low supply voltage (1.8 V). Design of the multi-stage power amplifier is discussed in Section III. Implementation of the on-chip magnetic components (e.g., baluns, power combiners, and transmission lines) and the active stages, including power dissipation and amplifier stability issues related to the physical layout, are described in Section IV. Amplifier testing and experimental results are presented in Sections V and VI of this paper. Although many of the design techniques are presented in the context of a power amplifier, they are widely applicable to other wireless transceiver circuit blocks operating beyond 20 GHz.
II. DESIGN CONSIDERATIONS FOR A 24-GHz SILICON POWER AMPLIFIER

A. Active Devices

To maximize the output power of the transistor, the optimal load value $R_{\text{opt}} \approx BV_{\text{CEO}}/I_{\text{MAX}}$ is required, where $BV_{\text{CEO}}$ is the collector-emitter breakdown voltage (i.e., base lead open-circuited), and $I_{\text{MAX}}$ is the current required to achieve peak gain from the transistor [20]. The breakdown voltages of silicon CMOS or SiGe transistors with $f_{\text{MAX}} > 70$ GHz for mm-wave (i.e., > 12 GHz$^2$) applications are typically less than 2 V, while III-V transistors with equivalent performance possess a breakdown greater than 8 V [19], [21]–[28]. Silicon-based power amplifiers require higher DC bias current (i.e., $I_{\text{MAX}}$) and therefore a smaller $R_{\text{opt}}$ in order to deliver the same amount of output power as their III-V counterparts. The relatively high bias current constrains the design of on-chip passive devices because of interconnect metal electromigration at high current densities. In addition, there is a lower bound on the value of $R_{\text{opt}}$ that can be transformed using practical impedance matching techniques to match the transistor collector and an off-chip 50 $\Omega$ load (e.g., an antenna) for maximum power transfer. For example, a 50:1 impedance transformation to synthesize $R_{\text{opt}}$ of 1 $\Omega$ or less is prone to component tolerances affecting bandwidth and loss in the matching network. On the other hand, an $R_{\text{opt}}$ of about 7 $\Omega$ is easier to match to 50 $\Omega$ [29].

It has been shown that only the time-averaged (i.e., DC) collector-emitter voltage of a SiGe transistor must be within $BV_{\text{CEO}}$ to avoid thermal runaway [30]. The peak collector-emitter voltage (of a time-varying signal) can approach the avalanche breakdown limit, which is higher than $BV_{\text{CEO}}$. Avalanche breakdown in a bipolar transistor is caused by impact ionization in the high electric field present at the base-collector junction, which causes holes to flow back toward the base. This current increases the base-emitter voltage, leading to a progressively larger collector current that eventually destroys the transistor [31]. To increase the transistor’s avalanche breakdown limit, the impedance connected in series with the base terminal must be low enough to shunt the hole current to ground without increasing the base-emitter voltage. In such cases, the maximum usable collector voltage approaches $BV_{\text{CEO}}$ [32], [33], which is typically a few volts higher than $BV_{\text{CEO}}$. However, for a common-emitter amplifier, a low impedance ($< 100$ $\Omega$) bias network in shunt with the input at the base attenuates the RF signal. Furthermore, a bias network synthesized by reactive components (e.g., capacitor, transmission lines) in order to realize a small impedance at the fundamental frequency may exhibit a higher impedance at the harmonic frequencies due to transmission line effects. Therefore, common-base stages are used in this work so that the base terminals are shorted to AC ground to maximize the transistor’s breakdown voltage. Fig. 1 shows the simulated $I$–$V$ relationship for such a common-base transistor. The output current $I_C$ is well-controlled by input current $I_E$ at both high (14 mA) and low (2 mA) levels. In particular, reducing $I_E$ (to 2 mA) also reduces $I_C$ (to 2 mA) despite the resulting high collector-emitter voltage. There is no sign of breakdown up to $V_{\text{CE}} = 4$ V (note that $BV_{\text{CEO}} = 1.8$ V).

B. Single-Ended Amplifier Topologies

From the previous discussion, it is clear that voltage and current density limitations of the transistor are constraints which guide the selection of a power amplifier architecture. There are several design difficulties when single-ended amplifiers are implemented in silicon technologies [34]–[37]. At 25 GHz, a bondwire inductance as low as 0.1 nH adds 15 $\Omega$ reactive impedance to the on-chip ground. In particular, silicon-based power amplifiers drive loads (i.e., $R_{\text{opt}}$) that are typically in the 7–15 $\Omega$ range. If the ground inductance is comparable to the load impedance, the emitter (or source) of the transistor and the on-chip ground see a voltage swing as large as the desired output voltage. The AC voltage on the semi-floating ground (i.e., a ground bounce) causes degeneration that reduces the gain and PAE of the power amplifier. In a multi-stage amplifier, on-chip ground bounce may cause undesirable feedback between stages and instability. To suppress local feedback via the ground connections, each gain stage can be isolated on-chip and connected to the off-chip ground plane with separate bondwires. However, more package ground pins are required. Also, losses for on-chip passive components (e.g., in an inter-stage matching network) increase when all on-chip current-return paths are grounded externally.

C. Differential Power Amplifier With an Output Balun

The difficulties of single-ended amplifier design are avoided when a differential topology is used [38]–[41]. An example is shown in Fig. 2(a) [39]. Here the transistors are driven differentially so that a virtual ground exists on-chip to minimize ground inductance. The matching network between the gain stages can be implemented using a monolithic transformer with a non-1:1 turns ratio [38], [39], $V_{\text{CC}}$ supply and DC bias to the transistors are directly connected to the center-taps (a virtual ground) of the transformer windings, eliminating RF bias chokes. A balun (balanced-to-unbalanced transformer) is needed to convert the differential signals from the final gain stage to a single-ended 50 $\Omega$ output (i.e., antenna). Baluns can be implemented off-chip [38], but on-chip L-C baluns are a compact and integrated solution in the mm-wave frequency range [39]. Asymmetry in the balun
response causes the transformed load impedance to appear inductive at one of the balanced output ports and capacitive at the other. Compensating for this effect requires careful matching of the transistors connected to each port of the balun. Therefore, a symmetric balun design is desirable for a mm-wave PA to ensure the transistors drive identical loads over a broad frequency range (i.e., >4 GHz).

**D. Transformer-Coupled Power Combining Balun**

When the required output power is larger than what is available from a single transistor, power combiners can be used to sum the power from several transistor units to achieve the desired overall output power over a wide bandwidth [42]. For example, the distributed active transformer (DAT) presented in [43] functions as an eight-way combiner which produces 2 W of output power at 2.4 GHz using CMOS driver transistors. Fig. 2(b) shows a simplified four-way power combiner to illustrate the concept. The power combiner consists of a one-turn primary coil divided into two sections and a one-turn secondary coil. Transistors Q1 to Q4 drive the four terminals of the primary coil in a push-pull fashion. The combined magnetic field induces a current in the secondary coil and drives the output load. For DC power consumption above 1 W, the primary coil conducts hundreds of milliamperes of DC current to the transistors. A single turn of metal is used to eliminate vias and thin metal underpasses connections that might restrict the DC current flow. With a 1:1 turns ratio and the segmented primary coil, the power combiner of Fig. 2(b) would ideally transform a 50-Ω output load into four equal loads of 12.5 Ω at each transistor.

Imperfect magnetic coupling between primary and secondary coils increases the loss of the power combiner. Monolithic transformers typically use multiple-turn primary and secondary coils to achieve a high ratio of mutual to self-inductance or magnetic coupling factor, $k$ (e.g., $k > 0.8$) [45]. For a single-turn planar transformer, the $k$-factor is limited to about 0.6, even when minimum metal spacing is used. This reduces the power from the transistor and defeats the purpose of the combiner, which is to combine power from multiple amplifier stages in order to realize a higher overall output. Consequently, a transformer combiner for power amplifier applications requires higher magnetic coupling ($k > 0.8$) to synthesize small load impedances and must also be able to supply sufficient DC current to bias the amplifier stages.

**E. Summary**

The SiGe power amplifier in this work adopts a differential topology to minimize the effects of ground inductance. Common-base gain stages biased at the maximum supply voltage ($V_{CC} = 1.8$ V) drive a relatively low load impedance ($R_{out}$ of 9 Ω) for maximum output power. With a minimum collector emitter voltage of 0.6 V when saturated, a single transistor is ideally able to deliver 19 dBm (80 mW) with this load impedance and supply voltage. In practice, a loss of 1 dB is anticipated for the output matching network between the final stage and the 50-Ω load, and the chip-to-circuit board interface also adds about 1 dB loss at 24 GHz. These losses reduce the maximum output power (for one transistor stage) to about 17 dBm. To increase the output power, a four-way balun combines the power from four transistors (each driving
Fig. 3. Power amplifier schematic.

$R_{\text{eq}} = 9\, \Omega$ load) to raise the total output power above 20 dBm. The power-combining balun also provides a single-ended output connection. The self-shielded balun [46] described in Section IV realizes magnetic coupling $k > 0.85$ in a single-turn design. The 15-dB gain requirement at the maximum output power is satisfied by using three stages of amplification.

III. THREE-STAGE SiGe POWER AMPLIFIER DESIGN

Fig. 3 illustrates the complete schematic of the 21–26-GHz SiGe power amplifier MMIC. It consists of two identical groups of three-stage differential amplifiers. The on-chip output balun T4 sums the (differential) power produced by these two amplifier groups to a single-ended 50-Ω output. Similarly, an on-chip power dividing input balun T1 produces two pairs of differential signals to drive the two amplifier groups from a single-ended 50-Ω input. Monolithic transformers (T2, T3, T5, and T6) also couple the common-base amplifier stages together. The design of this amplifier for 15-dB gain, 21-dBm output power, 21–26-GHz bandwidth, and 10%–19.5% PAE is described in the following subsections.

A. Selecting the Number of Gain Stages and Proper Transistor Technology

Aside from meeting the gain and output power requirement, high efficiency ($\eta$) and PAE are desired from the amplifier. $\eta$ is the ratio of RF output power to DC power consumption, while PAE is the ratio of RF power added to the signal by the amplifier (i.e., difference between input and output powers) and its DC power consumption. PAE is the preferred figure of merit for PAs because it accounts for the RF power consumed at the
input. Higher PAE implies less DC current consumption, higher overall reliability, and more relaxed requirements for dissipation of heat from the chip via the package to the surrounding environment.

Multi-stage amplifiers with relatively little gain per stage need more stages to satisfy the total gain requirements, which results in a lower PAE. Fig. 4 compares the required gain per stage versus the number of stages ($n$, for $1 \leq n \leq 5$) to satisfy the 15-dB gain requirement for this work. The total DC power consumed by each amplifier normalized to a single-stage design is also plotted in the same figure. Efficiency ($\eta$) for each gain stage is kept constant in this comparison to highlight the effect of gain per stage upon DC power in a multi-stage amplifier. For example, a single-stage Class-A design producing 100 mW of RF power (20 dBm) at the ideal 50% efficiency would consume 200 mW of DC power, and would require the highest performance (and likely most expensive) technology available. A five-stage amplifier requires only 3 dB per gain stage, and could be implemented in a low-cost technology such as CMOS with $f_T < 70$ GHz, but requires 94% more DC power than a single-stage solution (12.6, 25.2, 50.2, 100.2, and 200 mW DC power consumed by stages 1–5, respectively, for a total of 388.2 mW), again assuming 50% efficient (i.e., ideal) Class-A stages throughout. This simple example illustrates that the advantage of a low-cost technology is offset by the additional cost, complexity, and weight required to power the IC, especially for portable applications. In addition to poor PAE caused by higher DC power consumption, a cascade of many low-gain stages degrades the dynamic range of the RF signal as each stage contributes distortion and noise. The solution chosen is well-suited to the SiGe BiCMOS technology used for this work, which can achieve 15 dB gain using three gain stages.

### B. Transistor Size Ratio

The ratio of transistor sizes for the three amplification stages in the design is 1:3:12. This ratio was determined by considering the power gain added by each stage. Power gain of a 120-GHz $f_T$ silicon transistor with 2-dB gain compression at 24 GHz is approximately a factor of 4 or 7 dB (from simulation), including 1-dB loss for the input and output interstage matching networks. Assuming a gain of 4 in each stage and output power proportional to transistor size, a suitable ratio of transistor areas for the three gain stages is approximately 1:4:16. However, overestimating the gain results in undersized input devices that cannot drive the output transistors to compression or maximum power. On the other hand, underestimating the gain leads to oversized input devices, which drive the following stages with some margin at the cost of extra parasitic capacitance at the input. Losses close to the input have a relatively minor impact on the overall performance, but inadequate drive power (i.e., overall power output limited by the first stage) reduces the amplifier PAE and output power. For this reason, a slightly reduced ratio of transistor sizes (1:3:12) is used. The oversized input stage also compensates for power lost due to input impedance mismatch, which reflects input power back to the input signal source.

### C. Maximizing Gain, Output Power, and Efficiency

The performance of a multi-stage amplifier is limited by the weakest link in the cascade of stages from input to output. The amplifier can be optimized by recognizing the impact of gain, output power and PAE from each stage on the overall performance.

The overall PAE of a three-stage amplifier is given by

$$\text{PAE} \approx \frac{1}{\text{PAE}_3 + \frac{1}{G_3 \cdot \text{PAE}_2} + \frac{1}{G_2 \cdot G_3 \cdot \text{PAE}_1}}$$

where $G_{1-3}$ and $\text{PAE}_{1-3}$ are the respective gain and power-added efficiencies of each stage.

While gain is contributed equally by all the amplifier stages, (1) shows that the overall PAE is dominated by the final (i.e., third) stage. If $G_2$ and $G_3$ are both higher than 5 dB, the PAE
of stage one ($P_{AE1}$) has less than 10% influence on the overall PAE of the amplifier. Therefore, the first stage is biased Class-A (which requires a higher bias current) to maximize the overall gain without constraining the PAE of the amplifier. The second and third stages are biased with progressively less current (i.e., Class-AB) for higher overall efficiency.

This general principle also guided floor-planning of the amplifier layout. For example, the power combining balun at the output of the final stage is designed for the lowest possible loss, because losses incurred in the final stage penalize the overall output power and efficiency the most. The gain stages are placed as close to the amplifier output as possible, so that output power from the transistors sees the lowest possible path loss between the IC and the load. The amplifier can tolerate more passive component loss in the first stage, so on-chip transmission lines are used to transport the input signal to the first stage. This improves isolation by increasing the physical separation between input and output (i.e., the single-ended input and output are located on opposite sides of the amplifier layout).

D. Partitioning of Power Gain and DC Current Consumption Between Stages

Fig. 5 shows a snapshot of the differential amplifier simulated at 24 GHz to illustrate the design of the power amplifier based on the previous discussion. A half-circuit is shown in the figure for simplicity. Table I summarizes the operating conditions. Since the emitter and collector currents are approximately equal in magnitude, the power gain of the common-base transistor is the ratio of the real part of the impedances at the collector output and emitter input. Output power combiner $T_4$ transforms the 50-$\Omega$ load to 8.8 $\Omega$ for the third stage to drive. Transformer $T_3$ matches the 1.8 $+ 0.3$ j$\Omega$ input impedance of the third stage to a 19 $+ 27$ j$\Omega$ load for the second stage to drive. Note that the emitter input impedance of a common-base stage at 21–26 GHz has an inductive component. Similarly, transformer $T_2$ matches the second-stage emitter input impedance of 3.5 $+ 3.5$ j$\Omega$ to a 103 $+ 51$ j$\Omega$ load for the first gain stage.

The magnetizing and self-inductances of the transformer can be designed to cancel the shunt parasitic capacitance of the transistors (e.g., $C_{CB}$) and interconnects, so that the matched load impedance maximizes the gain at the desired operating frequency [44], [45]. However, to maximize the power amplifier’s bandwidth, the responses of transformers $T_2$ and $T_3$ were tailored to counter the low-pass characteristics of the transistors and passive devices in the multi-stage amplifier chain. For the common-base stage, which has a single-pole response, the transistor power gain increases by 6 dB for each 50% reduction in the operating frequency (i.e., current gain is inversely proportional to frequency) [47]. The transistor current gain at 26 GHz is therefore reduced by a factor of 1.24 compared to the current gain at 21 GHz (i.e., 26/21 = 1.24) in each stage. Over three stages, the cumulative current gain is reduced by about a factor of 2 (i.e., $[1.24]^3$) and power gain is reduced by 4 from 21 to 26 GHz. In addition, the circuit interconnections also have higher attenuation with increasing frequency. To compensate for this roll-off, transformers $T_2$ and $T_3$ must be designed to increase the collector load impedance at stages 1 and 2 by at least a factor of two so that the output power is equalized over the bandwidth from 21 to 26 GHz.

The magnitudes of the load impedances for the three stages are in a 12:3:1 ratio (i.e., approximately the inverse of the transistor size ratio) so that the magnitude of the voltage at the collector of each stage is approximately the same, and close to the maximum power output for each stage. The AC and DC requirements shown in Fig. 5 and Table I are guidelines for the circuit implementation, because their values vary under different operating conditions. For example, the input impedance of the transistor changes with input signal power level. This is most severe for the third (Class-AB) stage, where the input signal may switch the transistor almost completely on and off. Consequently, the input impedance of forward base-emitter junction (i.e., diode) changes substantially within a RF cycle, and its value cannot be easily predicted [48].

Finally, long-term reliability of the power amplifier is of equal or even higher priority than AC performance. A power amplifier that achieves excellent AC performance in the mm-wave regime and can operate reliably is a formidable challenge. To avoid interconnect failures due to electromigration and other problems caused by local heating, each 100 mA of DC current requires approximately 100 $\mu$m wide on-chip aluminum metal (1 $\mu$m thick) for conduction [49]. However, metal paths over 100 $\mu$m wide cannot be used to transport AC signals on-chip because of losses caused by capacitive coupling to the lossy silicon substrate. The interconnect for the power amplifier design must feed the two differential amplifier groups with a total of 460 mA for small-signal, and 600–660 mA of DC current at maximum output power. The solution proposed in this work utilizes self-shielded monolithic transformers [46] to couple the common-base stages, as described in the following section.
TABLE I
SPECTRE-RF SIMULATION RESULTS FOR THE AMPLIFIER OF FIG. 3 AT 24 GHz, P_{out} = 20.5 dBm

<table>
<thead>
<tr>
<th>Amplifier Class</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transistor Size Ratio</td>
<td>1x</td>
<td>3x</td>
<td>12x</td>
</tr>
<tr>
<td>$I_{DC}$ at no RF</td>
<td>12.5mA</td>
<td>30mA</td>
<td>72.5mA</td>
</tr>
<tr>
<td>$I_{DC}$ at $P_{out} = 20.5$ dBm</td>
<td>12.9mA</td>
<td>32.2mA</td>
<td>85.3mA</td>
</tr>
<tr>
<td>Collector Load Impedance, $Z_C$</td>
<td>103+54Ω</td>
<td>19+27μΩ</td>
<td>8.8Ω</td>
</tr>
<tr>
<td>Emitter Input Impedance, $Z_E$</td>
<td>10+7.5μΩ</td>
<td>3.5+3.5μΩ</td>
<td>1.8+0.3μΩ</td>
</tr>
<tr>
<td>Power Gain</td>
<td>10Ω/10Ω</td>
<td>19Ω/3.5Ω</td>
<td>8Ω/1.8Ω</td>
</tr>
<tr>
<td>$\text{Real}(Z_C)/\text{Real}(Z_E)$</td>
<td>~10.1dB</td>
<td>~7.3dB</td>
<td>~6.9dB</td>
</tr>
<tr>
<td>Total Transistor Gain</td>
<td>24.3dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of Input Power Dividing Balun, T1</td>
<td>~1.1dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of Transmission Line</td>
<td>~0.9dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of Transformers T2, T3</td>
<td>~2.2dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of Output Power Combining Balun, T4</td>
<td>~1.1dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss due to package I/O connections</td>
<td>~1.5dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplifier Power Gain</td>
<td>17.5dB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Transistor Gain - Passives Loss)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{CC}$ Supply Voltage</td>
<td>1.8V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total $I_{DC}$</td>
<td>522mA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total DC power</td>
<td>939mW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output Power, $P_{out}$</td>
<td>20.5 dBm (11.2mW)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAE</td>
<td>~11.7%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IV. TRANSFORMERS, POWER COMBINER, AND TRANSISTOR IMPLEMENTATION

A monolithic transformer has a unique advantage that cannot be equalled by transmission lines. When both DC and AC currents share the same current loop as in a transmission line, increasing the DC current rating implies making the conductors of the entire AC current loop wider. By contrast, a monolithic transformer has independent DC current loops (i.e., one for each winding). In a two-filament transformer, the primary and secondary windings act as AC current-return paths for each other and form one AC current loop. Therefore, the metal used to implement one winding can be made wider to satisfy a higher DC current rating, and the other winding with less DC current requirement can use a narrower metal if necessary. This gives the designer the freedom of optimizing the AC path separately from the DC path. Self-shielding design is applied to monolithic transformers in this work to minimize losses to the underlying substrate. The wider metal in one winding (carrying a high DC current) forms a shield to surround the other winding, so that the electric field of the AC current pair is confined internally. This permits both excellent AC performance and the ability to satisfy DC current rating constraints for on-chip interconnect reliability.

A. Self-Shielded Interstage Transformer

Fig. 6(a) shows the top view of the self-shielded monolithic transformer designed for interstage coupling. The transformer consists of a two-turn primary coil driven differentially by the first (or second) stage, and a single-turn secondary coil connected to the second (or third) stage. Since the transformer layout is symmetric, the center taps of the primary and secondary coils are virtual grounds. They are used as DC supply points for the collectors and emitters of the gain stages.

The primary coil driven by the collectors has a voltage swing from 0.5 to 3 V. On the other hand, the secondary coil is terminated by the emitters, which have a voltage swing of only ~0.1 to ~0.3 V. Therefore, as shown in the cross section of Fig. 6(b), the 6x lower voltage (and also lower impedance) present at the terminals of the secondary coil is used to form a coaxial shield around the primary coil. This self-shielded structure minimizes the losses caused by leakage of the electric field into the conductive silicon substrate without adding an explicit ground shield. Simulations show that parasitic coupling from the secondary coil of the transformer to the substrate has little effect on the overall performance of the amplifier. Due to the (as drawn) 2:1 turns ratio, and the different widths of the metal conductors.
used to implement the windings for the two coils, transformers T2 and T3 (see Fig. 5) provide impedance transformation ratios of 1:30 and 1:10, respectively. These transformers match the low-impedance (emitter) side of the common-base stage to the collectors of the previous stage. Since the secondary coil drives a stage with larger transistors, its coaxial shield structure supports higher DC current than the narrower primary coil, which carries the DC current for the (smaller) previous stage.

Fig. 7 shows the first-stage collector impedance transformed from the emitter inputs (about $3.5 + 3.5\, \text{fF}$) of the second stage by transformer T2 in Fig. 5. With an estimated parasitic capacitance of 57 fF between the first-stage collector to AC ground, the transformer passband is centered at 26.7 GHz (i.e., the upper end of the 21–26-GHz band). As the frequency increases from 21 to 26 GHz, the transformer response causes an increase in the collector load impedance (from 58 Ω to 140 Ω) to counter the decrease in transistor current gain. The coupling transformer between the second and third stages (T3) is designed to function in a similar fashion to T2.

The self-shielding technique can also be used to improve the mutual magnetic coupling of a transformer. The structure utilizes one transformer coil to form a coaxial shield around the second coil. By Ampere’s law, opposing AC current pairs flowing on the center and the outer coaxial coils have minimal magnetic field leakage. Therefore, the self-shielded transformer realizes a magnetic coupling factor $k \approx 0.9$ (extracted from electromagnetic simulation), using just a single-turn winding.
B. Self-Shielded Power Combining Balun

Fig. 8 shows the self-shielded four-way power combiner (T4) used in the power amplifier prototype. Since the secondary coil carries the combined power to the 50-Ω output, it is at a higher voltage. By using the 5× lower impedance (i.e., up to $R_{	ext{CPT}} \approx 9$ Ω) primary coil to form the outer coaxial shield around the secondary coil, the primary shields the higher voltage secondary coil and minimizes electric field leakage to the substrate. Note that the balun is symmetrical. The secondary coil is split into two windings to ensure that the voltage fluctuations on the secondary coil are coupled evenly to the four transistors. This minimizes mismatch in the load impedance seen at each port. Furthermore, the secondary coil is optimized by tapering the width from 14 to 6 μm. Along the secondary coil, its ground node (near 0 Ω) has minimal voltage but the highest current, and its output node (at 50 Ω) has the highest voltage but the lowest current. The wider (14 μm) end is connected to ground to minimize resistance to current flow, while the narrow (6 μm) end is at the output to minimize (voltage) coupling to the primary coil and the silicon substrate. Since it is a passive device, the power combiner is reciprocal. When the input and output terminals are exchanged, it can be used as a power dividing balun at the input (i.e., T1 in Fig. 5). T1 splits the single-ended amplifier input into two pairs of differential signals to drive the input stages.

Previous work on a power-combining balun prototype has shown that measurement and electromagnetic simulation for the self-shielded design are in good agreement [46]. Therefore, simulation can be used to investigate electrical behaviors of the combiner that cannot be measured directly. Fig. 9 shows the simulated impedance of the power combiner used in the power amplifier. The combiner is designed to resonate out the shunt capacitance (from final-stage transistors and metal wiring) at the transistor terminals. The 50-Ω output load is divided by the combiner into four equal loads of 8.8 Ω for the common-base output stages to drive in order to maximize the output power. The reactive load (in series) is within ±4 Ω from 21 to 26 GHz. Unlike a conventional lumped L-C balun where equal load impedances are presented to the transistors within only a narrow frequency band around resonance, the power combiner achieves less than ±2% load mismatch over the range from 15 to 40 GHz. The power amplifier is designed to be mounted as a flip-chip on a printed circuit board (PCB). With 0.1 nH of ground inductance (a worst case estimation) contributed by eight flip-chip stud bumps in parallel between the on-chip and PCB ground, the transistor loads increase to 11 Ω (without increased load mismatch between transistors).

The differential amplifier of the final stage draws up to 110 mA DC current for each half of the circuit under maximum output power. The second-to-third stage transformer (T3) and the power combiner (T4) are designed to handle 300 and 150 mA per transistor, respectively. Not only does the design margin in current conduction ensure reliability, the extra-wide transformer windings are also designed to remove heat from both of the collectors and emitters of the transistors to the off-chip PCB.

Differentially shielded interconnects are used on-chip to reduce attenuation of the input signal. The differential shield is implemented using floating metal that is placed underneath a differential transmission line. It reduces leakage of the electric
field and associated losses in the semiconducting silicon substrate. The shield is placed equidistant from the two conductors of the differential transmission line, so that the shielding strips are at a virtual ground [50], [51]. This shielding method does not require an explicit ground connection [52]–[54], which is difficult to define on-chip.

All of the passives were designed and simulated using Agilent’s Momentum 2.5D electromagnetic simulator. Two sheet metal layers are used to mimic the metals used to implement the passives. Lumped-element circuit models were extracted from the E-M simulation results for time-domain simulation of the power amplifier circuit (using Cadence Spectre-RF).

C. Common-Base Amplifier Physical Layout

Fig. 10 shows the high-level connection scheme of the differential common-base amplifier for the third gain stage. Each half of the amplifier consists of 16 unit transistors. Each transistor unit has two emitter stripes of 0.2 μm × 5μm. Therefore, the emitter areas of the two final gain stages are 128 μm² in total. Due to a negative temperature coefficient of current gain (β), the collector current of a SiGe HBT does not increase with rising temperature when the transistor base is biased by a current-source. By contrast, collector current increases with higher temperatures under a forced VBE bias [55]. To assume that the negative temperature coefficient for β automatically safeguards a SiGe power amplifier from thermal instability [56] is a misconception. In general, a power amplifier consists of multiple transistor units in parallel. Since they all share a common VBE voltage, a hotter transistor will draw even more DC current, leading to potential thermal problems. Even if a current source is used for biasing, it does not prevent a hotter unit from drawing more collector current than neighboring transistors, because only the total base current of all the units is kept constant.

To avoid unwanted local feedback that can cause stability problems (especially due to localized heating), the base terminals of the units within the same half of the differential amplifier are not connected together. The base of each unit is connected only to its symmetrical differential counter-part so that the midpoint of the interconnect is a virtual ground in the common-base configuration. Base ballast resistors (∼ 100 Ω, 16 in total) are connected to each of the virtual ground points in order to provide negative feedback to the transistor units that are undergoing localized heating and conducting more current than others. In general, the addition of ballast resistors in the signal path increases stability at the price of lower gain and efficiency [57]. Hence, the resistors, which are connected in the common-mode path at the virtual ground provide negative feedback only for common-mode signals. They do not increase the differential base resistance or degrade the performance of the differential amplifier.

For power-control purposes, the transistor units within the second and third stages are arranged into blocks that can be turned on/off with individual biases (e.g., V3a, V3b, V3c for the final stage). This allows the study of a common-base power amplifier optimally biased with a combination of Class-A, Class-AB and/or Class-B within the same gain stage to raise the efficiency under reduced output power condition.

D. Base Inductance Suppression and Equalization

One of the most problematic parasitics in the common-base stage at mm-wave frequency is base inductance, which can cause instability and oscillation. Base inductance originates from the interconnect between the virtual ground and the base terminal. In a power amplifier where relatively large transistors are employed, the problem is aggravated by two mechanisms. First, the base inductance is magnified by the transistor size. For example, when an interconnection with just 0.1 nH of inductance connects the base of the final stage (16 transistor units in parallel) to the virtual ground, the base of each unit will see 1.6 nH (16 times higher), or 240 Ω at 24 GHz instead of a virtual ground. Second, larger area transistors increase the base inductance if the length of the interconnection between the transistor units and the virtual ground increases.

Fig. 10. High-level connection scheme of differential common-base amplifier (third stage).

Fig. 11. Two methods to minimize base interconnect inductance (half-circuit shown). (a) Increase interconnect thickness. (b) Define a tightly coupled current-return path.
Two methods can be applied to suppress the inductance between the base and the virtual ground. The traditional approach [see Fig. 11(a)] is to make the interconnect thicker by using many metal layers in parallel to connect the base to the virtual ground. Doubling or tripling the conductor thickness can reduce the inductance by only about a factor of two. A better method [shown in Fig. 11(b)] is to provide a current-return path that is tightly coupled to the base interconnect (using first metal to an on-chip ground). This reduces the area of the current loop which gives rise to the inductance. Any AC base current flowing in M2 and M3 induces a current flowing in the opposite direction on M1 which cancels the overall magnetic field. As a result, the inductance of a 1-μm-wide base interconnect is reduced to about 100 pH per mm of interconnect length.

Finally, the 16 transistor units are physically located at different distances (about 14 to 70 μm) from the virtual ground. This causes up to a factor of 5 difference in the base resistance and inductance between the various transistor units. Consequently, their output signals are slightly misaligned in phase and magnitude, which would result in loss of power and efficiency when they are summed. Therefore, the interconnect path widths are adjusted as shown in Fig. 12 to equalize the base inductance and resistance of the units.

V. Flip-Chip Packaging for Millimeter-Wave Frequencies

The power amplifier is designed to be mounted as a flip-chip directly on a PCB for testing. Fig. 13 is a photomicrograph of the amplifier testchip. Double gold stud bumps are stacked onto each bondpad to give 50-μm clearance between PCB and the testchip, so that the PCB ground metal does not interfere with the magnetic flux linkage of the monolithic transformers. The output of the power combiner and the adjacent eight ground bumps form a coplanar wave guide (CPW) interface to the circuit board. The amplifier has a similar CPW interface for the input balun. The ground planes of the input and output CPWs are not connected on-chip in order to minimize undesired feedback from the output back to the input. On the circuit board, the grounds of the CPW join at the bottom solid ground plane with vias at a distance of over a quarter of a wavelength away from the middle of the flip-chip. This further increases the output-to-input isolation. V_DD and DC biases are placed around the perimeter of the amplifier testchip. The test IC consumes a total chip area of 2.45 × 2.45 mm². If the amplifier were integrated with a transmitter onto a single chip, the input could be driven differentially and the input balun eliminated. This would reduce the chip area by at least one-half.

VI. Experimental and Simulation Results

The amplifier is tested with a (V_DD) supply of 1.8 V, and the DC current drawn with no RF input signal applied is 460 mA. The gain, PAE, and output power versus input power at 22, 23, and 24 GHz are plotted in Fig. 15. The small-signal gain is approximately 19 dB. For the three frequencies shown in Fig. 15, PAE exceeds 10% at −3-dB gain compression. The maximum
PAE of 19.7% is obtained at 22 GHz, while maximum PAE at 24 GHz is 13%.

Fig. 16 shows the PAE, output power, and gain versus frequency when the amplifier is operating under saturated output power condition driven by a 6-dBm input. Peak output power of 23 dBm (200 mW) is achieved at 22 GHz, and over 20.8-dBm output power is available between 20 and 25 GHz. At 22 GHz, the final gain stage of the power amplifier achieves a maximum output power density (power per emitter area) of 2 mW/μm². PAE exceeds 12.5% between 20.3 and 24 GHz. The amplifier has over 15-dB gain from 21 to 25.5 GHz with 3-dB gain flatness. Over 15-dB gain at maximum output power level facilitates integration of the amplifier in a transceiver as fewer predriver stages are required. When biased at 230 mA (i.e., 50% of 460 mA DC current with no RF signal applied), there is less than 3-dB reduction in gain, output power, and slightly lower PAE, despite the transistor $g_{m}$ being reduced by half for all three stages. At maximum output power, the temperature of the amplifier flip-chip measured using a pyrometer aimed at the backside of the chip is 50°C (or 122°F). No heatsink or forced air cooling was used during testing.

Fig. 17 compares the measured and simulated results when the power amplifier is operating at saturated output power. The simulated and measured gain, output power, and PAE are all in good agreement.

S-parameters ($S_{11}$, $S_{21}$, and $S_{12}$) of the amplifier in the test fixture are plotted in Fig. 18. The small-signal gain is approximately 19 dB from 21 to 26 GHz. The $S_{11}$ indicates how well the input is matched to 50 Ω. $S_{11}$ is below −10 dB in the 24-GHz ISM band, but it was not designed for an input match to 50 Ω for the entire 21–26-GHz band. It should be noted that the requirement on $S_{11}$ can be relaxed if the amplifier is integrated with the transmitter stages. $S_{22}$ ranges from −8 dB.

\(^3\)Tested at a room temperature of 25 ± 5°C.
to $-1\,\text{dB}$, because the output power combiner is not designed for small-signal conjugate impedance match but for maximum output power (i.e., $R_{\text{in}}$ match). S12 indicates that the isolation from the output to the input is excellent, exceeding $-30\,\text{dB}$ from 20 to 30 GHz. Insertion loss ($\sim 1\,\text{dB}$) of a 1.5/8 CPW thru-line in the same fixture used to characterize the amplifier is also plotted in Fig. 18. This 1-dB loss originated from the thru-line is deembedded to obtain the measured performance in Figs. 15–17. The loss contributed by the flip-chip interface (between signal bondpads and circuit board) is included in the measured performance. The measured S-parameters can be used to calculate the necessary and sufficient conditions for unconditional stability (i.e., $k > 1$ and $B > 0$) [58]. The power amplifier satisfies both of these conditions within the operating range from 10 to 26 GHz.

To test intermodulation distortion, two tones separated by 5 MHz at 24 GHz were applied to the amplifier. At the output, the fundamental tone is $-3.17\,\text{dBm}$, and the third-order intermodulation is $48.17\,\text{dB}$ below the fundamental tone. Accounting for the 2-dB loss in the cable and the test fixture, the output third-order intercept point, OIP$_3$ is $+23\,\text{dBm}$. The OIP$_3$ is lower than the rule of thumb for OIP$_3$ estimation (i.e., OIP$_3$ 10 dB above $P_{-1\text{dB}}$) as the $P_{-1\text{dB}}$ compression point is about $+18.8\,\text{dBm}$ at 24 GHz. In general, a three-stage design has lower OIP$_3$ compared to a single stage alone, because the distortion accumulates along the amplification chain. Also, the final stage is a Class-AB amplifier which generates more distortion compared to Class-A due to its lower bias current.

Fig. 19 compares the performance of this work against other power amplifier designs. An experimental 24-GHz power amplifier in 0.18-$\mu\text{m}$ CMOS was reported to achieve 7-dB small-signal gain using a two-stage cascade structure [37]. At maximum output power of $+14.5\,\text{dBm}$, it has 4% PAE and 2.3-dB gain (i.e., 1.15-dB gain per stage). An 8–17-GHz 0.35- $\mu\text{m}$ SiGe power amplifier has 2%–16% PAE [39]. Commercial power amplifiers in 0.25-$\mu\text{m}$ GaAs MESFET and pHEMT technologies have PAEs of 11%, 15%, and 31%, respectively [59]–[61]. A 24-GHz three-stage power amplifier, designed for an automotive radar transceiver and fabricated in an 80-GHz $f_T$ SiGe bipolar technology on a high-resistivity (1 k$\Omega\cdot\text{cm}$) substrate, achieves 7.5-dB small-signal gain, and 6.5-dB gain at $+12\,\text{dBm}$ output power (i.e., 2.2-dB gain per stage) [15]. However, the PAE was not reported, and few circuit details were given. The three-stage power amplifier in this work has 15-dB gain (i.e., over 5-dB gain per stage at maximum output power), 20–23-dBm output power, 19.7% PAE at 22 GHz, and 13% PAE at 24 GHz. This performance compares favorably with GaAs pHEMT amplifiers in the 21–24-GHz range [59]–[61].

VII. CONCLUSION

A 21–26-GHz 1.8-V three-stage common-base power amplifier MMIC implemented in a 0.20-$\mu\text{m}$ SiGe technology ($f_T = 120\,\text{GHz}$, $BV_{CEO} = 1.8\,\text{V}$) has been presented. The self-shielded design principle for on-chip passive devices was utilized to optimize interstage and I/O coupling at mm-wave frequencies with high DC current ratings for the PA. Self-shielding minimizes substrate loss, current crowding due to skin effect and increases magnetic coupling of the interstage monolithic transformers and power-combining balun. Differential shielding was applied to the on-chip interconnect to minimize substrate loss without the need for an explicit on-chip ground reference. These techniques are simple to implement and are applicable to other passive devices (e.g., inductor, capacitors, etc.) for other RF and microwave circuit applications. With 19-dB small-signal gain, 15-dB gain at maximum output power, 20–23-dBm output power, 19.7% PAE at 22 GHz and 13% PAE at 24 GHz, the power amplifier in this work demonstrates that SiGe technology ($BV_{CEO} < 2\,\text{V}$) has the potential to compete with III-V GaAs technologies for medium-power amplifier applications at 20 GHz and potentially beyond.

ACKNOWLEDGMENT

The authors acknowledge technical support for testing provided by W. Straver, R. Klerks, and L. van Schie at TU Delft. Fabrication and technology support was provided by A. Joseph, Y. Tretiakov (now with RFMD), and D. Harame at IBM Microelectronics, Burlington, VT.


Tak Shun Dickson Cheung (S’04–M’05) was born in Hong Kong in 1972. He received the B.A.Sc. degree in electrical engineering from the University of Waterloo, Canada, in 1996 and the M.A.Sc. degree from the University of Toronto, Canada, in 1999, and is pursuing the Ph.D. degree at the same university.

John R. Long (S’77–A’78–M’83) received the B.Sc. degree in electrical engineering from the University of Calgary, Calgary, Canada, in 1984, and the M.Eng. and Ph.D. degrees in electronics from Carleton University, Ottawa, Canada, in 1992 and 1996, respectively.

...