6

Millimeter-wave Circuits and Applications

A. Mukherjee\textsuperscript{1}, W. Liang\textsuperscript{1}, M. Schröter\textsuperscript{1,2}, U. Pfeiffer\textsuperscript{3}, R. Jain\textsuperscript{3}, J. Grzyb\textsuperscript{3} and P. Hillger\textsuperscript{3}

\textsuperscript{1}Chair for Electron Devices and Integrated Circuits, Technische Universität Dresden, Germany
\textsuperscript{2}Department of Electrical and Computer Engineering, University of California at San Diego, USA
\textsuperscript{3}Institute for High-Frequency and Communication Technology (IHCT), University of Wuppertal, Germany

6.1 Millimeter-wave Benchmark Circuits and Building Blocks

A. Mukherjee, W. Liang and M. Schröter

The continuous progress of SiGe:C HBT BiCMOS process technology paves the way for high-volume low-cost mm-wave and sub-mm-wave applications. The design of the corresponding high-frequency (HF) integrated circuits requires accurate compact models for both active and passive devices. Especially, the compact models for active devices must cover many physical effects occurring in advanced process technologies and address a wide bias, temperature, and geometry range as well as high-frequency (HF) effects such as non-quasi-static delay and substrate coupling. Devices used in HF circuits typically operate at 3 to 10 times the circuit speed due to the harmonics generated within the circuits that ultimately determine the signal shape. The verification of compact models at such a high speed has become a major issue since device measurement capability has not kept pace with process and circuit development. While there has been some effort toward extending small-signal (S-parameter) measurement capability toward several 100 GHz, direct experimental verification of compact models for \textit{large-signal} operation at mm- and sub-mm-wave frequencies still appears illusive. For instance, load-pull measurements beyond 50 GHz are not only difficult and expensive but also do not provide any phase shift information, which is important for describing time-dependent large-signal switching correctly.
The demand for model accuracy to ensure one-pass success for saving R&D cost in mm-wave and sub-mm-wave circuit design forces compact models of transistors to undergo tests in a vast range of operating conditions instead of merely verifying typical device characteristics. Therefore, model verification has been extended to small circuits in which transistor operation can be tested under realistic application-relevant conditions. These circuits comprise benchmark blocks and small building blocks of larger systems.

Benchmark circuits on the one hand have to be sufficiently simple so as to avoid masking compact transistor model deficiencies by other effects, but should on the other hand resemble the typical transistor operation in related larger circuit building blocks. A well-selected set of benchmark circuits should allow the transistors (and their associated models) to be exercised in application-relevant operating modes beyond the typical standard device characteristics measured in a characterization lab. In addition, the same benchmark circuits can also be employed for evaluating process performance and for detecting processing issues in terms of the targeted applications during the process development phase.

The circuit building blocks are concerned with practical needs toward, e.g., lowering power consumption or utilizing the transistor non-linearity for harmonic power generation in mm-wave circuits. Here, transistor operation in extreme regions is of interest, e.g., at low collector–emitter voltages (i.e., at significantly forward-biased base–collector junction) or beyond the open-base breakdown voltage. The related building blocks target competitive figures of merit (FoMs) and serve also for demonstrating the process technology’s capability.

Using these relatively small circuits for the above-mentioned purposes has so far been hampered by various factors. On one side, modeling and process engineers lack the necessary circuit design expertise and on the other side circuit designers have little interest in designing, from their perspective, relatively simple circuits. In DOTSEVEN, for the first time, an attempt was started to better bridge these two worlds by fabricating a set of circuit blocks partially designed by the modeling community of the project. The experimental results of the various circuits were then compared with simulations in order to establish a solid understanding of the accuracy of the compact models under circuit-relevant constraints. Several examples are presented below.

6.1.1 Benchmark Circuits

A. Broadband amplifier using a Darlington pair

The broadband amplifier (BBA) is an integral part of both wireless and wireline communication systems. Figure 6.1 depicts two variants of the BBA
schematic, with their input and output matched to the 50 $\Omega$ system impedance. This type of amplifier generally shows a low-pass behavior, i.e., it provides its maximum gain at low frequencies. Here, the BBA topology is derived from the basic Darlington configuration [Mukh16, Gray08, Vera13, Vera14], but uses a modified Darlington pair consisting of an emitter follower and a common emitter transistor.

The degeneration resistance ($R_{E1}$) actually increases the terminal impedance of the stage and can at the same time be used to bias the transistor $Q1$. The advantage of the Darlington configuration over a simple single degenerated stage is that an appropriate choice of $R_{E1}$ yields a current gain bandwidth which can approach twice that of a single stage [Armil89].

The operating points of the transistors are adjusted, as shown in the schematic, through external bias-Tees and the resistor ($R_{E1}$). The feedback resistance ($R_F$) allows adjusting the gain flatness. One of the main aspects of benchmark circuits is the need to be able to quickly design them, preferably by modeling or process engineers. This requires the development of a generic design procedure [Mukh16], which may not achieve world-record performance but guarantees a working circuit that meets the circuit purposes. Such a procedure is given below:

1. The design starts with choosing $V_{CC}$ at or near $BV_{CEO}$ as specified by the corresponding process design kit (PDK) documentation.
2. Both the transistors are biased at $J_C(f_{T,peak})$ according to the $f_T$ – $J_C$ plot (cf. Figure 6.2). The corresponding $V_{BE}$ values determine $V_{BB} = V_{BE1} + V_{BE2}$.
3. As no explicit input and out matching network is used in the circuit, the emitter length of $Q1$ is adjusted so as to make the real part of the input...
impedance 50 Ω. For both transistors the minimum emitter width should be used.

4. Since Q1 is operated as emitter follower, the current through $R_{E1}$ is much larger than the base current into Q2 so that $R_{E1} \approx V_{BE2}/I_{C1}$.

5. The initial emitter length of Q2 can then be chosen similar to Q1, but needs to be adjusted according to its larger $V_{CE}$ to maintain operation at $J_C(f_{T,peak})$.

6. The initial value of $R_F$ can be obtained from, $R_F/[1 - S_{21}(f = 0)] = 50$ Ω.

7. To enhance the bandwidth of the amplifier, a peaking inductor ($L_P$) can be added in series to the input of Q1. The value of $L_P$ can be calculated from the resonance condition at the input of Q1, knowing its input impedance, and at the 3 dB frequency of the gain of the BBA without peaking inductance. The value can be calculated by $L_p = \text{Imag}(Z_{in})/(2\pi f_{3dB})$, where $f_{3dB}$ is the original 3 dB frequency of the amplifier.

8. Further optimization of the circuit is required after EM simulation of the entire circuit.

Two variants, with and without the peaking inductor, of the BBA were fabricated in IHP’s first DOTSEVEN technology run. The backend of that process offered seven metal layers with five thin metals and two thick top metal layers [IHP03]. Important transistor characteristics along with the comparison to the compact model data are shown in Figure 6.2.

The “topmetal2” of the process has been used to realize the inductor and the other required transmission line interconnects. Both amplifier versions were configured for on-wafer measurements using GSG probes along with DC biasing through bias-Tees. The die micrographs of the two amplifiers

![Figure 6.2](image-url)

**Figure 6.2** Comparison between measured data (symbols) and compact model HICUM/L2 results (lines). (a) Transit frequency $f_T$ vs. $J_C$ for different $V_{BC}$ values. (b) Transconductance $g_m$ vs. frequency for $V_{BC} = 0$V and at different $J_C = (1, 5, 10, 20)$ mA/µm².
are shown in Figure 6.3. The total die area of the individual amplifiers is just 0.05 \text{mm}^2 and fits into regular HF GSG pads also used for transistor characterization. The two amplifiers were biased with a single supply voltage of 1.8 V.

The S-parameters were measured using a Keysight PNA-L5235A with 110 GHz extenders. The measurement includes effects of pad parasitics, on-chip transmission lines connecting input and output and other components of the amplifier layout, i.e., no de-embedding of those elements was performed. The small signal gain $S_{21}$ in the frequency range of 0.5–110 GHz along with a comparison with post-layout simulation is shown in the Figure 6.4.

The measurement shows a bandwidth of 78 GHz for the BBA without peaking inductor and of 109 GHz for the amplifier with peaking inductor. The measured stability factors for both amplifier versions are shown in

![Figure 6.3](image)

(a) Die photograph of the BBA (a) with and (b) without peaking inductor. The chip size in both cases is $(0.245 \times 0.18) \text{mm}^2$.

![Figure 6.4](image)

(a) Small-signal gain and (b) stability factor of the BBA with and without peaking inductor: comparison between simulation (lines) and measurement (symbols).
Figure 6.4(b) and ensure unconditional stability. The simulations agree quite well with the measurements.

Transistor models for advanced SiGe HBTs have to cover many physical effects in order to achieve the desired accuracy for enabling first-pass design. The resulting model complexity makes it difficult to understand the impact of each physical effect or model parameter on circuit performance under realistic operating conditions. Knowing this dependence is important for model developers, circuit designers, and process engineers. Therefore, a sensitivity analysis was performed by analyzing the changes of the relevant circuit FoMs with respect to variations in model parameters. For the BBA here, the gain ($S_{21}$), the input and output reflection coefficients $S_{11}$ and $S_{22}$, the stability factor $k$, and the bandwidth were selected as the important FoMs. The model parameters of the two transistors (Q1 and Q2) were varied separately to identify the model parameters that are most influential on the above-mentioned FoMs. The Figure 6.5(a) displays the maximum relative sensitivity of the above mentioned FoMs for those model parameters of Q2 that cause changes of more than 5% in at least one of the FoMs as a response to a ±20% change in the respective model parameter. Figure 6.5(b) shows for Q1 the three model parameters that have the highest impact in at least one of the FoMs.

It can be observed in Figure 6.5 that the emitter resistance ($r_e$), the low-current transit time ($t_0$), and the thermal resistance ($r_{th}$) of Q2 have the biggest impact. $r_e$ causes a large change in input return loss as a small change in its value directly affects the corresponding BE-voltage of the transistor. The emitter resistance ($r_e$) of Q1 also shows considerable impact on input return loss but its contribution is masked by the large value of $R_{E1}$ in series.

Figure 6.5  Sensitivity of the series peaked BBA performance parameters with respect to HICUM model parameters for the transistors (a) Q1 and (b) Q2.
The influence of $t_0$ is caused by the associated diffusion capacitance ($S_{11}$), which dominates the input capacitance, and the base–collector voltage-dependent mobile charge in the transfer current, impacting the output conductance and thus $S_{22}$. Interestingly, the bandwidth is mostly impacted by $r_{th}$ of the second transistor Q2. The sensitivity analysis of the BBA without series peaking inductance shows a very similar trend and thus is not shown here.

**B. W-band low-noise amplifier (LNA)**

In this section, the design and implementation of a wide-band LNA for the frequency range of 90–110 GHz is described. The architecture employed here includes a shunt–shunt feedback resistor along with an input matching LC $\pi$-network and a post-cascode series peaking inductor [Lin07]. The basic idea of the $\pi$-network is to add a low-pass filter at the input side of the LNA [Lin07]. It enables the input impedance of the LNA to be matched with the source impedance (50 $\Omega$) when the input frequency is less than the cutoff frequency of the low-pass filter,

$$f \leq \frac{1}{2\pi\sqrt{L_B C_1}}$$  \hspace{1cm} (6.1)

Considering the equivalent circuit (EC) as shown in Figure 6.6(b), the input impedance ($Z_{in}$) can be written as,

$$Z_{in} = \frac{s^2 R_f C_\pi L_B + s L_B + R_f}{s^3 R_f C_{in} C_\pi L_B + s^2 L_B C_{in} + s R_f (C_{in} + C_\pi) + 1}.$$  \hspace{1cm} (6.2)

Setting $Z_{in} = 50 \Omega$ and with $s = j\omega$, the above equation can be solved for two frequencies with perfect input impedance matching.

![Figure 6.6](image.png)  
(a) LNA with $\pi$-network for input matching, (b) small signal equivalent circuit of the LNA with feedback resistance $R_{FB}$ and input matching elements.
\[
\omega_1 = 0 \tag{6.3}
\]
\[
\omega_2 = \sqrt{\frac{2}{L_B C_\pi} - \frac{1}{Z_0^2 C_\pi^2}} \tag{6.4}
\]

For the design of the LNA the cascode configuration is adopted due to its better reverse isolation and higher small-signal gain. The optimum DC bias voltage for the common-emitter transistor was found from the \(NF_{\text{min}}\) vs. \(J_C\) measurement (cf. Figure 6.7) of a separately measured single-emitter transistor.

Figure 6.8 shows the schematic of the implemented wide-band LNA.

**Figure 6.7** \(NF\) and \(NF_{\text{min}}\) versus \(J_C\) for a SiGe HBT with \(A_{E0} = 0.7 \times 0.9 \mu m^2\) for \(V_{CE} = 1.2\) V at \(f = 90\) GHz.

**Figure 6.8** Schematic of the single-stage wide-band (90–110 GHz) LNA.
At mm-wave frequencies, the layout of the circuit plays a pivotal role for the circuit performance. Initially both transistors were contacted up to top metal. The insertion of the series inductor $L_B$ and parallel capacitor $C_{\text{in}}$ compensates for the effect of the input capacitance $C_\pi$. This strategy produces a third-order ladder type low-pass filter network which can reduce the imaginary part of $Y_{11}$ and hence increase the input-matching bandwidth [Lin07]. The initial values of $L_B$ and $C_{\text{in}}$ were calculated with the help of Equation (6.4) and can be further adjusted for optimized performance. $C_{\text{in}}$ was implemented using the MIM capacitor between metal5 and topmetal1. At the desired high frequencies, $L_B$ was realized using the available topmetal2 of the process. A small degeneration inductor $L_E$ ($\sim 20$ pH), implemented as a metal line, was added to ensure good linearity and better stability [Ko96, Afsh06]. After several simulation iterations, the value of $R_F$ was fixed to 415 $\Omega$.

The output matching network consists of $L_C$ and $C_2$, the values of which were carefully chosen to provide $S_{22}$ matching over the frequency range of interest. A small peaking inductor $L_P$ was added to achieve a better small-signal gain ($S_{21}$) flatness. All interconnects between the passive elements were realized with topmetal2 and were EM-simulated to include the effects of the layout parasitics. The base biases of the two transistors were fed through 3 k$\Omega$ resistors that use unsalicided, p-doped gate polysilicon as resistor material [IHP03].

Depending on the biasing of this circuit different results are obtained. One goal here was to verify the compact model for low-power applications operating in saturation. Therefore, the DC bias values at the base terminals were chosen as $V_{B1} = 0.94$ V and $V_{B2} = 1.6$ V, respectively. Together with a supply voltage ($V_{CC}$) of 1.4 V this ensured that both transistors work in the “saturation region”, i.e., with positive external $V_{BC}$ of about 0.23 V.

The on-wafer S-parameter measurements were performed using a Keysight PNA-L5235A with 110 GHz extenders. The noise of the amplifier was measured at the IMS lab in Bordeaux, France. Figure 6.9 shows the S-parameters of the fabricated amplifier along with circuit simulation. The moderate performance of $S_{22}$ up to 80 GHz affects the output reflection coefficient of the measurement equipment. Generally, the agreement between measurement and simulation is quite satisfactory though.

The noise measurement was performed from 75 GHz to 90 GHz, which was the highest frequency range for which a noise source and respective measurement equipment were available. The measured $NF$ at 90 GHz is 5 dB,
Figure 6.9  (a) Small-signal results of the wide-band LNA: comparison between measurement (dashed lines) and simulation with HICUM/L2 (solid lines). (b) Corresponding frequency-dependent noise figure $NF$: comparison between measurements (symbols) and simulation of (blue line) and $NF_{\text{min}}$ (red dashed line).

which is slightly less than simulated. The simulation was performed with the noise-correlation model turned on.

Figure 6.10 shows the results of a sensitivity analysis for the designed LNA where the model parameters of the two transistors Q1 and Q2 were varied separately. In case of transistor Q1, the impacts of only those model parameters are shown that cause the maximum relative sensitivity of the relevant FoMs to vary at least 4% in response to a parameter variation in the range of $\pm 20\%$. In case of Q2 the three most influential model parameters are shown.

Figure 6.10  Sensitivity of the wideband LNA performance parameters with respect to HICUM model parameters for the transistors (a) Q1 and (b) Q2.
It can be observed from the above figure that the low-current transit time \( (t_0) \) and the associated delay time of the transfer current \( (alit*t_0) \) of Q1 and Q2 have the biggest impact on mainly the noise figure (NF). The delay time enters the sensitivity through the noise correlation. The gain of the cascode amplifier is basically controlled by the CE transistor (Q1) rather than the transistor in CB mode (Q2). The emitter resistance \( (r_e) \) of Q1 mostly impacts the input reflection and gain. Generally, the FoMs are less sensitive though to the parameters of Q2 compared to those of Q1.

It must be mentioned here that in this sensitivity study the model parameters are varied individually, i.e., their correlation through process and structural parameters of the transistor were ignored \[Schr05\]. A study including the correlations can be done using a special transistor scaling tool \[Schr99\].

The performance of LNAs can be compared through the following FoM,

\[
F_{oM_{LNA}} = \frac{\text{Bandwidth (GHz)} \times \text{Gain (dB)}}{(F_{avg} - 1) \times P_{DC}(mW)},
\]

where, \( F_{avg} \) is the average noise factor within the band and \( P_{DC} \) is the DC power dissipation of the circuit. In this FoM, the gain in decibels used as the power consumption is proportional to gain in decibels \[Sato10\]. Table 6.1 compares the performance of this LNA with other state-of-the-art broadband mm-wave LNAs reported recently. Despite operation in saturation, the performance compares reasonably to the other designs and its FoM is only exceeded by an 80 nm HEMT amplifier.

### 6.1.2 Circuit Building Blocks

So far, some results going along this direction have been reported. In \[Seth11, Inan14\], LNAs for the 8–12 GHz and 10–22 GHz bands were designed with reduced supply voltages. At a higher frequency, a 65 GHz LNA was implemented in a 130 nm SiGe HBT process in \[Agar14\]. A 53.5 GHz SiGe HBT oscillator with only 0.5 V supply voltage was reported in \[Sah14\]. In the following context, the design of a W-band low-power LNA is presented, which uses an ultra-low supply voltage \( (V_{CC} = 0.5 \text{ V}) \). This work aims to give an example showing how far the DC power consumption can be reduced while maintaining meaningful circuit performance for a mm-wave LNA with HBT transistors biased in the saturation region. The impact of varying transistor series resistances on voltage gain, minimum noise figure
<table>
<thead>
<tr>
<th>Reference</th>
<th>Tech $f_T/f_{MAX}$ (GHz)</th>
<th>Topology</th>
<th>3 dB BW (GHz)</th>
<th>Gain (dB)</th>
<th>$NF$ (dB)</th>
<th>$OP_{1dB}$ (dBm)</th>
<th>DC (mW)</th>
<th>FOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Kiss10]</td>
<td>SiGe:C HBT 220/285</td>
<td>2-stage CE</td>
<td>55–77 (33%)</td>
<td>20</td>
<td>5.8 (est)</td>
<td>3</td>
<td>40</td>
<td>3.92</td>
</tr>
<tr>
<td>[Gilr11]</td>
<td>0.18µ BiCMOS 200/200</td>
<td>5-stage CE</td>
<td>69–95 (31%)</td>
<td>20</td>
<td>&lt;12</td>
<td>–</td>
<td>63</td>
<td>0.917</td>
</tr>
<tr>
<td>[May10]</td>
<td>0.12µ BiCMOS 200/265</td>
<td>5-stage CE</td>
<td>82–100 (20%)</td>
<td>27</td>
<td>8*</td>
<td>–</td>
<td>27.6</td>
<td>3.31</td>
</tr>
<tr>
<td>[Chen12]</td>
<td>0.18 µ BiCMOS 200/180</td>
<td>4-stage cascode</td>
<td>86–106 (21%)</td>
<td>25</td>
<td>&lt;9</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>[Sato10]</td>
<td>80 nm InP HEMT 380/283</td>
<td>3-stage CG</td>
<td>68–110 (47%)</td>
<td>18</td>
<td>3.5</td>
<td>–4</td>
<td>12</td>
<td>50.85</td>
</tr>
<tr>
<td>[Koch10]</td>
<td>100 n InAlAs mHEMT 200/300</td>
<td>4-stage CS</td>
<td>115–150 (26%)</td>
<td>15</td>
<td>5–6 (est)</td>
<td>–</td>
<td>35–40</td>
<td>5.42</td>
</tr>
<tr>
<td>[Zhan12]</td>
<td>0.13 µ BiCMOS</td>
<td>4-stage</td>
<td>132–160 (19%)</td>
<td>21</td>
<td>&lt;9.5*</td>
<td>–</td>
<td>14.5</td>
<td>5.84</td>
</tr>
<tr>
<td>[Liu13]</td>
<td>0.25 µ BiCMOS 180/220</td>
<td>2-stage cascode</td>
<td>47–77 (48%)</td>
<td>22.5</td>
<td>&lt;7.2</td>
<td>4.5*</td>
<td>52</td>
<td>4.35</td>
</tr>
<tr>
<td>[Liu13]</td>
<td>0.13 µ BiCMOS 250/300</td>
<td>2-stage cascode</td>
<td>70–140* (66%)</td>
<td>25</td>
<td>&lt;7* &lt;9*</td>
<td>1*</td>
<td>54</td>
<td>6.10</td>
</tr>
<tr>
<td>Thiswork</td>
<td>0.13 µ BiCMOS 505/720</td>
<td>1-stage cascode</td>
<td>67–117 (54%)</td>
<td>12</td>
<td>&lt;9.6* 5 @90 GHz</td>
<td>0.49*</td>
<td>12</td>
<td>12.5</td>
</tr>
</tbody>
</table>

Table 6.1 Comparison of LNA related FoMs for different technologies and topologies (*simulation; #78–110 GHz estimated)
(\(NF_{\text{min}}\)), and third-order input intercept point (IIP3) is also investigated for this low-power LNA.

Furthermore, the design of a W-band frequency tripler with 0.5 V supply voltage is presented, aiming at an output signal at 96 GHz from a 32 GHz input signal with as low as possible DC power consumption. This frequency tripler could be used as a candidate for generating W-band signals, together with a fundamental-tone oscillator located at a much lower frequency, to alleviate the problem of directly designing a W-band oscillator with satisfactory performance. Comparison between simulated and experimental results is given to verify the accuracy of HICUM model parameters at W-band when transistors are used to design non-linear mm-wave circuits with a reduced supply voltage.

### 6.1.2.1 W-band low-noise amplifier (LNA) with 0.5 V supply voltage

IHP SG13G2 SiGe HBT technology was used for designing the LNA in this section. Figure 6.11(a) shows that a collector current density of more than 10 mA/\(\mu\)m\(^2\) is needed to bias the HBT from this technology at its peak transit frequency (\(f_T\)). The corresponding base–emitter DC bias voltage (\(V_{\text{BE}}\)) can be obtained from Figure 6.11(b), which shows that at least 0.85 V is needed to achieve such a collector current density. Therefore, if the design of an LNA with ultra-low supply voltage is targeted (like \(V_{\text{CC}} = 0.5\) V), then the transistor has to be biased in the saturation region (\(V_{\text{BC}} > 0.35\) V). Figure 6.11(a) also implies that the HBT in this technology can still provide

---

**Figure 6.11**  (a) Transit frequency of an HBT from IHP SG13G2 versus its collector current density with \(V_{\text{BC}} = -0.5, 0,\) and 0.5 V. (b) Corresponding transfer characteristics.
an acceptable value for peak $f_T$ when $V_{BC}$ equals 0.5 V (around 260 GHz, compared with the value of around 320 GHz with $-0.5$ V $V_{BC}$ in the normal forward-active case). In other words, with a supply voltage as low as 0.5 V, this transistor still retains a decent speed for designing mm-wave circuits.

The topology of the LNA, shown in Figure 6.12, consists of three stages of common-emitter configuration with emitter–collector transformer feedback to improve reverse isolation and stability at high frequencies. Besides stability, the emitter series inductor also serves as part of the impedance matching network. The amplifier is biased with $V_b = 0.89$ V and $V_{cc} = 0.5$ V, while the total power consumption of the three stages is only 2.79 mW (1.86 mA for each stage). The topmost metal layer provided by the technology (TopMetal 2) is used to fabricate the transmission lines, whereas the lower metal layers (TopMetal 1, and Metal 5/4/3/2) are used for transitions going through different low-level layers. The bottom metal layer (Metal 1) is used as the ground plane all over the layout of the circuit. The LNA is designed with the aid of constant available power gain circles and constant NF circles of each stage.

Figure 6.13 illustrates the constant available power gain circles and constant NF circles of the transistor used in the first stage of the amplifier (without transformer feedback) at 94 GHz. The source impedance posed to the base of the transistor (transformed from a 50-$\Omega$ signal source by the input matching network) is selected as close to the center of the constant available power gain circles as possible for higher power gain, while the source impedance is also chosen as close to the center of the constant NF circles as possible for lower noise mismatch (leading to lower noise contribution by the
corresponding amplifier stage). Therefore, in practice a compromise has to be made between power matching and noise matching in an LNA design.

The measured and simulated S-parameter results of the three-stage amplifier are shown in Figure 6.14(a). With only 0.5 V collector supply voltage, this LNA can still provide 14.38 dB peak power gain at 91 GHz and more than 10 dB power gain over a frequency range from 86 GHz

Figure 6.14 Measured (symbols) and simulated (lines) results of the W-band ultra-low power three-stage LNA: (a) S-parameters and (b) noise figure.
to 100 GHz. The measured and simulated NFs of the LNA are shown in Figure 6.14(b), where the correlated noise has been turned on (flcono = 1) and off (flcono = 0), respectively. The measured NF at 90 GHz is 5.44 dB. The measured input-referred 1 dB compression point is –21 dBm at 91 GHz. Fairly good agreement between measurement and simulation (especially for the S-parameter results) has been achieved, which verifies the accuracy of the compact model (HICUM/L2) with a forward-biased BC junction at high frequencies (W-band). Figure 6.14(b) also shows that including noise correlation is not negligible, when there is a demand to accurately capture the noise performance of amplifiers at W-band. Note that noise correlation increases with frequency.

To investigate the impact of transistor series resistances on the circuit performance of this LNA ($S_{21}$, $NF_{\text{min}}$, and IIP3), a sensitivity analysis was performed at 92 GHz for the emitter, base, and collector series resistances ($\pm 30\%$ variation) as shown in Figure 6.15. The corresponding model

![Figure 6.15](image_url) (a) Absolute sensitivity of LNA FoMs w.r.t. to series resistance variation. Detailed variation of (b) $S_{21}$, (c) minimum noise figure, (d) input referred third-order intercept point, all w.r.t. to series resistance variation.
Table 6.2  Performance summary of the W-band LNAs

| Reference | Technology | Freq (GHz) | \(20 \times \log_{10} |S_{21}|\) (dB) | \(NF\) (dB) | \(IIP_3\) (dBm)* | \(P_{DC}\) (mW) |
|-----------|------------|------------|---------------------------------|-----------|----------------|-------------|
| [Ceti12]  | 45 nm CMOS | 95         | 10.7                            | 6         | 14.6           | 52          |
| [Vigi16]  | 28 nm CMOS | 90         | 28                              | 7         | –2.7           | 31.3        |
| [Sev10]   | 130 nm SiGe | 95         | 9                               | 8.6       | –5.3           | 13          |
| [May10]   | 120 nm SiGe | 95         | 23                              | 8         | N/A            | 28          |
| [Yang13]  | 90 nm SiGe | 90         | 19                              | 5.1       | –10.4          | 43          |
| [Ina14]   | 90 nm SiGe | 94         | 10                              | 4.2       | –1.9           | 8.8         |
| This work | 130 nm SiGe | 90         | 14.3                            | 5.44      | –9.1           | 2.79        |

*The listed \(IIP_3\) results are estimated from the reported input-referred 1 dB compression points.

parameters are the zero-bias internal base resistance \(r_{\text{Bi0}}\), the external base resistance \(r_{\text{Bx}}\), the emitter resistance \(r_E\), and the external collector resistance \(r_{\text{Cx}}\). Regarding the sensitivity of the input-referred \(IIP_3\), one would expect \(r_E\) to have the largest impact on the linearity of an amplifier due to the series negative feedback introduced by this resistance. A reason for the less-than-expected change of \(IIP_3\) may be that the impact of \(r_E\) (8.16 \(\Omega\)) is masked by that of the inductor in series at the emitter (cf. schematic in Figure 6.12), which has an impedance (\(\omega L\)) of 8 \(\Omega\) at 92 GHz. The detailed variations of \(S_{21}\), \(NF_{\text{min}}\), and \(IIP_3\) with regard to the variations of series resistances are shown in Figures 6.15(b–d). The fact that \(r_{\text{Bx}}\) has the biggest impact on the FoMs confirms that, at least for the considered process technology, the maximum oscillation frequency is the more relevant standard device FoM.

The performance of this LNA, along with the comparison with other reported LNAs operating around 90 GHz, is summarized in Table 6.2. The results of this work clearly imply the option of operating transistors with very low collector supply voltage (0.5 V) while maintaining a reasonable power gain and NF performance at W-band.

6.1.2.2 W-band low-power frequency tripler

In this section, the design of a W-band low power frequency tripler is introduced. This frequency tripler is also designed in IHP SG13G2 SiGe HBT technology. The schematic of the core part of this frequency tripler is shown in Figure 6.16. The transistors used in the core harmonic generation cells have an emitter size of 0.07 \(\mu m \times 0.9 \mu m \times 3\). The total DC power consumption (including the buffer amplifier) is 4.66 mW with a 0.5 V supply voltage [Lia17].
The topology of the tripler consists of two parts: the harmonic generation part and the output buffer amplifier part. Differential configuration is used in the harmonic generation part to suppress the even-order harmonic signal, with the use of on-chip baluns for single-ended-to-differential conversion. Extensive electromagnetic (EM) simulation was performed during this design as shown in Figure 6.17. The small-signal input/output return loss results are measured from one break-up harmonic generation cell as shown in Figure 6.18(a), which implies that the strongest output signal occurs at around 93 GHz.
6.1 Millimeter-wave Benchmark Circuits and Building Blocks

Figure 6.18 (a) Input and output return losses of the core part of the frequency tripler; (b) Conversion gain (actually loss) of the frequency tripler.

96 GHz with an input at around 32 GHz, which is as expected for the correct function of a W-band frequency tripler. The simulated and measured conversion loss results of the frequency tripler are shown in Figure 6.18(b), which shows a minimum conversion loss of 3.79 dB when generating a 96 GHz output signal. These conversion loss results are measured with only –10 dBm input signal over the frequency range of 26–36 GHz.

The performance of this frequency tripler is summarized in Table 6.3 along with the performance of some other reported W-band frequency triplers. The work in [Yeh13] has also demonstrated an ultra-low power frequency tripler using the injection-locking mechanism, but the required input signal power can be as high as 6 dBm, which will impose significant additional power consumption and design effort on the preceding circuit blocks. Looking at Table 6.3, it seems that the design of the frequency tripler in this work has proved the potential for utilizing high-speed SiGe HBT technology biased in the saturation region to implement a competitive

<table>
<thead>
<tr>
<th>Reference</th>
<th>Technology</th>
<th>Freq (GHz)</th>
<th>Pin (dBm)</th>
<th>Peak Conv. Gain (dB)</th>
<th>Harmonic Rejection (dB)</th>
<th>$P_{DC}$ (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Chen10]</td>
<td>65 nm CMOS</td>
<td>85–95.2</td>
<td>0</td>
<td>–13.5</td>
<td>&gt;30</td>
<td>19.8</td>
</tr>
<tr>
<td>[Wang12]</td>
<td>180 nm SiGe</td>
<td>96</td>
<td>0</td>
<td>–7</td>
<td>&gt;20</td>
<td>75</td>
</tr>
<tr>
<td>[Hung10]</td>
<td>150 nm mHEMT</td>
<td>72–114</td>
<td>14.5</td>
<td>–20</td>
<td>–</td>
<td>120</td>
</tr>
<tr>
<td>[Vish12]</td>
<td>90 nm CMOS</td>
<td>90–115</td>
<td>8</td>
<td>–2</td>
<td>–</td>
<td>17</td>
</tr>
<tr>
<td>[Yeh13]</td>
<td>90 nm CMOS</td>
<td>94</td>
<td>–1</td>
<td>–</td>
<td>&gt;20</td>
<td>3</td>
</tr>
<tr>
<td>This work</td>
<td>130 nm SiGe</td>
<td>88.5–103.5</td>
<td>–10</td>
<td>–3.79</td>
<td>&gt;30</td>
<td>4.66</td>
</tr>
</tbody>
</table>
W-band frequency tripler while greatly reducing the DC power consumption. Also note the relaxed and thus cost-efficient process node (130 nm) of the SiGe technology used here.

### 6.2 Millimeter-wave and Terahertz Systems

*U. Pfeiffer, R. Jain, J. Grzyb and P. Hillger*

With DOTSEVEN technology it becomes conceivable to realize high-speed circuits operating up to fundamental frequencies of 300 GHz and with utilization of higher harmonics (sub-harmonic operation) even beyond the intrinsic cutoff frequency of the active device. This is the portion of the electromagnetic spectrum, where millimeter-wave and terahertz-systems meet, and where advanced SiGe HBT technologies have a wide-range potential. For instance, the RF bandwidth in communication systems is typically in the order of 10% of the carrier frequency, at 300 GHz, this provides a wide absolute bandwidth of 30 GHz, enabling data-rates in the order of tens of gigabits per second. Similarly, future high-precision radars will profit from the abundant bandwidth at frequencies above 200 GHz and terahertz 3D computed tomography (CT) imagers can be entirely implemented in a silicon process technology. The design, simulation, and performance of this emerging application space are described in the following.

The section “240 GHz SiGe Chipset” describes a 240 GHz SiGe chipset for ultra-high data-rate communication at frequencies above 200 GHz. The high \( f_{\text{MAX}} \) achieved in IHPs DOTSEVEN technology enabled the design of high-performance fundamentally operated 240 GHz transmitter (Tx) and receiver (Rx) chip-set fully packaged including an on-chip primary antenna coupled to a secondary low-loss hyper-hemispherical silicon lens antenna. A record data-rate of 40 Gbps for QPSK modulation was demonstrated.

The section “210–270 GHz Circularly Polarized Radar” describes a 240 GHz circularly polarized FMCW radar demonstrator in IHPs DOTSEVEN technology. It shows the highest operational bandwidth and range resolution reported for any silicon-based radar system. The proposed circular polarization concept additionally increases the SNR by 6 dB when compared to conventional radar implementations.

The process improvements in Infineon’s DOTSEVEN technology made it possible to implement an all-silicon terahertz 3D imager demonstrator presented in the section “0.5 THz Computed Tomography.” The main driving motor for this development was to showcase the potential of free-running triple-push oscillator source at around 500 GHz for high-quality absorption
measurements of hidden objects. The sources have been used together with custom asymmetric terahertz detectors to build a 3D terahertz CT system. This demonstrator is able to reconstruct 3D volume renders of hidden objects with an optically limited voxel resolution of around 2 mm × 2 mm. Contrary to previously demonstrated terahertz CT systems that typically use bulky and expensive III–V sources, the demonstrator is comprised solely of hardware fabricated in SiGe HBT technology from the DOTSEVEN project.

### 6.2.1 240 GHz SiGe Chipset

From an application perspective, the frequency upscaling above 200 GHz comes with a lot of benefits. A higher fractional bandwidth and a finer diffraction-limited spatial resolution both benefit the numerous applications ranging from high-data rate communication, RADAR imaging, and even spectroscopic characterization of the materials. The implementation of such systems requires wideband RF front-end components and wideband on-chip antennas. In this section, we present a generic 240 GHz Tx and Rx chipset which was developed under the DOTSEVEN project. This differential chipset operates in the quadrature mode, and the frequency of 240 GHz refers to the center frequency of the local oscillator (LO) signal, which was designed to be very wideband and tunable to make the chipset useful for a plethora of applications.

The block diagrams for both the Tx and Rx are shown in Figure 6.19 [Sarm16, Sarm16b]. The LO generation network consists of an active balun, a ×16 frequency multiplier followed by a three-stage power amplifier (PA), and a differential 90° hybrid. The active balun is used for single-ended to differential conversion of the single-ended low-frequency signal (13.75–17.25 GHz) applied from an external frequency synthesizer which drives the succeeding ×16 stage. The ×16 frequency multiplier circuit forms the core of the LO generation network and it consists of four cascaded frequency-doubler stages, which are staggered tuned in frequency to increase the operational bandwidth [Sarm13, Sarm14]. The LO signal thus generated is amplified with a three-stage PA and then passed through a passive wideband 90° hybrid coupler to generate the quadrature signal. In the Tx, the quadrature LO signal is mixed with external quadrature IF signal to generate a wideband RF which is boosted in power with a four-stage PA and is then radiated through an on-chip ring antenna into a hyper-hemispherical silicon lens and subsequently to the free space. Similarly, at the Rx, the RF signal travels from the lens-antenna to a three-stage PA and subsequently to IF down-conversion mixers. The quadrature LO generation network at the Rx is similar to that
6.2.1.1 Wideband LO signal generation
Let us start discussing the design details for this chipset, starting with the LO generation circuitry which forms the core for both the Tx and the Rx. In this design, the frequency multiplication technique is used instead of...
a HF voltage-controlled oscillator (VCO) for LO-signal generation above 200 GHz. This preference was made based on the following reasons:

1. Frequency multipliers offer higher tuning range, higher usable bandwidth, and a flexible phase noise performance compared to the VCOs. At high-frequencies, the overall VCO tuning range is limited by the vector parasitics [Chi13].

2. A wideband tunable LO is also needed to realize a generic chipset which can be used across a spectrum of applications such as high-speed communication, material characterization, imaging, and frequency-modulated continuous wave (FMCW) RADAR [Sarm16, Grz16].

3. The often stated and major drawback of multiplier chains is that they have an overall higher power consumption. However, VCO-based LO sources also need additional frequency dividers, which increase their overall power consumption as well. A free running VCO is otherwise limited to the on–off keying (OOK) modulation with a poor spectral efficiency [Sarm14].

An expanded block diagram of the LO generation network is shown in Figure 6.20. The $\times 16$ multiplier chain is composed of four cascaded doubler stages, each of which is based on the common Gilbert-cell topology where the RF and LO ports are supplied with the same signal for in-phase multiplication to extract the second harmonic. The three-stage PA is composed of pseudo-differential cascode topology which shall be discussed later.

The LO generation circuitry for a generic transceiver chipset must fulfill the following performance requirements:

1. To ensure the generic nature of the chipset, the LO signal must have a large bandwidth. While a system based on fixed or narrowband LO along with wideband mixers can amplifiers is suitable enough for communication applications, other applications such as FMCW radar require a wideband tunable LO source [Grz16].

![Figure 6.20](image-url)
2. As mentioned above, the power-hungry nature of the multiplier chains is often a major concern. The design must limit the power dissipation of the multiplier chain to as low as possible.

3. As the LO generation is based on harmonic extraction; the spectral purity of the multiplier chain is very important. Any spurious tones from the doubler stages reaching the mixer may start corrupting the IF thereby limiting the IF bandwidth.

4. The mixers need a minimum LO drive power of around 1 mW (0 dBm). Therefore, the generated LO signal power must be high enough to manage this power level along with the additional losses in the passive hybrid coupler.

For these reasons, the LO generation sub-system was designed to generate a power of at least 5 dBm over a 3 dB bandwidth of 40 GHz [Sarm14]. To understand the design further, we need to have a look at the circuit description of each component individually.

(A) x16 frequency multiplier

(i) Gilbert-cell frequency doubler

The circuit schematic for Gilbert-cell based unit doubler stage is shown in Figure 6.21. This topology is chosen due to an inherent differential operation and a high conversion gain (CG) as compared to a conventional class-B bias multiplier topology [Sarm11, Hung05, Oje11]. The capacitance $C_{in}$ couples the differential input signal from the transconductance stage (Q1, Q2) to the switching quad (Q3–Q6).

The inductors $L_b$ and $L_c$ are part of the input and output matching networks. These are implemented on-chip with shielded microstrip lines in the top most metal layer with lengths $l_b$ and $l_c$, respectively. Shielded microstrip lines limit the electric field coupling between different parts of the circuits. The values of the matching network elements are also provided in Figure 6.21. As the input and output of each doubler stage from D1 to D4 progressively shift to higher frequency, the design of each stage is optimized along the following guidelines [Sarm14, Sarm16]:

1. The early stages D1 and D2 operate at lower frequencies and therefore the effect of parasitics is less pronounced. This allows for saving of some chip area by omitting the bias inductor $L_b$ entirely in favor of a resistor $R_b$.

2. The transistor stages D1 + D2 are optimized for a high CG, which allows them to operate with a lower LO input power. This minimizes
the LO leakage associated with the inherent asymmetry of Gilbert-cell based frequency doublers, which may otherwise produce the spurious harmonics at the output of the multiplier chain.

3. The transistor sizing is determined at D4. The maximum transistor size is limited to $4 \times (0.96 \times 0.12) \mu m^2$ as any further scaling will require an accurate synthesis of a very small $L_c (< 10 \mu H)$ which is very difficult for an on-chip BEOL environment.

4. For all the doubler stages, the transistor sizes are kept constant (same as D4) to maintain a sufficient interstage drive power. The large transistors benefit the stages D1 and D2 as they lower the inductance $L_c$ required to tune out transistor parasitic capacitance, saving further chip area.

(ii) Interstage matching network

The design of interstage matching network among the stages D1–D4 is very crucial for achieving the desired wideband operation. The matching network must be tuned to the second harmonic of interest from the preceding stage for a doubler operation. Also, higher order even harmonics must be sufficiently attenuated; otherwise, they would exist in the pass-band of subsequent stages.

Another concern is the center frequency alignment between the stages. If the center frequencies of stages D1–D4 are perfectly aligned to consecutive
second-order harmonics with similar relative bandwidth, then the overall multiplier chain frequency roll-off becomes much sharper (as in the case of higher order filters) and thus the net bandwidth is reduced. The overall bandwidth $BW_{\text{overall}}$ of $N$ cascaded stages is related to the bandwidth $BW$ of single stage as $BW_{\text{overall}} = BW \times \sqrt{2^{1/N} - 1}$ [Ana04]. This implies that for a four-stage network, the overall 3 dB bandwidth corresponds to a mere 0.75 dB bandwidth of the individual stages, or the 3 dB bandwidth of the individual stages equates to the overall 12 dB bandwidth.

One trick to mitigate this limitation is to use a staggered frequency tuning, where the stages are deliberately misaligned for an overall smoother frequency roll-off [Sarm13, Sarm14]. The detailed interstage matching network used in this design is shown in Figure 6.22. The doublers D1 and D3 are tuned higher while D2 and D4 are tuned lower and this resulted in a much smoother roll-off beyond the 3 dB point. As shown in Figure 6.23, the peak CG at the output of D1, D2, D3, and D4 are at the frequencies of 35, 55, 130, and 230 GHz respectively. For D1–D2 and D2–D3, the interstage matching is such that the optimum impedance is transformed at the output of D1 and D2 at the second harmonic of interest. Additionally, it ensures that the impedance is low at the fourth and eighth harmonics for D1 (passband of D3 and D4) and at the fourth harmonic for D2 (passband of D4). For this design, the D3 output and the D4 input were matched to a 100-Ω differential impedance for the ease of breakout characterization and interfacing with other circuits. The simulated (large signal) output impedance at the output of each doubler considering the loading of the succeeding stages is shown in Figure 6.24. The low impedance at the undesired harmonics ensures sufficient harmonic rejection. The stagger

![Interstage matching between the doubler stages of the multiplier chain for LO generation. The tuning inductance $L_c$ connected to the collector output is not shown, after [Sarm16]. Other parameter values are mentioned in the table in Figure 6.21.](image-url)
Figure 6.23  Simulated: (a) output power and (b) CG of the individual doubler stages. The doublers D1 and D3 are tuned higher, while D2 and D4 are tuned lower, and this resulted in an overall flat response, after [Sarm16].

Figure 6.24  Simulated impedance at the collector outputs of D1–D4 derived from the large signal S-parameter simulations, after [Sarm16].

frequency tuning between the stages resulted in an overall simulated 3 dB bandwidth of 50 GHz (210–260 GHz) with a peak output power of –2.8 dBm at 240 GHz. The overall multiplier chain along with the active balun at the input consumes about 720 mW of power [Sarm16].

(B) Power amplifier (PA)
The ×16 frequency multiplier is cascaded with a three-stage PA. A detailed schematic of the single stage of this PA is shown in Figure 6.25. The circuit architecture is based on pseudo-differential cascode amplifier. The cascode
topology is a popular choice in high-frequency amplifier design. The use of common-base stage as a load to the common-emitter transistor in a cascode reduces the Miller capacitance, therefore improving the reverse isolation, stability and the ease of impedance matching. Also, the voltage swing for the PA is limited by the base–collector breakdown voltage ($B V_{CBO}$), which is larger than the collector–emitter breakdown voltage ($B V_{CEO}$) in a common-emitter configuration. A higher voltage swing allows for a higher saturated output $P_{sat}$ from the PA [Kerh15].

A general design outline for the PA is provided in [Sarm13b]. The differential configuration leads to a virtual ground at the base of the common-base stage of the cascode amplifier, which enables efficient and compact on-chip layout by relaxing the need for extensive on-chip decoupling capacitors. However, while a differential topology is expected to have a good common mode rejection, the design of a true differential amplifier requires an active tail current source. At frequencies reaching 200 GHz, the impedance of such current source becomes very low rendering it ineffective. Inductor-based current sources are also challenging as their low self-resonance frequency (SRF) limits the maximum synthesizable inductance, and the quarter wave transmission line-based inductors are inherently narrow band. Therefore, in this design, a pseudo-differential architecture is used where the common-emitter terminal is grounded and the base biasing is provided through the current mirrors. It becomes very challenging to implement the switching PAs at frequencies extending beyond 100 GHz due to the transistor parasitics.
In this design, a low power gain of the device at the high operating frequency (beyond $f_{\text{MAX}}/2$) necessitates the use of class-A biasing for the PA.

In Figure 6.25, the emitter area of each of the transistors Q1–Q4 is $8 \times (0.96 \times 0.12) \, \mu\text{m}^2$. The emitter area and subsequently the power gain is therefore limited by the device parasitics. For larger device size, the required matched tuning inductor becomes too small (less than 10 pH) for on-chip implementation. The microstrip line-based inductor TL1, capacitor C1, and the coupled microstrip line CLIN2 are part of the output match, while CLIN1 is part of the input match [Sarm16]. A decoupling capacitor of 100 fF is used at the common base of transistors Q3–Q4 (not shown here) [Sarm16c].

To maximize the power, the optimum load impedance at the collector node ($R_{\text{opt}}$) must ensure a simultaneous maximization of the voltage and current swing. This can be derived from the loadline analysis and depends on the breakdown voltage and the maximum allowable current density [Ref]. The output resistance of the cascode $R_o$ should also be as high as possible to maximize the power delivered to the load [Ref]. However, due to the internal device parasitics, even when the reactance at the output node is tuned out, the output resistance shows a sharp reduction with frequency ($R_o \propto 1/f^2$) as shown in [Sarm13b]. In this case the loadline impedance match becomes inefficient and therefore instead a conjugate matching is used in this design.

For the multistage PA design, the device sizing is generally scaled from input to the output stages for handling progressively increasing RF power levels. However, this also requires a modified interstage matching network for each subsequent stage. In this design, the transistor sizes for all the stages are kept identical, which provides the flexibility to cascade multiple stages based on the gain requirements without a need for altering the interstage matching network. Identical stages also reduce the probability of frequency misalignment between different stages. Note that a four-stage variant of this PA is used in the Tx after the up-conversion mixer to increase the transmit power.

For the interstage matching, the capacitor-coupled LC resonator technique is used [Shek06, Ana04, Nel32]. Here, the coupling capacitor C1 between the stages introduces an additional zero in the passband which improves the overall bandwidth.

The multiplier chain with the three-stage PA was characterized separately in a breakout structure using WR03 220–325 GHz ground-signal-ground (GSG) waveguide probes along with an on-chip wideband (210–280 GHz) Marchand Balun at the output. A DC–40 GHz GDG probe was used to provide a low-frequency (<20 GHz) input signal using an external frequency
synthesizer. An Erickson Calorimeter equipped with a WR03 waveguide taper was used for the absolute output power measurement for two different input LO power levels (–10 dBm and 0 dBm), and the results are shown in Figure 6.26 [Sarm14, Sarm16c]. For a –10 dBm input LO power, the peak output power is 6.4 dBm at 230 GHz, and the 3 dB RF bandwidth is 50 GHz (215–265 GHz). The output power remains constant for an input LO power of 0 dBm below the input frequency of 13.75 GHz due to some spurious harmonic generation. The LO generation network also shows a phase noise degradation of around 25 dB and consumes about 0.74 W of DC power.

(C) On-chip wideband quadrature coupler
For the quadrature operation, the LO form the PA is provided to a 3 dB 90° coupler. A simplified geometry of the on-chip quadrature coupler is shown in Figure 6.27. It is implemented using three buried metal layers. The coupler exploits a combination of broadside coupling between the strip conductors located on different metallization layers and edge coupling between adjacent strip conductors located on the same layers. The design is sized to minimize the propagation loss and equalization of the propagation speed for all propagating modes to ensure maximum operation bandwidth. The coupler operates along with differential 100-Ω grounded coplanar stripline feeds implemented on a thick top metal layer. The coupler was originally designed

![Figure 6.26](image-url)
for a duplex FMCW radar chipset [Grz15], and unlike the conventional couplers, the required circuit connections here favor the placement of Through and Coupled ports on the same side. Therefore, the EM structure is carefully optimized to minimize the layout asymmetry.

When employed in the transceiver chipset, one of the input ports of the coupler is terminated with a matched 100-Ω differential load impedance. This does not result in any additional losses, as the isolation between the input ports is more than –23 dB. The input return loss is less than –26 dB for a wide 160–340 GHz bandwidth (Figure 6.28). The simulated phase and amplitude imbalance between the quadrature output ports within this frequency band are better than 3° and 1.6 dB respectively.

![Figure 6.27](image.png) Simplified metal-level multi-layer geometry for the differential quadrature coupler, after [Grz15].

![Figure 6.28](image.png) Simulated input match at all four ports of the quadrature coupler and isolation between the input ports for a differential excitation. All ports are referred to a 100-Ω differential impedance, after [Grz15].
6.2.1.2 Transmitter building blocks

Other than the common LO generation network, an up-conversion mixer at the Tx converts an external IF signal into the RF signal which is then boosted with a four-stage PA before radiating through the antenna. The relevant design criteria for the Tx thus are wideband RF and IF operation, as well as sufficient output power.

(A) Up-conversion mixer

The up-conversion mixer is based on the double-balanced Gilbert-cell topology due to its inherent differential operation, LO rejection, and high CG [Voin13]. In the circuit shown in Figure 6.29, the switching quads (Q1–Q4, Q7–Q10) are driven by the quadrature LO signal while the transconductance stages (Q5–Q6, Q11–Q12) are driven by the quadrature IF signal. The center taps for both inductors $L_b$ and $L_c$ are used for the DC biasing, where $L_b$ forms the part of LO matching network. Both the inductors $L_b$ and $L_c$ are implemented in TM1 (second from the top) metal layer as microstrip transmission lines [Sarm16c].

Since the LO drive power is limited, minimum-sized transistors were used at the switching quad stages to allow for a stronger switching and to reduce the parasitic capacitances. Here, the transistors with emitter areas $2 \times (0.96 \times 0.12) \ \mu m^2$ and $1 \times (0.96 \times 0.12) \ \mu m^2$ were used for the

![Figure 6.29](image-url) Up-conversion mixer schematic with additional buffer stages for wideband 50-Ω IF matching, after [Sarm16b].
transconductance stages and the switching quads, respectively. Also, buffer stages with a shunt resistance of 50 Ω were added for a wideband match to the external IF. For these buffer amplifiers, the transistors with an emitter area of $4 \times (0.96 \times 0.12) \, \mu m^2$ are used for linearity reasons.

Note that the mixers are not characterized as separate breakouts, and it requires high-power LO above 200 GHz from external signal sources. To simulate this mixer, a 25 MHz, –5 dBm signal was applied at one of the IF channels. For a 240 GHz LO with 5 dBm power at the input of the 90° coupler, the simulated peak CG is 0.7 dB, $P_{sat}$ is –5 dBm, and $OP_{1dB}$ is –8 dBm. The simulated 3 dB IF and RF bandwidths are 38 GHz and 50 GHz, respectively [Sarm16c].

(B) Four-stage PA
The four-stage PA extends on the three-stage PA used in the LO generation network, with one identical additional stage. This PA was characterized as a separate breakout. At 230 GHz, the measured peak small-signal gain is 26 dB, 3 dB RF bandwidth is 28 GHz, and $S_{11} \leq -10$ dB between 215 and 255 GHz. For large signal measurements at 240 GHz, the PA provides a gain of 12.5 dB at compression, and the 1 dB compression points for input and output are –16.5 dBm and 3.7 dBm, respectively. Also, the measured $P_{sat} > 6$ dBm for the 220–260 GHz frequency range. Note that since the drive power of Tx PA is large, the large signal bandwidth is more applicable to communication links, and this is usually larger than the small signal bandwidth. The measured peak power-added efficiency (PAE) at 240 GHz is 1% [Sarm14].

6.2.1.3 Receiver building blocks
Noise is the primary concern for the Rx sensitivity. Therefore, along with a wideband IF and RF, a low NF at the Rx is very much desirable. At the Rx, the three-stage variant of the PA is used as a preamplifier. The center frequencies of both the three-stage and four-stage PAs are similar, and therefore the probability of frequency misalignment between the Tx and Rx becomes very low [Sarm16b]. Also, the small signal bandwidth of the preamplifier decides the Rx bandwidth when it is driven with a low power signal (as in the case of a communication system with Tx and Rx separated over a wide distance resulting in a large path loss). A three-stage PA shows a larger small signal bandwidth as compared to the four-stage PA and therefore it is preferred at the Rx, ensuring a wideband operation.
(A) Down-conversion mixer

The down-conversion mixer at the Rx end, also implemented with a double-balanced Gilbert-cell topology, is shown in Figure 6.30. The RF current at the transconductance stage (Q9–Q10) is fed by the input RF signal through the antenna and the pre-amplification PA, and this is shared between the I and Q switching quads (Q1–Q4 and Q5–Q8, respectively), which in turn are supplied with the quadrature LO signal from the on-chip wideband 90° coupler. Common-emitter buffers with a series resistance of 50 Ω are added at the IF outputs for a wideband match. The transistor sizes for the mixer and the buffers are identical to those of the up-conversion mixer at the Tx. The choice of load resistor $R_c$ determines the trade-off between the CG and the RC time constant-limited IF bandwidth at the collector output. The load resistor of 200 Ω used in this design corresponds to a simulated 3 dB IF bandwidth of 35 GHz and a peak CG of –0.5 dB. For a 33 MHz IF and 240 GHz, 5 dBm LO at the input of the quadrature coupler, the simulation predicts a 3 dB RF bandwidth of 52 GHz and a minimum NF of 14.2 dB. Also, both the simulated minimum LO power and IP$_{1dB}$ are around –2.5 dBm [Sarm16b, Sarm16c].

6.2.1.4 Antenna design

The linearly polarized on-chip antenna in the Tx and the Rx chipset is topologically similar to the differential wire ring topology [Grz12]. It consists of

![Figure 6.30 Schematic for the down-conversion mixer. Here, additional buffer stages were added to have 50-Ω input impedance required for wideband IF matching, after [Sarm16b].](image-url)
two wire semi-rings connected along the center feed. For wideband operation, the feed is non-uniformly tapered using step-wise approximation [Grz17]. It is designed to illuminate a 9 mm-diameter silicon hyper-hemispherical lens through the chip backside. The lens reduces the influence of surface waves on the radiation efficiency and radiation patterns and inherently delivers a high gain to compensate for the high free-space propagation loss. The backside radiation offers significant advantages over the front-side radiation. The bandwidth is no longer limited by the distance of the ground plane (few $\mu\text{m}$) as in the case of front-side radiation. The form-factor reduction is by a factor of 3.3 ($\sqrt{11}$ for silicon), which is 39% less than in the case of front-side radiation ($\sqrt{3.9}$ for silicon-dioxide). Moreover, the ability to mount external silicon lens of different sizes gives the ability to have flexible application specific directivity. The lens extension is chosen to be close to the elliptical position with extension to radius ratio of 34.4%. The antenna provides a differential impedance of 100 $\Omega$ over a very wide bandwidth ($S_{11} < -20$ dB over 180–330 GHz) [Sarm16]. The simulated cross-polarization is below 20 dB for differential operation. By providing a low impedance (4–5 $\Omega$) for the common mode, radiation from the parasitic common-mode signal is minimized. It is also optimized for the minimization of mode conversion (differential to the common mode) and the simulated mode conversion is below 40 dB. The overall directivity of the antenna with the lens is 26.4 dBi at 240 GHz.

6.2.1.5 Packaging and high-speed PCB design
The chip-on-board (COB) technology is used for both Tx and Rx packaging. The entire chip-on-lens assembly is accommodated inside a recess on a PCB, and chip pads are connected to the PCB bond pads through the wirebond process. A heat sink with direct thermal contact to the silicon lens is also added to improve the heat dissipation away from the chip (Figure 6.33).

The high-frequency IF signal is the major concern while designing the PCB. The PCB material should have a low dispersion and low dielectric loss. Therefore, materials with low relative permittivity and low loss tangent are favored. Here, the ROGERS 4350B material from Rogers Corporation is used. This material is designated for high-frequency applications and it shows a permittivity $\varepsilon_R$ and a loss tangent $\tan \delta$ of 3.66 and 0.0037, respectively over DC-20 GHz frequency range. The choice of PCB thickness is a trade-off between the mechanical stability and the maximum allowed line width for the lowest impedance microstrip lines. Since the differential and quadrature IF routing lines must be highly symmetrical, very large trace widths cannot be
tolerated within a reasonable PCB area. Here, a PCB thickness of 0.388 mm is found to be optimum, with a 50-Ω microstrip line corresponding to a 0.718 mm width.

The wirebond connecting the chip to the PCB also limits the IF bandwidth. Therefore, a phase linear wideband matching filter needs to be implemented on the PCB. While the Bessel filters show the most linear phase response, the feasible component values limit the choice to maximally flat or Butterworth filter topology which still provides a more linear phase response as compared to the Chebyshev filters. For this purpose, an 8-section lumped parameter LC filter is synthesized using Richard transformation on an iterative basis [Pozar09], which takes into account the wire bond inductance (Figure 6.31) [Sarm16b].

Figure 6.32 shows the results from the full EM simulation of the filter. For this, the PCB and the chip ground pads are included to accurately model the ground return current. This simulation predicts an insertion loss ($S_{21}$) of $-0.5$ dB in the passband with a 3 dB bandwidth of 15 GHz. The input return loss ($S_{11}$) is less than $-10$ dB for up to 14 GHz and the group delay variation is less than 10% up to 9 GHz.

### 6.2.1.6 Tx and Rx characterization

The chip-micrograph of the Tx and Rx chipset with the on-chip antenna is shown in Figure 6.34 [Sarm16b]. For on-wafer characterization, the Tx
and Rx have an auxiliary balun instead of an on-chip antenna, and are not shown here. A WR03 GSG waveguide probe for the 220–325 GHz band is used to measure the RF output from the Tx and to supply RF to the Rx. Using the Short-Open-Load (SOL) calibration, the loss due to the probe and the waveguide is estimated to be 7.5 dB in this band. A GSG probe (DC-40 GHz) is used to couple the low-frequency LO signal with $-10$ dBm output power to the chip. A 25 MHz signal from a function generator is along with external $90^\circ$ and $180^\circ$ hybrids are used for the differential quadrature IF signal generation at the Tx side, and the output RF power is measured using an Erickson calorimeter. For the Rx characterization, a WR03 VNA extension module is used in the transmit mode as an RF source, and its output power
Figure 6.34 Chip-micrograph of the Tx and Rx chipset. The total chip area including the pads is (a) Tx: 1.613 mm$^2$ (b) Rx: 1.522 mm$^2$. For the on-wafer measurements, an auxiliary balun with an estimated 2.5 dB loss has been added at the output, after [Sarm16b].

is calibrated using the Erickson calorimeter. The down-converted IF power at 33 MHz is measured with the spectrum analyzer.

The on-wafer characterization results are shown in Figure 6.35 [Sarm16b]. The Tx can deliver up to 6 dBm output power at 240 GHz and the 3 dB bandwidth is 40 GHz. The peak CG for the Rx is 11 dB and the NF is 15 dB while the 3 dB RF bandwidth is 28 GHz. The NF is calculated using the direct method [Oje12] under the assumption that the input noise floor is $-174$ dBm/Hz (thermal noise at the room temperature).

For the IF bandwidth characterization, the packaged Tx and Rx (with on-chip antenna and the lens) are placed back to back with a distance of 90 cm. The IF inputs of the Tx and the Rx are connected to a VNA and swept IF

Figure 6.35 On-chip characterization results for (a) the Tx, and (b) the Rx, after [Sarm16b].
measurements are done for a fixed LO input of 240 GHz. The results indicate a 6 dB IF bandwidth of 13 GHz as shown in Figure 6.36 [Sarm16b].

6.2.1.7 Ultra-high data rate wireless communication
The measurement setup for ultra-high data rate wireless communication system is shown in Figure 6.37 [Pedro17, Grz17b]. The fully integrated and packaged Tx and Rx modules were separated by a link distance of 1 m. The differential inputs of the Tx are connected to an arbitrary waveform generator (Tektronix AWG70001A) and the IF outputs of the Rx are connected to a real-time oscilloscope (Tektronix DPO77002SX) using phase-matched cables.

Figure 6.36 Measurement results from the IF bandwidth characterization over a link distance of 90 cm. For this measurement, the LO is fixed at 240 GHz and the measured 6 dB IF bandwidth is 13 GHz, after [Sarm16b].

Figure 6.37 Measurement setup for the high data rate wireless communication with arbitrary waveform generator (Tektronix AWG70001A) and real-time oscilloscope (Tektronix DPO77002SX), after [Pedro17].
The external LO inputs for both Tx and Rx are driven by the same 15 GHz signal from an external synthesizer with a power splitter.

With no channel equalization applied, maximum transmission speeds of 30 Gbps with an EVM of 26% and 50 Gbps with an EVM of 29% were demonstrated for BPSK and QPSK, respectively, using PRBS9 binary sequence. To increase the reliability of the link with a smaller EVM, a second test was performed for reduced transmission speeds, resulting in an EVM of 11% for 25 Gbps and an EVM of 22% for 40 Gbps for BPSK and QPSK, respectively (Figure 6.38). The limitations in the board bandwidth as well as in the Rx RF/LO bandwidth influence the achievable EVM for the tested modulation speeds.

6.2.2 210–270 GHz Circularly Polarized Radar

Contrary to other imaging techniques, high-resolution radar-based imagers are capable of providing significant improvements in the imaging quality thanks to their range-gating capabilities [Coop08, Lian14, Dick04, Quas09, Graj15]. Similar to general-purpose transceivers operating beyond 200 GHz, they feature low integration and are thus not commonly used because they become expensive and space-inefficient [Coop08, Esse08, Bryl13]. Considering the increasing popularity of radar sensors in various high-volume consumer and industrial markets such as health care [Li13], autonomous navigation in robotic platforms [Chen08, Moal14], non-destructive testing [Karp05], and automotive systems [Maur11], the implementation costs with low weight and small form-factor are more and more relevant. By suitable

![Figure 6.38](image-url) Measured eye diagrams for: 25 Gbps BPSK modulation (left); 40 Gbps QPSK modulation (right), after [Pedro17].
combination of microelectronic packaging with silicon technologies, high-integration levels of the complete radars become a reality and will develop in the future into the solution of choice for such sensors. Currently, most of Si-integrated radars are operated below 100 GHz [Shen12, Maur11] in view of the technology limitations.

Within the frame of DOTSEVEN project, a complete highly integrated FMCW homodyne monostatic radar system operating around 240 GHz with a 60 GHz bandwidth and a state-of-the-art 2.57 mm-range resolution was developed. Its RF front-end is implemented in the form of a single chip in a 0.13-µm SiGe HBT technology with $f_T/f_{MAX}$ of 300/450 GHz from IHP. To facilitate a low-cost packaging scheme, the chip further includes a wideband lens-integrated on-chip annular-slot antenna [Grz15, Grzyb15] and is wire-bonded onto a low-cost FR4 printed-circuit board. Despite the expected lower sensitivity [Graj15] of the homodyne monostatic architecture, this radar topology was selected for implementation due to low costs of the accompanying baseband chain and the highest possible integration level on a single chip. As opposed to classical linearly polarized monostatic radar front-ends [Jahn12, Jaes13, Jaes14], this radar employs circular polarization [Kim05, Statn15] for multiple reasons. The first reason is the absence of on-chip circulators separating Tx and Rx paths. This issue is typically solved by using equivalent quasi-circulators made of on-chip directional hybrid couplers such as rat-race [Jahn12]. Such a solution suffers from an excessive 6 dB loss in SNR because of some additional power loss in the terminating loads [Jahn12, Kim05]. As shown in [Statn15], this loss can be gained back by means of a circularly polarized architecture. Circular polarization may further increase detection probability in the presence of wave depolarization [Moal14, Nash16] or reduce the influence of Rx jamming while operating multiple radar sensors simultaneously [Lian07] and of ghost targets in indoor environments [Moal14].

Figure 6.39 presents the radar chip micrograph and its block diagram. Circular polarization is provided by a broadband annular-slot antenna supporting two orthogonal polarizations [Grz15, Grzyb15] driven from a wideband quadrature coupler [Grz15]. The transceiver is implemented in a fully differential configuration. To achieve the fundamental radar operation in a wide frequency range of 210–270 GHz, the LO-generation path is realized with the $\times 16$ multiplier-chain architecture because of the missing appropriate tuning varactors devices. Both Tx and Rx paths share the same LO-generation chain which is driven around 13.1–16.9 GHz at a power level of around 0 dBm.
The LO drive is provided from the printed circuit-board level as a single-ended signal which is then converted to a differential topology by means of an active balun in front of the multiplier chain. The $\times 16$ multiplication factor was selected in view of the limited RF performance of the regular mm-long wire-bonded interconnects. The multiplier chain comprises four cascaded Gilbert-cell frequency doublers [Sarm16] which inherently provide differential operation. The output signal from the LO-path is equally split by a novel differential Gysel power divider to drive both Tx and Rx paths. Compared to the Wilkinson divider, the chosen Gysel power-splitter is capable of providing an improved isolation between its two output ports at the operation frequency. Each of two outputs drives a four-stage power amplifier with a small-signal gain of 14 dB and a $P_{\text{sat}}$ of around 7 dBm [Sarm16]. One of the amplifiers is connected directly to the Tx port of the circularly polarized antenna whereas the other drives the down-conversion mixer. Due to the similar impedance range at the inputs of both amplifiers, the power splitter imbalance can be minimized. Furthermore, the power amplifiers are useful in providing an improved TX-to-RX isolation from the LO-chain side. The down-converting mixer is operated fundamentally and implemented as...
a double-balanced Gilbert-cell topology. In order to minimize the influence of excessive mixer noise on the Rx NF, the receive signal from the antenna output port is pre-amplified with a three-stage PA. Here, the power amplifier was used instead of a regular LNA [Statn15] to maximize both the radar operation bandwidth and the linearity with similar noise performance metrics to that achievable with a silicon-integrated LNA at the operation frequency [Statn15]. Please note that the Rx linearity is crucial for the radar operation because of its monostatic architecture suffering from the TX-to-RX leakage. From previous measurements of the similar Rx paths [Sarm16], an input-referred 1 dB Rx compression point of around –9 dBm sets a reference value for finding the minimum required antenna input match to avoid Rx compression. More advanced adaptive leakage power cancellation techniques [Brook05, Beas90] can be considered in the future to improve the radar performance.

For free-space operation, the transceiver chip is mounted on the back of a high-resistivity hyper-hemispherical silicon lens with the primary on-chip feed antenna aligned with the lens optical axis. Then, the entire chip-on-lens assembly is in turn mounted in a recess of a regular FR-4 PCB surrounded by a large metal plane, as shown in Figure 6.40. The lens volume is crucial for thermal control for the chip dissipating around 1.6 W. Here, the heat is

![Figure 6.40](image.png) Complete radar transceiver module with a copper heat sink and a 9 mm silicon lens; after [Sta15]. The incorporated IR-image indicates that the chip-on-lens assembly is at around 29°C.
transferred to a copper heatsink through the lens attached to the PCB bottom side which stays in direct contact with the lens. The lens further allows an in-door operation of the complete radar module with no additional external optical components because of a significant increase of the antenna effective gain. Moreover, the pointing-direction errors present in a lens-integrated 2-antenna system (bistatic radar) radiating at an angular offset from the lens optical axis are eliminated in the monostatic radar architecture relying on a single on-chip antenna aligned with the lens center. The current radar implementation features a 9 mm-diameter hyper-hemispherical lens with a 1.3 mm extension to maximize the antenna directivity [Fili93]. Such a lens size provides enough volume for cooling the chip to 29°C, as shown in Figure 6.40.

The integration of a high-performance antenna on a silicon chip is one of the most challenging tasks because the typical cross section of a silicon chip comprises only few metal layers embedded in a low-refraction-index BEOL (Back-End-of-Line) dielectric stack, typically only few micrometers thick, which is located on the top of a lossy bulk silicon substrate. With such a dielectric stack, there are basically only two options for implementing on-chip antennas. The first and the most straightforward approach is to use a ground-plane support between the BEOL stack and the substrate to realize the classical microstrip-type antenna radiating to the top of a silicon chip. This solution, however, results in low radiation efficiency and narrow operation bandwidth [Jaes13, Jaes14]. With the ground-support eliminated, electromagnetic waves start penetrating the complete volume of a lossy and electrically thick substrate launching various parasitic modes such as surface waves. This leads to a very poor prediction of radiation characteristics over a large RF bandwidth, parasitic inter-element coupling, and low radiation efficiency. An alternative solution is to apply the so-called lens-integrated on-chip antennas for mm-wave and THz applications [Fili93, Jha14] which was the preferred option also for this work.

For this particular case, an isolation between the transmit path and the receive path over a wide operation frequency range is additionally required which is provided by a circularly polarized antenna. Circular polarization is achieved by suitable combination of a broadband differential quadrature coupler [Grz15] and a novel dual-polarization circular-slot antenna [Grzyb15]; as shown in Figure 6.41. Here, a left-handed polarization (LHCP) is implemented in the Tx path (‘Tx out’ in Figure 6.41), whereas the receive signals reflected in free-space at an odd number of times are directed to the receive port (‘Rx in’ in Figure 6.41) as RHCP waves.
Figure 6.41 A 3-D EM simulation model of the packaged radar module with a silicon chip mounted on the back of a 9-mm lens; after [Sta15]. The chip-on-lens assembly is placed inside a rectangular recess in the PCB and surrounded by a large ground plane. The slot antenna with the differential quadrature coupler in the BEOL dielectric stack of a silicon chip is shown in the magnified view. The transmit port and the receive port are denoted as ‘Tx out’ and ‘Rx in’, respectively.

The implemented slot antenna is capable of supporting two orthogonal polarizations launched by two orthogonal pairs of patch probes located along the slot circumference (inset of Figure 6.41). The antenna is embedded in a 12-µm thick BEOL stack with seven aluminum layers on top of a 150-µm thick lossy substrate with a bulk resistivity of 50 Ωcm. Its transmit and receive ports are interconnected to the corresponding differential outputs of the quadrature coupler with two intermediate 900-µm long T-line sections implementing the mode conversion from a microstrip line configuration on the antenna side to a grounded-coplanar stripline feed on the coupler side. A 10 dB-defined input-impedance operation bandwidth of the standalone slot antenna is very broad and spans between 150 GHz and 500 GHz. The quadrature coupler driving the antenna (inset of Figure 6.41) is realized by exploiting both the side coupling between two adjacent strips and the broadside coupling between two other strips located on different metallization layers. A total coupling length is only 110 µm. In order to ensure broadband operation of the hybrid coupler [Grz15], its layout is implemented by means of three buried metal layers of the BEOL stack to equalize propagation speeds of the relevant modes. For the considered radar operation bandwidth of 210–270 GHz,
the simulated phase and amplitude imbalance of the coupler are within ±0.75° and 0.2 dB, respectively, whereas the isolation between the Tx and Rx ports is superior to –23 dB within 160–340 GHz. A radiation efficiency of around 62–67% within 200–300 GHz was simulated for the complete circularly polarized antenna, comprising the slot radiator and the coupler, radiating through a 150-μm thick lossy Si-substrate into a silicon half-space [Tong94]. The parameters 3.8 \times 10^7 and 0.02 were assumed for metal conductivity and dielectric loss tangent of the BEOL stack, respectively. The complete packaged chip-on-lens assembly, as shown in Figure 6.41, was further EM-simulated in the transmit mode to study the leakage between the Tx and Rx ports through the antenna path in the presence of reflections at the lens aperture. This leakage may potentially cause nonlinear effects in the Rx, an increase in the noise level, or even lead to the Rx saturation [Brook05, Stov92, Ondr81, Pipe95]. From Figure 6.42, the simulated TX-to-RX leakage and return loss at the TX and RX ports are better than –21 dB and –23 dB, respectively. Considering the previously estimated Rx P_{1dB} compression point and the Tx output power, such antenna isolation is not expected to result in the Rx compression. Moreover, the simulated radiation efficiency of the entire packaged lens-integrated antenna is only a few percent lower than that for the corresponding silicon half-space because of the transmit mode of operation.

![Figure 6.42](image.png)  Simulated return loss at the TX port and the TX-to-RX leakage for the complete chip-on-lens packaged assembly from Figure 6.41; data from [Sta15].
The complete radar module was characterized in free-space exploiting the Friis-transmission equation in the antenna far-field zone. The following key parameters were measured: radiation patterns with antenna directivity and axial ratio, transmitted power, and Rx CG and NF. The measurements were conducted for both operation modes: transmit and receive. The antenna directivity in the TX and the RX operation mode was measured to be 25.8–27.8 dBi and 25.9–27 dBi, respectively, for the radar operation frequency range of 210–270 GHz. An exemplary chosen radiation pattern at 270 GHz for the radar module in the transmit mode is shown in Figure 6.43. The pattern shows good beam rotational symmetry with a side-lobe level of around –17 dB. An axial ratio of 1–1.45 and 1–1.35 for the transmit and the receive mode of operation, respectively, was further measured at boresight for 210–290 GHz. The frequency-dependent radiated output power is presented in Figure 6.44, where a different power level for two module orientations (see Figure 6.39 for orientation definition) can be recognized. This difference is predominantly related to the non-ideal antenna axial ratio. A peak radiated power is around 5 dBm and a –10 dB-defined RF bandwidth is 46 GHz (217–263 GHz) [Sta15].

Figure 6.43  Azimuthal view of the antenna co-polar radiation pattern at 270 GHz for the radar module operating in the transmit mode.
Figure 6.44  Frequency-dependent radiated power; data from [Sta15]. Both the total power and the power levels for two orthogonal antenna orientations (‘A-plane’ and ‘B-plane’) are plotted. For plane orientation, please, refer to Figure 6.39.

The CG and the NF of the radar module operating in the receive mode were measured with the pre-calibrated reference power source from OML. It should be noted that this characterization was conducted in the presence of leakage from the Tx chain operating simultaneously, resulting in the noise floor increase of the receive path. The NF was calculated indirectly from the measured CG and the noise power spectral density at the radar baseband outputs because of missing noise standards in the lab equipment at the operation frequency. The frequency dependence of the measured CG and NF is plotted in Figure 6.45. A peak CG of 12.1 dB with the corresponding minimum NF
of 21.1 dB was measured. A −10 dB-defined RF operation bandwidth for the radar receive mode is around 46.3 GHz (214.8–261.1 GHz).

Besides the 210–270 GHz transceiver module, the complete radar system includes an external in-house-developed linear-frequency chirp generator, a set of differential IF amplifiers with a data-acquisition unit, and a MATLAB code for signal post-processing. Its architecture is shown in Figure 6.46.

For maximum range resolution [Meta07], fast and wideband sawtooth up-chirps of high linearity from 13.1 GHz to 16.9 GHz driving the input port of the radar RF module are first generated. The chirps can be made as short as 100 µs but the overall system was optimized for a chirp duration of 2 ms. The chirp generator relies on a hybrid architecture consisting of a direct digital synthesizer (DDS), a phase-locked loop (PLL), and a VCO. Here, the chirp signal from the DDS serves as a reference for the integer-N PLL which then up-converts the reference DDS frequency to the required 13.1–16.9 GHz. The chosen chirper topology combines the benefits of both the PLL [Zhiy13] and the DDS in terms of fast-chirp generation and spectral purity. In particular, the spurious tones from DDS are suppressed by the PLL loop-filter. In comparison with a fractional-N PLL, the DDS-based architecture offers finer frequency resolution [Stel05]. In the current implementation, the chirper is realized as a set of off-the-shelf PCB-mounted components with the DDS circuit clocked at 1 GHz from the signal generator Agilent E8257D. In the PLL, a high tuning range (13–20 GHz) VCO from Sivers IMA is used. In order to minimize the overall PLL phase noise, a comparison frequency of the phase-frequency detector was selected to be as high as possible, resulting in the PLL loop division ratio, N, of 192 (4 × 48). A third-order active filter with a 1.8 MHz low-noise operational amplifier implements the PLL loop filter. With the aid of a behavioral simulation model, the PLL loop bandwidth

Figure 6.46  Architecture of the complete radar under test with a metallic plate located at a distance R; after [Sta15].
was set to 479 kHz with an appropriately high phase margin of the open-loop transfer function (around 58°) for minimum total integrated phase noise of the chirp generator.

In Figure 6.47, an exemplary chosen phase noise at the frequency chirper output around 15 GHz is presented. The total jitter integrated from 10 Hz to 100 MHz is 2.62° rms. The linearity of the generated frequency ramp was verified by means of Hilbert transform [Thro84, Grei93] but only indirectly at the divider output due to limitations in the lab equipment. With this approach, a complex-value analytic signal is first obtained from the real-value time-domain train acquired at the radar IF output. The instantaneous phase of such an analytic signal is then compared with the ideal phase trajectory of a linear frequency chirp and the phase deviation between the two can be computed. The corresponding frequency error results from the rate of change of this phase deviation and its root-mean square value is defined over the chirp duration time. In particular, for a frequency ramp of 210–270 GHz swept over 2 ms, the rms frequency error is below 9 kHz.

In the current radar implementation, beat signals at the IF ports are digitized by an external 16 bit sampling card from National Instruments and then post-processed using MATLAB routines. The card sampling rate is limited to 2 MHz which results in 4,000 samples for 2 ms long chirp period. For the consecutive experiments, a Hanning window will be predominantly applied.

![Figure 6.47 Phase noise of the frequency chirp generator at 15 GHz driven from the frequency synthesizer Agilent E8257D at 1 GHz; after [Sta15]. A frequency of 15 GHz corresponds to 240 GHz for the up-converted signal at the radar RF output.](image-url)
in the FFT data post-processing as a fair compromise between selectivity and resolution [Harr78]. For a 2 ms long frequency chirp, this window corresponds to an equivalent noise bandwidth of 750 Hz.

In order to minimize the influence of the RF front-end non-idealities on the performance of the complete radar system, a 2-step calibration procedure was conducted with the aid of a metallic plate as a single-target reflector at a distance of 80 cm from the radar module, as shown in Figure 6.46. Such a target shows the $1/\lambda^2$ frequency-dependent radar cross section, where $\lambda$ is the free-space wavelength. The first step of the calibration aims at removing the influence of the close-in returns resulting from the limited isolation between Tx and Rx paths which do not depend on the imaged objects and appear as low-frequency IF signals at the Rx output port. This step does not require any reference target to be placed in front of the radar antenna. However, to mimic this ‘no-target’ radar response in the presence of close-proximity reflections in the lab environment coming from insufficient absorber attenuation, the metal plate was tilted by 45° to the boresight of the radar antenna. In the next step, the metallic plate was set back to its perpendicular position with respect to the radar boresight and the Hilbert transform on the acquired calibration signal from the plate was applied to calibrate the influence of parasitic phase and amplitude modulations resulting from the non-ideal RF front-end characteristics. A frequency dependence of the normalized power received from the plate after the Hilbert-transformed time-domain IF calibration train is plotted in Figure 6.48. Considering the $1/\lambda^2$ frequency-dependent radar

![Figure 6.48](image)

**Figure 6.48** Normalized frequency-dependent power received from the calibrating metallic plate after the Hilbert-transformed time-domain IF calibration train; data from [Sta15].
cross section of the metallic-plate, a $-10$ dB-defined bandwidth of around 45 GHz can be estimated for the complete radar transceiver combining the characteristics of both Tx and Rx paths.

For verification purposes of the applied calibration procedure, the beat-signal time-domain trains were further acquired for different positions of the metal plate. Exemplary, the calibrated frequency-dependent radar response to the plate at a distance of 60 cm is presented in Figure 6.49.

![Figure 6.49](image-url)  
**Figure 6.49** Radar response to a metallic plate located at a distance of 60 cm from the radar module; data from [Sta15]. (a) Magnitude response after amplitude calibration only, (b) instantaneous beat frequency after the amplitude and phase corrections. The beat frequency de-embedded from the peak in the IF power spectrum of the return signal is 124.1 kHz (see also Figure 6.50). Both frequency and time units are shown due to duality of the sweep time and the actual RF frequency for a linear FMCW radar.
From the amplitude-corrected magnitude response, it can be noticed that the envelope of the acquired beat signal is almost constant in the frequency range of around 60 GHz but it starts deviating below 220 GHz and beyond 260 GHz. It was verified that two parasitic harmonics with the $\times 14$ and $\times 18$ multiplication factor leaking from the multiplier chain-based LO path are mainly responsible for this behavior which could not be appropriately calibrated. Similarly, the extracted instantaneous beat frequency after amplitude and phase correction steps is influenced by the same harmonic spurs. These harmonic distortions and not the Rx noise floor are primarily limiting the achievable spurious-free dynamic range (SFDR) and the operational bandwidth of the currently implemented radar.

The corresponding FFT-computed IF power spectra of the calibrated beat signal for two RF bandwidths of 60 GHz and 45 GHz are shown in Figure 6.50. The Hanning window was applied in the computation for low spectral leakage and large dynamic range (DR) [Harr78]. Here, similar to the plots from Figure 6.49, the influence of $\times 14$ and $\times 18$ harmonic spurs at 109 kHz and 140 kHz, respectively, located around the main peak at 124.1 kHz can be recognized. Please note that for a reduced bandwidth of 45 GHz, the radar SFDR achieves around $-40$ dBc.

From Figure 6.50, the achievable radar range resolution can be further extracted by means of the so-called point spread function (PSF). For a 60 GHz operation bandwidth, a theoretical range resolution of $c/2B = 2.5$ mm can be calculated for the currently implemented radar, where $B$ is the RF bandwidth and $c$ is the speed of light. This theoretical resolution is, however, of limited practical use. In practice, the ability of distinguishing between two close-proximity targets is more relevant. In this case, the main-lobe full-width at $-6$ dB of the PSF becomes the parameter of interest. For a rectangular weighting function promising the best resolution but with the highest side-lobe level of $-13$ dB, it results in 3 mm. With the Hanning window, commonly applied in imaging for low spectral leakage, this number becomes $2c/2B = 5.0$ mm [Harr78]. A main-lobe full-width at $-6$ dB of 5.14 mm was extracted for the radar implemented here operating with the maximum considered bandwidth of 60 GHz after amplitude and phase corrections and the Hanning window applied, which is close to a theoretical limit of 2.5 mm.

A very simple 2-D scanning optical setup comprising two elliptical collimating and refocusing mirrors, as shown in Figure 6.51, was applied to demonstrate the radar 3-D imaging capabilities. Here, the scanned objects were placed in the focal point of one of the mirrors. A lateral optical resolution of around 1 mm was estimated for this setup with the aid of a
Figure 6.50 The FFT-computed IF spectrum of the calibrated beat signal corresponding to the metallic plate spaced by 60 cm from the radar module for two different operational RF bandwidths: (a) 60 GHz and (b) 45 GHz. For comparison purposes, the chirp duration was varied for both bandwidths to arrive at the same beat frequency of 124.1 kHz. For 60 GHz, it was set to 2 ms whereas for 45 GHz it was reduced accordingly. The influence of harmonic spurs can be identified around 109 kHz and 140 kHz.

simple aluminum pin-type heatsink as a resolution target. For a set of the consecutive imaging experiments, the previously mentioned Hanning weighting function was replaced by the Hamming window offering a slightly improved resolution of 4.65 mm [Harr78] for the full 60 GHz operation bandwidth.

As a scanning object, a 12 cm $\times$ 6 cm large cardboard box with a hidden blister pack of drugs and two missing tablets was chosen. The object was meander-scanned in both directions (X and Y), as sketched in Figure 6.52, and the acquired data was post-processed using 3-D data matrix routines from
6.2 Millimeter-wave and Terahertz Systems

![2-D optical scanning test setup](image)

Figure 6.51 2-D optical scanning test setup for demonstration of the radar 3-D imaging capabilities. A total path length between the radar module and the object is around 780 mm.

MATLAB. Each IF signal burst took 2 ms and was sampled at two MSPS. The range profile was sampled with a resolution of 0.5 mm, whereas the pixel lateral pitch was set to 0.25 mm after interpolation, resulting in an image size of $200 \times 480 \times 240$ voxels for a range-gated distance of ±50 mm around the image center position.

Figure 6.52 presents the exemplary chosen 2-D image of the normalized power received for an object-to-radar distance of 780 mm altogether with the range profiles for two lateral positions across the cardboard which correspond to the present and the missing tablet. It can be noticed that the signal returns correlating with the positions of the lidding seal of aluminum foil, the plastic cavity, and the cardboard box can be identified due to the radar appropriate range resolution and finally the missing tablets can be detected. The corresponding 3-D surface reconstruction of the scanned object can be found in Figure 6.53. Here, the image is formed with a peak-search algorithm which identifies the positions of the highest reflected power for each lateral position and the color-scale represents the normalized received power.

6.2.3 0.5 THz Computed Tomography

Increasing the transistors $f_{\text{MAX}}$ deep into the terahertz frequency range does not only lead to a significant performance improvement for mm-wave and sub-mm-wave circuits, it also contributes to the vision of closing the THz gap with silicon-based circuits. Traditional compound semiconductor-based
Figure 6.52  3-D imaging experiment with the implemented radar module. (a) Cardboard box with a blister pack of drugs with two missing tablets as the scanned object. (b, c) 2-D scan of the normalized power received for an object-to-radar distance of 780 mm altogether with the range profiles for two different X–Y positions across the cardboard. The positions correspond to the present and the missing tablet, respectively. Both acquired range profiles show a DR of around 50 dB.
THz imaging systems tend to be bulky and expensive and thus suffer from a poor price–performance ratio. The advances in SiGe-HBT and CMOS technology continuously increase the device power generation and detection capabilities in the THz frequency range and thus may ultimately leverage the commercial interest in THz imaging systems. Three-dimensional THz imaging based on the principle of computed tomography (THz-CT) is one of the applications that may potentially be explored in commercial environments. THz-CT offers volumetric object reconstruction with an image contrast based on the characteristic THz absorption of the illuminated material. Since THz radiation is non-ionizing and thus requires no dedicated safety measures, THz-CT represents an interesting alternative to X-ray technology for low-cost industrial quality control.

The THz-CT system implemented in this work is solely based on components built in silicon technology that was developed in the frame of DOTSEVEN. Figure 6.54 shows an illustration of the THz-CT scanner. The radiation of a SiGe-HBT source is focused in the object plane and refocused
Figure 6.54 Illustration of the THz-CT scanner. The system comprises a 490 GHz SiGe-HBT source, an NMOS detector, and an optical train based on four $f\# = 2$, 50 mm PTFE-lenses. The object is rotated ($\phi$) and stepped in the 2D object plane ($y,z$).

Figure 6.55 Photograph of the THz-CT scanner.

to an NMOS detector with low-cost PTFE lenses. The transmission through the object is measured along the $y$-axis and for different projection angles to form the sinograms of the object. This process is repeated along the $z$-axis to allow full 3D reconstruction based on a filtered back-projection algorithm. In order to facilitate measurements at multiple projection angles and positions, the location of the object at is computer-controlled by $x \times y \times \phi$ stepper motors. For each position the detector output signal is sampled with a data-acquisition system.

6.2.3.1 Components
There are two components that define the quality of THz-CT systems. Increasing the DR, which is defined as the relation between maximum received signal without an object and the integrated noise over the readout bandwidth, relaxes the trade-off between object thickness, material composition, and scanning time. Secondly, the achievable image resolution is
inversely proportional to the beam spot size in the imaging plane, which is defined by the wavelength and effective aperture and focal length of the optics. Since the output power of silicon-based radiation sources drops significantly when going beyond $f_{\text{MAX}}$, the trade-off between achievable DR and operational frequency becomes the bottleneck for silicon-based THz-CT systems. This fact stresses the need for a high-quality source design that maximizes the radiated power while providing sufficient directivity and Gaussicity.

**A. Source design**

The source in this work was implemented in Infineon DOTSEVEN 0.13 $\mu$m SiGe BiCMOS technology with an $f_T/f_{\text{MAX}}$ of 260/350 GHz [Hill15]. It comprises a broadband lens-integrated circular slot antenna coupled to a single-ended triple-push Colpitts oscillator. Figure 6.56 shows the schematic and the micrograph of the source. An operation frequency of 490 GHz was chosen as a compromise between resolution and power generation capability of the technology. However, the output frequency is still significantly higher than the $f_{\text{MAX}}$ of the technology, necessitating the use of harmonic generation techniques. In this design, a triple-push topology is used to extract the third harmonic at the base terminal of three Colpitts oscillators. The circular slot antenna connected to the common base node loads the circuit at the fundamental oscillation frequency and thus forces the oscillators to run $120^\circ$ out of phase [Tang01]. In this mode, the third harmonic currents of all three oscillators add in phase, while the currents at the fundamental frequency superimpose destructively. Note that the symmetry of the physical design is very important since it directly impacts the extraction efficiency of the third harmonic and the rejection of the fundamental.

![Figure 6.56 Schematic (a) and micrograph (b) of the 490 GHz SiGe-HBT radiator, after [Hill15].](image)
The design procedure for the Colpitts oscillator can be summarized as follows. First, small-signal simulations were used to optimize the transistor size and the feedback capacitor $C_e$ by maximizing the negative resistance at the design frequency of 163 GHz. After that, the tank inductance $L_{T_b}$ was sized to tune out the imaginary part of the transistor input impedance. The series-series feedback introduced by $C_e$ increases the reverse transmission behavior of the circuit and thus the impact of the collector load impedance on the third harmonic matching at the base terminal [Pfei14]. In this design, the maximum third harmonic output power is realized by avoiding the lossy collector via stack and by using an ideal short circuit at the collector terminal. The broadband harmonic idler is realized with three capacitors ($C_{mom}$, $C_{mim1}$, $C_{mim2}$) that are self-resonant at the first, second, and third harmonic.

The lossy silicon die and the strict design rules that are usually enforced upon modern silicon technologies make silicon chips a very unfavourable environment for integrated antennas. At the same time, the requirements for the antenna system in a THz-CT system with a free-running source are high. Process variations and modelling inaccuracies can lead to a shift in the oscillation frequency after manufacturing which makes a broadband antenna design inevitable if a fist-pass design is needed. Additionally, the antenna needs to have a directivity of around 20 dBi to be compliant with low-cost optical components, i.e., 5 mm-diameter PTFE-lenses. These requirements call for a broadband lens-integrated antenna system that is composed of an on-chip primary antenna and a secondary hyper-hemispherical silicon lens. In this design a multi-layer linearly polarized circular slot antenna was used to illuminate a 4 mm-diameter silicon lens. Figure 6.57 pictures the HFSS model used for 3-D EM simulation of the antenna system. The simulation results show a directivity of 22.5 dBi and 86% radiation efficiency.

The source module was fully characterized in free space. The radiation frequency of the source was measured with an 18th harmonic mixer. For the biasing conditions that optimize the radiated power for all supply voltages, the oscillation frequency is close to constant at 490 GHz. The radiated power was measured with a photo-acoustic power meter (TK). Figure 6.58 shows that the source delivers an output power of up to 38 $\mu$W with a DC-to-RF of up to 0.059%.

6.2.3.2 Detector design
The terahertz detector is a zero-bias NMOS detector fabricated in IHP’s 0.13-$\mu$m SiGe-BiCMOS technology. The asymmetric NMOS device is derived from the standard 1.2 V NMOS by a modified layout of the source/drain
The drain side comprises the normal high-dose n+ extension (HDD) and halo implants while the source side is implanted with the low-dose n-extension (LDD) from the 3.3 V I/O devices. Due to the absence of halo implants at the source side, a reduced threshold voltage of the transistor shifts the optimal gate bias point for highest sensitivity to zero volts. Similar to the source design, the detector antenna is comprised of a primary on-chip antenna and a secondary silicon-lens. Figure 6.59 shows the...
measured voltage responsivity and noise-equivalent power (NEP) of the zero-bias NMOS detector for different gate bias voltages. The responsivity peaks at zero bias with 450 V/W and the detector shows an NEP of 80 pW/√Hz.

6.2.3.3 THz-CT results

The THz-CT system can be operated in two modes. A rapid acquisition mode with CW illumination with continuous object rotation and a rotary encoder-based angle allocation can be used for rapid data acquisition. Furthermore, an acquisition mode with a chopped source and stepped object rotation can be used when high accuracy and DR are needed. Figure 6.60 shows the measured DR for different source chopping frequencies for 1 ms lock-in time constant. The system offers 38 dB DR in CW mode and around 60 dB for chopping frequencies higher than 1 kHz. The spot size and related image resolution was measured using the knife edge method [Gonz13]. A knife was mounted to a high-precision translation stage and was moved into spot to block off a 2D plane from further propagation through the optical train. Figure 6.61 shows the normalized measured power received by the detector for y- and z-axis knife translation in the focus point and the corresponding result of the Gaussian fit obtained with

\[ P = \frac{P_{\text{max}}}{2} [1 + \text{erf}(\sqrt{2} \frac{(x - x_0)}{w})], \]
Figure 6.60  Dynamic range of the THz-CT system for different chopping frequencies measured with a lock-in amplifier with 1 ms time constant.

Figure 6.61  Measured and fitted normalized power at the detector for a knife translation in y- and z-directions. The Gaussian beam waists are 2.54 mm in y-direction and 2.40 mm in z-direction.

Figure 6.62  Tomographic reconstruction of a Y-shaped hook driver inside a polyethylene container. The image was recorded with 1 mm spatial and 9° angular resolution within a 250 min acquisition time.
where $P_{\text{max}}$ is the maximum received power, $x - x_0$ is the relative position from the beam center, and $w$ is the $1/e^2$ beam radius [Gonz13]. The resulting Gaussian beam waists in $y$- and $z$-directions are 2.54 mm and 2.40 mm. The values closely resemble the theoretically estimated value of 2.26 mm for an effective lens aperture of 30 mm. Figure 6.62 shows the result of the tomographic reconstruction of a Y-shaped hook driver that is hidden inside a polyethylene container with a size of 26 mm × 48 mm. The image was recorded in stepped acquisition mode with 1 kHz chopping and a spatial and angular resolution of 1 mm and 9°, respectively. The total scanning time in this scanning time is still quite high with 250 min. However, with a continuous object rotation and CW detection, the overall acquisition time can be reduced to as low as 20 min.

References


[Inan14] Inanlou, F., Coen, C. T., and Cressler, J. D. (2014). A 1.0 V, 10–22 GHz, 4 mW LNA utilizing weakly saturated SiGe HBTs for singlechip,


References


