

# Physical design optimization of MOSFETs for millimeter wave and sub-millimeter wave circuits

Qun Jane Gu · Zhiwei Xu · Jenny Yi-Chun Liu

Received: 28 March 2014 / Revised: 28 January 2015 / Accepted: 29 January 2015  
© Springer Science+Business Media New York 2015

**Abstract** Deep-scaled CMOS technologies have provided ultra-high speed devices and enabled the possibility to achieve integrated millimeter wave and sub-millimeter wave system-on-a-chip (SoC). However, the drawbacks and constraints of deep-scaled technologies require not only creative circuit design ideas but also accurate active device models and optimum physical design. The paper presents a systematic design approach to optimize active device physical design and build the corresponding model, which is exemplified through two layout examples. With the optimum active device physical design and new circuit design ideas, we have successfully demonstrated several key circuits in mm-wave/sub-mm-wave frequency ranges: a 450 GHz voltage controlled oscillator, a 200 GHz frequency divider, and a 200 GHz amplifier all in 65 nm bulk CMOS technologies. The availability of these key building blocks would benefit the realization of mm-wave/sub-mm-wave SoCs in the future.

**Keywords** Amplifier · Frequency divider · Millimeter-wave · Oscillator · Sub-millimeter wave · Terahertz

## 1 Introduction

Millimeter wave and sub-millimeter wave circuits and systems attract increasing interest due to their high potential in various applications, such as wireless sensing and imaging, high speed wireless communications, and all-weather radars [23]. mm-wave/sub-mm-wave electronics in III–V technologies have demonstrated impressive performances in key

circuit components and systems [3, 12, 17, 19, 25]. To reduce small form factor and materialize wide deployments, silicon based processes hold high potentials due to their high integration capabilities. Today's advanced technologies provide high speed devices to make the silicon based solutions possible and therefore attract lots of research interest, including high output power signal generation beyond 200 GHz [1, 15, 20, 27], THz frequency synthesizer [2], and silicon based mm-wave/sub-mm-wave circuits and systems [6–9, 14, 16, 22, 24, 26, 28]. All these researches aim to enable integrated mm-wave/sub-mm-wave systems.

To achieve the goal of integrated mm-wave/sub-mm-wave SoC, various key building blocks with low power and small form factor are prerequisite, including mm-wave/sub-mm-wave signal generators and amplifiers. Different circuits have different design goals and may employ devices of various sizes, which necessitates different optimization schemes and approaches. Heydari [10] and Liang [13] have demonstrated several layout design techniques for individual circuits. To generalize the approach for different devices, this paper presents a systematic active device physical design approach and provides design insight for mm-wave/sub mm-wave circuits.

This paper is arranged as follows. Section 2 focuses on the systematic active device layout design and optimization. Section 3 demonstrates several mm-wave/sub-mm-wave circuits with the presented active device physical design and new circuit design ideas. Section 4 concludes the paper.

## 2 MOSFET design optimization

Deep scaled CMOS technologies provide ever-increasing speed devices. However, the speed benefits have been

---

Q. J. Gu · Z. Xu (✉) · J. Y.-C. Liu  
Davis, CA, USA  
e-mail: xuzhw@yahoo.com

significantly offset by large extrinsic parasitics from layout due to continuously reduced channel length. The ratio of extrinsic parasitic capacitance versus device intrinsic capacitance keeps increasing with technology generations [13]. These parasitics significantly reduce device’s cut-off frequencies and possible circuit operating speeds. Therefore, active device physical design and optimization are very challenging to support circuits in mm-wave/sub-mm-wave frequencies, which is close to device cut-off frequencies.

Different circuits normally have different design targets, which may result in different device sizes and device performance metrics. For example, in oscillator design, the switching devices tend to be small for high operating frequencies and good phase noise concern by employing high impedance tank. In power amplifier design, the amplification device is normally large to enable the delivery of large output power. Nevertheless, in mm-wave/sub mm-wave design, there is one common requirement, which is to maximize device cut-off frequencies for ultra high frequency operations. There are two types of cut-off frequency definitions:  $f_T$ , unit current gain frequency, and  $f_{MAX}$ , unit power gain frequency, which are associated with device parameters and can be represented as [29]:

$$f_T \approx \frac{1}{2\pi} \frac{g_m}{C_{gg} + C_{par} + C_{ov}} \tag{1}$$

$$f_{MAX} \approx \frac{f_T}{2\sqrt{(R_i + R_g)(g_{ds} + \pi f_T C_{ov})}} \tag{2}$$

where  $g_m$  is the device transconductance,  $C_{gg}$ ,  $C_{par}$  and  $C_{ov}$  are the gate input capacitance, parasitic gate-bulk capacitance, and gate-drain overlap capacitance.  $R_i$  and  $R_g$  are gate-charging and gate input resistance.  $g_{ds}$  is the output conductance.  $R_i$  and  $g_{ds}$  are determined by device intrinsic characteristics and are layout independent, while  $R_g$  is the device extrinsic resistance and has layout-dependency.  $f_{MAX}$  is more closely correlated to the circuit highest operating frequency than  $f_T$ . For example, there are a few oscillators [5, 18], which demonstrate fundamental oscillation frequency higher than device  $f_T$ , while still lower than  $f_{MAX}$ . It suggests that device  $f_{MAX}$  is indeed the operating frequency limit for circuits and needs to be maximized by minimizing the layout dependent parasitics, which is the focus of this paper.

To provide accurate design insights in mm-wave/sub-mm-wave regime, we need to build layout-aware device models to enable the proper choice of devices in the design since the models from foundries are often not countable in such frequency ranges. On the other hand, it may not be efficient to re-build the device model for each device size. To facilitate accurate circuit design, we have formed a procedure to build a scalable and layout-aware active device model by adding extrinsic parasitics on top of the device core model from foundries. The additional extrinsic parasitics are

extracted from parasitic extraction tools such as Calibre, or EM simulation tools like HFSS and Momentum, and then are modeled based on the device physical layout.

Before building device models, it is beneficial to achieve design insights in how to optimize devices for high performance. The following discussion focuses on this topic. Equation (2) specifies the parameters which determine the device  $f_{MAX}$ . Some of them are determined by the device physical layout. Unfortunately, these parameters normally do not drop together with the constraints of physical design. For example, in order to reduce poly gate resistance, more fingers are often adopted, which in turn results in higher coupling capacitance  $C_{gd}$  and  $C_{gs}$ . In addition, a large number of fingers elongates gate interconnection, which essentially increases the gate resistance and offsets the reduction in the poly gate resistance of each finger. Clearly, a trade-off exists. To better assist design, we need to conduct quantitative analysis on these parameters and derive the corresponding trend. Figure 1(a) shows a device layout with single gate connection and multiple fingers. The gate resistance can be represented as.

$$R_g = R_{acc} + K_{1,Rg}f_n + K_{2,Rg} \frac{R_{cont,p} + R_{via}}{f_n} + \frac{l_{end} + l_{ext}}{l_f f_n} R_{sq,p} + \frac{R_{sq,p} w_f}{3 l_f f_n} \tag{3}$$

where  $R_{cont,p}$ ,  $R_{via}$ ,  $R_{sq,p}$  are poly to metall per contact resistance, metall to metal 2 per via resistance and poly gate sheet resistance per square, and  $f_n$ ,  $l_f$  and  $w_f$  are the device number of fingers, finger length and finger width, respectively. The first term  $R_{acc}$  in Eq. (3) accounts for the metal access resistance, the second term  $K_{1,Rg}f_n$  represents the resistance of the metal parallel with poly and is proportional to the number of fingers  $f_n$ , the third term corresponds to the resistance involved with vias and contacts and inversely proportional to  $f_n$ . The fourth term represents the poly gate access resistance.  $l_{end}$  is the poly length from contact edge to poly extension edge and  $l_{ext}$  is the poly extension length. This resistance contribution is also inversely proportional to  $f_n$ . The fifth term corresponds to the poly resistance related to the channel, which is typically referred to the gate resistance. The number “3” in the denominator accounts for distributed poly gate resistance.

The source and drain resistance can be represented as

$$R_s = w_f \frac{K_{1,Rs}}{\left|\frac{l_n+2}{2}\right|_n} + \frac{R_{cont,d}}{\left|\frac{w_f}{S_{cn}}\right|_n \left|\frac{l_n+2}{2}\right|_n} + \frac{S_{tg}}{w_f f_n} R_{sq,d} \tag{4}$$

$$R_d = w_f \frac{K_{1,Rs}}{\left|\frac{l_n+1}{2}\right|_n} + \frac{R_{cont,d}}{\left|\frac{w_f}{S_{cn}}\right|_n \left|\frac{l_n+1}{2}\right|_n} + \frac{S_{tg}}{w_f f_n} R_{sq,d} + \frac{R_{via}}{\left|\frac{w_f}{S_{cn}}\right|_n \left|\frac{l_n+1}{2}\right|_n} \tag{5}$$

where  $R_{cont,d}$ ,  $R_{via}$ ,  $R_{sq,d}$  are diffusion area to metal1 per contact resistance, metal 1 to metal 2 per via resistance and diffusion area sheet resistance per square, respectively.  $S_{cn}$  and  $S_{tg}$  are the diffusion contact pitch distance and the contact edge to gate distance, respectively, shown in Fig. 1. The operation  $\lfloor x \rfloor_n$  is to obtain the round integer number of the inside value  $x$ . The first terms in Eqs. (4) and (5) correspond to the metal access resistance, which are proportional to finger width and the number of source or drain regions. The second terms are the resistance related to the diffusion to metal1 contact, which are inversely proportional to  $w_f$  and  $f_n$ . The third terms are the resistance corresponding to diffusion area resistance, which is proportional to the product of  $w_f f_n$ . The fourth term in Eq. (5) represents the via resistance from metal1 to metal2 or to the used top metal of the drain. Since  $w_f f_n$  is the total device width, this equation indicates that the source resistance is strongly correlated with the total device width, while demonstrates limited adjustable range when changing the number of fingers.

The extrinsic gate drain, gate source, and drain source capacitance can be represented as:

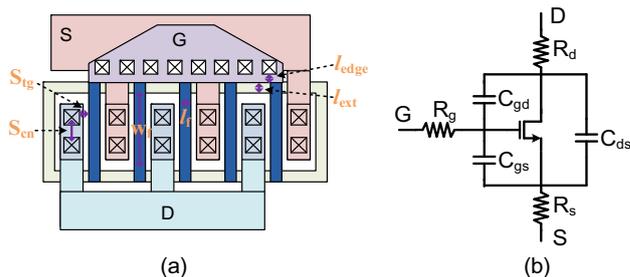
$$C_{gd,ext} = K_{1,Cgd} w_f f_n + K_{2,Cgd} l_f \left| \frac{f_n + 1}{2} \right|_n \quad (6)$$

$$C_{gs,ext} = K_{1,Cgs} w_f f_n + K_{2,Cgs} l_f \left| \frac{f_n + 2}{2} \right|_n \quad (7)$$

$$C_{ds} = K_{1,Cds} W. \quad (8)$$

Both  $C_{gd,ext}$  and  $C_{gs,ext}$  have two similar terms, where the first term represents the coupling capacitance parallel with the channel and the second term stands for the coupling capacitance at the end of channels.  $C_{ds,ext}$  is proportional to device total width  $W = w_f f_n$ . In Eqs. (3)–(8), all the K values are layout and process dependent, which will be derived through fitting. There are also other parameters such as capacitance between gate to bulk, drain to bulk, and source to bulk which are also related to device layout and should be included in the analysis for more accurate analysis.

Given these parameters, we can build a scalable and layout-aware model, as shown in Fig. 1(b), to assist circuit design. There are two ways to obtain these parameters. Approach one is through direct calculation according to the design manual. This approach is not very accurate at high frequencies due to skin effect and generates distributed parasitic network which is too complex for design insights; Approach two is to leverage post layout extraction and EM tools to derive parasitic values to form scalable models. The second approach is adopted in our design. To derive these variables, different device size layouts have been extracted to form simultaneous equations. After the



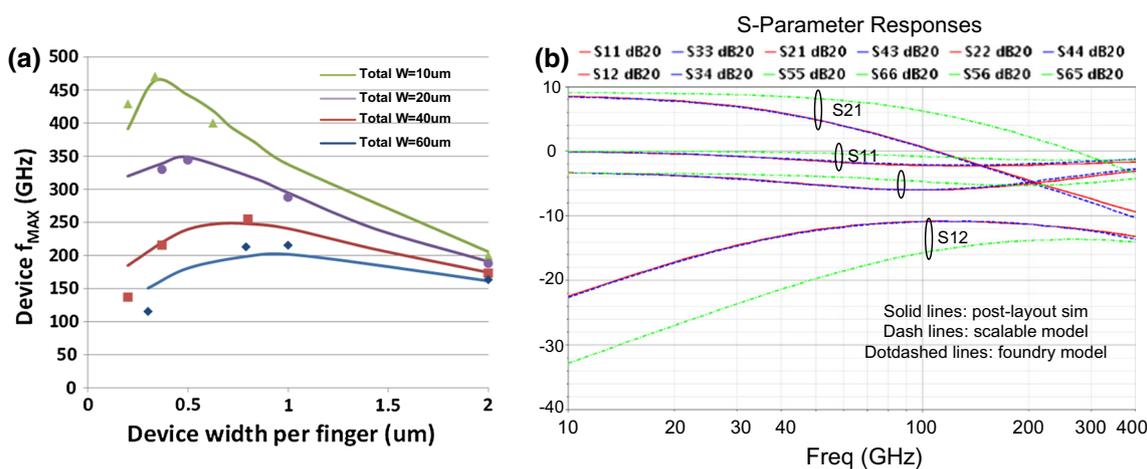
**Fig. 1** **a** A layout representation of a single gate connection MOS device and **b** the corresponding scalable model with the annotated extrinsic parasitics to be determined by the proposed method

parameters are extracted, the parameters are then input into the device parasitic components to form a more accurate model. Since this model is scalable with device size, finger number, it can be used similarly as pcells from foundries to assist circuit design. The parasitic inductances are omitted in the above equations assuming the device size is still significantly smaller than the operating signal wavelength.

The steps to extract device parasitics for scalable models are described as follows:

1. With a device layout, we eliminate all the layers except the extended poly layer and above metal and via layers. This structure constitutes the extrinsic parasitics enclosing the core device.
2. This structure is then evaluated through EM simulation tools (ADS Momentum or HFSS) to obtain an N-port network S-parameter matrix.
3. To achieve the first order estimation of capacitance, we employ the similar derivative equations as Eqs. (1)–(6) in Ref. [13]. First order resistance estimation is based on Eqs. (3)–(5).
4. Starting from the first order estimation, we derive the accurate parasitic network by fitting with the EM simulated N-port network matrix.
5. The steps from (1) to (4) are repeated for two other size devices so that three equations are formed for each parasitic parameter. Given that, all the unknown fitting numbers (up to three) for each parasitic parameter can be derived.

Subsequently, a scalable device model with an extrinsic parasitic parameter network is constituted to facilitate mm-wave/sub-mm-wave circuit design. With these parameters, we can also obtain the optimum device arrangement, such as the number of fingers. To validate the scalable device model, we compare its  $f_{MAX}$  with that of actual devices through post layout extraction and EM simulation characterization on the parasitics. Figure 2(a) presents the simulated device  $f_{MAX}$  from both the derived scalable model and the extracted post-layout circuit elements. Figure 2(b) shows a broad-band S-parameter simulation



**Fig. 2** **a** Device  $f_{MAX}$  based on the created scalable model (lines) and direct extraction value from layout (symbols). **b** S-Parameter simulation results comparison among post-layout extraction based

device (solid lines), scalable model (dash lines), and the core model (dot-dashed lines) from the foundry

results of one device size, 20  $\mu\text{m}/60$  nm, among the three cases: post-layout simulation results, the proposed scalable model simulation results and the core model from the foundry simulation results. It indicates that the proposed model results agree very well the post-layout extracted results, while the results from the core device model are quite different. This proves the effectiveness to use the scalable model in circuit design. Compared with the design procedure starting from inaccurate core device model then iterating the design and device option with post layout extraction and EM simulation tools, the derived scalable model provides more accurate circuit performance estimation during initial design stage, leads to the right optimization direction, and reduces design number-of-iterations significantly. Figure 2(a) also presents the trend that a smaller the device prefers a smaller finger width to achieve the highest  $f_{MAX}$ . This trend is consistent with the analysis from Eqs. (3) to (8). Intuitively, multiple finger structure is employed to reduce the poly gate resistance with the price of increased parasitic capacitance. Therefore, an optimum finger number exists. When the device is small, to maintain the optimum finger number leads to a small finger width. When the device is large, optimum finger width becomes larger. In addition, the  $f_{MAX}$  versus finger width becomes flatter when the device total width increases.

### 3 Millimeter and submillimeter wave circuits

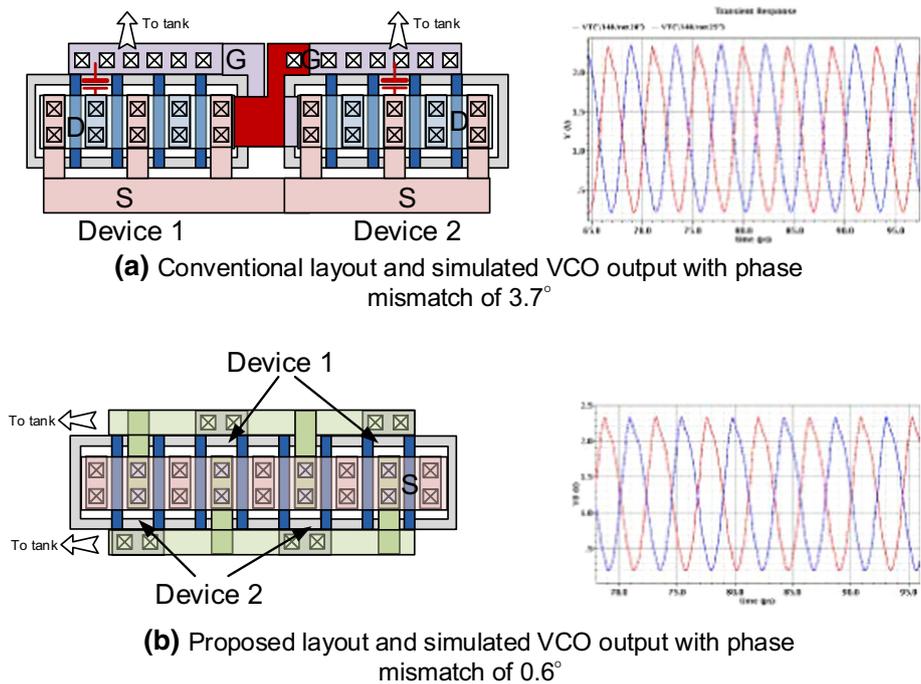
#### 3.1 Terahertz VCO

Different from commercial off-the-shelf (COTS) components that often utilizes 50  $\Omega$  as characteristic impedance, many integrated on-chip high frequency circuits can adopt

high characteristic impedance. Instead of handling power, these on-chip circuits often process signals in voltage domain. Such circuit may prefer small active device size in some circumstances due to several reasons. First, small devices present small loading to itself and the previous stage for a higher operating frequency. Second, small devices allow using large passive components, especially inductors, which normally brings the advantages of high parallel impedance to reduce power consumption and achieve high performance such as phase noise and voltage gain. Third, while small devices cannot deliver large power, they can still provide large voltage swing with a high impedance load, which is a major merit for many on-chip circuits. An oscillator is one of such circuits and prefers small size devices.

In oscillator design, the cross-coupled devices provide the needed negative impedance for oscillation, whose characteristics greatly affect the oscillator performance. Therefore, their physical design needs a special attention. Figure 3(a) shows a conventional cross coupled device layout, which has two issues that may degrade circuit performance. First, the large coupling capacitance between gates and drains erodes capacitive loading room. Second, the crossing lines among the gate and the drain are hard to be matched. One side has smaller resistance  $<1$   $\Omega$ , the other side has much larger resistance, highlighted in red. This large resistance line passes along a narrow trace and goes through only a couple of contacts, which constitutes large parasitic resistance  $>$  a few ohms. These large parasitic resistance and capacitance not only drop  $f_{MAX}$ , but also cause large mismatches. Figure 3(a) shows the simulated voltage controlled oscillator (VCO) output signal phase mismatch of about  $3.7^\circ$ . In order to overcome these issues, we propose a mingled layout structure for the

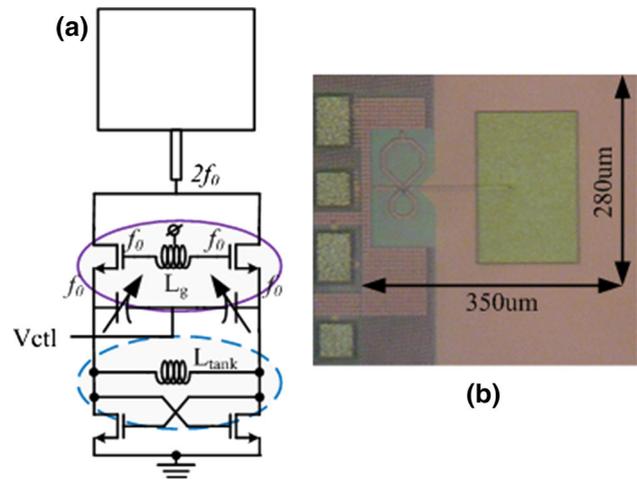
**Fig. 3** **a** Conventional cross coupled active device layout and the VCO output waveforms based on it, which presents phase mismatch of  $3.7^\circ$ . **b** the proposed mingled active device layout, and the VCO output waveforms based on it, which presents phase mismatch of  $0.6^\circ$



cross-coupled pairs, as shown in Fig. 3(b). The two cross coupled devices are mingled into one active area region and grouped together for every two fingers. With this configuration, the crossing connection can be conveniently built with drains directly connecting to gates without crossing interconnect, as shown by green lines in Fig. 3(b). The proposed mingled device layout structure greatly reduces parasitics for high frequency operations and generates symmetrical layout. The VCO based on the proposed layout reduces the output signal phase mismatch to  $0.6^\circ$ . As result, the VCO has a stronger 2nd harmonic, which is the target in our design.

Typically, the circuits' highest operating speed is limited up to a fraction of the device cut-off frequencies. An optimum device physical design in 65 nm CMOS can provide the device's  $f_{MAX}$  around 300 GHz, which may be still not enough to support higher than 200 GHz operations. Therefore, new circuit design techniques are needed to break through device speed limitations. In the oscillator design, we proposed a new oscillator architecture which utilizes a frequency selective negative resistance (FSNR) tank in parallel with a conventional tank to boost both operation frequency and loop gain to achieve a high fundamental oscillation frequency [5].

Figure 4(a) shows the circuit schematic, with the chip photo shown in Fig. 4(b). The oscillator consists of a primary tank with  $L_{tank}$  shunt between the drains of the bottom cross coupled pair (blue dash circle) and a parallel FSNR tank with  $L_g$  shunt between cascode devices (pink solid circle). The FSNR is to provide an equivalent



**Fig. 4** The proposed 450 GHz push-pull VCO with FSNR tank, **a** schematic, **b** the die photo in 65 nm CMOS

inductor together with a negative impedance. This unique feature occurs at high frequencies and matches with our high operating frequency requirements. Therefore, the overall oscillation frequency is higher than each individual tank oscillation frequency and the overall tank impedance is also boosted by the FSNR's negative resistance. This architecture not only allows large  $L_{tank}$  and  $L_g$  inductances to facilitate on-chip design, but also combines them via cascode circuit to form a hybrid tank with a low overall inductance for terahertz operations. The boosted overall tank impedance eases the demand on the cross coupled device's  $g_m$  to permit smaller devices for further higher

oscillation frequencies. In addition, the added FSNR tank vertically shares the same current with the conventional tank and thus does not consume extra power. With the added FSNR, the oscillator can generate a fundamental oscillation frequency higher than device  $f_T$ .

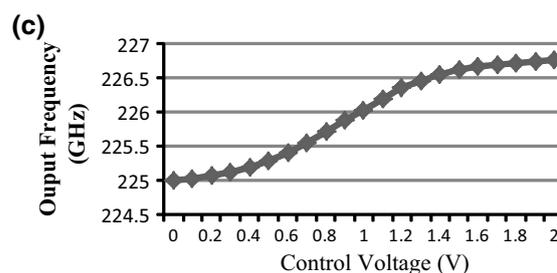
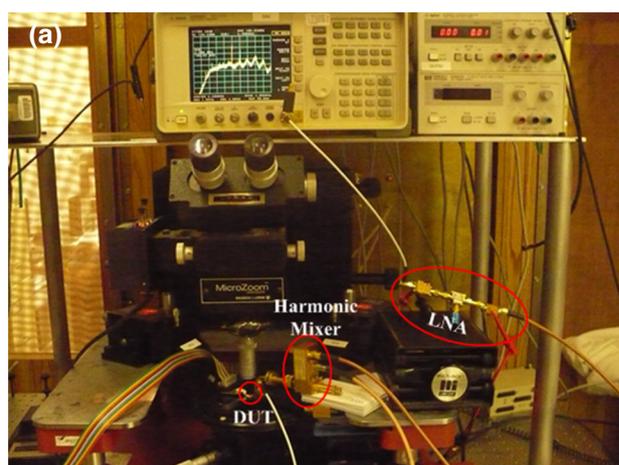
To further push the operation frequency into THz range, push–pull structure is adopted to generate the 2nd order harmonic signal at the drain common mode output [11, 21]. The second order harmonic signal from the output will feed the on-chip patch antenna to radiate over-the-air. A differential varactor is inserted among the differential outputs, which aims for frequency tuning with a higher quality factor  $Q$  than that of single ended connection.

Two measurement approaches are used to characterize the VCO: electronic methods and quasi-optic methods. Electronic methods, as shown in Fig. 5(a), are used to measure the fundamental oscillation frequency and the tuning range. The oscillator output, mixed with a high order harmonic of an external LO, is down-converted to an IF, which is then fed into a low noise amplifier and finally to a spectrum analyzer. By shifting the external LO and measuring the corresponding IF signal shifting frequency, the used LO harmonic order for down-conversion can be derived, which leads to the measured VCO output frequency [5]. As shown in Fig. 5(b), as LO shifts by 5 MHz, the measured IF shifts by 105 MHz. It indicates LO's 21<sup>st</sup> harmonic (i.e.  $21 \times 10.924 = 229.404$  GHz) has been utilized to down-convert the fundamental signal at 225.006 GHz to the IF at 4.398 GHz. Figure 5(c) presents the VCO tuning range, whose fundamental oscillation frequency is from 225 GHz to 226.7 GHz with 1.7 GHz tuning range, by using a small size varactor of  $0.74 \mu\text{m} \times 0.06 \mu\text{m}$ . This VCO draws 5 mA current from 1.4 V power supply.

Interferometer-based quasi-optical approaches are adopted for the 2nd order harmonic signal measurement. As shown in Fig. 6(a), the signal, radiating from the vertically mounted VCO's on-chip antenna, passes through the interferometer and then is detected by a bolometer. The output signal spectrum, recovered through FFT, is shown in Fig. 6(b). The measured spectrum represents un-calibrated power, of which the 2nd order harmonic has a large attenuation from setup and oxygen absorption than the fundamental frequency does. Therefore, the actual output power from the second order harmonic should be large. The fundamental frequency signal radiation may be from the on-chip inductors, which are essentially loop antennas.

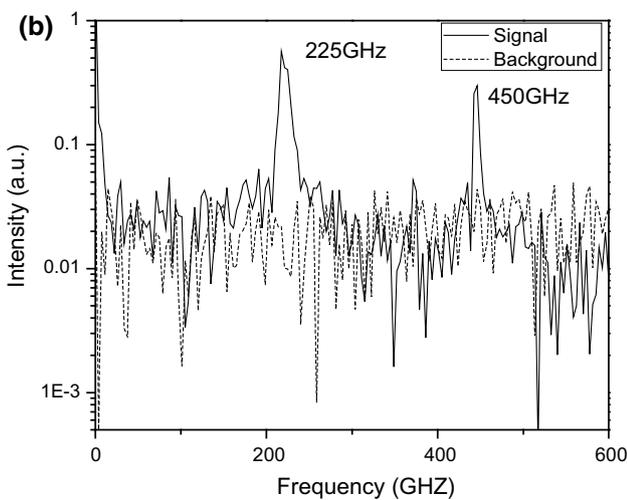
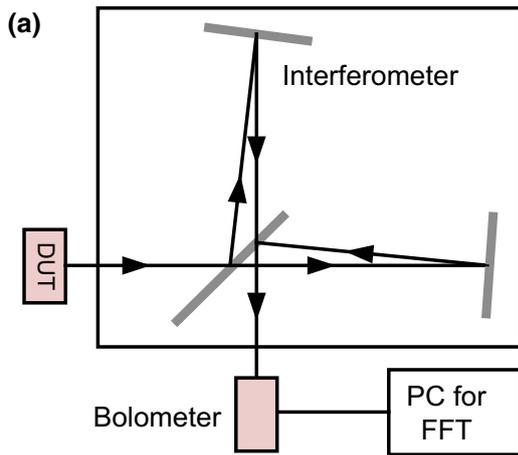
### 3.2 200 GHz frequency divider

To realize stable and controllable signal generation, closed loop synthesizers are needed. In synthesizer circuitries, the first frequency divider following the oscillator is one of the



**Fig. 5** **a** The electronic measurement approach testing setup, **b** the measured IF output with shifting external LO by 5 MHz, and **c** the measured tuning frequency from 225 to 226.7 GHz

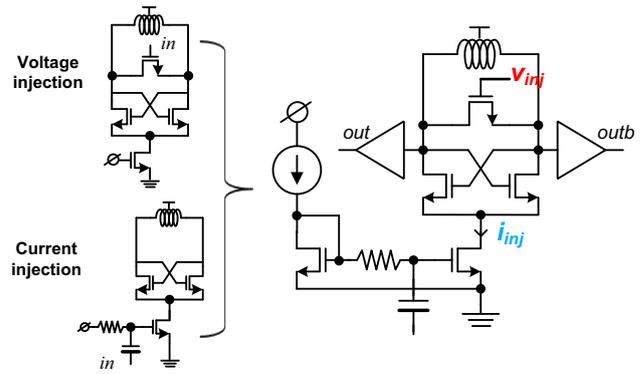
most challenging blocks due to the requirements of both high operating frequency and wide locking range. To achieve larger than 150 GHz frequency division, dynamic structure is typically employed. An injection locking based scheme is a viable approach for high frequency operations. Since an injection locking divider by itself is an oscillator typically formed with cross coupled pairs, it has the same design considerations as those for oscillators. The layout structure of cross coupled pairs in injection locking dividers shares the same approach as the one shown in Fig. 3(b).



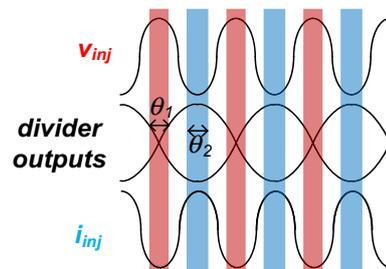
**Fig. 6** **a** Interferometer-based quasi-optical approaches to measure the 2nd order harmonic output, **b** the measured output spectrum with background noise represented by *dotted line*

Besides device physical design optimization, we create a new time-interleaved injection locking architecture to achieve simultaneously high operating frequency and wide locking range [18]. As illustrated in Fig. 7, the input signal is injected to both the top voltage mixing device and the bottom current source device to attain extended injection angles, which leads to a higher oscillation frequency and a wider locking range simultaneously.

Voltage and current injection methods, traditionally exclusive from each other, are integrated by this time-interleaved scheme. Through working at different time periods, these two types of injection are complement to each other to boost injection efficiency, as shown in Fig. 8. For voltage injection, the input signal  $V_{in}$  is injected at the gate of the NMOS mixer that shunts outputs of the cross coupled pair. As the injection voltage increases and the  $V_{gs}$  starts to exceed the device threshold, the mixer turns on and introduces a low impedance path to pull its source and



**Fig. 7** Proposed time-interleaved injection locking scheme based prescaler topology to boost locking range



**Fig. 8** The illustration of boosted injection angles of the time-interleaved injection locking frequency divider, which shows the timing relationship between divider voltage output and injection voltage and current

drain (or the cross coupled outputs) voltages closer. As a result, when voltage injection occurs at an instance outside of the output voltage crossing time period of the prescaler, the voltage injection tends to pull output voltage toward each other. If the prescaler’s natural oscillation frequency is close to half of the injection frequency, such an effect will ultimately align the prescaler’s output frequency and its phase with the voltage injection signal, as shown by the orange area in Fig. 8 with the defined injection angle of  $\theta_1$ . On the other hand, a current signal  $I_{inj}$  is injected via the current source of the cross coupled pair. During the positive (or negative) current injection cycle, the increased (or decreased) source current would split unequally to the resonant tank and increase (or decrease) the voltage difference between prescaler outputs. When the prescaler’s natural oscillation frequency is close to half of the current injection frequency, the prescaler’s output maximum (or minimum) points is synchronized with effective current injection time zones, represented by the current injection angle  $\theta_2$ , blue area in Fig. 8.

With this proposed time-interleaved dual-injection locking scheme, the overall injection strength is enhanced by two means. First is the added injection strength due to both voltage and current injections. Second, and more

importantly, the interleaving injection renders smaller current during voltage injection period that is equivalent to lower the oscillator current  $I_{osc}$ , thus increasing the  $I_{inj}/I_{osc}$  ratio for an extended locking range. This indicates a preferred phase relationship between voltage injection and current injection. Using factors  $\alpha$  and  $\gamma$  to represent this phase relationship, the time-interleaving locking range can be elaborated as

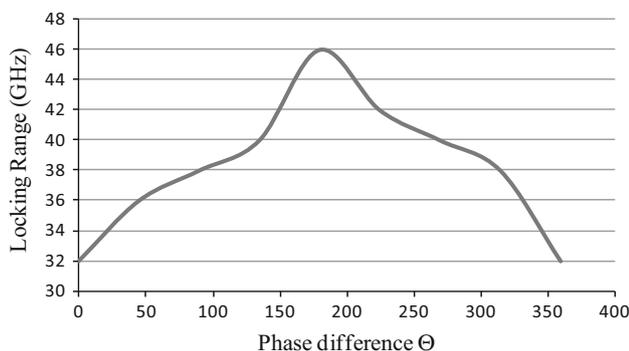
$$\Delta\omega \approx \frac{\omega_o}{2Q} \left( \frac{I_{inj-v}}{\gamma \times I_{osc}} + \frac{\alpha \times I_{inj-i}}{I_{osc}} \right). \quad (9)$$

When voltage and current injections are 180 degree out of phase,  $\alpha/\gamma$  reach their maximum/minimum values respectively, which leads to the maximum divider locking range. When the phase difference deviates from 180 degree, the factors  $\alpha$  and  $\gamma$  deviate from their optimum values and the injection efficiency decreases so that the locking range drops [4]. Figure 9 shows the locking range versus the phase relationship between voltage and current injections.

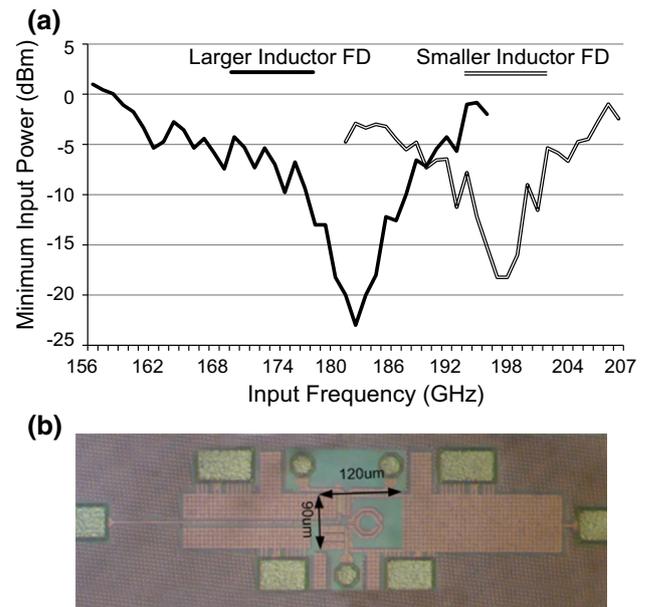
Two prescalers with different inductor values (about 120 and 150 pH, respectively) are implemented in 65 nm CMOS technology. The measured input sensitivities of both prescalers are elucidated by drawing the minimum input power versus the input frequency, shown in Fig. 10(a). The demonstrated locking ranges are over 37 GHz (158–195 GHz, or 21 %) with <0 dBm input power and 27 GHz (181–208 GHz, or 14 %) with <-1 dBm input power, respectively. A chip photo is shown in Fig. 10(b) with the core chip area 0.12 mm × 0.09 mm and the power consumption of 2.4 mW.

### 3.3 200 GHz amplifier

Power amplifiers normally require large size devices to achieve high power delivery and power gain. Therefore, their device optimization is different from the case for small size devices. There are two typical layout structures:



**Fig. 9** Simulated prescaler locking range versus phase relationship between voltage and current injection

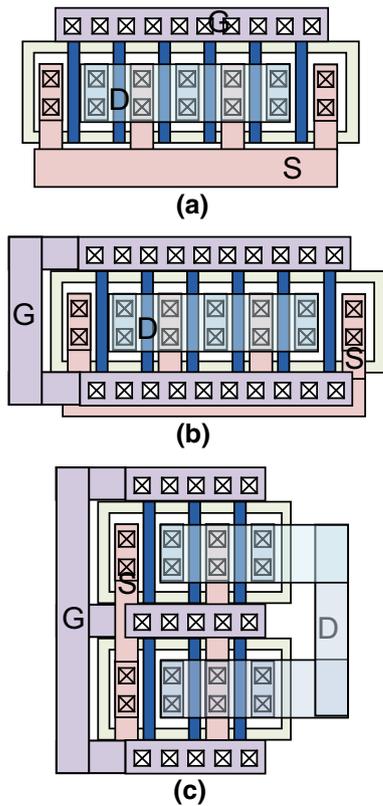


**Fig. 10** **a** Measured input sensitivities of the two prescalers, **b** die photo of the proposed CMOS prescaler in 65 nm CMOS

signal gate connection and double gate connection. Figure 11(a) shows a single gate connection layout structure, and the relative high gate poly resistance along channels hurts circuit performance. Figure 11(b) illustrates a double gate connection, which reduces gate poly resistance along channels. However, the long gate interconnection introduces additional gate parasitic resistance. To address this issue, we propose the third approach, as shown in Fig. 11(c), which folds the device to shorten the gate interconnection parasitic resistance without sacrificing other parameters compared with single gate or double gate connection structures. With 40  $\mu\text{m}$  device, the folded double gate layout structure provides  $f_{\text{MAX}}$  about 285 GHz, versus 218 GHz in simple double gate connection and 243 GHz in single gate connection.

We choose the folded double gate connection layout structure for amplifier circuits and build a scalable device model. Based on the model, we have conducted the first step design and optimization for a 200 GHz amplifier. After that, post layout simulation is utilized for final optimization. The layout-aware scalable model greatly reduces design number-of-iterations. The 200 GHz amplifier schematic is shown in Fig. 12. It features five-stage cascode amplification with pseudo-differential structure to have high noise immunity with the target for system-on-a-chip (SoC) implementations. Transformers T1, T2, T3 and T4 serve as inter-stage matching networks.

The on-chip baluns not only transfer between single ended off-chip signals and differential on-chip signals to facilitate testing, but also serve as 50  $\Omega$  matching networks. Each stage's gate and drain biases are set through



**Fig. 11** a A single gate connection, b a double gate connection, and c the proposed folded double gate connection for large size devices

the transformer/balun center taps individually for independent optimization.

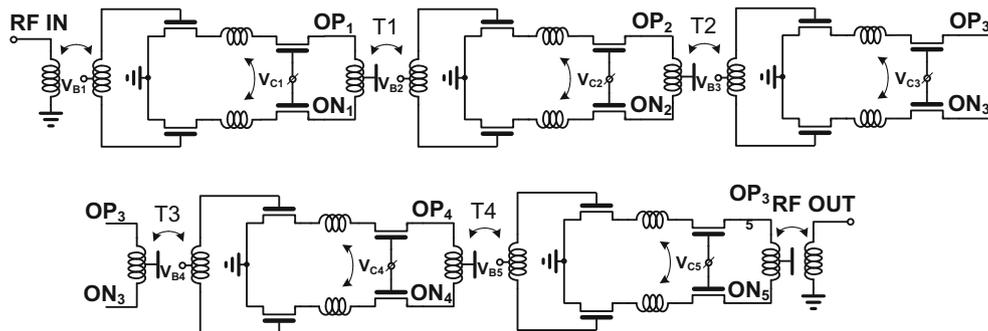
A cascode structure improves stability by increasing the reverse isolation. Unfortunately, it creates a short path to ground through the stray capacitors from the amplification device drain, cascode device source and interconnect parasitics between them. This short path forms a lossy route and degrades amplifier gain and power efficiency significantly in mm-wave/sub-mm-wave frequencies. To mitigate this issue, a series inductor is then inserted between cascode devices to construct a  $\pi$ -matching network together with the device stray capacitors to boost amplifier power gain and efficiency.

The amplifier characterization in such high frequencies is a major challenge due to delicacy of test setups and lacking of instruments. An up to 220 GHz vector network analyzer (VNA) with frequency extension modules is used to characterize this amplifier small signal features. Figure 13 shows the measured small signal S-parameters, which demonstrates a highest gain about 8 dB at around 200 GHz. The positive power gain (larger than 0 dB) frequency range is from 184 to 206 GHz. The S11 is less than  $-10$  dB at around 200 GHz, S12 is less than  $-20$  dB over the entire G band, and the S22 is better than  $-35$  dB for frequency larger than 190 GHz [30].

To measure large signal characteristics, we have used a frequency multiplier chain to generate a 200 GHz signal and a power sensor to detect signal strength. A linearly adjustable attenuator is used to sweep the input power. Figure 14(a) shows the both measured and simulated output power and gain versus input signal power after de-embedding the setup loss. The measured gain is about 7 dB, which is consistent with small signal characterization through a VNA. The Psat and OP1 dB are  $>-10.3$  dBm under 2 V supply, which are mainly limited by the small input signal source and large setup losses. Figure 14(b) shows the measured and simulated PAE. The results in Fig. 14(a) and (b) present consistency between measurements and simulations, which validates the accuracy of the device model. Figure 14(c) presents the die photo of this 200 GHz pseudo differential CMOS amplifier. It is fabricated in 65 nm CMOS technology and occupies  $0.875 \times 0.333$  mm<sup>2</sup> and  $0.68 \times 0.085$  mm<sup>2</sup> with and without pads. The 200 GHz amplifier draws 54 mA from a 2 V supply.

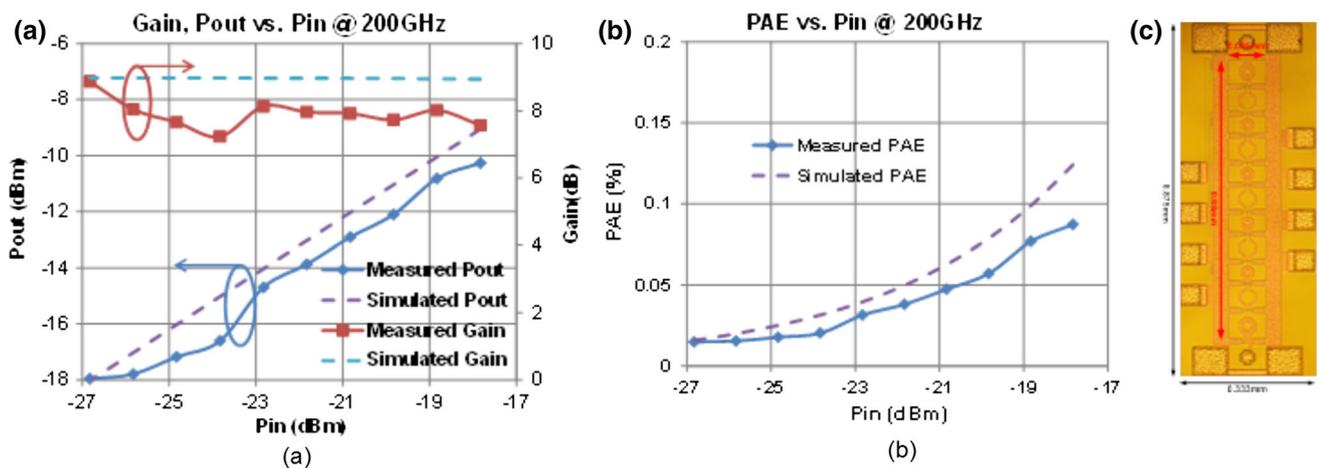
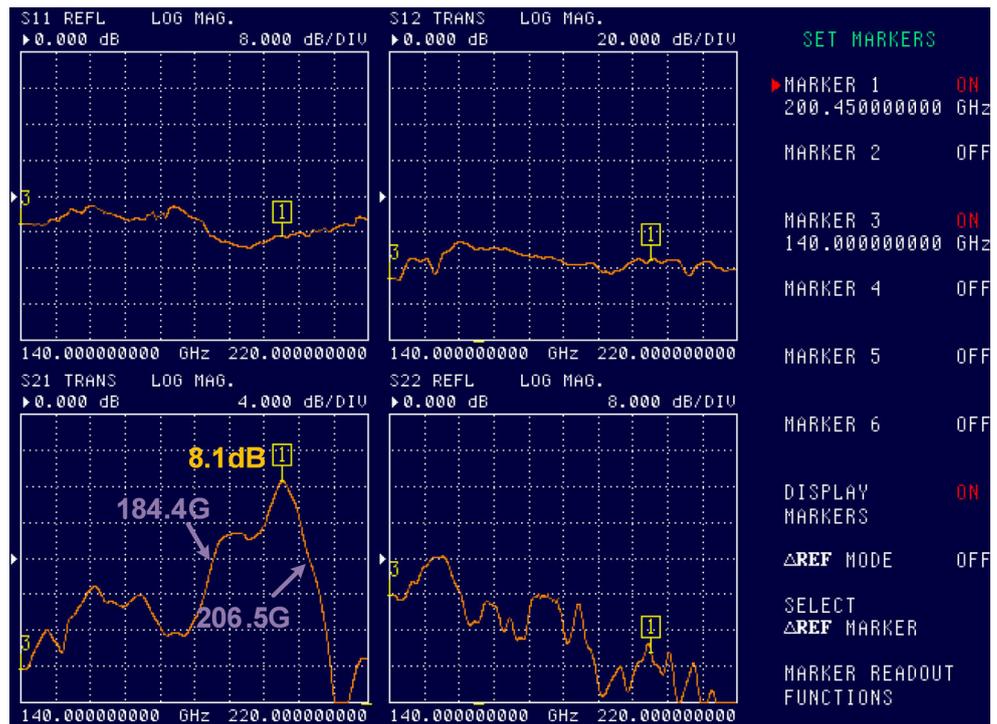
### 4 Conclusions

This paper presents a systematic active device model and layout optimization approach to guide mm-wave/sub-mm-wave circuit design in CMOS technologies. Specifically, the layout-aware active scalable model assists more



**Fig. 12** The 200 GHz fully differential CMOS amplifier schematic

**Fig. 13** Measured S-parameters for the 200 GHz amplifier



**Fig. 14** **a** Measured 200 GHz amplifier pout and gain versus input power (*solid lines*) compared with simulated results (*dash lines*), **b** measured 200 GHz amplifier PAE versus input power (*solid lines*) compared with simulated results (*dash lines*), and **c** its die photo in 65 nm CMOS

accurate design optimization and reduces the number of design iterations between circuit optimization and physical layout. Layout-aware model also facilitates device layout optimization for different circuit blocks. The proposed active device optimization approach is validated by several key mm-wave/sub-mm-wave building blocks in 65 nm CMOS technologies. The 450 GHz VCO, using mingled device layout and the FSNR tank technique, achieves fundamental oscillation frequency higher than device  $f_T$ . The 200 GHz frequency divider, using the created time-interleaved injection locking technique, achieves the ultra-wide locking range. A 200 GHz amplifier is successfully

demonstrated by optimizing device layout to achieve signal gain close to device intrinsic cut-off frequency. It suggests that active device physical design optimization is critical for high performance mm-wave/sub-mm-wave circuits.

## References

- Adnan, M., & Afshar, E. (2014). "A 247-to-263.5 GHz VCO with 2.6 mW peak output power and 1.14 % DC-to-RF efficiency in 65 nm bulk CMOS," *Proceedings of IEEE International Solid-State Circuits Conference (ISSCC)* (pp. 262–263).

2. Chiang, P.-Y., Wang, Z., Momeni, O., & Heydari, P. (2014). "A 300 GHz frequency synthesizer with 7.9 % locking range in 90 nm SiGe BiCMOS," *Proceedings of IEEE International Solid-State Circuits Conference (ISSCC)* (pp. 260–261).
3. Deal, W.R., Leong, K., Zamora, A., Radisic, V., & Mei, X. B. (June 2014). "Recent progress in scaling InP HEMT TMIC technology to 850 GHz," *International Microwave Symposium, IMS2014*.
4. Gu, Q. J., Jian, H.-Y., Xu, Z., Wu, Y.-C., Chang, F., Baeyens, Y., & Chen, Y.-K. (July 2011). CMOS prescaler(s) with maximum 208 GHz dividing speed and 37 GHz time-interleaved dual-injection locking range. *IEEE Transactions on Circuits and Systems-II*, 58(7), 393–397.
5. Gu, Q. J., Xu, Z., Jian, H.-Y., Xu, X., Chang, F., Liu, W., & Fetterman, H. (June 2010). "Generating terahertz signals in 65 nm CMOS with negative-resistance resonator boosting and selective harmonic suppression," *IEEE Symposium on VLSI Circuit* (pp. 109–110).
6. Gu, Q. J., Xu, Z., & Chang, M.-C. F. (2011). "Millimeter wave and sub-millimeter wave circuits for integrated system-on-a-chip," *2011 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT)*.
7. Gu, Q. J., Xu, Z., Jian, H.-Y., Tang, A., Chang, M.-C. F., Huang, C.-Y., & Nien, C.-C. (2011). A 100 GHz integrated CMOS passive imager with >100 MV/W responsivity, 23fW/√Hz NEP. *IET Electronics Letters*, 47(9), 544–545.
8. Han, R., & Afshari, E. (Dec. 2013). A CMOS high-power broadband 260 GHz radiator array for spectroscopy. *IEEE Journal of Solid-State Circuits*, 48(12), 3090–3104.
9. Han, R., & Afshari, E. (June 2012). "A broadband 480-GHz passive frequency doubler in 65-nm bulk CMOS with 0.23 mW output power", *IEEE RFIC*.
10. Heydari, B., Bohsali, M., Adabi, E., & Niknejad, A. M. (Dec. 2007). Millimeterwave devices and circuit blocks up to 104 GHz in 90 nm CMOS. *IEEE Journal of Solid-State Circuits*, 42(12), 2893–2903.
11. Huang, D., LaRocca, T. R., Samoska, L., Fung, A., & Frank Chang, M.-C. (Feb. 2008). "324 GHz CMOS frequency generator using linear superposition technique," *ISSCC Digestive Technology Papers* (pp. 476–477).
12. Lewark, U. J., Zwick, T., Tessmann, A., Massler, H., Leuther, A., & Kallfass, I. (June 2014). "Active 600 GHz frequency multiplier-by-six S-MMICs for submillimeter-wave generation," *International Microwave Symposium, IMS2014*.
13. Liang, C. K., & Razavi, B. (Feb. 2009). Systematic transistor and inductor modeling for millimeter-wave design. *IEEE Journal of Solid-State Circuits*, 44(2), 450–457.
14. May, J. W., & Rebeiz, G. M. (May 2010). Design and characterization of W-band SiGe RFICs for passive millimeter-wave imaging. *IEEE Transactions on Microwave Theory and Techniques*, 58(5), 1420–1430.
15. Ojefors, E., Grzyb, J., Zhao, Y., Heinemann, B., Tillack, B., & Pfeiffer, U. R. (Feb. 2011). "A 820 GHz SiGe chipset for terahertz active imaging applications," *IEEE International Solid-State Circuits Conference* (pp. 224–225).
16. Pfeiffer, U. R., Zhao, Y., Grzyb, J., Al Hadi, R., Sarmah, N., Förster, W., Rucker, H., & Heinemann, B. (2014). "A 0.53THz reconfigurable source array with up to 1 mW radiated power for terahertz imaging applications in 0.13 μm SiGe BiCMOS," *Proceedings of IEEE International Solid-State Circuits Conference (ISSCC)* (pp. 256–257).
17. Radisic, V., Deal, W. R., Leong, K. M. K. H., Yoshida, W., Liu, P. H., Uyeda, J., et al. (July 2010). A 10 mW submillimeter wave solid state power amplifier module. *IEEE Transactions on Microwave Theory and Techniques*, 58(7), 1903–1909.
18. Razavi, B. (Feb. 2011). A 300-GHz Fundamental oscillator in 65-nm CMOS technology. *IEEE Journal of Solid-State Circuits*, 46(4), 894–903.
19. Samoska, L., Fung, A., Pukala, D., Dawson, D., Kangaslahti, P., Lai, R., Sarkozy, S. Mei, X. B., & Boll, G. (June 2011). "On-wafer measurements of S-MMIC amplifiers from 400-500 GHz," *IEEE Symposium of Microwave Theory and Techniques*.
20. Sengupta, K., & Hajimiri, A. (Feb. 2012). "A 0.28 THz 4 × 4 power-generation and beam-steering array," *ISSCC Digestive Technology Papers*.
21. Seok, E., Cao, C., Shim, D., & Arenas, D. J. et. al. (Feb. 2008). "A 410 GHz CMOS push-push oscillator with an on-chip patch antenna," *ISSCC Dig. Tech. Papers* (pp.472–473).
22. Seok, E., Shim, D., Mao, C., Han, R., Sankaran, S., Cao, C., et al. (Aug. 2010). Progress and challenges towards terahertz CMOS integrated circuits. *IEEE Journal of Solid-State Circuits (JSSC)*, 45(8), 1554–1564.
23. Siegel, P. H. (March 2002). Terahertz technology. *IEEE Transactions Microwave Theory Techniques*, 50(3), 910–928.
24. Tang, A., Xu, Z., Gu, Q. J., Wu, Y.-C., & Chang, M.-C. F. (2011). "A 144 GHz 2.5 mW multi-stage regenerative receiver for mm-wave imaging in 65 nm CMOS," *IEEE RFIC symposium*.
25. Tessmann, A., Leuther, A., Massler, H., Hurm, V., Kuri, M., Zink, M., Riessle, M., Stulz, H. P., Schlechtweg, M., & Ambacher, O. (June 2014). "A 600 GHz low-noise amplifier module," *International Microwave Symposium, IMS2014*.
26. Tomkins, A., Garcia, P., & Voinigescu, S. P. (October 2010). A passive W-band imaging receiver in 65-nm bulk CMOS. *IEEE Journal Solid-State Circuits*, 45(10), 1981–1991.
27. Tousi, Y., & Afshari, E. (2014). "A scalable THz 2D phased array with +17 dBm of EIRP at 338 GHz in 65 nm bulk CMOS," *Proceedings of IEEE International Solid-State Circuits Conference (ISSCC)* (pp. 258–259).
28. Uzunkol, M., Golcuk, F., Cetinoneri, B., Atesal, Y. A., Gurbuz, O. D., Edwards, J. M., & Rebeiz, G. M. (June 2012). "Millimeter-wave and terahertz sources and imaging systems based on 45 nm CMOS technology," *International Microwave Symposium, IMS2012*.
29. Woerlee, P. H., Knitel, M. J., van Langevelde, R., Klaassen, D. B. M., Tiemeijer, L. F., Scholten, A. J., & Zegers-van, A. T. A. (Aug. 2001). Duijnhoven, "RFCMOS performance trends". *IEEE Transactions on Electron Devices*, 48(8), 1776–1782.
30. Xu, Z., Gu, Q. J., & Frank Chang, M. C. (2011). A 200 GHz CMOS amplifier working close to device ft. *IET Electronics Letters*, 47(11), 639–641.



**Qun Jane Gu** (S'00-M'07) received the B.S. and M.S. from Huazhong University of Science and Technology, Wuhan, China, in 1997 and 2000, the M.S. from the University of Iowa, Iowa City, in 2002 and the Ph.D. from University of California, Los Angeles in 2007 all in electrical engineering. She received UCLA fellowship in 2003 and Dissertation Year Fellowship in 2007. After graduation, she worked as senior design engineer in Wionics Realtek research group and staff

design engineer in AMCC on CMOS mm-wave and optic I/O circuits. Most recently, she was a postdoctoral researcher in UCLA. From August 2010 to August 2012, she joined University of Florida as assistant professor. After that, she is now an assistant professor at University of

California, Davis. Her research interest spans high efficiency, low power interconnect, mm-wave and sub-mm-wave integrated circuits and SoC design techniques, as well as integrated THz imaging systems.



**Zhiwei Xu** (S'97-M'03-SM'10) received the B.S. and M.S. degree from Fudan University, Shanghai, China and Ph.D. from University of California, Los Angeles, all in Electrical Engineering. He held industry positions with G-Plus Inc., SST communications, Conexant Systems and NXP Inc., where he led development for wireless LAN and SoC solutions for proprietary wireless multimedia systems, CMOS cellular transceiver, Multimedia over Cable (MoCA) system and TV tuners.

He is currently with HRL laboratories, working on various aspects of millimeter and sub-millimeter wave integrated circuit and system, software defined radios, high speed ADC and analog VLSI. He has published in various journals and conferences, one contribution to the encyclopedia of wireless and mobile communications, about ten granted and pending patents.



**Jenny Yi-Chun Liu** received the B.S. degree in Electronics Engineering from National Chiao Tung University, Taiwan, in 2005, and the M.S. and Ph.D. degrees in Electrical Engineering from University of California, Los Angeles (UCLA), in 2008 and 2011, respectively. She is a postdoctoral scholar at UCLA from 2011 to 2012. She joins the department of Electrical Engineering at National Tsing Hua University as an assistant professor in 2012. She interned at

TSMC in 2009 designing RF/millimeter-wave front-end circuits. Her current research interests include millimeter-wave and THz devices, circuits and systems. She was the recipient of the Outstanding Contribution Prize of the Asia-Pacific Microwave Conference, 2010. She was also the recipient of the prestigious 4-year Elite Fellowship for doctoral study from National Science Council, Taiwan, from 2005 to 2009. Physical Design Optimization of MOSFETs for millimeter wave and sub-millimeter wave circuits.