A Systematic Approach to Designing Broadband Millimeter-Wave Cascode Common-Source With Inductive Degeneration Low Noise Amplifiers

Yaolong Hu, *Graduate Student Member, IEEE*, and Taiyun Chi, *Member, IEEE*

***Abstract*— This paper presents a design methodology that can effectively extend the bandwidth of a cascode common-source with inductive degeneration low noise amplifier (LNA), which is one of the most popular LNA topologies in the millimeter-wave bands. Specifically, this methodology addresses how to broaden the input matching bandwidth by realizing dual-resonant S_{11}, and how to extend the gain bandwidth by synthesizing a transformer-based second-order bandpass output network. As a proof of concept, a 27–46 GHz LNA is implemented in the GlobalFoundries 45-nm CMOS SOI process, achieving 25.5–50 GHz 3-dB gain bandwidth, 27–46 GHz return loss bandwidth, 21.2 dB peak gain, 2.4 dB minimum noise figure, and −9.5 dBm peak IIP_3, under 25.5 mW DC power consumption. Consistent performance is measured across multiple samples, demonstrating the robustness of the presented design methodology.**

***Index Terms*— 5G, bandpass network, bandwidth, broadband, CMOS, impedance transformation, inductive degeneration, input matching, low noise amplifier (LNA), millimeter-wave (mmWave), transformer.**

I. Introduction

THERE is a growing interest in developing instantaneously broadband mmWave transceivers that can concurrently support multiple 5G NR bands from 24 to 43 GHz (band n257 – band n261) [1], [2], [3], [4], [5], [6]. This trend is motivated by emerging communication needs such as inter-band carrier aggregation to increase the overall data throughput, global multi-standard coverage to support international roaming, and agile frequency hopping to avoid user interference and congestion. In addition to high-speed wireless communications, the last few years have seen the rise of wireless sensing at mmWave frequencies, which uses electromagnetic transmission and reception for sensing environmental variables, such as gesture estimation [7] and heart rate [8] and respiratory rate monitoring [9]. Wireless sensing is being discussed as an additional function to be supported in cellular networks as smaller wavelengths of mmWave carriers can achieve higher sensing resolution. Much like communications, emerging wireless sensing applications also favor instantaneously broadband transceivers, as the range resolution is inversely proportional to the bandwidth (BW).

As the first stage of the receiver (RX), LNA plays an important role in defining the RX noise figure (NF) and BW. A few broadband mmWave LNA designs have been reported recently. In [3], a resistive feedback technique is presented, achieving 20–40 GHz BW and 2.5–3 dB NF with 18 mW DC power. In [10], a three-stage staggered gain tuning technique is demonstrated, achieving 24–44 GHz BW and 4.2–5.5 dB NF with 58 mW DC power. A dual-path noise cancellation LNA is introduced in [11]. By using a common-gate path and a resistive feedback common-source path, it achieves 22.9–38.2 GHz BW and 2.65–4.62 dB NF. In [12], a 22–32 GHz LNA is presented based on a multistage transformer-based noise matching technique, achieving 1.7 dB minimum NF in a 22-nm FDSOI process. Another 22-nm FDSOI LNA with a similar BW is reported in [13]. It also utilizes transformer-based input matching to enhance the BW and demonstrates 3.1–3.7 dB NF with 20.5 mW DC power.

The key contribution of this paper is to present a systematic yet intuitive design approach that can turn a conventional cascode common-source with inductive degeneration LNA into a broadband implementation. The presented approach incurs minimal design overhead and NF degradation and can be readily adapted to guide broadband LNA designs in other frequency bands. Specifically, we introduce two circuit innovations to enhance the LNA BW [14] and present a detailed study of their design space. First, we extend the input matching BW by realizing dual resonances for the input reflection coefficient (S_{11}). This is made possible by exploring the intrinsic gate-to-drain parasitic capacitance of the input transistor and the frequency-dependent behavior of the first-stage load impedance – both are often ignored in the conventional input matching analysis of the common-source with inductive degeneration topology. Second, we extend the gain BW by constructing a wideband second-order bandpass output network that can be miniaturized into a single transformer footprint. It naturally absorbs the transformer's non-ideal magnetic coupling, finite winding inductances, and parasitic capacitances while achieving a uniform transimpedance gain across a wide frequency range.

This paper is organized as follows. In Section II, the design procedure and design equations to realize the dual-resonant input matching are discussed. In Section III, the synthesis flow of the transformer-based second-order bandpass output

Manuscript received 21 November 2022; accepted 18 January 2023. Date of publication 30 January 2023; date of current version 31 March 2023. This work was supported in part by the National Science Foundation under Grant CNS-1956297. This article was recommended by Associate Editor H. Sjoland. *(Corresponding authors: Yaolong Hu; Taiyun Chi.)*

The authors are with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005 USA (e-mail: yh72@rice.edu; taiyun.chi@rice.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSI.2023.3239396.

Digital Object Identifier 10.1109/TCSI.2023.3239396

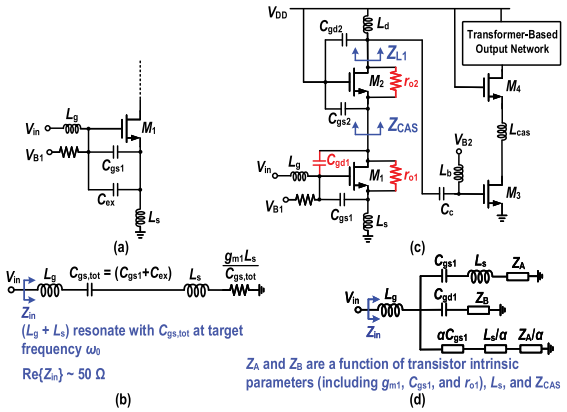


Fig. 1. (a) Schematic of the common-source with inductive degeneration LNA and (b) its input matching equivalent circuit. (c) Schematic of a two-stage cascode common-source with inductive degeneration LNA with C_{gd1}, r_{o1}, and r_{o2} highlighted, which are often ignored in the conventional input matching analysis. (d) Its re-derived input matching equivalent circuit.

network is introduced. Section IV presents a 27–46 GHz LNA design example. Its measurement results are presented in Section V. Section VI concludes this paper.

II. Dual-Resonant Input Matching

A. Common-Source With Inductive Degeneration LNA and Its Bandwidth Bottleneck

One of the most widely used LNA topologies at RF and mmWave is the common-source with inductive degeneration (Fig. 1) [15], [16], [17], [18], [19], [20], [21], [22], [23], [24]. Its equivalent circuit consists of the gate inductor L_g, the gate-to-source capacitor $C_{gs,tot}$ (including the parasitic capacitor C_{gs1} and an explicit capacitor C_{ex}), the source degeneration inductor L_s, and a frequency-independent real part $g_{m1}L_s/C_{gs,tot}$. Ignoring the gate-to-drain parasitic capacitance C_{gd1} and the output impedance r_{o1} of the input transistor, the input impedance is derived as [25]

$$Z_{in} = \frac{1}{j\omega C_{gs,tot}} + j\omega(L_g + L_s) + \frac{g_{m1}L_s}{C_{gs,tot}} \tag{1}$$

The input matching BW is inherently limited since the equivalent circuit only results in a single LC resonance at the target frequency ω_0.

As mentioned earlier, C_{gd1} is usually neglected in the input matching analysis. This is a reasonable assumption at low-GHz radio frequencies, especially in advanced technology nodes, since an explicit capacitor C_{ex} is often needed to increase L_s and decrease L_g [25] so that their values become realizable for on-chip or on-package integration. As a result, the ratio of C_{gd1}:$C_{gs,tot}$ = C_{gd1}:(C_{gs1}+C_{ex}) is quite small, and thus, ignoring C_{gd1} does not compromise the accuracy of the analysis.

However, it is a common practice *not* to add C_{ex} at mmWave, since the values of L_g and L_s can be directly accommodated on-chip. Removing C_{ex} can also achieve a better NF [25]. As such, C_{gd1} becomes comparable to $C_{gs,tot}$, and transistor-level simulations start to deviate from the prediction made by (1). In fact, including the parasitic capacitance of the routing to higher metal layers, C_{gd1}:C_{gs1} is only about 1:2.

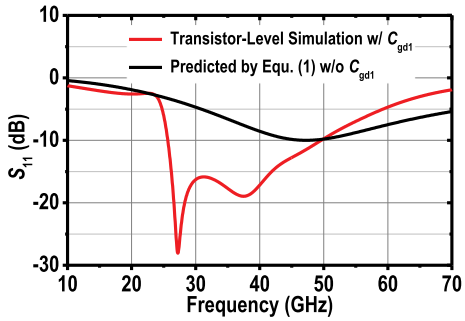


Fig. 2. Transistor-level S_{11} simulation including C_{gd1} [14] and the predicted S_{11} based on (1) without considering C_{gd1}.

Figure 2 shows the predicted S_{11} based on (1) without considering C_{gd1} and the transistor-level simulation including C_{gd1}, using the component values from our reference design [14]. A large discrepancy can be clearly seen in this comparison.

B. Input Matching Equivalent Circuit Including the Gate-to-Drain Parasitic Capacitance of the Input Transistor

To bridge the gap between transistor-level simulations and hand calculations, we re-derive the input matching equivalent circuit to include C_{gd1}, as shown in Fig. 1(d). The equivalent circuit consists of three parallel branches in series with L_g. The impedances Z_A and Z_B in Fig. 1(d) are derived as

$$Z_A = \frac{\omega^2 L_s^2 + g_{m1}L_s r_{o1}/C_{gs1}}{r_{o1} + Z_{CAS} + j\omega L_s} \tag{2}$$

$$Z_B = \frac{r_{o1} + j\omega L_s}{r_{o1} + Z_{CAS} + j\omega L_s} Z_{CAS} \tag{3}$$

where Z_{CAS} is the impedance looking into the source terminal of the cascode transistor M_2. Including the channel-length modulation effect of M_2, Z_{CAS} is given as

$$Z_{CAS} = \frac{r_{o2} + Z_{L1}}{1 + g_{m2}r_{o2}} \tag{4}$$

where Z_{L1} is the load impedance of the first stage.

In Fig. 1(d), the first parallel branch is the same as in the conventional input matching equivalent circuit when r_{o1} is ignored. The second branch models the feed-forward current through C_{gd1}. The third branch is a scaled version of the first branch with a coefficient α, which is derived as

$$\alpha = \frac{\frac{g_{m1}Z_{CAS}}{j\omega C_{gs1}} - Z_{CAS}\frac{\frac{g_{m1}Z_{CAS}}{j\omega C_{gs1}} + j\omega L_s + \frac{g_{m1}L_s}{C_{gs1}}}{r_{o1}+j\omega L_s + Z_{CAS}}}{\frac{1}{j\omega C_{gd1}} + Z_{CAS}\frac{r_{o1}+j\omega L_s}{r_{o1}+j\omega L_s+Z_{CAS}}} \tag{5}$$

Although the expression of α looks quite complicated, α is essentially a function of the transistor intrinsic parameters (including g_{m1}, C_{gs1}, C_{gd1}, and r_{o1}), which are frequency independent, and Z_{CAS}, which is frequency dependent. Our key observation is that Z_{CAS} and the resulting α provide additional degrees of freedom to shape the input impedance Z_{in} over frequency, and thus, it becomes possible to synthesize dual-resonant S_{11} by controlling the frequency response of Z_{CAS}.

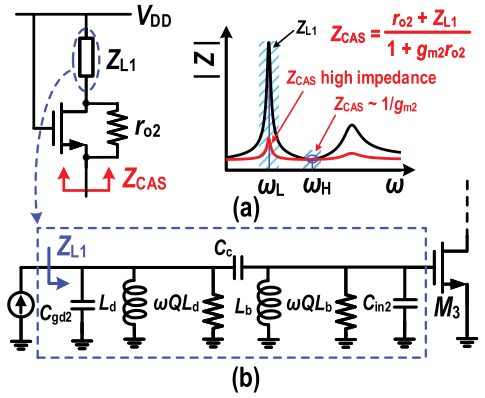


Fig. 3. (a) Z_{CAS} behaves as a high impedance at ω_L but a low impedance at ω_H by properly designing Z_{L1}. (b) Z_{L1} is the impedance looking into the inter-stage capacitively coupled resonator.

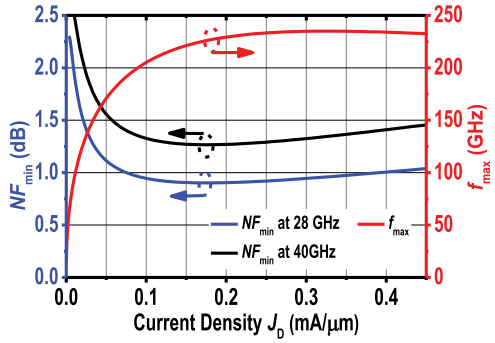


Fig. 4. Simulated NF_{min} and f_{max} versus current density J_D for a floating-body transistor in the GlobalFoundries 45-nm CMOS SOI process.

C. Achieving Dual-Resonant S_{11}

In this sub-section, we present a systematic approach to realizing dual-resonant S_{11} at two target frequencies ω_L and ω_H, respectively. To validate our analysis, we also include numerical calculations and simulation results in this sub-section, based on our reference design [14] using the GlobalFoundries 45-nm CMOS SOI process. Our design target is $\omega_L/2\pi = 27$ GHz and $\omega_H/2\pi = 41$ GHz.

As mentioned earlier, our key idea is to differentiate the value of Z_{CAS} at the two S_{11} resonances, allowing us to optimize the two S_{11} resonances sequentially. Specifically, Z_{CAS} is implemented as a high impedance at ω_L but a low impedance at ω_H [Fig. 3(a)]. This can be achieved by properly designing Z_{L1}, which is the impedance looking into the capacitively coupled resonator between the first stage and second stage, as shown in Fig. 3(b). The detailed design procedure to arrive at the dual-resonant S_{11} is presented as follows.

Step 1: Determine the optimal biasing current density $J_{D,\,opt}$ and the size $(W/L)_1$ of the input transistor M_1.

This step is quite similar to other mmWave LNA designs presented in literature [16]. In our reference design, we first simulate the minimum NF (NF_{min}) and f_{max} against the biasing current density J_D of a 45-nm floating-body NMOS transistor, as shown in Fig. 4. From the simulation, we choose $J_{D,\,opt} = 0.2$ mA/μm, achieving a low NF_{min} and a high f_{max} simultaneously. Under this biasing condition, the maximum transistor size can be determined based on the DC power budget. In our reference design, the width of M_1 is chosen as 55 μm, resulting in 11 mA DC current.

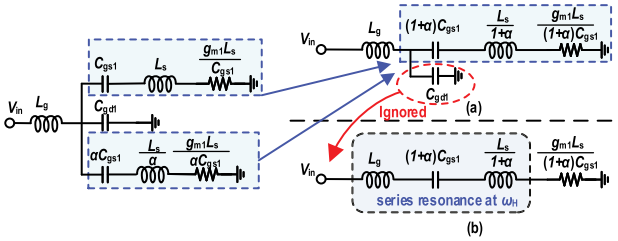


Fig. 5. The input matching equivalent circuit in Fig. 1(d) can be consolidated around ω_H. (a) Including the second branch leads to better accuracy, while (b) removing the second branch simplifies the analysis.

Note that once the biasing condition and the size of M_1 are chosen, its intrinsic transistor parameters, i.e., g_{m1}, r_{o1}, C_{gs1}, and C_{gd1}, are determined. The equations to extract these parameters from the simulated or measured Y-parameters can be found in [26].

Step 2: Synthesize the High-Frequency S_{11} Resonance

The goal of this step is to achieve the high-frequency S_{11} resonance at the target frequency ω_H and to determine the required $(W/L)_2$, L_s, and L_g.

Since Z_{L1} is designed to be a low impedance at ω_H based on our assumption [Fig. 3(a)], Z_{CAS} can be approximated as

$$Z_{L1} \approx 0 \Rightarrow Z_{CAS} \approx \frac{1}{g_{m2}} \text{ at } \omega_H \tag{6}$$

where g_{m2} is the transconductance of the cascode transistor M_2. The output impedance of the input transistor r_{o1} has little effect on the coefficient α at ω_H as $1/g_{m2}$ is much smaller than r_{o1}. Ignoring r_{o1}, the expression of α in (5) can be simplified as

$$\alpha \approx \frac{g_{m1}Z_{CAS}}{j\omega C_{gs1}}/(\frac{1}{j\omega C_{gd1}} + Z_{CAS}) \approx \frac{g_{m1}/g_{m2}}{C_{gs1}/C_{gd1}} \tag{7}$$

From (7), α is purely real around ω_H, so we can consolidate the first and third branches in Fig. 1(d) as a single branch. A low Z_{CAS} also leads to a low impedance for Z_B based on (3) and thus, Z_B in the second branch can be ignored without compromising the accuracy. The input matching equivalent circuit in Fig. 1(d) can then be simplified around ω_H, as shown in Fig. 5(a).

Within the frequency of interest of our reference design (25–50 GHz), the impedance of the second branch $1/j\omega C_{gd1}$ is at least 1.8× higher than that of the first branch. To simplify the input impedance analysis around ω_H and develop design insights, we temporarily ignore the second branch, as shown in Fig. 5(b). In this case, the input impedance presents a single *RLC* series resonance, as

$$Z_{in} = j\omega L_g + \frac{1}{j\omega C_\alpha} + j\omega L_\alpha + R_\alpha \tag{8}$$

where

$$L_\alpha = \frac{L_s}{1+\alpha} \tag{9}$$

$$C_\alpha = (1+\alpha)C_{gs1} \tag{10}$$

$$R_\alpha = \frac{g_{m1}L_s}{(1+\alpha)C_{gs1}} \tag{11}$$

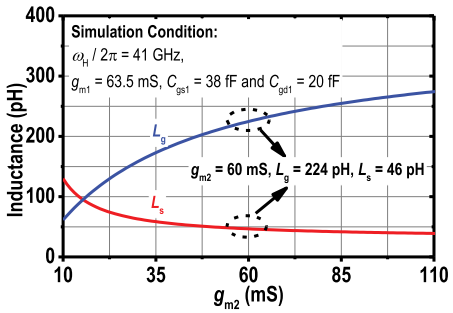


Fig. 6. Calculated L_s and L_g against g_{m2} based on (12) and (13) for our reference LNA design assuming $\omega_H/2\pi = 41$ GHz.

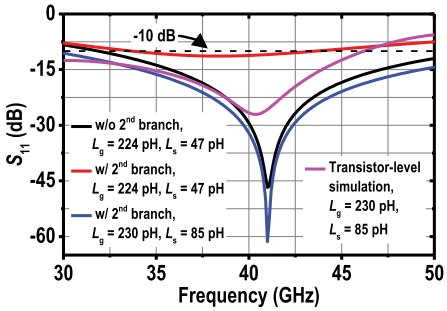


Fig. 7. Predicted S_{11} around ω_H based on the equivalent circuit in Fig. 5 and the transistor-level simulation of our reference LNA design.

Note that the equivalent circuit in Fig. 5(b) is almost identical to that of the conventional common-source with inductive degeneration topology, except the impedances of C_{gs}, L_s, and $g_m L_s/C_{gs}$ are scaled by a factor of $(1+\alpha)$. To realize impedance matching, R_α should be set to be R_s, and the resonant frequency should be set to be ω_H. Based on these two conditions and the simplified expression of α in (7), the unknown inductors L_s and L_g can be solved, as

$$L_s = \frac{R_s C_{gs1}}{g_{m1}}(1+\alpha) \approx R_s\left(\frac{C_{gd1}}{g_{m2}} + \frac{C_{gs1}}{g_{m1}}\right) \tag{12}$$

$$\begin{aligned} L_g &= \frac{1/\omega_H^2}{(1+\alpha)C_{gs1}} - \frac{L_s}{(1+\alpha)} \\ &\approx \frac{(1 - L_s C_{gs1}\omega_H^2)C_{gs1} g_{m2}}{C_{gs1}\omega_H^2(g_{m2}C_{gs1} + g_{m1}C_{gd1})} \end{aligned} \tag{13}$$

For a target resonant frequency ω_H, L_g and L_s are only a function of g_{m2}, since g_{m1}, C_{gs1}, and C_{gd1} are already determined in the *Step 1*. To illustrate this relationship, we plot the required L_s and L_g against g_{m2} based on (12) and (13) in Fig. 6, assuming $\omega_H/2\pi = 41$ GHz. In our reference LNA design, g_{m2} is chosen as 60 mS, which is very close to g_{m1}.

Depending on the target operating frequency, ignoring the second branch as in Fig. 5(b) may lead to compromised accuracy, especially in higher mmWave bands where the impedance of the second branch, i.e., $1/\omega C_{gd1}$, becomes lower. We then perform a more rigorous analysis by including the second branch as in Fig. 5(a). In this case, Z_{in} is given as

$$Z_{in} = j\omega L_g + \frac{1}{j\omega C_{gd1}} \| \left(\frac{1}{j\omega C_\alpha} + j\omega L_\alpha + R_\alpha\right) \tag{14}$$

The real and imaginary parts of Z_{in} are:

$$\mathrm{Re}\{Z_{in}\} = \frac{R_\alpha/(\omega C_{gd1})^2}{R_\alpha^2 + (\omega L_\alpha - 1/\omega C_\alpha - 1/\omega C_{gd1})^2} \tag{15}$$

$$\begin{aligned} \mathrm{Im}\{Z_{in}\} = \omega L_g + \Big[&\frac{-R_\alpha^2}{\omega C_{gd1}} + \frac{2L_\alpha}{\omega C_\alpha C_{gd1}} - \frac{\omega L_\alpha^2}{C_{gd1}} \\ &- \frac{1}{\omega^3 C_\alpha^2 C_{gd1}} - \frac{1}{\omega^3 C_\alpha C_{gd1}^2} + \frac{L_\alpha}{\omega C_{gd1}^2}\Big] \\ &\times \Big[R_\alpha^2 + \Big(\omega L_\alpha - \frac{1}{\omega C_\alpha} - \frac{1}{\omega C_{gd1}}\Big)^2\Big]^{-1} \end{aligned} \tag{16}$$

The required L_s and L_g to realize the input matching at ω_H can be analytically derived by setting (15) to be R_s and (16) to be zero, respectively. However, the calculations can be quite complex. Instead, we can rely on numerical solvers to find L_s and L_g. Similar as shown in Fig. 6, L_s and L_g are only a function of g_{m2}.

Based on the simplified schematic [Fig. 5(b)] and the more rigorous schematic [Fig. 5(a)], we apply the analyses described above to guide the design of our reference LNA and determine the required L_s and L_g. First, we choose the design point according to the simplified schematic and (12)–(13). The predicted S_{11} is plotted in Fig. 7 (the black curve). As we expected, a deep S_{11} resonance is realized at the desired frequency of 41 GHz. Next, we plot the S_{11} using the same L_g and L_s but including the second branch (the red curve in Fig. 7). The resonant frequency down-shifted from 41 GHz to 38.3 GHz, and the depth of S_{11} becomes worse. This aligns with the trade-off we mentioned earlier – the hand calculations and design equations do get simplified when we use the simplified schematic, at the cost of compromised accuracy. To restore a deep S_{11} notch at the target frequency, additional CAD optimizations are needed to fine adjust the values of L_g and L_s. Still, it is always helpful to use the simplest possible analysis to arrive at an initial estimation of circuit parameters and then refine them with optimizations in practical designs.

Alternatively, we can include the second branch in the analysis from the beginning [Fig. 5(a)], if the goal is to achieve an accurate calculation of the design parameters. When we start with the equivalent circuit in Fig. 5(a), the predicted S_{11} (the blue curve in Fig. 7) is very close to the transistor-level simulation (the pink curve) in terms of the frequency and depth of S_{11}. The slight mismatch is due to the finite Z_{L1} in practice, which we have assumed to be zero to simplify our analysis, as shown in (6). Nevertheless, using the equivalent circuit in Fig. 5(a) requires more exhaustive calculations to find the desired L_s and L_g based on (15) and (16).

Step 3: Synthesize the Low-Frequency S_{11} Resonance

Up to this point, we have determined the parameters of the input transistor M_1, the parameters of the cascode transistor M_2, the gate inductor L_g, and the source degeneration inductor L_s. The only undecided circuit parameters in the first stage are the passive elements in the inter-stage capacitively coupled resonator, i.e., C_c, L_d, and L_b in Fig. 1(c) and Fig. 3(b). Their component values determine both the input impedance (Z_{L1}) and the transimpedance gain of the capacitively coupled resonator. The Q in Fig. 3(b) models the quality factor of the inductors, and C_{in2} models the input capacitance of the second stage.

To realize the desired S_{11} resonance at ω_L, the required Z_{L1} can be solved based on the equivalent circuit in Fig. 1(d).

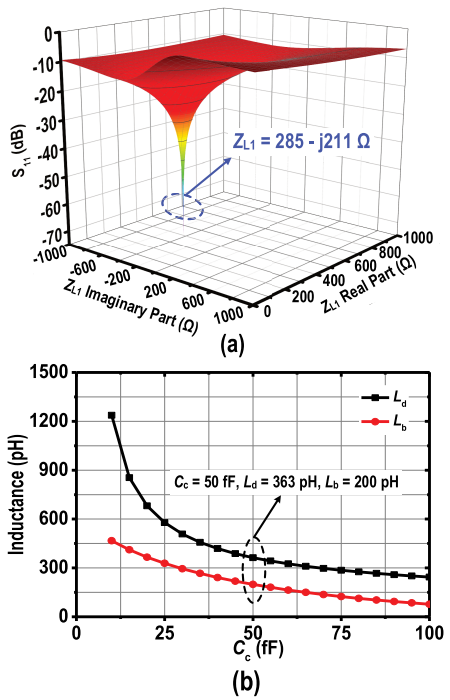


Fig. 8. (a) Calculated S_{11} at ω_L against the real and imaginary parts of Z_{L1} based on the input matching equivalent circuit in Fig. 1(d). (b) Calculated L_d and L_b against C_c to realize the optimum Z_{L1}.

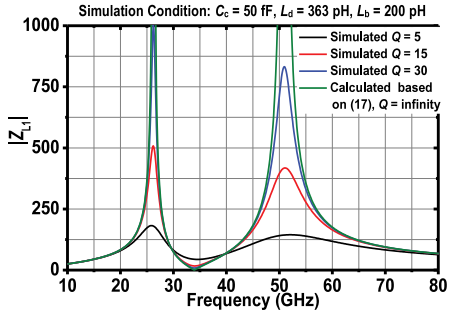


Fig. 9. $|Z_{L1}|$ against frequency for different Q. The two parallel resonant frequencies $\omega_{p1,2}$ and the series resonant frequency ω_s are shown in (18) and (19).

Using our reference LNA design as an example, we plot the S_{11} against the real and imaginary parts of Z_{L1} in Fig. 8(a), assuming $\omega_L/2\pi = 27$ GHz. Based on the optimum Z_{L1}, which is 285 - j211 Ω for our reference design, we can get a family of solutions for L_d, L_b, and C_c, as plotted in Fig. 8(b).

In addition to achieving the desired S_{11} resonance, the selection of L_d, L_b, and C_c is also crucial in shaping the voltage gain of the first stage. This is because a capacitively coupled resonator can realize two parallel resonances and one series

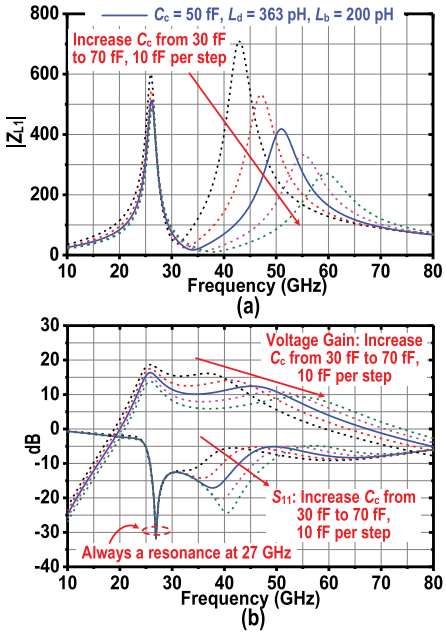


Fig. 10. (a) Simulated Z_{L1} magnitude versus different (C_c, L_d, L_b) solutions. (b) Simulated first-stage voltage gain and S_{11} versus different (C_c, L_d, L_b) solutions.

resonance for its input impedance [27]. If we assume Q is infinite, Z_{L1} can be expressed as (17), shown at the bottom of this page. By setting the denominator and numerator of (17) to be zero, the two parallel resonant frequencies $\omega_{p1,2}$ and the series resonant frequency ω_s can be derived as (18) and (19), shown at the bottom of this page, respectively. Note that Z_{L1} can no longer reach infinity or zero with a finite Q, as illustrated in Fig. 9. Due to the parallel and series resonances, the transimpedance gain and the resulting first-stage LNA voltage gain present two peaks and one dip in between [Fig. 10(b)]. We can see that as C_c becomes larger, the gain difference at ω_{p1} and ω_s also gets larger, and the gain at ω_{p2} becomes lower. Although this gain fluctuation can be compensated using the staggered tuning technique [10], to simplify the network design of the following stages, we only consider (C_c, L_d, L_b) solutions with a <6-dB gain ripple, so that we can assign ω_{p1} as the lower 3-dB cutoff frequency of the following stages to reduce the gain ripple to be <3 dB. This sets an upper bound for C_c. On the other hand, a smaller C_c brings $\omega_{p1,2}$ closer; as a result, we can no longer realize a low impedance for Z_{L1} at ω_H [Fig. 10(a)], and the high-frequency S_{11} resonance disappears [Fig. 10(b)], setting a lower bound for C_c. Considering this trade-off, we choose

$$Z_{L1} = \frac{\omega L_d[\omega^2(C_c + C_{in2})L_b - 1]}{\omega^4 L_d L_b(C_{gd2}C_c + C_{gd2}C_{in2} + C_cC_{in2}) - \omega^2[L_d(C_{gd2} + C_c) + L_b(C_c + C_{in2})] + 1} \tag{17}$$

$$\omega_{p1,2}^2 = \frac{L_d(C_{gd2} + C_c) + L_b(C_c + C_{in2}) \mp \sqrt{[L_d(C_{gd2} + C_c) + L_b(C_c + C_{in2})]^2 - 4L_dL_b(C_{gd2}C_c + C_{gd2}C_{in2} + C_cC_{in2})}}{2L_dL_b(C_{gd2}C_c + C_{gd2}C_{in2} + C_cC_{in2})} \tag{18}$$

$$\omega_s^2 = \frac{1}{L_b(C_c + C_{in2})} \tag{19}$$

C_c to be 50 fF in our reference LNA design, yielding L_d = 363 pH and L_b = 200 pH. The simulated voltage gain peaks at 25.8 GHz and at 45.2 GHz, respectively.

D. Noise Analysis of the Proposed Dual-Resonant S_{11} Technique

For a two-stage cascode common-source with inductive degeneration LNA, its noise is dominated by the channel noises of three transistors – the input transistor M_1, the cascode transistor M_2, and the second-stage common-source transistor M_3. For M_1, only half of its noise current flows to the output when the input matching is realized [25]. Thus, the noise factor of M_1 can be approximated as

$$F_1 \approx \gamma g_{do1} R_s (\frac{\omega}{\omega_T})^2 \tag{20}$$

where γ is the excess noise coefficient, g_{do1} is the zero-bias conductance of M_1, R_s is the source impedance, and ω_T is the angular cutoff frequency. Since the proposed dual-resonant S_{11} technique ensures a good input matching over a wide BW, F_1 is similar to that of a typical narrowband cascode common-source with inductive degeneration LNA.

For the cascode transistor M_2, its noise is typically ignored in low-GHz LNA analysis, because it is degenerated by the output impedance of M_1. However, the noise of M_2 becomes more pronounced as frequency increases. As shown in [28], [29], the noise factor of the M_2 can be approximated as

$$F_2 \approx \gamma g_{do2} R_s (\frac{\omega^2 C_x}{\omega_T g_{m2}})^2 \tag{21}$$

where g_{do2} is the zero-bias conductance of M_2, C_x is the parasitic capacitance at the source of M_2. As the frequency increases, F_2 becomes larger due to C_x. Although the cascode topology provides better reverse isolation, it does have a larger NF than the common-source LNA, especially in high mmWave bands.

For the second-stage common-source transistor M_3, its noise factor is attenuated by the first-stage voltage gain A_{V1} [25], as

$$F_3 \approx \frac{\gamma g_{do3}}{R_s g_{m3}^2 A_{V1}^2} \tag{22}$$

F_3 manifests itself when A_{V1} is low, which happens around the series resonant frequency of the inter-stage network ω_s. As a result, a slight noise penalty is expected around ω_s.

In summary, with the proposed dual-resonant S_{11} technique, we anticipate the NF to be comparable to that of a classic narrowband cascode common-source with inductive degeneration LNA, except for the frequency around ω_s, where a slight NF degradation is expected due to the increased noise contribution of M_3. A generally increased NF over frequency is also expected due to the M_2 noise becoming more significant as frequency goes higher. A detailed NF simulation including the noise matching and noise summary of our LNA prototype is further elaborated in Sec IV.

E. Summary of the Dual-Resonant S_{11} Design Flow

In summary, the design procedure to achieve dual-resonant S_{11} consists of three steps, as shown in Fig. 11. First, the biasing condition and device size of the input transistor are determined based on the DC power budget and simulated NF_{min} and f_{max}, as illustrated in Fig. 4. The intrinsic parameters of the input transistor are also decided after this step. Second, the biasing and device size of the cascode transistor and the values of inductors L_g and L_s are determined based on the target high-frequency S_{11} resonance. Initial parameter estimation can be derived using the simplified schematic in Fig. 5(b) and (8)–(13). Alternatively, a more accurate calculation can be performed using the schematic in Fig. 5(a) and equations (14)–(16). Third, the values of C_c, L_d, and L_b are determined based on the target low-frequency S_{11} resonance. The design equations and trade-offs are summarized in Fig. 8, Fig. 10, and (17)–(19).

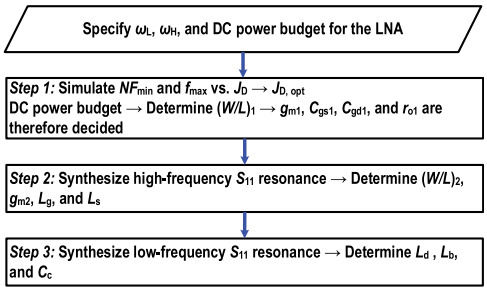


Fig. 11. Design flowchart to realize dual-resonant S_{11}.

Note that the analysis described above ignores the parasitic capacitances of the inductors. This is because different inductor layout styles (i.e., different numbers of turns, radii, metal stack options, etc.) may end up with the same inductance but very different parasitic capacitances; ignoring all parasitic capacitances allows us to simplify the analysis without losing the design intuition and to stay generic without worrying about layout-dependent effects. Once the initial values of the inductors are decided according to the proposed design flow, they can be laid out based on the chip floorplan, and their parasitic capacitances can be extracted and added back to the equivalent input matching schematic to re-derive a new set of parameters. The final components can be arrived at after a few iterations. Meanwhile, CAD optimizations can be performed to optimize the component values.

We'd like to emphasize that our key idea is to leverage the intrinsic gate-to-drain parasitic capacitance of the input transistor C_{gd1} and the frequency-dependent behavior of the first-stage load impedance Z_{L1}, which are often overlooked in the conventional input matching analysis. The presented approach only requires component value updates without the need to modify the cascode common-source with inductive degeneration LNA topology. As such, it introduces minimal design and area overhead when transforming an existing narrowband mmWave LNA design into a broadband implementation.

III. Transformer-Based Second-Order Bandpass Output Network

The effective BW of an LNA is defined as the intersection of the −10-dB S_{11} BW and the 3-dB gain BW. As such, achieving a flat gain within the frequency of interest is equally important as expanding the input matching BW for wideband

LNAs. Although the inter-stage capacitively coupled resonator provides the desired Z_{L1} over frequency to realize a dual-resonant S_{11}, it inevitably results in a gain dip in the middle of the BW. Therefore, the design goal of the second stage is to compensate for the first-stage gain dip and in turn, realize a flat overall gain across the operating frequency. When using a cascode amplifier as the output stage, it can be generally modeled as a high-impedance current source in parallel with the device parasitic capacitance. Therefore, the gain shape of the second stage is dominated by its output network.

A popular design methodology to realize broadband networks is to synthesize a high-order bandpass response and include the parasitic capacitance as part of the network [30], [31], [32]. High-order bandpass networks can also enable a few useful functionalities, such as providing low-impedance DC feeds [33], [34], [35], [36] and impedance up- or down-transformation [37], [38]. In this paper, we focus on second-order bandpass networks and their miniaturization into a transformer.

A canonical second-order bandpass network is shown in Fig. 12, which can be transformed from a low-pass prototype [39]. In particular, the coefficients g_1 and g_2 set the desired network response, ω_0 is the center frequency, which is the geometric mean of the lower cutoff frequency ω_1 and the higher cutoff frequency ω_2, and Δ is the fractional BW. As the circuit model of a physical on-chip transformer contains two inductors – a series leakage inductor and a shunt magnetization inductor [40], it is possible to miniaturize a canonical second-order bandpass network into a single transformer footprint, achieving a size reduction of roughly $2\times$.

To compensate for the gain dip of the first stage, ω_1 is chosen to be the same as the first parallel resonant frequency of the inter-stage network ω_{p1}. This ensures that the first gain peak is effectively attenuated by 3 dB. Since the component values of the inter-stage network (i.e., C_c, L_d, and L_b) are chosen to ensure the gain difference between the first-stage peak and dip to be <6 dB, the overall two-stage gain variation remains within 3 dB. Additionally, since the second gain peak of the first stage is usually insignificant due to the degraded Q at higher frequencies, the higher cut-off frequency of the output network ω_2 is chosen as high as possible to extend the overall LNA gain BW.

In this section, we present two network synthesis methods to realize such network miniaturization and discuss their pros and cons. After showing the detailed synthesis procedure with design equations, we present a design example to illustrate the proposed design procedure.

A. Miniaturizing a Second-Order Bandpass Network Into a Single Transformer Footprint Using One-Step Norton Transformation

Starting with a canonical second-order bandpass network shown in Fig. 13, we first perform an inductive Norton transformation on the shunt-series inductors (L_1 and L_2). Norton transformation is a powerful technique in matching network designs to topologically swap a series inductor with a shunt inductor while maintaining the BW of the network. Here, the

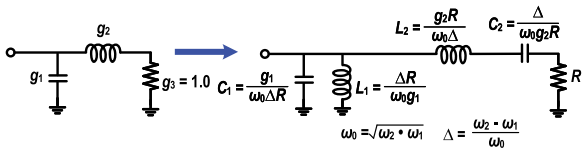


Fig. 12. Transforming a normalized second-order low-pass prototype to a bandpass network.

inductive Norton transformation ratio n_L can be found as [41]

$$n_L = \frac{L_1 + L_2}{L_1} \tag{23}$$

Next, we insert an ideal transformer with a turn ratio of $k{:}n$ between the shunt inductor and the series capacitor. The ideal transformer and the two inductors (L_3 and L_4) can be replaced by a physical on-chip transformer if the following condition is satisfied, as

$$\frac{L_3}{L_4} = \frac{L_2}{L_1} = \frac{1-k^2}{k^2} = \frac{g_1 g_2}{\Delta^2} \tag{24}$$

From (24), it can be seen that the required transformer magnetic coupling coefficient k is determined once the network prototype (indicated by the coefficients g_1 and g_2) and the fractional BW Δ are given. This is an important conclusion, which is further elaborated in Sec III-B.

The device parasitic capacitance C_{dev} and the transformer parasitic capacitance of the primary winding C_{par1} can be absorbed by the shunt capacitor C_3. However, one critical drawback of this approach is that there is no budget for the transformer parasitic capacitance of the secondary winding C_{par2}. As such, the frequency response of a practical transformer-based implementation would deviate from that of the original second-order bandpass network even when the network loss is not taken into consideration. Such a deviation would become more significant as the frequency gets higher.

To address this drawback, we present another network miniaturization approach based on two Norton transformations in the next sub-section.

B. Miniaturizing a Second-Order Bandpass Network Into a Single Transformer Footprint Using Two-Step Norton Transformation

As shown in Fig. 14, we first split the capacitor C_2 into two series capacitors C_3 and C_{2a} and then perform a series-to-parallel conversion on the capacitor C_{2a} and the load resistor R. The quality factor of this series-to-parallel conversion Q_s is calculated as

$$Q_s = \frac{1}{\omega_0 C_{2a} R} \tag{25}$$

Note that there exists an upper bound for Q_s since C_{2a} has to be greater than C_2. This upper bound is the loaded quality factor of the series section of the bandpass prototype and is given as

$$Q_s < Q_{\text{prototype, series}} = \frac{1}{\omega_0 C_2 R} = \frac{g_2}{\Delta} \tag{26}$$

Next, we apply two Norton transformations on the series-shunt capacitors (C_3 and C_4) and the shunt-series inductors (L_1 and L_2), respectively. n_c is the capacitive Norton

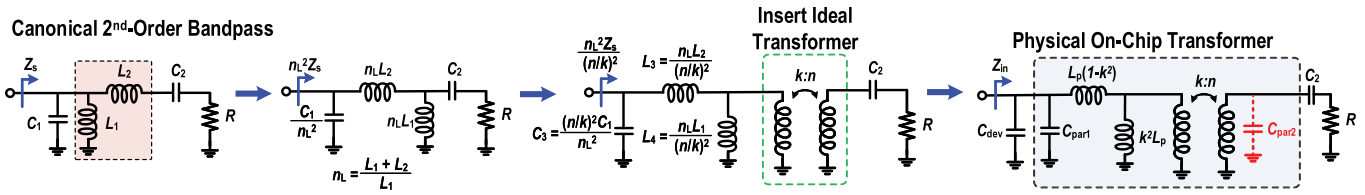


Fig. 13. Converting a canonical second-order bandpass network to a transformer-based network using one-step Norton transformation.

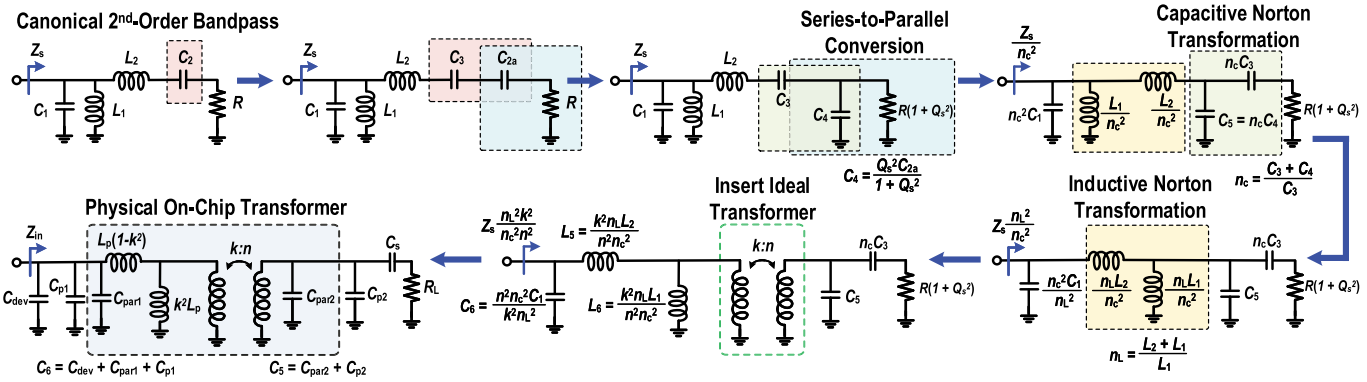


Fig. 14. Converting a canonical second-order bandpass network to a transformer-based network using two-step Norton transformation.

transformation ratio and n_L is the inductive Norton transformation ratio. Since the capacitive Norton transformation down-converts the impedance on its left, whereas the inductive Norton transformation up-converts the impedance, the input impedance Z_{in} is scaled by a factor of n_L^2/n_c^2 after the two Norton transformations.

Finally, we insert an ideal transformer with a turn ratio of *k:n* between the shunt inductor and the shunt capacitor C_5. If the following two conditions are met:

$$L_5 = \frac{k^2 n_L L_2}{n^2 n_c^2} = L_p(1-k^2) \tag{27}$$

$$L_6 = \frac{k^2 n_L L_1}{n^2 n_c^2} = L_p k^2 \tag{28}$$

the network highlighted in light blue can be implemented as a physical transformer with an actual turn ratio of *1:n*, a magnetic coupling coefficient of k, and a primary-winding self-inductance of L_p. The required k can be found as

$$k = \sqrt{\frac{\Delta^2}{g_1 g_2 + \Delta^2}} \tag{29}$$

Compared to the network synthesis approach presented in Sec III-A, the additional capacitive Norton transformation is particularly important as it provides the capacitance budget to absorb the transformer secondary-winding parasitic capacitance C_{par2}. On the primary side, the shunt capacitor C_6 includes the device parasitic capacitance C_{dev}, the transformer primary-winding parasitic capacitance C_{par1}, and if needed, an explicit capacitor C_{p1}.

It turns out that the series-to-parallel conversion quality factor Q_s is a crucial design parameter. Once the center frequency ω_0, fractional BW Δ, and network prototype (coefficients g_1 and g_2) are known, all the circuit parameters (except for k) in Fig. 14 can be derived analytically as a function of Q_s:

$$L_p = \frac{(g_1 g_2 + \Delta^2) R_L}{\omega_0 g_1 n^2 \Delta (1+Q_s^2)[1 + \frac{Q_s}{1+Q_s^2}(g_2/\Delta - Q_s)]^2} \tag{30}$$

$$C_5 = C_{par2} + C_{p2} = \frac{Q_s}{\omega_0 R_L} \times [1 + \frac{Q_s}{1+Q_s^2}(g_2/\Delta - Q_s)] \tag{31}$$

$$C_6 = C_{dev} + C_{par1} + C_{p1} = \frac{n^2 \Delta g_1 (1+Q_s^2)}{\omega_0 R_L (g_1 g_2 + \Delta^2)} \times [1 + \frac{Q_s}{1+Q_s^2}(g_2/\Delta - Q_s)]^2 \tag{32}$$

$$C_s = \frac{1+Q_s^2}{\omega_0 R_L (g_2/\Delta - Q_s)} \times [1 + \frac{Q_s}{1+Q_s^2}(g_2/\Delta - Q_s)] \tag{33}$$

Here, R_L is the load impedance of the network, which is single-ended 50 Ω or differential 100 Ω for a stand-alone LNA test chip, or models the input impedance of the following stage in a complete RX frontend.

The input impedance of the network Z_{in} is also a function of Q_s, as

$$Z_{in} = \frac{R_L}{n^2(1+Q_s^2)[1 + \frac{Q_s}{1+Q_s^2}(\frac{g_2}{\Delta} - Q_s)]^2} \times (\frac{g_1 g_2}{\Delta^2} + 1) \tag{34}$$

A larger Z_{in} is generally preferred as it leads to a higher transimpedance gain of the output network.

1) Summary of the Design Procedure: With all the design equations derived, the design procedure to miniaturize a conical second-order bandpass network into a single transformer footprint is summarized as follows.

First, given the target overall LNA gain and BW, once the design of the first stage is ready by following the dual-resonant input matching design procedure presented in Sec II, the frequency response of the output network, i.e., its ω_0, ω_1, Δ, and maximum tolerable in-band ripple (indicated by g_1 and g_2) can be decided.

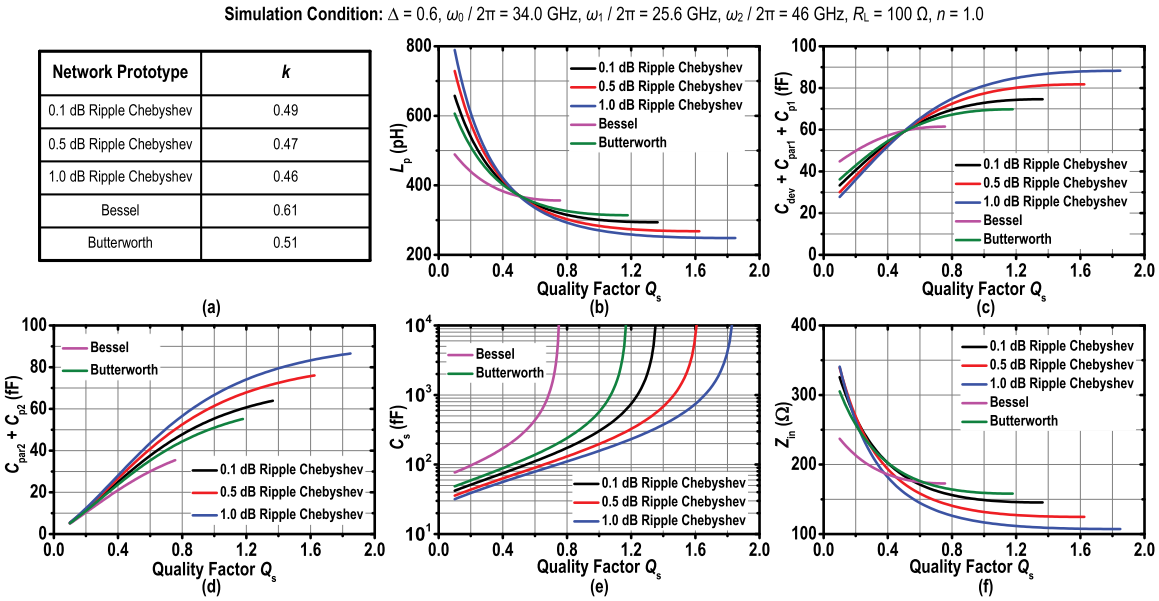


Fig. 15. Design space for different second-order bandpass network prototypes with calculated (a) k, (b) L_p vs. Q_s, (c) $(C_{par1} + C_{dev} + C_{p1})$ vs. Q_s, (d) $(C_{par2} + C_{p2})$ vs. Q_s, (e) C_s vs. Q_s, and (f) Z_{in} vs. Q_s.

Second, the required k can be calculated based on (29). It can be seen that k increases monotonically as the desired fractional BW Δ becomes larger. For on-chip transformers, there typically exists an upper bound for the achievable k, which in turn, sets the upper limit of Δ that can be practically realized. Additionally, the design curves for L_p, $(C_{dev} + C_{par1} + C_{p1})$, $(C_{par2} + C_{p2})$, C_s, and Z_{in} can be plotted as a function of Q_s based on (30)–(34). A lower Q_s is generally preferred as it leads to smaller parasitic capacitances, which in turn, results in a larger Z_{in} and larger Δ.

Finally, a physical on-chip transformer needs to be constructed to satisfy all the parameters. If these design parameters do not result in a practically achievable physical transformer, then we need to either increase Q_s to have more budget for the parasitic capacitances or relax the network specifications such as BW or in-band ripple. Additional EM optimizations may be required to fine-tune the transformer geometry and the values of the passive components.

We'd like to point out that the proposed network synthesis approach is fundamentally different from conventional transformer-based networks with two shunt capacitors at the primary and secondary windings [42]. From the network topology perspective, our approach has a series capacitor at the secondary winding, whereas the conventional designs always have a shunt capacitor at the secondary winding. This is because the starting points of the proposed network synthesis approach and prior reported transformer-based networks are quite different. Conventional broadband transformer-based networks are also known as magnetically coupled resonators. Although they can realize a dual-peaking frequency response, a critical limitation is that they cannot decouple the resonant frequencies and the gain ripple. In other words, having the two peaking frequencies more spread out inevitably leads to a larger gain ripple in between [43]. On the contrary, our approach starts with a canonical second-order bandpass network, consisting of a shunt LC branch and a series LC branch. As a result, the bandwidth and gain shape are completely decoupled, meaning that we can synthesize arbitrary frequency responses (such as Butterworth, Chebyshev, Bessel, etc.) following the classic "Insertion Loss" method of filter design [39].

	Butterworth	Butterworth after tuning	EM Implementation
L_p	352 pH	352 pH	351 pH
k	0.51	0.51	0.45
n	1.0	1.0	0.93
$C_{par1} + C_{dev} + C_{p1}$	62 fF	62 fF	66 fF (C_{dev} = 36 fF, C_{par1} = 30 fF, C_{p1} = 0)
$C_{par2} + C_{p2}$	35 fF	47 fF	58 fF (SE C_{par2} = 36 fF, SE C_{p2} = 80 fF)
C_s	141 fF	160 fF	160 fF

(a)

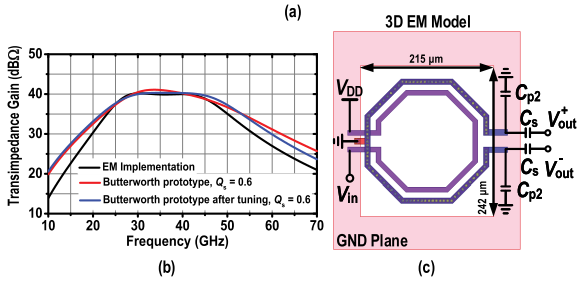
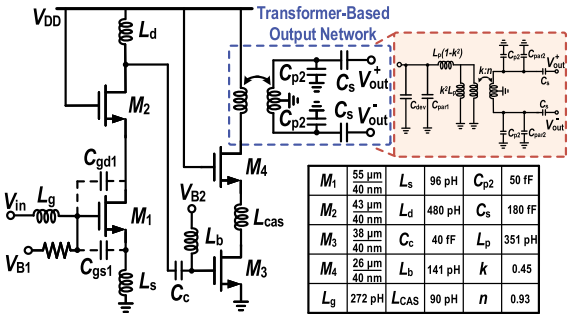


Fig. 16. (a) Design parameters based on Fig. 15 and after tuning to accommodate the network loss. The transformer parameters derived from the 3-D EM simulation are also listed in the table. (b) Simulated network transimpedance gain over frequency. (c) 3-D EM model of the transformer.

2) A Design Example: In this sub-section, we use our reference LNA design as an example to illustrate the design procedure. The target center frequency $\omega_0/2\pi$ is 34 GHz, the fractional BW Δ is 60%, the differential load impedance R_L is 100 Ω, and the transformer turn ratio n is 1. For different second-order network prototypes, the required k is listed in Fig. 15 (a), and the design curves for L_p, $(C_{dev} + C_{par1} + C_{p1})$, $(C_{par2} + C_{p2})$, and C_s against Q_s are plotted in Fig. 15(b)-(e). Note that there exists an upper bound of Q_s introduced in the series-to-parallel conversion, as indicated in (26). The input impedance Z_{in} is also plotted in Fig. 15(f), showing a monotonic decrease with respect to Q_s.

For our reference LNA design, we choose Butterworth as the desired network response and $Q_s = 0.6$. Based on the design curves in Fig. 15, the required transformer parameters and the network transimpedance gain over frequency are shown in



M_1	55 μm / 40 nm	L_s	96 pH	C_{p2}	50 fF
M_2	43 μm / 40 nm	L_d	480 pH	C_s	180 fF
M_3	38 μm / 40 nm	C_c	40 fF	L_p	351 pH
M_4	26 μm / 40 nm	L_b	141 pH	k	0.45
L_g	272 pH	L_{CAS}	90 pH	n	0.93

Fig. 17. LNA schematic with component values.

Fig. 16(a) and (b), respectively. Since practical transformers are lossy with a quality factor ∼15 in our target frequency bands, the network transimpedance gain starts to deviate from the ideal maximally flat response. We slightly fine-tune the shunt capacitance C_{p2} and series capacitance C_s on the secondary side of the transformer so that its in-band ripple is again minimized when the loss is taken into consideration [the blue curve in Fig. 16(b)]. Next, we implement a physical transformer with the goal of satisfying all the design parameters listed in Fig. 16(a). The 3-D EM model of the transformer and its dimension are shown in Fig. 16(c). Based on the EM-simulated S-Parameters, its k, n, L_p, C_{par1}, and C_{par2} are extracted and are reasonably close to our design target. The transimpedance gain based on the 3-D EM model [the black curve in Fig. 16(b)] is also very close to our desired frequency response.

IV. A 27–46-GHz LNA Implementation Example

Following the design procedure to achieve dual-resonant input matching in Sec II-B and the network synthesis approach to achieve broadband yet miniaturized output network in Sec III-B, a proof-of-concept 27-46-GHz broadband LNA is presented in this section as an implementation example. It is fabricated in the GlobalFoundries 45-nm CMOS SOI process. The design goal is to cover multiple mmWave 5G NR bands centered around 26, 28, 39, and 41 GHz (band n257 – band n261).

The LNA schematic is shown in Fig. 17, consisting of two stages. Both stages are biased with $J_{D,\,opt} = 0.2$ mA/μm to achieve a low NF_{min} and a high f_{max} simultaneously. The output transformer has differential outputs, making the LNA easier to be employed in a receiver chain, since the following blocks such as the mixer, variable gain amplifier, and phase shifter, are usually designed in a differential manner. In the design phase, the component values are first derived based on the analysis in Sec II and III and then optimized by 3-D EM simulations to accommodate the parasitic effects.

The simulated first-stage, second-stage, and overall LNA voltage gain is shown in Fig. 18. The first stage exhibits two gain peaks due to the inter-stage capacitively coupled resonator, as discussed in Sec II-C. To realize a flat overall gain, the low cutoff frequency of the output network is aligned with the first parallel resonant frequency of the inter-stage network, which is 26.3 GHz. As shown in Fig. 18, the simulated overall peak voltage gain is 22.7 dB at 41.3 GHz with an in-band ripple of 1.4 dB.

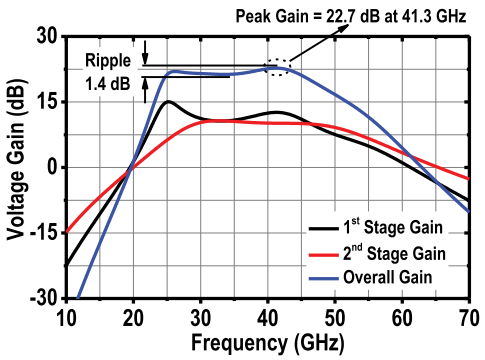
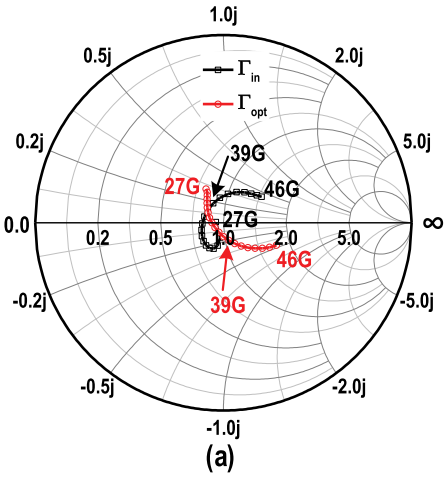


Fig. 18. The first-stage, second-stage, and overall LNA voltage gain.



Frequency (GHz)	M_2 Noise Contribution / M_1 Noise Contribution	M_3 Noise Contribution / M_1 Noise Contribution
27	39.8%	55.2%
30	53.9%	111.2%
35	75.6%	133.6%
45	76.0%	22.3%

(c)

Fig. 19. (a) Simulated Γ_{in} and Γ_{opt} from 27 to 46 GHz. (b) Simulated LNA NF. (c) LNA noise summary based on simulations.

The simulated input reflection coefficient Γ_{in} and the optimum noise reflection coefficient Γ_{opt} over frequency are shown on the Smith Chart in Fig. 19(a). There exists a slight mismatch between Γ_{in} and Γ_{opt} trajectories above 40 GHz, resulting in a 0.34 dB increase in the simulated NF at 45 GHz, as shown in Fig. 19(b). The higher NF at higher frequencies is due to two reasons. First, the cascode transistor M_2 contributes more noise at higher frequencies, which can be seen from the

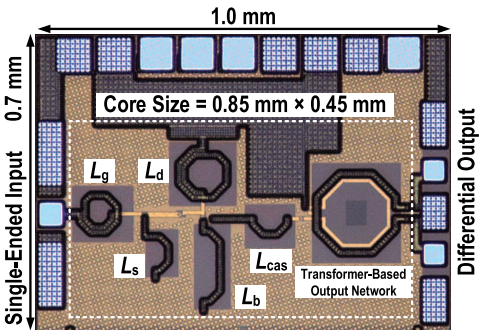


Fig. 20. Chip micrograph.

LNA noise summary in Fig. 19(c). This is a typical issue for mmWave cascode LNAs [44]. Second, the first stage has a gain dip between 30 and 35 GHz, where the noise of the second-stage common-source transistor M_3 starts to manifest itself, as shown in Fig. 19(c).

Since the proposed design approach uses the same narrowband cascode common-source with inductive degeneration circuit topology with only component value updates, we expect minimal overhead in terms of chip area. In our original LNA design, we did not pay special attention to minimizing the chip area, especially the y dimension of the design. As can be seen in the chip micrograph (Fig. 20), if we reduce the y dimension of inductors L_d, L_s, and L_b, the chip area can be significantly reduced. Therefore, we present an updated layout of L_d, L_s, and L_b in Fig. 21(a), leading to a significantly reduced LNA core area of 0.21 mm^2. Meanwhile, the simulated S-parameters remain almost the same as those of our original design, as shown in Fig. 21(b).

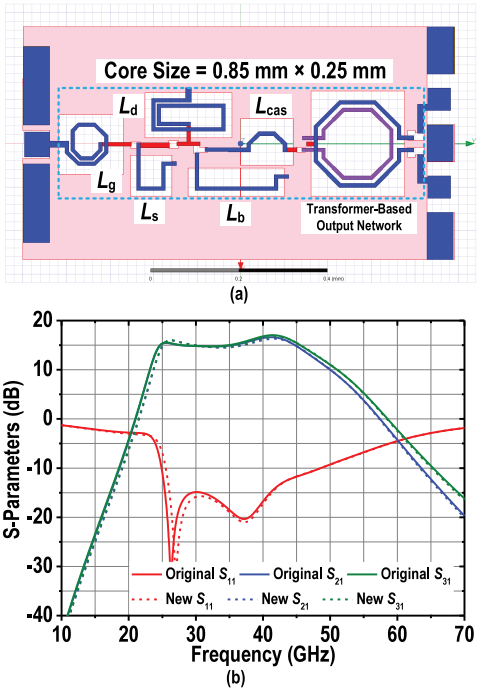


Fig. 21. (a) Optimize the layout of L_d, L_s, and L_b to reduce the y dimension and the resulting chip area. The updated layout is simulated in HFSS. (b) Simulated S-parameters based on the new layout and the original design.

V. Measurement Results

This section presents the S-parameters, NF, and linearity measurement results, which are all based on the probing of three samples. The supply voltage V_{DD} of the LNA is 1.3 V, and the DC power consumption is 25.5 mW.

A. S-Parameters Measurement

Since this LNA has a single-ended input and a differential output, to characterize its small-signal performance, a three-port S-parameters measurement is performed using the Keysight N5225B four-port vector network analyzer. The measured and simulated S_{11}, S_{21}, and S_{31} are shown in Fig. 22(a). The LNA achieves a peak single-ended / differential gain of 18.2 / 21.2 dB at 37.8 GHz, with a 3-dB gain BW of 25.5 to 50 GHz. The measured S_{11} is lower than -10 dB from 27 to 46 GHz; its dual-resonant behavior can be clearly seen. The effective bandwidth of our prototype, which is defined as the intersection of the -10-dB S_{11} BW and the 3-dB gain BW, is limited by the input matching. The measured differential gain and S_{11} of three samples are plotted in Fig. 22(b), achieving negligible sample-to-sample variations.

The simulated and measured geometrically derived stability factors for the load (μ) and for the source (μ') are shown in Fig. 22(c). They are always larger than 1, proving that the LNA is unconditionally stable over the frequency.

The measured balancing between the two differential outputs is summarized in Fig. 23. Within the effective BW of 27 to 46 GHz, the phase mismatch is less than 3°, and the amplitude mismatch is smaller than 1 dB, showing well-balanced differential outputs. The common-mode rejection ratio (CMRR) is >24 dB.

B. Noise Figure Measurement

The NF is measured using the Keysight N9040B spectrum analyzer and the Keysight 346CK01 noise source. The NF measurement results are summarized in Fig. 24(a). The minimum NF is 2.4 dB at 27.7 GHz, and the NF remains below 4.2 dB within the effective BW. The measured NF of the three samples is consistent.

C. Linearity Measurement

The measured input-referred 1-dB compression point (IP_{1dB}) and differential output-referred 1-dB compression point (OP_{1dB}) of the three samples are shown in Fig. 24(b). Due to the frequency range limitation of our signal generator, the 1-dB compression point measurement is performed only up to 40 GHz. The LNA achieves -25.6 to -17.4 dBm IP_{1dB} and -4.6 to $+2.9$ dBm OP_{1dB} from 27 to 40 GHz.

For the third-order intercept point (IP_3) measurement, two tones with a 100 MHz separation are applied to the LNA input, and their power levels are varied from -41.6 to -36.6 dBm. The measured input-referred IP_3 (IIP_3) and differential output-referred IP_3 (OIP_3) from 27 to 40 GHz are plotted in Fig. 24(c). The best IIP_3 is -9.5 dBm at 32 GHz and the best OIP_3 is 12.1 dBm at 34 GHz, showing state-of-the-art LNA linearity.

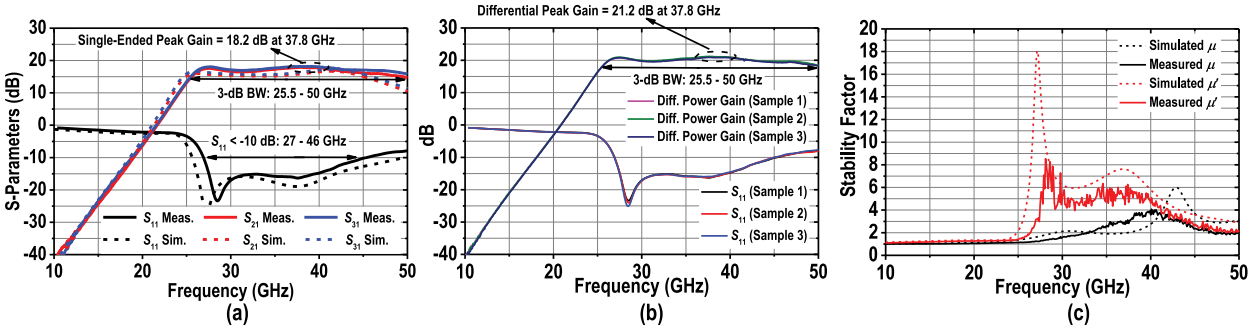


Fig. 22. (a) Measured and simulated S_{11}, S_{21}, and S_{31}. (b) Measured differential power gain and S_{11} of three samples. (c) Measured and simulated stability factors μ and μ'.

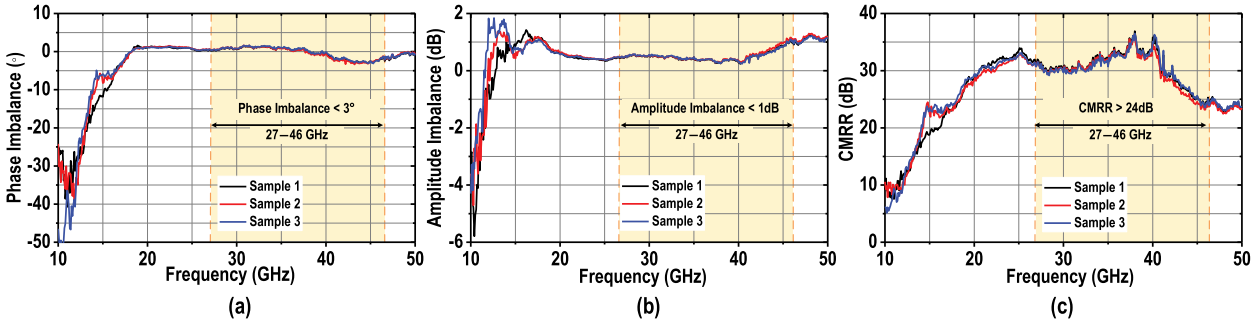


Fig. 23. (a) Measured phase imbalance, (b) amplitude imbalance, and (c) CMRR of three samples.

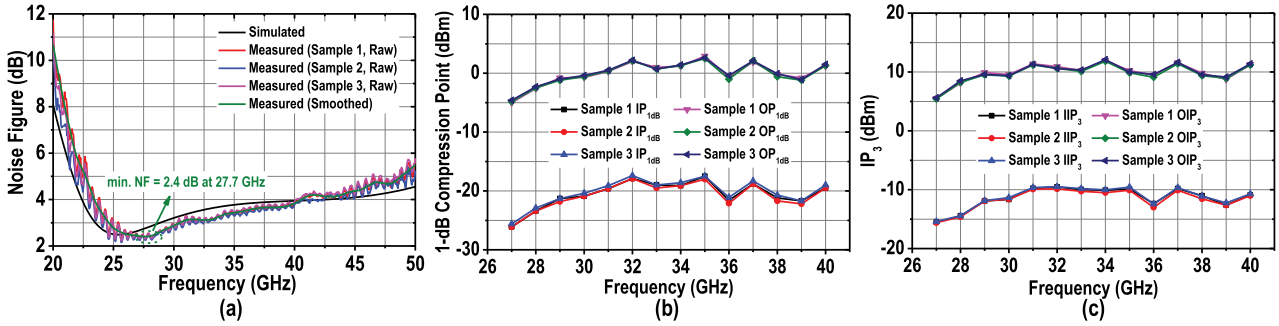


Fig. 24. (a) Simulated NF and Measured NF of three samples. (b) Measured IP_{1dB} and differential OP_{1dB}. (c) Measured IIP_3 and differential OIP_3.

TABLE I

PERFORMANCE COMPARISON TABLE

Reference	BW* (GHz)	3-dB BW (GHz)	Peak Gain (dB)	NF (dB)	IIP_3 (dBm)	P_{DC} (mW)	Core Size (mm^2)	FoM**	Technology
This Work	**27–46**	**25.5–50.0**	**21.2**	**2.4–4.2**	**-9.5**	**25.5**	**0.38** / **0.21**§	**424**	**45-nm CMOS SOI**
MWCL 2018 [45]	26–33	26–33	27.1	3.3–4.3	-12.1	31.4	0.26	211	40-nm CMOS
			18.4	3.4–4.4	-6.1†	21.5		159	
GSMM 2018 [10]	27–47.5†	24–47.5	20	4.2–5.5	-9.4††	58	0.2	69	45-nm CMOS SOI
RFIC 2019 [13]	24–29†	24–29†	19.1†	3.1–3.7	-13.2	20.5	0.22	34	22-nm CMOS SOI
	37–42†	37–42†	23		-19			15	
TMTT 2020 [3]	20–40	20–40	21.1	2.5–3	-7	18	0.67	1297	45-nm CMOS SOI
JSSC 2020 [12]	22–32	19–36	21.5	1.7–2.2	-13.4	17.3	0.05	293	22-nm CMOS SOI
	22–32	20–36	17.9	2.1–2.9	-14.4	5.6		242	
TMTT 2020 [46]	21–48†	17–48	19.5†	2–3.2	-7.9††	25	0.63†	840	45-nm CMOS SOI
JSSC 2021 [11]	22.9–38.2	22.9–38.2	14.5	2.6–4.6	-3.6	18.9	0.16	400	28-nm CMOS
RFIC 2021 [47]	22.2–43	21.8–43	21.1	3.5–5.3	-3.0	22.3	0.22	1573	28-nm CMOS

* Intersection of 3-dB gain BW and 10-dB return loss BW.

§ Based on updated layout.

† Graphically estimated.

†† Estimated using IP_{1dB} + 9.6 dB.

** $FoM = \frac{10^3 \times Gain[\frac{W}{W}] \times BW_{eff}[GHz] \times IIP_3[mW]}{P_{DC}[mW] \times (NF[linear]-1) \times f_c[GHz]}$, f_c is the geometric mean, peak gain and minimum NF are taken in FoM calculation.

VI. Conclusion

A comparison with state-of-the-art broadband LNAs at a similar frequency range is shown in Table I. To benchmark their performance, a Figure-of-Merit (FoM) involving the power gain, BW, IIP_3, DC power, NF, and the center frequency is adopted, as

$$FoM = \frac{10^3 \times Gain[\frac{W}{W}] \times BW_{eff}[GHz] \times IIP_3[mW]}{P_{DC}[mW] \times (NF[linear] - 1) \times f_c[GHz]} \tag{35}$$

Here, the BW is defined as the intersection of the 3-dB gain BW and the 10-dB return loss BW. As shown in Table I, the presented LNA achieves state-of-the-art BW and a competitive FoM. The reported FoM is lower than those in [3], [46], and [47] mainly because our measured IIP_3 is lower. Although linearity enhancement is not the major focus of this paper, the proposed dual-resonant input matching and broadband output network techniques are compatible with a few well-established IIP_3 enhancement techniques, such as the multi-gate transistor (MGTR) and derivative superposition (DS) [48]. Combining these techniques can potentially improve our IIP_3 by a few dB and in turn, result in a higher FoM.

In conclusion, two design techniques are presented in this paper to broaden the BW of mmWave LNAs. First, we extend the input matching BW by synthesizing dual-resonant input matching. Second, we extend the gain BW by constructing a second-order bandpass output network that can be miniaturized into a single transformer footprint. A proof-of-concept 27–46 GHz LNA is implemented in Globalfoundries 45-nm CMOS SOI process, achieving 21.2 dB peak gain, 2.4 dB minimum NF, and -9.5 dBm IIP_3 with 25.5 mW DC power. This design could be readily integrated with wideband RX frontends for multi-band 5G communications and high-resolution wireless sensing applications.

References

[1] M.-Y. Huang, T. Chi, S. Li, T.-Y. Huang, and H. Wang, "A 24.5–43.5-GHz ultra-compact CMOS receiver front end with calibration-free instantaneous full-band image rejection for multiband 5G massive MIMO," *IEEE J. Solid-State Circuits*, vol. 55, no. 5, pp. 1177–1186, May 2020.

[2] A. Alhamed, G. Gultepe, and G. M. Rebeiz, "A multi-band 16–52-GHz transmit phased array employing 4 × 1 beamforming IC with 14–15.4-dBm P_{sat} for 5G NR FR2 operation," *IEEE J. Solid-State Circuits*, vol. 57, no. 5, pp. 1280–1290, May 2022.

[3] M. Lokhandwala, L. Gao, and G. M. Rebeiz, "A high-power 24–40-GHz transmit–receive front end for phased arrays in 45-nm CMOS SOI," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 11, pp. 4775–4786, Nov. 2020.

[4] J. Pang et al., "A power-efficient 24-to-71 GHz CMOS phased-array receiver utilizing harmonic-selection technique supporting 36dB inter-band blocker rejection for 5G NR," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 65, Feb. 2022, pp. 434–436.

[5] S. Mondal and J. Paramesh, "A reconfigurable 28-/37-GHz MMSE-adaptive hybrid-beamforming receiver for carrier aggregation and multi-standard MIMO communication," *IEEE J. Solid-State Circuits*, vol. 54, no. 5, pp. 1391–1406, May 2019.

[6] L. Gao, Q. Ma, and G. M. Rebeiz, "A 20–44-GHz image-rejection receiver with >75-dB image-rejection ratio in 22-nm CMOS FD-SOI for 5G applications," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 7, pp. 2823–2832, Jul. 2020.

[7] S. Trotta et al., "2.3 SOLI: A tiny device for a new human machine interface," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2021, pp. 42–44.

[8] C. Li et al., "A review on recent progress of portable short-range noncontact microwave radar systems," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 5, pp. 1692–1706, May 2017.

[9] T.-Y. J. Kao, Y. Yan, T.-M. Shen, A. Y.-K. Chen, and J. Lin, "Design and analysis of a 60-GHz CMOS Doppler micro-radar system-in-package for vital-sign and vibration detection," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 4, pp. 1649–1659, Apr. 2013.

[10] V. Chauhan and B. Floyd, "A 24–44 GHz UWB LNA for 5G cellular frequency bands," in *Proc. 11th Global Symp. Millim. Waves (GSMM)*, May 2018, pp. 1–3.

[11] Z. Deng, J. Zhou, H. J. Qian, and X. Luo, "A 22.9–38.2-GHz dual-path noise-canceling LNA with 2.65–4.62-dB NF in 28-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 56, no. 11, pp. 3348–3359, Nov. 2021.

[12] B. Cui and J. R. Long, "A 1.7-dB minimum NF, 22–32-GHz low-noise feedback amplifier with multistage noise matching in 22-nm FD-SOI CMOS," *IEEE J. Solid-State Circuits*, vol. 55, no. 5, pp. 1239–1248, May 2020.

[13] L. Gao and G. M. Rebeiz, "A 24–43 GHz LNA with 3.1–3.7 dB noise figure and embedded 3-pole elliptic high-pass response for 5G applications in 22 nm FDSOI," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2019, pp. 239–242.

[14] Y. Hu and T. Chi, "A 27–46-GHz low-noise amplifier with dual-resonant input matching and a transformer-based broadband output network," *IEEE Microw. Wireless Compon. Lett.*, vol. 31, no. 6, pp. 725–728, Jun. 2021.

[15] V. Jain, S. Sundararaman, and P. Heydari, "A 22–29-GHz UWB pulse-radar receiver front-end in 0.18-μm CMOS," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 8, pp. 1903–1914, Aug. 2009.

[16] T. Yao et al., "Algorithmic design of CMOS LNAs and PAs for 60-GHz radio," *IEEE J. Solid-State Circuits*, vol. 42, no. 5, pp. 1044–1057, May 2007.

[17] H. Hashemi and A. Hajimiri, "Concurrent multiband low-noise amplifiers-theory, design, and applications," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 1, pp. 288–301, Jan. 2002.

[18] P. Heydari, "Design and analysis of a performance-optimized CMOS UWB distributed LNA," *IEEE J. Solid-State Circuits*, vol. 42, no. 9, pp. 1892–1905, Sep. 2007.

[19] T.-K. Nguyen, C.-H. Kim, G.-J. Ihm, M.-S. Yang, and S.-G. Lee, "CMOS low-noise amplifier design optimization techniques," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 5, pp. 1433–1442, May 2004.

[20] T. Chi, J. S. Park, S. Li, and H. Wang, "A millimeter-wave polarization-division-duplex transceiver front-end with an on-chip multifeed self-interference-canceling antenna and an all-passive reconfigurable canceller," *IEEE J. Solid-State Circuits*, vol. 53, no. 12, pp. 3628–3639, Dec. 2018.

[21] M. Y. Huang, T. Chi, F. Wang, T. W. Li, and H. Wang, "A 23-to-30 GHz hybrid beamforming MIMO receiver array with closed-loop multistage front-end beamformers for full-FoV dynamic and autonomous unknown signal tracking and blocker rejection," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2018, pp. 68–70.

[22] H. Mohammadnezhad, R. Abedi, and P. Heydari, "A millimeter-wave partially overlapped beamforming-MIMO receiver: Theory, design, and implementation," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 5, pp. 1924–1936, May 2019.

[23] R. Garg et al., "A 28-GHz beam-space MIMO RX with spatial filtering and frequency-division multiplexing-based single-wire IF interface," *IEEE J. Solid-State Circuits*, vol. 56, no. 8, pp. 2295–2307, Aug. 2021.

[24] S. Li, T. Chi, D. Jung, T.-Y. Huang, M.-Y. Huang, and H. Wang, "4.2 an E-band high-linearity antenna-LNA front-end with 4.8dB NF and 2.2 dBm IIP3 exploiting multi-feed on-antenna noise-canceling and Gm-boosting," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 1–3.

[25] R. Behzad, *RF Microelectronics*. New York, NY, USA: Prentice-Hall, 2012.

[26] O. Inac, M. Uzunkol, and G. M. Rebeiz, "45-nm CMOS SOI technology characterization for millimeter-wave applications," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 6, pp. 1301–1311, Jun. 2014.

[27] F. Vecchi et al., "A wideband receiver for multi-Gbit/s communications in 65 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 46, no. 3, pp. 551–561, Mar. 2011.

[28] L. Gao, E. Wagner, and G. M. Rebeiz, "Design of E- and W-band low-noise amplifiers in 22-nm CMOS FD-SOI," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 1, pp. 132–143, Jan. 2020.

[29] S. Guo, T. Xi, P. Gui, D. Huang, Y. Fan, and M. Morgan, "A transformer feedback G_m-boosting technique for gain improvement and noise reduction in mm-Wave cascode LNAs," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2080–2090, Jul. 2016.

[30] H. Wang, C. Sideris, and A. Hajimiri, "A CMOS broadband power amplifier with a transformer-based high-order output matching network," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2709–2722, Dec. 2010.

[31] M. Bassi, J. Zhao, A. Bevilacqua, A. Ghilioni, A. Mazzanti, and F. Svelto, "A 40–67 GHz power amplifier with 13 dBm P_{SAT} and 16% PAE in 28 nm CMOS LP," *IEEE J. Solid-State Circuits*, vol. 50, no. 7, pp. 1618–1628, Jul. 2015.

[32] M. Vigilante and P. Reynaert, "On the design of wideband transformer-based fourth order matching networks for E-band receivers in 28-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 52, no. 8, pp. 2071–2082, Aug. 2017.

[33] Y. Zhao and J. R. Long, "A wideband, dual-path, millimeter-wave power amplifier with 20 dBm output power and PAE above 15% in 130 nm SiGe-BiCMOS," *IEEE J. Solid-State Circuits*, vol. 47, no. 9, pp. 1981–1997, Sep. 2012.

[34] T. Chi, M.-Y. Huang, S. Li, and H. Wang, "17.7 a packaged 90-to-300 GHz transmitter and 115-to-325 GHz coherent receiver in CMOS for full-band continuous-wave mm-wave hyperspectral imaging," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2017, pp. 304–305.

[35] S. Li, T. Chi, J.-S. Park, H. T. Nguyen, and H. Wang, "A 28-GHz flip-chip packaged Chireix transmitter with on-antenna outphasing active load modulation," *IEEE J. Solid-State Circuits*, vol. 54, no. 5, pp. 1243–1253, May 2019.

[36] Y. Hu, X. Zhang, and T. Chi, "A 28 GHz hybrid-beamforming transmitter array supporting concurrent dual data steams and spatial notch steering for 5G MIMO," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Apr. 2021, pp. 1–2.

[37] A. Bevilacqua, "Fundamentals of integrated transformers: From principles to applications," *IEEE Solid State Circuits Mag.*, vol. 12, no. 4, pp. 86–100, Nov. 2020.

[38] S. Li, T. Chi, T.-Y. Huang, M.-Y. Huang, D. Jung, and H. Wang, "A buffer-less wideband frequency doubler in 45-nm CMOS-SOI with transistor multiport waveform shaping achieving 25% drain efficiency and 46–89 GHz instantaneous bandwidth," *IEEE Solid-State Circuits Lett.*, vol. 2, no. 4, pp. 25–28, Apr. 2019.

[39] D. Pozar, *Microwave Engineering*, 4th ed. Hoboken, NJ, USA: Wiley, 2011.

[40] J. R. Long, "Monolithic transformers for silicon RF IC design," *IEEE J. Solid-State Circuits*, vol. 35, no. 9, pp. 1368–1382, Sep. 2000.

[41] L. Besser and R. Gilmore, *Practical RF Circuit Design for Modern Wireless Systems* (Passive Circuits and Systems), vol. 1. Boston, MA, USA: Artech House, 2003.

[42] A. Mazzanti and A. Bevilacqua, "Second-order equivalent circuits for the design of doubly-tuned transformer matching networks," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 12, pp. 4157–4168, Dec. 2018.

[43] C.-H. Li, C.-N. Kuo, and M.-C. Kuo, "A 1.2-V 5.2-mW 20–30-GHz wideband receiver front-end in 0.18-μm CMOS," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 11, pp. 3502–3512, Nov. 2012.

[44] Z. Wang, P.-Y. Chiang, P. Nazari, C.-C. Wang, Z. Chen, and P. Heydari, "A CMOS 210-GHz fundamental transceiver with OOK modulation," *IEEE J. Solid-State Circuits*, vol. 49, no. 3, pp. 564–580, Mar. 2014.

[45] M. Elkholy, S. Shakib, J. Dunworth, V. Aparin, and K. Entesari, "A wideband variable gain LNA with high OIP3 for 5G using 40-nm bulk CMOS," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 1, pp. 64–66, Jan. 2018.

[46] L. Gao and G. M. Rebeiz, "A 22–44-GHz phased-array receive beamformer in 45-nm CMOS SOI for 5G applications with 3–3.6-dB NF," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 11, pp. 4765–4774, Nov. 2020.

[47] A. Ershadi, S. Palermo, and K. Entesari, "A 22.2–43 GHz gate-drain mutually induced feedback low noise amplifier in 28-nm CMOS," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2021, pp. 27–30.

[48] H. Zhang and E. Sánchez-Sinencio, "Linearization techniques for CMOS low noise amplifiers: A tutorial," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 1, pp. 22–36, Jan. 2011.

Yaolong Hu (Graduate Student Member, IEEE) received the B.Eng. degree from Zhejiang University, Hangzhou, China, in 2018. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Rice University, Houston, TX, USA. From May 2022 to August 2022, he was an Analog Design Intern with Kilby Labs, Texas Instruments Inc., Dallas, TX, USA. His research interests include millimeter-wave integrated circuits and systems. He was a recipient of the IEEE Custom Integrated Circuits Conference (CICC) Best Student Paper Award in 2021 and the IEEE International Microwave Symposium (IMS) Advanced Practice Paper Award Finalist in 2021.

Taiyun Chi (Member, IEEE) received the B.S. degree (Hons.) from the University of Science and Technology of China (USTC), Hefei, China, in 2012, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, USA, in 2017.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA. His research interests include RF/millimeter-wave/terahertz integrated circuits and integrated bio-sensors and bio-actuators. His research group received the 2022 IEEE RFIC Symposium Best Student Paper Award Finalist, the 2021 IEEE Custom Integrated Circuits Conference (CICC) Best Student Paper Award, the 2021 IEEE International Microwave Symposium (IMS) Advanced Practice Paper Award Finalist, and the 2021 Texas Wireless Symposium Student Research Competition Award (1st Place). He was also a recipient of the Sigma Xi Best Ph.D. Thesis Award (Georgia Tech Chapter) in 2018, the IEEE CICC Best Paper Award in 2017, the IEEE Solid-State Circuits Society (SSCS) Predoctoral Achievement Award in 2017, the Microwave Theory and Techniques Society (MTT-S) Graduate Fellowship for Medical Applications in 2016, and the USTC Guo Moruo Presidential Scholarship in 2012. He is a Technical Program Committee (TPC) Member of the IEEE IMS and the TPC Co-Chair of the IEEE Texas Symposium on Wireless and Microwave Circuits and Systems.