Geometry and Bias Current Optimization for SiGe HBT Cascode Low-Noise Amplifiers

Qingqing Liang, Guofu Niu, John D. Cressler, Stewart Taylor[1], and David L. Harame[2]

Alabama Microelectronics Science and Technology Center, Electrical and Computer Engineering Department, 200 Broun Hall, Auburn University, Auburn AL 36849, USA
Tel: (334) 844-1853 / Fax: (334) 844-1888 / E-mail: qingqing@eng.auburn.edu
[1]Maxim Integrated Products, 7250 NW Evergreen Parkway, Hillsboro, OR 97124, USA
[2]IBM Microelectronics, Essex Junction, VT 05452, USA

Abstract— **This work presents a new design methodology for inductively-degenerated cascode low-noise amplifiers using advanced epitaxial-base SiGe HBTs. Noise figure, gain, and IIP3 are calculated using calibrated linear circuit analysis and a Volterra series methodology as a function of the two major design variables: emitter geometry and biasing current. An optimum SiGe HBT LNA design point which balances input impedance match, high IIP3, noise figure, gain, and power consumption is obtained from calculated noise figure, gain, and IIP3 contours as a function of bias current and geometry. Simplified analytical expressions of IIP3, gain, and noise figure are presented to give additional insight. The optimum LNA design point for the 50 GHz SiGe HBT process technology under study yields a 2 GHz LNA with 15.8 dBm IIP3, 18 dB gain, 1.15 dB noise figure, and a $|s11|$ less than −20 dB for a biasing current of 7.5 mA. The calculated results show good agreement with HP Advanced-Design-System simulations. The design tradeoffs illuminated by this optimization methodology are highlighted and discussed.**

I. Introduction

Low-noise amplifiers (LNA) are typically used in the first stage of an RF receiver to provide low-noise amplification. Besides providing gain while adding as little noise as possible, an LNA should minimize signal distortion, present a specific impedance (e.g., 50 Ω) to the input source, and consume as small a power as possible. The purpose of this work is to explore the optimum LNA design that balances input impedance match, gain, noise figure, linearity, and power consumption for a state-of-the-art SiGe HBT technology. The circuit topology chosen is the popular inductively-degenerated cascode LNA (shown in Fig. 1). The devices were fabricated using a 50 GHz self-aligned, deep-trench isolated, epitaxial-base SiGe HBT technology [1]. The standard emitter stripe width is 0.5 μm for this technology and is held fixed for all designs. The measured transistor characteristics for a 0.5 × 20 × 2 μm^2 SiGe HBT are summarized in Table I. Fig. 2 shows the measured f_T and f_{max} as a function of collector current density J_C.

TABLE I
MEASURED SiGe HBT PARAMETERS FOR A_E = 0.5 × 20 × 2 μm^2.

Peak β	110
Peak f_T	51 GHz
τ_{ec}	2.7 ps
r_{bb} @ I_C = 10 mA	8.9 Ω
BV_{ceo}	3.3 V

II. Design Methodology

Fig. 1 shows the schematic of the cascode amplifier [2]. The first amplifier stage acts as a transconductor, while the second amplifier stage produces a unity current gain. The nonlinearity of the circuit mainly

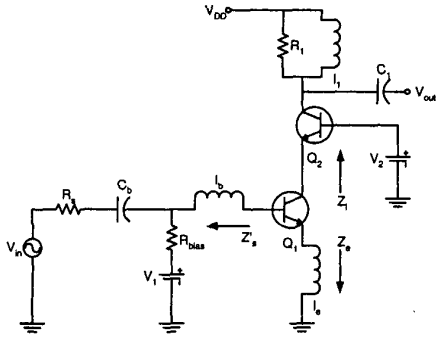

Fig. 1. An inductively degenerated cascode LNA. l_e and l_b are used to match the input impedance to 50 Ω.

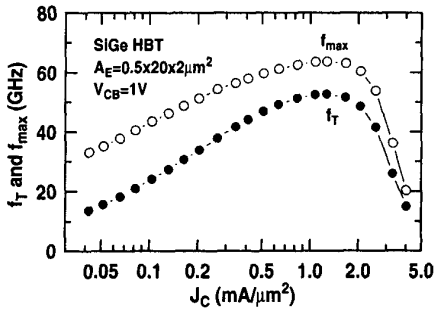

Fig. 2. Measured f_T and f_{max} as a function of current density for a A_E = 0.5×20×2 μm^2 transistor.

results from the first stage. For a given bias current and device geometry (emitter length), the values of emitter and base inductances are optimized to match a 50 Ω source impedance (R_s). The large-signal circuit is linearized and solved using nodal analysis. The emitter and base inductances are first optimized numerically for input impedance match. Noise figure and gain are then calculated using standard linear circuit analysis. IIP3 is then calculated using a calibrated Volterra series based approach [3], which has the unique advantage of distinguishing the individual physical nonlinearities in the device compared to other methods. Each physical nonlinearity can be turned on or off individually in the IIP3 calculation, providing a means of investigating the

0-7803-7246-8/02/$10.00 © 2002 IEEE

cancellation between the different device nonlinearities. In all cases, the calculations are based on actual measured device parameters from this 50 GHz SiGe HBT technology, thus ensuring the predictive nature of this optimization methodology. The IIP3 calculation method is further verified by simulation results obtained using ADS. Fig. 3 shows the IIP3 versus biasing I_C for a single transistor amplifier obtained using our Volterra-series approach and ADS. At an input power less than −20 dBm, the error of calculated IIP3 is less than 0.05 dB, demonstrating the good precision offered by our approach.

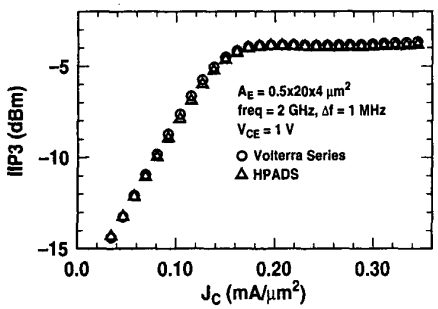

Fig. 3. IIP3 as a function of bias current in a single transistor. $A_E = 0.5 \times 20 \times 2\ \mu m^2$, $V_{ce} = 1$ V. The tone spacing is 1 MHz.

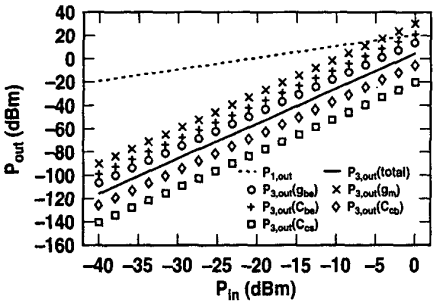

Fig. 4. Comparison of the calculated P_{out} versus P_{in} at 2 GHz for a cascode LNA. $A_E = 0.5 \times 20 \times 2\ \mu m^2$, $I_{ce} = 7$ mA, and $V_{ce} = 1.5$ V. The tone spacing is 1 MHz.

Fig. 4 shows the calculated intermodulation output power as a function of input power for an emitter length of 40 μm with a collector current of 7 mA at 2 GHz. The intermodulation result obtained with all of the transistor nonlinearities turned on is shown together with the results obtained by turning on the individual nonlinearities: g_m (×) represents the nonlinear transconductance ($I_C - V_{CE}$), g_{be} (o) represents the nonlinear EB conductance ($I_B - V_{BE}$), C_{be} (+), C_{cb} (◇) and C_{cs} (□) represent the nonlinear emitter-base, collector-base, and collector-substrate capacitances, respectively.

Impact ionization and CB capacitance, the two dominant nonlinearities in a single transistor amplifier [3], are no longer the dominant distortion mechanisms in this cascode amplifier. For the same supply voltage, the V_{CB} across both transistors is reduced compared to a single transistor amplifier, thus effectively suppressing the effects of impact ionization. The reduced Miller effect of the cascode architecture helps to reduce the nonlinearity due to the CB capacitance. Therefore, the load dependence of IIP3 is weak, as expected. Moreover, the overall third-order distortion output power is smaller than the distortion caused individually by g_m (×), g_{be} (o) or C_{be} (+). As shown below, this is due to nonlinearity cancellation [4]. The degree of cancellation is both bias and geometry dependent.

III. Optimization

To achieve an optimized LNA design which balances input impedance match, noise figure, gain, linearity, and power consumption, we have calculated the noise figure, gain, and linearity contours as a function of device emitter length and biasing current. For each emitter length and current bias combination, l_e and l_b are determined for input impedance matching by numerical optimization. IIP3, gain and noise figure are calculated using an in-house program written in MATLAB.

The upper limit of the collector current is set by the power consumption constraint. A frequency of 2 GHz was used in this optimization, but the approach is valid for all relevant RF frequencies.

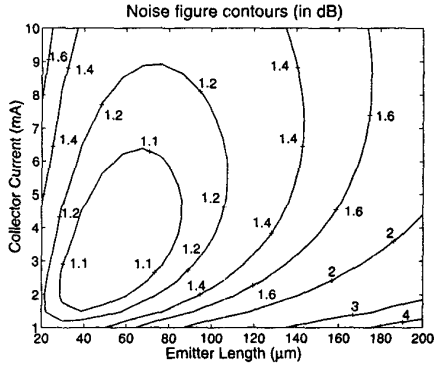

Fig. 5. Noise figure contours as a function of emitter length and collector current for an input impedance matched LNA at 2 GHz.

Fig. 5 shows noise figure contours as a function of emitter length and collector current. The noise figure is minimum at a collector current of about 4 mA and an emitter length of 60 μm. As shown below, this design point is optimum for noise figure, but is not, however, optimum for IIP3.

Fig. 6 shows gain contours as a function of emitter length and collector current. For a given gain requirement (e.g. 15 dBm), the collector current and emitter length should be located above the corresponding gain contour. Gain increases with current density because f_T and f_{max} increase with bias current. This occurs as transistor length decreases for a fixed bias current or as bias current increases for a fixed transistor length.

Fig. 7 shows IIP3 contours as a function of emitter length and collector current. The design space for $NF \leq 1.2$ dB is within the dashed line, and the design space for $gain \geq 15$ dB is above the dash-dot line. Within the design space that meets both $NF \leq 1.2$ dB and $gain \geq 15$ dB, IIP3 changes from approximately −5 dBm to 15 dBm. The optimum design point for a maximum IIP3 is emitter length = 80 μm and collector current = 7.5 mA. The maximum IIP3 is then above 15 dBm with a resultant noise figure of 1.15 dB.

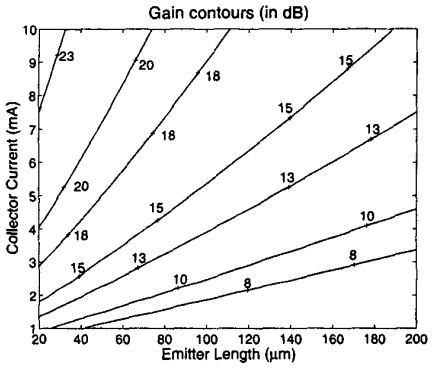

Fig. 6. Gain contours as a function of emitter length and collector current for an input impedance matched LNA at 2 GHz.

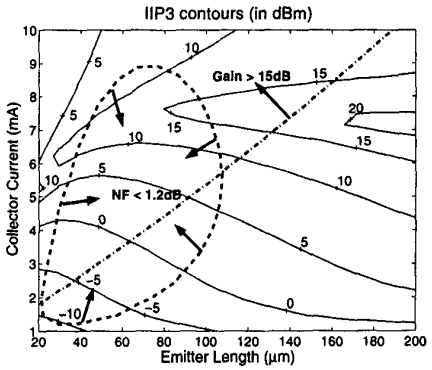

Fig. 7. IIP3 contours as a function of emitter length and collector current for an input impedance matched LNA at 2 GHz. The dashed line is the noise figure contour at 1.2 *dB*, and the dash-dot line is the gain contour at 15 *dB*.

In contrast, IIP3 at the design point optimum for noise figure (I_C = 4 mA, emitter length= 60 μm) is only 0 dBm. The design point optimum for IIP3 is a better choice, because noise figure is still near its minimum, while IIP3 is significantly higher (by 15 dBm). The disadvantage of this approach, however, is that the required bias current is 3.5 mA higher. For an optimized design (80 μm and 7.5 mA), the resultant $IIP3 = 15.8$ dBm, $gain = 18$ dB, $NF = 1.15$ dB, and $|s11| < -20$ dB. If the power consumption constraint is tightened to $I_C \leq 5.5$ mA, an IIP3 of 5 dBm can be obtained at $I_C = 5.5$ mA, emitter length= 50 μm, with a near minimum noise figure of 1.08 dB.

IV. Discussion

To gain a better insight into the impact of emitter length and I_C on LNA design, new forms of analytical equations for IIP3, gain, and noise figure are derived. Distortion cancellation, which dominates IIP3, is written as an explicit expression. Noise figure is decomposed into three terms which better demonstrate the impact of emitter length and I_C. The first-order derivations assume the condition of an input-impedance matched cascode structure.

For the present analysis, the first stage of the cascode structure can be simplified, as shown in Fig. 8. First, the approximated values of l_e and l_b for the input-impedance match are

$$l_e \simeq \frac{R_s}{2\pi f_T} - \frac{2\pi f_T(1/g_{be} - R_s)}{\omega^2\beta^2} \quad (1)$$

$$l_b \simeq \frac{2\pi f_T(1/g_{be} - R_s)}{\omega^2\beta} - l_e$$

$$l_b \simeq \frac{2\pi f_T(1/g_{be} - R_s)}{\omega^2\beta} - \frac{R_s}{2\pi f_T} \quad (2)$$

where $\omega = 2\pi f$ is the operation frequency. g_{be} is neglected in the traditional input impedance matching equations, assuming $1 << \omega\beta/2\pi f_T$ and $R_s << 1/g_{be}$. This assumption, however, is no longer valid here because f_T is much higher than the operational frequency for SiGe.

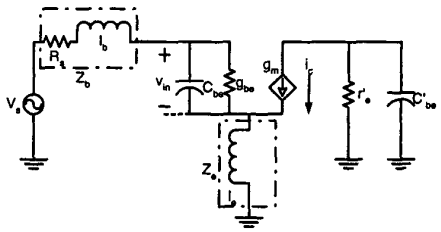

Fig. 8. Simplified circuit for the first stage of cascode LNA.

The third-order intermodulation can be derived under the assumption that nonlinearities due to C_{cs}, impact ionization, and CB capacitance are negligible. Applying Volterra series to the simplified circuit, one finds

$$IM3 \simeq \left| \frac{3}{4} \cdot C(s_1, s_2) \cdot L(s) \cdot (1 - G(2s_1) - 2G(s_1 - s_2)) \right| \quad (3)$$

where $s_1 = j\omega_1$, $s_2 = j\omega_2$, and $C(s)$ are

$$C(s_1, s_2) = \frac{1}{6V_t^2} \cdot v_{in,1}(s_1) \cdot v_{in,1}(-s_2) \quad (4)$$

$$\simeq \frac{1}{6V_t^2} \cdot |v_{in,1}^2(s_1)| \quad (5)$$

where $v_{in,1}$ is the 1st order AC voltage on the EB junction, and

$$L(s) = \frac{V_t \cdot K(s) \cdot G(s)}{I_C} \quad (6)$$

where $L(s)$ is proportional to the current gain $G(s)$ at frequency s, and where

$$G(s) = \frac{A(s) \cdot I_C}{B(s) + A(s) \cdot I_C} \quad (7)$$

$$K(s) = \frac{B(s)}{A(s) \cdot V_t} \quad (8)$$

$$A(s) = \left(s \cdot \tau_f + \frac{1}{\beta} \right) \cdot (Z_b(s) + Z_e(s)) + Z_e(s) \quad (9)$$

$$B(s) = V_t \cdot [1 + s \cdot C_{te}(Z_b(s) + Z_e(s))] \quad (10)$$

$$Z_b(s) = s \cdot l_b + R_s \quad (11)$$

$$Z_e(s) = s \cdot l_e \quad (12)$$

$C(s_1, s_2)$ and $L(s)$ are monotonic functions of emitter length and I_C. The peak IIP3 is obtained while varying the emitter length or I_C as

the cancellation term $|(1 - G(2s_1) - 2G(\Delta s))|$ is minimized. Fig. 9 shows the three terms: $|C(s_1, s_2)|$, $|L(2s_1 - s_2)|$, and $|(1 - G(2s_1) - 2G(\Delta s))|$ together with total IM3, and IIP3 for the input-impedance matched amplifier. The minimum value of the cancellation term, IM3, and the maximum IIP3 occur at the same value of I_C, meaning that the cancellation term dominates IM3.

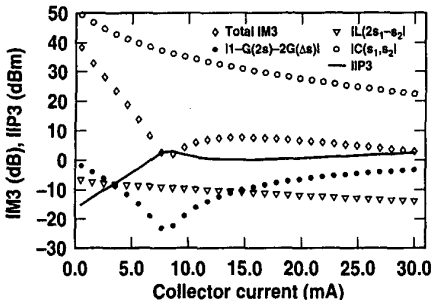

Fig. 9. IM3, IIP3, and the 3 component terms as a function of I_C in the input-impedance matched mode.

The gain of the input impedance matched cascode LNA can be derived under the assumption that the CB capacitance in the first stage is negligible and the current gain of the common-base stage is close to unity at 2 GHz. Since the input impedance is matched to the source resistance R_s, the current injected into the EB junction is constant. Thus, the current gain of the circuit is equal to

$$\frac{\beta g_{be}}{g_{be} + j\omega C_{be}} = \frac{\beta}{1 + j\omega\beta/2\pi f_T}$$

and the power gain can be written as

$$G = \frac{\beta^2 R_{load}}{[1 + (\omega\beta/2\pi f_T)]^2 R_s} \tag{13}$$

At low injection, as j_C increases, f_T increases, and thus the gain increases. Hence gain increases with increasing I_C and fixed emitter length, and decreases with increasing emitter length and fixed I_C, is expected.

A simplified noise model can be obtained based on the assumption that the base resistance (the base resistance is considered only as a thermal noise source), CB capacitance of the first stage, and the Early effect are negligible. The noise contributed by the shot noise source $\bar{i}_b$, $\bar{i}_c$ and thermal noise source $\bar{v}_b$ can be written as

$$NF = 10\log_{10}(1 + n_{ib} + n_{ic} + n_{vb}) \tag{14}$$

$$n_{ib} = \frac{(g_{be}R_s)^2 + [B(1 - g_{be}R_s)]^2}{2g_{be}R_s} \tag{15}$$

$$n_{ic} = \frac{4(g_{be}R_s)^2 + [g_{be}R_s/B + B(1 - g_{be}R_s)]^2}{2g_mR_s} \tag{16}$$

$$n_{vb} = \frac{r_b}{R_s} \tag{17}$$

where n_{ib}, n_{ic} and n_{vb} are the ratios of the output power caused by $\bar{i}_b$, $\bar{i}_c$ and $\bar{v}_b$ to the output power caused by thermal noise $\bar{v}_s$, $B = 2\pi f_T/\omega\beta$. Under a power consumption limitation, I_C normally is less than 10 mA, meaning $g_{be}R_s << 1$. Rewriting the equation above, one obtains

$$n_{ib} = \frac{(g_{be}R_s)^2 + B^2}{2g_{be}R_s} \tag{18}$$

$$n_{ic} = \frac{4(g_{be}R_s)^2 + [g_{be}R_s/B + B]^2}{2g_mR_s} \tag{19}$$

$$n_{vb} = \frac{r_b}{R_s} \tag{20}$$

At low injection conditions and 2 GHz operational frequency, $B << 1$. B is an increasing function of j_C, g_{be} and g_m are increasing functions of I_C, and r_b is a decreasing function of emitter length. For the condition of varying I_C and fixed emitter length: 1) we see that for small I_C, n_{ic} dominates the noise figure, and decreases as I_C increases; 2) at larger I_C, n_{ib} dominates the noise figure, and increases as I_C increases. For the condition of varying emitter length and fixed I_C: 1) we see that for small emitter length n_{rb} dominates the noise figure, and decreases as emitter length increases; 2) at larger n_{ic} dominates noise figure, and increases as emitter length increases. Thus the emitter length and I_C value for minimum noise figure can be obtained by using the equations above.

Our analytical equations nicely capture the results calculated by numerical method describe above. Fig. 10 shows the IIP3 as a function I_C calculated by both our numerical method and analytical equations. A good agreement between these two methods is achieved. Based on the analytical equations, one can thus roughly locate the optimum design point without complex computations.

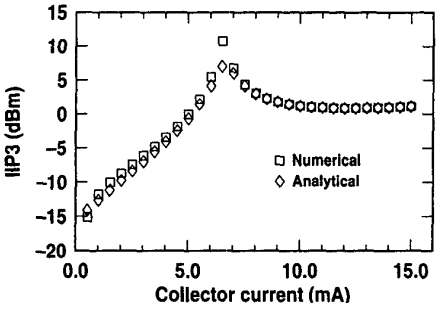

Fig. 10. IIP3 as a function of I_C with emitter length of 40 μm at 2 GHz, calculated using numerical and analytical methods.

V. ACKNOWLEDGMENT

This work was supported by an IBM University Partner Award, SRC under #2000-HJ-769 and #2001-NJ-937, and National Science Foundation under ECS-0119623 and ECS-0112923. The wafers were fabricated at IBM Microelectronics, Essex Junction, VT. We would like to thank D. Herman, S. Subbanna, A. Joseph, and B. Meyerson for their support and contributions.

REFERENCES

[1] D.L. Harame *et al.*, "Si/SiGe epitaxial-base transistors–Part I: Materials, physics, and circuits," *IEEE Trans. Electron Devices*, vol. 42, pp. 469-482, 1995.

[2] D. Shaeffer *et al.*, "A 1.5V, 1.5GHz CMOS Low Noise Amplifier," *IEEE Journal Solid-State Circuits*, vol. 32, pp. 745-759, 1997.

[3] G.F. Niu *et al.*, "RF linearity characteristics of SiGe HBTs," *IEEE Trans. Microwave Theory and Techniques*, vol. 49, pp. 1558-1565, 2001

[4] S. Maas *et al.*, "Intermodulation in heterojunction bipolar transistors," *IEEE Trans. Microwave Theory and Techniques*, vol. 40, pp. 442-448, Mar. 1992.