# LINEARIZATION OF CMOS LNA'S VIA OPTIMUM GATE BIASING

*Vladimir Aparin, Gary Brown*

QUALCOMM Inc.
5775 Morehouse Drive
San Diego, CA 92121, USA

*Lawrence E. Larson*

University of California in San Diego
9500 Gilman Drive
La Jolla, CA 92093, USA

**ABSTRACT**

A FET linearization technique based on optimum gate biasing is investigated at RF. A novel bias circuit is proposed to generate the gate voltage for zero 3rd-order nonlinearity of the FET transconductance. The measured data show that a peak in $IIP_3$ occurs at a gate voltage slightly different from the one predicted by the dc theory. The origins of this offset are explained based on a Volterra series analysis and confirmed experimentally. The technique was used in a 0.25$\mu$m CMOS cellular-band CDMA LNA. At the optimum bias, the amplifier achieved a NF of 1.8dB, an $IIP_3$ of +10.5dBm, and a power gain of 14.6dB with a current consumption of only 2mA from 2.7V supply.

## 1. INTRODUCTION

The single-tone desensitization requirement for CDMA phones demands a very high LNA $IIP_3$ in combination with low NF, high gain and low current consumption. This design challenge requires the use of linearization techniques. A simple technique based on low-frequency low-impedance termination of the LNA input has been shown to work reliably in Si BJT and SiGe HBT LNAs [1]. But it is not applicable to FETs. In [2], feed-forward linearization was used to achieve a very high $IIP_3$ of a CMOS LNA. Since this technique is very sensitive to mismatches between the main and auxiliary gain stages and errors in the signal scaling, it is rarely used in high-volume production. A FET can also be linearized by biasing at a gate-source voltage ($V_{\rm GS}$) at which the 3rd-order derivative of its dc transfer characteristic is zero [3]. One of the major drawbacks of this technique is that a significant $IIP_3$ improvement occurs in a very narrow $V_{\rm GS}$ range (about 10-20mV) around the optimum voltage and there are no known bias means to automatically generate this voltage. Therefore, a bias circuit has to be manually tuned to this voltage, which makes $IIP_3$ to vary significantly with process and temperature variations.

This paper proposes a novel bias circuit that generates the gate bias voltage at which the 3rd-order nonlinearity of the FET transconductance is zero and automatically tracks the process and temperature variations. The effect of the optimum gate biasing on linearity at RF frequencies is analyzed using Volterra series. We show that there exist an offset between $V_{\rm GS}$ at which $IIP_3$ peaks and $V_{\rm GS}$ at which the 3rd-order nonlinearity of the FET transconductance is zero. The origins of this offset are explained. Measured data collected on a 0.25$\mu$m CMOS LNA designed for cellular CDMA applications confirm the theoretical findings.

## 2. DC THEORY OF OPTIMUM GATE BIASING

Consider a common-source FET biased in saturation. Its small-signal output current can be expanded into the following power series in terms of the small-signal gate-source voltage $v_{gs}$ around the bias point

$$i_d(v_{gs}) = g_1 v_{gs} + g_2 v_{gs}^2 + g_3 v_{gs}^3 + \cdots \quad (1)$$

where $g_1$ is the small-signal transconductance and the higher-order coefficients ($g_2$, $g_3$ etc.) define the strengths of the corresponding nonlinearities. Among these coefficients, $g_3$ is particularly important because it controls the 3rd-order intermodulation distortion ($IM_3$) at low signal levels and, thus, determines $IIP_3$, which is given by

$$A_{\rm IP3} = \sqrt{\frac{4}{3}\left|\frac{g_1}{g_3}\right|} \quad [\mathrm{V}]. \quad (2)$$

The power series coefficients generally depend on the dc gate-source and drain-source voltages, $V_{\rm GS}$ and $V_{\rm DS}$. However, the dependence on $V_{\rm DS}$ for a FET in saturation can be neglected. Then the coefficients of (1) can be found as

$$g_1(V_{\rm GS}) = \frac{\partial I_D}{\partial V_{\rm GS}}, \quad (3a)$$

$$g_2(V_{\rm GS}) = \frac{1}{2}\frac{\partial^2 I_D}{\partial V_{\rm GS}^2} = \frac{1}{2}\frac{\partial g_1(V_{\rm GS})}{\partial V_{\rm GS}}, \quad (3b)$$

$$g_3(V_{\rm GS}) = \frac{1}{6}\frac{\partial^3 I_D}{\partial V_{\rm GS}^3} = \frac{1}{3}\frac{\partial g_2(V_{\rm GS})}{\partial V_{\rm GS}}. \quad (3c)$$

As an example, we measured the dc transfer characteristic of a 350$\mu$m/0.25$\mu$m NFET used in the LNA described

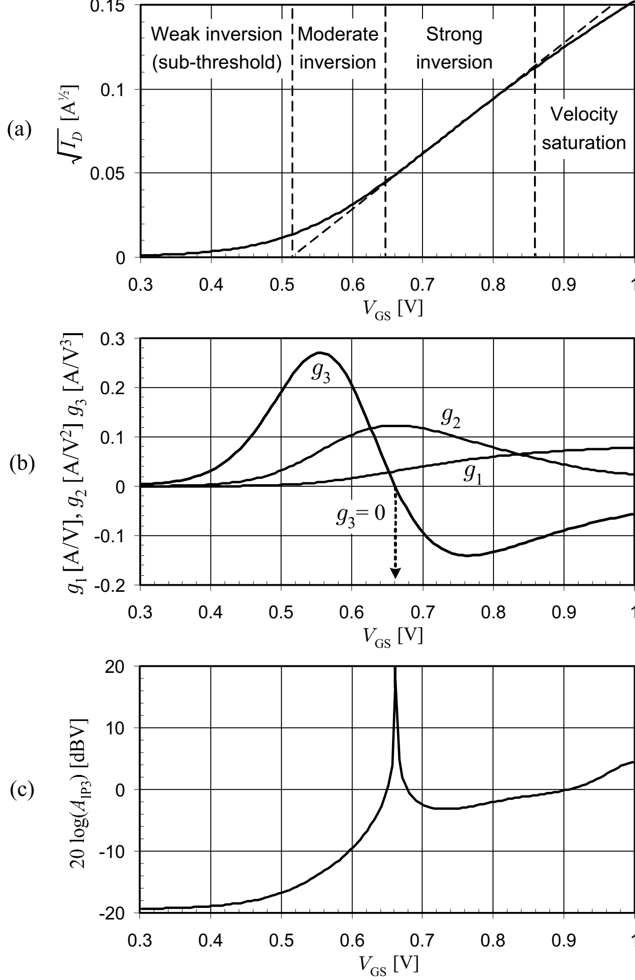


**Fig. 1**. 350$\mu$m/0.25$\mu$m NFET. (a) Measured dc transfer characteristic. (b) Power series coefficients. (c) Theoretical $A_{\rm IP3}$.

later here. Then, we extracted its power series coefficients and calculated $A_{\rm IP3}$ using (2). The results are presented in Fig. 1. As can be seen, in the transition region from moderate to strong inversion, there is a $V_{\rm GS}$ at which $g_3 = 0$ and $A_{\rm IP3} = \infty$. The peak in $A_{\rm IP3}$ is very narrow and requires an accurate biasing within $\pm$10mV of the optimum voltage. Manually tuned bias circuits are not capable of such precision in the presence of process and temperature variations.

## 3. BIAS CIRCUIT FOR ZERO $G_3$

In order for a bias circuit to generate and automatically maintain the bias voltage for zero $g_3$, it should produce a dc voltage or current proportional to $g_3$ and have a dc feedback to set it to zero. The insight into how to design such a bias

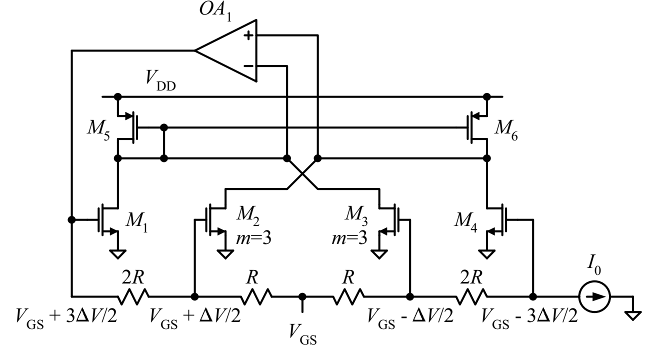


**Fig. 2**. Bias circuit for zero $g_3$.

circuit can be gained if equations (3a)-(3c) are rewritten in terms of small deviations of the dc voltages and currents:

$$g_1(V_{\rm GS}) = \frac{I_D(V_{\rm GS}+\Delta V/2) - I_D(V_{\rm GS}-\Delta V/2)}{\Delta V}, \quad (4a)$$

$$g_2(V_{\rm GS}) = \frac{g_1(V_{\rm GS}+\Delta V/2) - g_1(V_{\rm GS}-\Delta V/2)}{2\Delta V} = \frac{I_D(V_{\rm GS}+\Delta V) + I_D(V_{\rm GS}-\Delta V) - 2I_D(V_{\rm GS})}{2\Delta V^2}, \quad (4b)$$

$$g_3(V_{\rm GS}) = \frac{g_2(V_{\rm GS}+\Delta V/2) - g_2(V_{\rm GS}-\Delta V/2)}{3\Delta V} = \frac{1}{6\Delta V^3}\left[\begin{matrix} I_D(V_{\rm GS}+3\Delta V/2) + 3I_D(V_{\rm GS}-\Delta V/2) \\ -3I_D(V_{\rm GS}+\Delta V/2) - I_D(V_{\rm GS}-3\Delta V/2)\end{matrix}\right]. \quad (4c)$$

If the term in the brackets of (4c) is zero, then $g_3$=0. The bias circuit that generates this term and automatically tunes $V_{\rm GS}$ for zero $g_3$ is shown in Fig. 2. FETs $M_1$-$M_4$ are scaled versions of the main transistor (not shown). They are built of the same unit cells. $M_2$ and $M_3$ have three times more unit cells than $M_1$ and $M_4$. $M_1$ generates the first current in the brackets of (4c). $M_2$ generates three times the third current. $M_3$ generates three times the second current and $M_4$ generates the fourth current. The currents of $M_1$ and $M_3$ (as well as $M_2$ and $M_4$) are added by connecting their drains together. The common drains are biased through the current mirror $M_5$/$M_6$ where $M_5$ and $M_6$ are PFETs of equal size. If the drain voltages of $M_5$ and $M_6$ are equal, their currents are equal and the term in the brackets of (4c) is zero. This balance is ensured by the operational amplifier $OA_1$ that senses the difference between the drain voltages and generates the input voltage for the resistor chain $2R/R/R/2R$ creating a feedback loop. The generated bias voltage for the main FET is tapped at the center of the resistor chain.

The generated $V_{\rm GS}$ is fairly insensitive to the value of $\Delta V$ (=$2I_0R$) as indicated in Fig. 3. However, it is very sensitive to mismatches between $M_1$-$M_4$. Its deviation from the optimum $V_{\rm GS}$ is maximum if the threshold voltage of

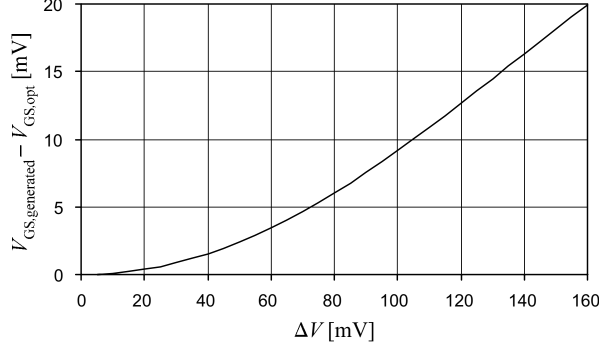


**Fig. 3**. Simulated deviation of generated $V_{\rm GS}$ in Fig. 2 from the optimum voltage for zero $g_3$ as a function of $\Delta V$.

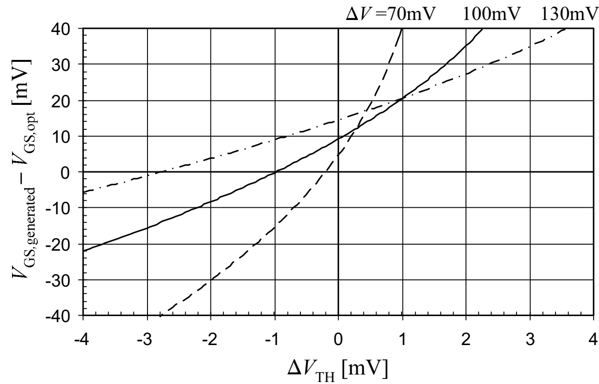


**Fig. 4**. Simulated deviation of generated $V_{\rm GS}$ from the optimum voltage for zero $g_3$ as a function of $\Delta V_{\rm TH}$.

$M_1$ and $M_3$ shifts in one direction and that of $M_2$ and $M_4$ shifts in the opposite direction. For simplicity, we will assume that $M_1$ is matched to $M_3$ and $M_2$ to $M_4$. Fig. 4 plots the deviation of the generated $V_{\rm GS}$ from the optimum $V_{\rm GS}$ as a function of the threshold voltage mismatch between $M_1$/$M_3$ and $M_2$/$M_4$ for different values of $\Delta V$. Positive $\Delta V_{\rm TH}$ corresponds to lower $V_{\rm TH}$ of $M_1$/$M_3$ than that of $M_2$/$M_4$. As can be seen, a higher threshold of $M_1$/$M_3$ compensates for the systematic error in $V_{\rm GS}$ due to $\Delta V$. The standard deviation of $\Delta V_{\rm TH}$ can be estimated as [4]

$$\sigma_{\Delta V_{\rm TH}} \approx \frac{A_{V_{\rm TH}}}{\sqrt{WL_g n_f}} \quad [\text{mV}] \tag{5}$$

where $A_{V_{\rm TH}} \approx$1mV$\mu$m per 1nm of the gate oxide thickness, $W$ is the gate finger width in $\mu$m, $L_g$ is the gate length in $\mu$m, and $n_f$ is the total number of gate fingers in $M_1$ and $M_3$ (or $M_2$ and $M_4$). For the given 0.25$\mu$m CMOS process, the oxide thickness is 5nm and, thus, $A_{V_{\rm TH}} = 5$mV$\mu$m. With $W = 10\mu$m, $L_g = 0.25\mu$m and $n_f = 20$, we get $\sigma_{\Delta V_{\rm TH}} = 0.71$mV. As can be seen from Fig. 4, to minimize

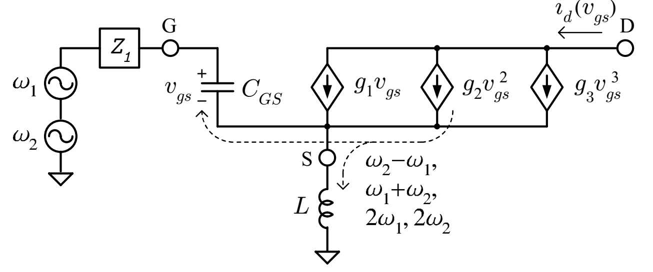


**Fig. 5**. Small-signal nonlinear equivalent circuit of a common-source FET.

the $V_{\rm GS}$ error within 2-3 times $\sigma_{\Delta V_{\rm TH}}$, $\Delta V$ should be more than 100mV. This error is still larger than the $V_{\rm GS}$ range within which a significant $IIP_3$ improvement occurs (see Fig. 1(c)). To further reduce the $V_{\rm GS}$ error, a process with thinner gate oxide can be selected.

## 4. RF THEORY OF OPTIMUM GATE BIASING

Consider a small-signal nonlinear equivalent circuit of a FET shown in Fig. 5 where $Z_1$ is the transformed output impedance of the signal generator and $L$ is the source degeneration inductance. Here we neglect $C_{GD}$ and the nonlinearities of the order higher than three. In this *weakly nonlinear* case, the $IM_3$ distortion would be generated entirely by the $g_3 v_{gs}^3$ component of the drain current if $L$ was zero. The source degeneration inductance creates a feedback path for the drain current to $v_{gs}$. This feedback is particularly strong for high frequency spectral components of $i_d$. For example, the 2nd harmonics generated by $g_2 v_{gs}^2$ are fed back across the gate and source adding to the fundamental components of $v_{gs}$. These spectral components are then mixed in $g_2 v_{gs}^2$ to produce the responses at $2\omega_1 \pm \omega_2$ and $2\omega_2 \pm \omega_1$. Thus, the 2nd-order nonlinearity of $i_d$ also contributes to $IM_3$.

Reusing the Volterra series results for the circuit in Fig. 5 from [1] and assuming that $\Delta\omega$ $(= \omega_2 - \omega_1)$ is much smaller than $\omega_1$ and $\omega_2$ such that $j\Delta\omega L \approx 0$ and the signal generator is conjugately matched to the FET input at $\omega_1$ and $\omega_2$, we get the following expression for $IIP_3$:

$$IIP_3 = \frac{4g_1^2\omega^2 L C_{\rm GS}}{3|\varepsilon(\Delta\omega, 2\omega)|} \tag{6}$$

where

$$\varepsilon(\Delta\omega, 2\omega) = g_3 - \frac{2g_2^2/(3g_1)}{1 + \dfrac{1}{j2\omega L g_1} + \dfrac{j2\omega}{\omega_T} + \dfrac{Z_1(2\omega)}{\omega_T L}}, \tag{7}$$

$\omega \approx \omega_1 \approx \omega_2$ and $\omega_T = g_1/C_{\rm GS}$. A peak in $IIP_3$ occurs when $g_3$ is equal to the real part of the second term in (7). It can be shown that this real part is positive and,

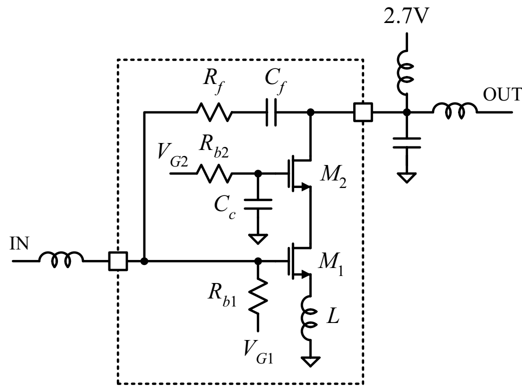


**Fig. 6**. Simplified schematic diagram of LNA.

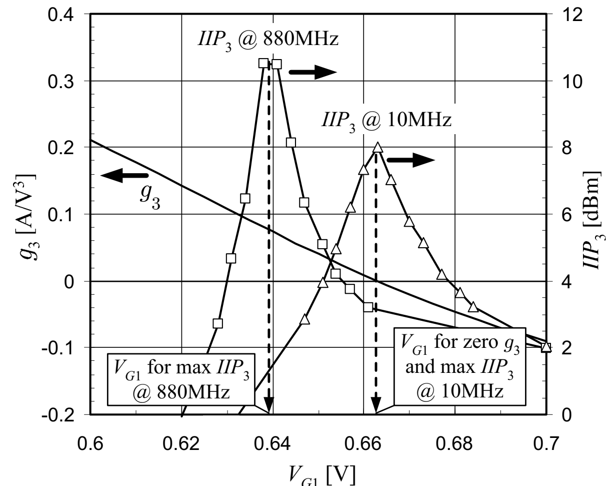


**Fig. 7**. Measured LNA $IIP_3$ and extracted $g_3$ of $M_1$ as functions of $M_1$ gate bias.

therefore, $g_3$ must be positive as well for the real part of $\varepsilon(\Delta\omega, 2\omega)$ to be zero. As a result, the peak in $IIP_3$ occurs below the optimum bias voltage for zero $g_3$. The imaginary part of $\varepsilon(\Delta\omega, 2\omega)$ limits the magnitude of this peak. For a relatively large degeneration inductance, $IIP_3$ may not peak at all. We found that the imaginary part of the second term in (7) can be tuned to zero by optimizing the load impedance of the FET drain. However, the real part of this term will remain nonzero causing a shift in the $IIP_3$ peak.

## 5. LNA DESIGN AND MEASURED RESULTS

To prove our theory, a cellular-band CDMA LNA was designed and fabricated in $0.25\mu$m CMOS technology. Its schematic is shown in Fig. 6 where the dashed box indicates the chip boundaries. The source degeneration inductor is just a bondwire ($\approx 0.45$nH). The load impedance presented to the drain of $M_1$ is optimized by tuning $C_c$ such that the imaginary part of (7) is nearly zero. The measured $IIP_3$ as a function of $V_{G1}$ with two tones centered at 880MHz and separated by 1MHz is shown in Fig. 7 together with $g_3$ of $M_1$. The $IIP_3$ exhibits a sharp peak of +10.5dBm at a voltage that is 22mV below $V_{G1}$ at which $g_3$ is zero. The latter gives $IIP_3$ of only +3.2dBm. To verify that the offset between $V_{G1}$ for peak $IIP_3$ and $V_{G1}$ for zero $g_3$ is due to the LNA reactances, $IIP_3$ of the same LNA was also measured with the two tones centered at 10MHz and its dependence on $V_{G1}$ is plotted in Fig. 7. As can be seen, $IIP_3$ in this case peaks exactly at $V_{G1}$ at which $g_3$=0. The measured LNA gain and NF are 14.6dB and 1.8dB respectively at the peak-$IIP_3$ bias in the cellular frequency band. The current consumption is 2mA from 2.7V excluding the bias circuit.

## 6. CONCLUSIONS

We have shown that biasing a common-source FET at $V_{\mathrm{GS}}$ for zero $g_3$ causes a significant improvement in $IIP_3$ at low frequencies. We proposed a bias circuit that automatically generates this optimum voltage and showed that its precision can be satisfactory, provided that the mismatch between the reference FET's is reduced by increasing their total gate area and selecting a process with a thin gate oxide. We analyzed the effect of the optimum gate biasing on $IIP_3$ at RF and showed that the circuit reactances that introduce feedback can shift the peak in $IIP_3$ away from the bias voltage for zero $g_3$ and reduce the magnitude of this peak. We proposed a method based on tuning the drain load impedance to cancel the effect of the reactances on the maximum $IIP_3$. However, the offset between $V_{\mathrm{GS}}$ at which $IIP_3$ is maximum and $V_{\mathrm{GS}}$ at which $g_3 = 0$ can not be cancelled. Thus, a manual bias tuning is required to achieve a significant $IIP_3$ improvement at RF making this technique sensitive to process and temperature variations.

## 7. REFERENCES

[1] V. Aparin and L. E. Larson, "Linearization of monolithic LNAs using low-frequency low-impedance input termination," *Europ. Solid-State Circ. Conf.*, pp. 137-140, Sept. 2003.

[2] Y. Ding and R. Harjani, "A +18dBm IIP3 LNA in 0.35um CMOS," *IEEE ISSCC*, pp. 162-163, 2001.

[3] G. Qu and A. E. Parker "Analysis of intermodulation nulling in HEMTs," *Conf. on Optoelectr. & Microelectr. Mat. & Dev.*, pp. 227-230, Dec. 1996.

[4] M. J. M. Pelgrom, H. P. Tuinhout, and M. Vertregt, "Transistor matching in analog CMOS applications," *IEDM Dig.*, pp. 915-918, Dec. 1998.