Enhanced Gm3 Cancellation For Linearity Improvement in CMOS LNAs

Mallesh Rajashekharaiah, Parag Upadhyaya, Deukhyoun Heo
School of Electrical Engineering and Computer Science
Washington State University
Pullman, USA

Abstract—An enhanced linearity improvement technique based on the third order intermodulation distortion cancellation or $gm^3$ cancellation technique is presented in this paper. By identifying the issues related to the conventional $gm^3$ cancellation method, the proposed technique is used to design a 0.18um CMOS LNA using Jazz semiconductor's BiCMOS process. With an IIP3 of more than +16dBm and gain of more than 15dB, the enhanced linearity LNA or EL2NA with a current consumption of 9mA from a 1.8V supply, provides an efficient way of improving linearity of 5G range direct conversion receivers.

I. INTRODUCTION: LINEARITY PHENOMENON IN CMOS LNAs

Linearity is one of the most critical features of a front component like the low noise amplifier (LNA) and directly impacts the linearity of the whole receiver module. A front end with good linearity performance can minimize several RF circuit issues like harmonic generation, intermodulation distortion, cross talk and desensitization [1]. Odd order distortions are of particular importance due to the difficulty and complexity associated with reducing their effect. Chief among them is the third order intermodulation distortion or IMD3 [2]. A popular measuring parameter for intermodulation distortion levels is the third order intercept point or IP3. The IP3 can be measured using a simple two-tone test in which the signal amplitude, $A$, is chosen to be sufficiently small so that higher order nonlinear terms are negligible and gain is relatively constant. As $A$ increases, the fundamentals increase in proportion to $A$, while the third order components increase in proportion to $A^3$. When plotted on a logarithmic scale, the slope of third order IM products is three times that of fundamental (See Fig 1). The point at which the extrapolated curves meet is called the third order intercept point or IP3. The vertical coordinate of this point is called the output IP3 or OIP3 and the horizontal coordinate is called the input IP3 or IIP3.

The IIP3 is a very effective measure of the circuit linearity, more than simple IM measurement. If the magnitude of IM products is used as a measure of linearity, then the input amplitude needs to be specified [1]. On the other hand, the IIP3 by itself can be a unique means of comparing linearity of different circuits. Higher the IIP3, more linear is the circuit.

II. LINEARITY ENHANCEMENT TECHNIQUES

There are several methods proposed in the literature to improve an LNA’s linearity or third order inter-modulation performance. One is the use of a multi-tanh topology, which is popular in the BJT domain (Fig.2) [4]. The basic principle involved in this technique is illustrated in Fig. 2. By superposition of the offset $g_{m}$ characteristics, using either voltage offset or current offset, the resultant $gm$ characteristic can be made such that it is linear over a wider range of input voltages, as seen in the figure. This principle can be extended further to include more transistor pairs to further increase the linear region area. It should be noted however, that the area under the $g_{m}(Vin)$ curve remains constant, irrespective of the change in shape.
The additional stages for multi-tanh topology are, however, detrimental to the noise figure of the circuit. A similar technique is present for the CMOS domain, although it targets the third order transconductance of the devices rather than the first. This technique, called the \( g_{m3} \) cancellation technique, will be explained in the following sections. Another major issue related with this method is the increased power consumption. For state-of-the-art DCR solutions, the low power consumption is an important requirement and hence the multi-tanh principle is not very efficient.

Source degeneration is another effective linearity enhancement technique. As seen in Fig.3, the topology consists of a source degenerated transistor, whose \( g_m \) characteristics reduce to a linear ratio of source and drain loads. This linear ratio is what forms the core of this method. However, there is a direct trade-off of the IIP3 with the gain of the amplifier.

Active and passive feedback is another useful method for improving the linearity of an LNA (See Fig.4). The enhancement in linearity due to feedback topologies is not significant. Also, passive feedback can be detrimental to noise and also may pose circuit stability issues due to the presence of uncontrolled feedback. Active feedback on the other hand provides controllable feedback and aids in reducing harmonic feedback in the gate-drain region, thus resulting in reduced third order intermodulation. However, the improvement is balanced down by increased noise at the input and increased power consumption.

III. \( GM3 \) CANCELLATION TECHNIQUE

An effective method of reducing third order intermodulation is to cancel the third order transconductance \( (g_{m3}) \). The drain current of a common source FET can be expressed by the Taylor series as,

\[
i_{DS} = I_{dc} + g_m v_{gs} + \frac{g_{m2}}{2!} v_{gs}^2 + \frac{g_{m3}}{3!} v_{gs}^3
\]

Here, \( v_{gs} \) is the small signal gate-to-source voltage and \( g_{mn} \) is the \( n \)th order transconductance. The third order transconductance \( g_{m3} \) plays a very important role in the linearity of a device. A plot of \( g_{m3} \) (See Fig.5) shows a negative peak in the voltage range of 0.5V to 0.9V, which is the normal operating gate voltage to achieve optimum power consumption with good matching and noise figure. Using the transconductance of an additional parallel transistor, M2, operating in the triode region, this negative peak can be compensated for (See Fig. 5). This effectively leads to improved IIP3 and hence, higher linearity [5]. But again increased power consumption is an issue with an additional current path. To keep the power consumption low while using an additional transistor, this secondary transistor can to be operated in the sub-threshold region.
Although the $g_{m3}$ cancellation technique works well in theory, in practice there are several issues involved. Earlier designs have all used negative feedback, in the form of a degeneration inductor, for linearization in addition to nonlinearity cancellation [5]. The major drawback associated with this technique is the contribution of second order nonlinearity to the IMD3. The source degeneration inductance creates a feedback path for the drain current to the gate-source voltage, $v_{gs}$. This feedback is particularly strong for high frequency spectral components of the drain current [5]. For example, the 2nd harmonics generated by $g_{m3}v_{gs}$ are fed back across the gate and source adding to the fundamental components of $v_{gs}$. These spectral components are then mixed in $g_{m3}v_{gs}$ to produce the responses at $2v_{gs}$ and $2v_{gs}+v_{gs}$. Thus, use of a degeneration inductor introduces unwanted spectral components leading to degradation of system linearity.

IV. ENHANCED GM3 CANCELLATION TECHNIQUE

In order to overcome the mentioned issues, we introduce the use of shunt feedback instead of series feedback, as in the case of a degeneration inductor. Thereby, the advantages of negative feedback to reduce the harmonic feedback are retained, while the contribution of second order nonlinearities to the overall IMD3 is reduced considerably. Further, unlike series feedback which increases the impedance at the input gate of M1, shunt feedback does not [6]. This effectively translates to a lower noise contribution. Active feedback is an option that can be pursued, but the downside is the increase in power consumption due to additional current in the feedback path [7]. Fig.6 shows the proposed architecture for the Enhanced Linearity Low Noise Amplifier (EL2NA).

Simulation results show an IIP3 of +15dBm with 9mA of current drawn from a 1.8V supply while providing a gain of 15dB at 5.6GHz frequency. Input and output reflection coefficients are shown in Fig.7 and are less than -15dB. The noise figure is below 2dB, and it needs to be mentioned here that the $g_{m3}$ cancellation topology inherently has a higher noise contribution [5]. The use of the enhanced $g_{m3}$ cancellation technique gives an improvement of at least 10dBm over conventional methods of IIP3 improvement. All simulation results are shown in Fig.7.
3.00 5.00 7.00 9.00 11.00 13.00 15.00
4.0 4.5 5.0 5.5 6.0 6.5 7.0
Freq (GHz)
S21 (dB)
1.50 2.00 2.50 3.00 3.50 4.00 4.50 5.00
NF (dB)
1.50 2.00 2.50 3.00 3.50 4.00 4.50 5.00
S11 S22 (dB)

(b) S21 and NF

(c) S11 and S22

Figure 7. Simulation results for EL2NA (a) IIP3 (b) S21 and NF (c) S11 and S22

The proposed LNA was designed using the 0.18um CMOS process from Jazz Semiconductor’s SiGe BiCMOS technology. All simulations were carried out in Agilent’s Advanced Design System (ADS) and layout was done in Cadence Virtuoso Layout editor. The silicon layout of the EL2NA is shown in Fig 8. The total area of the LNA is less than 0.8mm<sup>2</sup>, including the test pads.

The EL2NA, when compared to the previously published high-linearity LNA structures in the 5GHz frequency range, has the best overall figure of merit (FOM). The FOM is defined as,

\[
FOM = \frac{OIP3}{(NF-1)\times(power\ consumption)}
\]  

(2)

The state-of-the-art in 5GHz highly linear LNA designs is [8] and although it displays very good IIP3, the presence of multiple modules in the circuit, would lead to high power consumption. Fabricated LNA chip from Jazz foundry is awaited to carry out the silicon characterization of the proposed circuit. In order to aid in the silicon characterization process, the matching was implemented partly on-chip and partly off-chip. On the chip, multiple shunt capacitors have been added instead of a single large one at the input and output. Depending on the matching requirements, one or more of the capacitors can be disconnected using a laser cutter. This introduces a certain amount of flexibility in the matching process while maintaining the advantages associated with on-chip passive components. The off-chip part of the matching networks provides larger degree of flexibility.

![Figure 8. Enhanced Linearity LNA layout](image)

**ACKNOWLEDGMENT**

The authors would like to thank CDADIC for funding the work. The authors would also like to thank Jeff Dykstra of Motorola and Matt Miller of Freescale Semiconductor for their continued support.

**REFERENCES**