A wideband IM3 cancellation technique for CMOS attenuators

W. Cheng, M.S. Oude Alink, A.J. Annema, G. J.M. Wienk, B. Nauta

University of Twente, Enschede, Netherlands

Contact author: Wei Cheng;
Mailing address: Room 2009, Building Carre, Hallenweg 15, University of Twente, 7522NH Enschede, The Netherlands;
Telephone: office +31534892727, mobile +31616291178;
Fax: +31534891034;
Email: w.cheng@utwente.nl;

Abstract:

A highly linear Π attenuator system using a wideband IM3 cancellation technique is presented that provides 4 discrete attenuation levels with 6dB spacing for DC-5GHz. For the whole band, S11<-14dB, attenuation flatness<1.6dB, +10dBm input P1dB and +26dBm IIP3 are achieved. For the TV band (0.1GHz-1.2GHz) +30dBm IIP3 is achieved. The active area is 0.0054mm² in a standard 0.16um bulk CMOS process.

Text:

In the receiver path and in spectrum analyzers typically gain control blocks are used to limit the incident power to the level that the receiver circuitry can handle without degrading the linearity; in the transmitter path stringent power control is also desirable. Although variable-gain amplifiers (VGAs) traditionally implement the gain control block, attenuators based on FET transistors show superior performances on linearity, power handling capability and power consumption.
Much effort [1-3] has been devoted to improving the linearity and power handling capability of attenuators. The adaptive bootstrapped body biasing [1] is used in a cascaded $\Pi$ attenuator to suppress the body-related parasitic effects and improve $P_{1\text{dB}}$. A $\Pi$ attenuator with parallel branches designed for discrete attenuation steps achieving $+23\text{dBm}$ IIP3 is shown in [2]. The stacked FET technique used in [3-4] distributes the signal swing among many FETs in series to reduce the drain-source voltage swing for each FET and hence reduce the IM3 distortion. However, the large transistors required by this technique bring in large parasitic capacitances, which lower the bandwidth and increase the minimum insertion loss at high frequency. Moreover, the capacitive nonlinearities introduced by large parasitic capacitances will limit the highest achievable IIP3. Therefore, this technique is mainly effective in SOI CMOS [3].

This paper proposes a wideband IM3 cancellation technique for bulk CMOS $\Pi$ attenuators where the IM3 distortion currents of transistors within the attenuator cancel each other. In the $\Pi$ attenuator in Fig. 1, the voltage swing across the nonlinear output conductances of transistor $M_1$, $M_2$ and $M_3$ generate IM3 distortion currents $i_{D}^{M_1}$, $i_{D}^{M_2}$ and $i_{D}^{M_3}$ defined from drain to source. The distortion current of $M_1$ ($i_{D}^{M_1}$) is directed towards $R_{\text{load}}$, while the distortion currents of $M_2$ and $M_3$ ($i_{D}^{M_2}$ and $i_{D}^{M_3}$) are directed out of $R_{\text{load}}$. This suggests that the distortion due to $M_1$ can cancel the distortion generated by $M_2$ and $M_3$. Using the weakly nonlinear distortion analysis approach in [5], for a certain attenuation value $0 \leq A \leq 1$, the IM3 output voltage is given by

$$v_{\text{out}}^{\text{IM3}} \propto \frac{3xV_{\text{in}}^3}{(1+A)^4} \times \left[ \frac{16A^4}{W_{M_1}^3} - \frac{(1-A)^4A^3}{W_{M_2}^3} - \frac{(1-A)^4A^3}{W_{M_3}^3} \right]$$ (1).

In deriving (1) it is assumed that the transistors’ third-order output conductance nonlinearity dominantly contributes to IM3. Equation (1) suggests that in the $\Pi$ attenuator, the IM3 current of
M_1 at least partially cancels the IM3 current of M_2 and M_3 at the output. When \( \frac{16A^3}{W_{M_1}^3} = \frac{(1-A)^4A^2}{W_{M_2}^3} + \frac{(1-A)^4}{W_{M_3}^3} \) there is full IM3 cancellation within the \( \Pi \) attenuator; then IIP3 is limited by mechanisms such as capacitive nonlinearities and some conductive distortion terms of the transistors that become dominant once the third-order output distortion term is cancelled.

To verify (1) we implemented the \( \Pi \) attenuator shown in Fig. 1 with four parallel branches in a standard 0.16um bulk CMOS process. For this attenuator, each parallel branch is designed for -12dB attenuation \( (A = 0.25) \) and has different width for M_1 that yields either full or partial IM3 cancellation. For the measurements, we enable M_2 and M_3 while selectively enabling one parallel branch at a time. This mimics a \( \Pi \) attenuator with a selectable W_{M_1} for fixed W_{M_2} (23um) and W_{M_3} (20um). For a two-tone input signal centered at 1GHz with 3.2MHz spacing, the IIP3 is extrapolated from an input power range of -15dBm to -10dBm.

The measured and simulated IIP3 (using PSP model) as a function of W_{M_1} is shown in Fig. 2, which verifies this IM3 cancellation theory. Note that the calculated optimal W_{M_1} is 18um while it is 20um from simulations. For small W_{M_1}, M_1 is dominant for the IM3. As W_{M_1} increases its distortion decreases and hence IIP3 increases until its maximum at full IM3 cancellation. For even larger W_{M_1} the IIP3 is dominated by M_2 and M_3 yielding a saturated sub-optimum value because of the fixed width of M_2 and M_3. Fig. 2 also shows the measured IIP3 for four different W_{M_1} by sweeping the center frequency from 0.1GHz to 3GHz. For the \( \Pi \) attenuator with full IM3 cancellation (W_{M_1}=20um), the IIP3 decreases as the frequency increases because of the phase difference between the distortion current of M_1 and that of M_2 and M_3 increasingly deviates from
180° due to parasitic capacitances. As a result, the IM3 cancellation degrades as the frequency increases. Since this IM3 cancellation technique involves no extra devices but only properly dimensioning the transistors’ width, there is no need to trade large transistor for high linearity as in [3].

A four-branch attenuator system with selectable attenuation, see Fig. 3, using this IM3 cancellation technique was fabricated in a standard 0.16um CMOS process. Each branch contains one Π attenuator that is designed for one specific attenuation setting (-6dB, -12dB, -18dB and -24dB) and optimized for full IM3 cancellation. A 3-to-8 digital decoder is used to provide the controlling voltage (1.8V for enabling and 0V for disabling). The active area for the attenuator and the digital decoder is 50x30 um$^2$ and 60x65 um$^2$ respectively. For each of the forenamed attenuation settings, only one branch is enabled. For minimum signal attenuation, the series transistors in all four branches are enabled, and the shunt transistors in all four branches are disabled, yielding an additional -1.8dB attenuation setting. For isolation and ac-bootstrapping purposes, the gate of transistor $M_1$ in the Π attenuators is connected to the controlling voltage via a 20kΩ resistor; while the bulk is connected to GND via a 20kΩ resistor. Performance was measured by on-wafer probing. The two-tone spacing is 3.2MHz for all measurements.

Fig. 3 shows the measured and simulated S11/S21 for all attenuation settings. Due to a mistake in the decoder design, the minimum attenuation setting (-1.8dB) cannot be enabled. For this setting we only show the simulated S11/S21. Due to unaccounted parasitics, the measured S21 for $f_{tr}>5$GHz deviates >1.6dB from simulation. Since the proposed IM3 cancellation does not require large transistors, 5GHz bandwidth is achieved with S11<-14dB and with S21 variation.
<1.6 dB. The difference between the measured NF and measured insertion loss (|S21|) is within 0.1dB up to 10GHz, which shows little noise is introduced by the attenuator.

Fig. 4 shows the measured HD1 and IM3 output at 1GHz. The compression/expansion P_{1dB} is above 10dBm for all attenuation settings. The IIP3 are respectively 31dBm, 33dBm, 38dBm and 36dBm for attenuation settings -6dB, -12dB, -18dB and -24dB. Due to higher-order nonlinearities, the IM3 curves start to show 5^{th} order behavior for input power levels higher than -8dBm.

Fig. 5 shows the measured IIP3 by sweeping the input frequency f_{rf} from 0.1GHz to 10 GHz. We achieve +30dBm IIP3 for the TV band (DC-1.2GHz) and +26dBm for DC-5GHz. At higher f_{rf}, extra phase shifts caused by the parasitic capacitances degrades the IM3 cancellation. The measured IIP3 of ten dies for f_{rf}=1GHz is also shown in Fig. 5. The IIP3 variation is within 1dB, which shows good robustness of this IM3 cancellation technique.

The benchmarking results in Fig. 6 shows that the presented attenuator achieves very high linearity for a very low active area in fully standard bulk CMOS. In conclusion this IM3 cancellation provides a robust IIP3 improvement with minimum active area consumption.

**Acknowledgements:**

We thank NXP for chip fabrication, and G. van der Weide, M. C. M. Soer and H. de Vries for assistance.
References:


Figure captions:

Figure 1: Concept of the wideband IM3 cancellation for Π attenuator.

Figure 2: Measured and simulated IIP3 illustrating the IM3 cancellation technique for the Π attenuator.

Figure 3: S11/S21 for the four-step attenuator. Line for measurement results, symbol for simulation results.

Figure 4: Measured output HD1 and IM3 vs input power for two tone input signals centered at 1GHz with 3.2 MHz spacing.

Figure 5: Measured IIP3 as a function of the two tone center frequency, with 3.2 MHz frequency spacing and (inset) measured IIP3 of ten samples for $f_{\text{rf}}=1\text{GHz}$.

Figure 6: Comparison of state-of-art attenuators.

Figure 7: Micrograph of the chip fabricated in a standard 160nm bulk CMOS.
Figure 1: Concept of the wideband IM3 cancellation for Π attenuator.
Figure 2: Measured and simulated IIP3 illustrating the IM3 cancellation technique for the Π attenuator.
Figure 3: S11/S21 for the four-step attenuator. Line for measurement results, symbol for simulation results.
Figure 4: Measured output HD1 and IM3 vs input power for two tone input signals centered at 1GHz with 3.2 MHz spacing.
Figure 5: Measured IIP3 as a function of the two tone center frequency $f_{\text{rf}}$, with 3.2 MHz frequency spacing and measured IIP3 of ten samples for $f_{\text{rf}}=1\text{GHz}$. 
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>0.18um CMOS</td>
<td>65nm CMOS</td>
<td>0.13um CMOS</td>
<td>CMOS SOI</td>
<td>0.18um CMOS</td>
<td>0.16um CMOS</td>
</tr>
<tr>
<td>VDD [V]</td>
<td>1.8</td>
<td>1.2</td>
<td>1.2</td>
<td>5</td>
<td>N/A</td>
<td>1.8</td>
</tr>
<tr>
<td>Chip area[mm²]</td>
<td>0.28</td>
<td>0.05</td>
<td>0.7</td>
<td>N/A</td>
<td>0.5</td>
<td>0.0054</td>
</tr>
<tr>
<td>Bandwidth[GHz]</td>
<td>0.4-3.7</td>
<td>0.4-0.8</td>
<td>DC-2.5</td>
<td>0.05-4</td>
<td>DC-14</td>
<td>DC-5</td>
</tr>
<tr>
<td>IIP3[dBm]</td>
<td>+15</td>
<td>+23</td>
<td>+10</td>
<td>+47</td>
<td>29@10GHz</td>
<td>+30(0.1-1.2GHz) +26(0.1-5GHz)</td>
</tr>
<tr>
<td>P1dB[dBm]</td>
<td>+6</td>
<td>N/A</td>
<td>+2.5</td>
<td>30</td>
<td>15@10GHz</td>
<td>+10</td>
</tr>
<tr>
<td>Attenuation flatness[dB]</td>
<td>2.6</td>
<td>N/A</td>
<td>2.6</td>
<td>3</td>
<td>0.5</td>
<td>1.6</td>
</tr>
<tr>
<td>Max. attenuation[dB]</td>
<td>33</td>
<td>48</td>
<td>42</td>
<td>40</td>
<td>31.5</td>
<td>24</td>
</tr>
<tr>
<td>Minium attenuation[dB]</td>
<td>0.96-2.9</td>
<td>N/A</td>
<td>0.9-3.5</td>
<td>2.4-4</td>
<td>3.7-10</td>
<td>1.8-2.4 (sim)</td>
</tr>
<tr>
<td>Return loss[dB]</td>
<td>&gt;9</td>
<td>&gt;12</td>
<td>&gt;8.2</td>
<td>&gt;14</td>
<td>&gt;9</td>
<td>&gt;14</td>
</tr>
<tr>
<td>Control mode</td>
<td>Linear-in-dB</td>
<td>Discrete step</td>
<td>Linear-in-dB</td>
<td>Linear-in-dB</td>
<td>Discrete step</td>
<td>Discrete step</td>
</tr>
</tbody>
</table>

Figure 6: Comparison of state-of-art attenuators.
Figure 7: Micrograph of the chip fabricated in a standard 160nm bulk CMOS.