A $g_m/I_D$ Based Methodology for the Design of CMOS Analog Circuits and Its Application to the Synthesis of a Silicon-on-Insulator Micropower OTA

F. Silveira, D. Flandre, and P. G. A. Jespers

Abstract—A new design methodology based on a unified treatment of all the regions of operation of the MOS transistor is proposed. It is intended for the design of CMOS analog circuits and especially suited for low power circuits where the moderate inversion region often is used because it provides a good compromise between speed and power consumption. The synthesis procedure is based on the relation between the ratio of the transconductance over dc drain current $g_m/I_D$ and the normalized current $I_D/(W/L)$. The $g_m/I_D$ indeed is a universal characteristic of all the transistors belonging to the same process. It may be derived from experimental measurements and fitted with simple analytical models. The method was applied successfully to the design of a silicon-on-insulator (SOI) micropower operational transconductance amplifier (OTA).

I. INTRODUCTION

The present trend toward portable equipment as well as the increasing circuit density and size of integrated systems tend to make low power consumption a primary concern [1], [2]. In CMOS analog circuits the minimum power consumption is achieved when MOS transistors operate in the weak inversion region [3]. However, the best compromise in terms of consumption and speed is achieved in moderate inversion.

On the contrary, complex synthesis tools based on SPICE-like models and simulators coupled with optimization routines, analytical or quick hand methods for synthesizing analog circuits give an insight on the design problem. However, these procedures generally suppose that the MOS transistors are either in strong or weak inversion. Mainstream methods assume generally strong inversion and use the transistor gate voltage overdrive (GVO) as the key parameter, where GVO = $V_G - V_T$, $V_G$ being the gate voltage and $V_T$ the threshold voltage [4]–[6]. Micropower design techniques, on the other hand, exploit known weak inversion models [3], [7].

This paper presents a new design methodology that allows a unified synthesis methodology in all regions of operation of the MOS transistor. It provides an alternative taking full advantage of the moderate inversion region to obtain a reasonable speed-power compromise. The method exploits the transconductance over dc drain current ratio ($g_m/I_D$) relationship versus the normalized current $[I_D/(W/L)]$.

An example of the proposed methodology is presented in the paper considering an experimental silicon-on-insulator (SOI) micropower CMOS OTA. SOI technology was chosen because it offers significant assets for low-power circuit design as discussed in [8]–[10]. This is further exemplified from the perspective of the synthesis methodology as well as from the results gathered from the experimental SOI OTA.

The paper is organized as follows. In Section II, we introduce the proposed design methodology and its main features. The application to the design of an SOI micropower CMOS OTA and validation through experimental results are presented in Sections III and IV. Section V demonstrates the usefulness of the proposed technique for the quick comparison of the performances of various CMOS technologies. The SOI CMOS results presented in Section IV are compared to those achieved with bulk CMOS and the advantages of SOI technology over bulk CMOS for low-power analog circuits are further underlined.

II. THE $g_m/I_D$ METHOD

In the proposed method, we consider the relationship between the ratio of the transconductance $g_m$ over dc drain current $I_D$ and the normalized drain current $I_D/(W/L)$ as a fundamental design tool. The choice of $g_m/I_D$ is based on its relevance for the three following reasons.

1) It is strongly related to the performances of analog circuits.

2) It gives an indication of the device operating region.

3) It provides a tool for calculating the transistors dimensions.

In order to illustrate this, let us consider the so-called “intrinsic gain stage” as a simple example. The intrinsic gain stage consists of a single transistor in the common source configuration loaded by an ideal current source (delivering a dc current $I_D$) and a capacitance $C_L$. We call $g_m$ the small-signal transconductance and $V_A$ the Early voltage which controls the transistor small-signal output conductance, $g_A = I_D/V_A$.

The dc gain ($A_0$) and transition unity-gain frequency ($f_T$) are given by

$$A_0 = -\frac{g_m}{I_D} V_A \quad (1)$$

and

$$f_T = \frac{1}{2\pi C_L} g_m \quad (2)$$

The $g_m/I_D$ ratio is a measure of the efficiency to translate current (hence power) into transconductance; i.e., the greater the $g_m/I_D$ value, the greater the transconductance we obtain at a constant current value. Therefore, the $g_m/I_D$ ratio is
 Sometimes interpreted as a measure of the "transconductance generation efficiency" [6].

The relation of the \( g_m/I_D \) ratio with the transistor operating mode can be observed from the fact that this ratio is equal to the derivative of the logarithmic of \( I_D \) with respect to \( V_G \) as shown below.

\[
\frac{g_m}{I_D} = \frac{1}{I_D} \frac{\partial I_D}{\partial V_G} = \frac{\partial (\ln I_D)}{\partial V_G} = \frac{\partial \left[ \frac{I_D}{W/L} \right]}{\partial V_G}. \tag{3}
\]

This derivative is maximum in the weak inversion region where the \( I_D \) dependence versus \( V_G \) is exponential while it is quadratic in strong inversion, becoming almost linear deeply in strong inversion because of the velocity saturation. The maximum is equal to \( 1/(nU_T) \) where \( n \) is the subthreshold slope factor and \( U_T \) the thermal voltage. The \( g_m/I_D \) ratio decreases as the operating point moves toward strong inversion when \( I_D \) or \( V_G \) are increased as shown in Fig. 1. Therefore, the \( g_m/I_D \) ratio is also an indicator of the mode of operation of the transistor.

Let us now consider the dependence of \( g_m/I_D \) on transistor size. The normalized current \( I_C \) is independent of the transistors size. According to (3) the \( g_m/I_D \) ratio is also size independent. Therefore, the relationship between \( g_m/I_D \) and the normalized current is a unique characteristic for all transistors of the same type (nMOS or pMOS) in a given batch. Of course, this statement must be revised when dealing with short channel transistors.

The "universal" quality of the \( g_m/I_D \) versus \( I_C \) curve can be extensively exploited during the design phase, when the transistor aspect ratios (\( W/L \)) are unknown. Once a pair of values among \( g_m/I_D, g_m, \) and \( I_D \) has been derived, the \( W/L \) of the transistor can be determined unambiguously. An example of this procedure is described in the next section.

**III. APPLICATION TO THE SYNTHESIS OF A SILICON-ON-INSULATOR MICROPPOWER CMOS OTA**

The proposed methodology was applied to the synthesis of the cascoded OTA [12] shown in Fig. 2. We assume that the total supply current (\( I_{SUM} \)) is known \textit{a priori} to 2 \( \mu \text{A} \). Furthermore, we assume that the load capacitor \( (C_L) \) is equal to 10 \( \text{pF} \) and the supply voltage \( (V_{DD}) \) equal to 3 \( \text{V} \). The design procedure is illustrated here aiming at the best performances in terms of: dc open loop gain \( (A_0) \), transition frequency \( (f_T) \), phase margin \( (PM) \), and slew rate \( (SR) \). However, it can straightforwardly be modified to take into account other performance aspects (like noise or common mode rejection) that may be relevant to the particular application, as long as they are directly related to the unified \( g_m/I_D \) versus \( I_C \) relationship, i.e., to the current and small-signal parameters. For large-signal related performance aspects such as signal swing, an \( I_D-V_G \) or \( g_m/I_D-V_G \) relationship.
is required. This extension of the methodology was applied to predict the output swing.

The methodology is graphically illustrated in Fig. 3, where all the data to be provided by the designer are shown in bold italics. To evaluate $A_0$, $SR$, and $f_T$ we use the following expressions:

- for the gain $A_0$:
  \[
  A_0 = \left( \frac{g_m}{I_D} \right)_1 \cdot \left( \frac{g_m}{I_D} \right)_2 \cdot \frac{1}{V_{A6} + V_{A7} + V_{A9} + V_{A10}} \quad (4)
  \]
  where $(g_m/I_D)_1$, $(g_m/I_D)_2$ are the $g_m/I_D$ ratio of the input and cascode transistors, respectively, $V_{A6}, V_{A7}, V_{A9}$, and $V_{A10}$ are the Early voltages of $T_6, T_7, T_9$, and $T_{10}$. The Early voltages will be considered proportional to the transistor length with a typical value of the constant of proportionality of 7 V/\mu m. Our experimental results have confirmed this approximation to be satisfactory for $L$ values in our range of interest from 3–12 \mu m.

- for $SR$ and the first order approximation of $f_T$:
  \[
  SR = \frac{B \cdot I_{D1}}{C_L} \\
  f_T = \frac{B \cdot g_m1}{2 \Pi \cdot C_L} \quad (5)
  \]
  where $B$ is the current mirror gain shown in Fig. 2. $I_{D1}$ is the current through the input pair transistors, and $g_m1$ is the input transistors transconductance.

The sequence followed for the synthesis is as follows.

1) The drain current of each transistor is determined from the specified total current $I_{tot}$ and the choice of the

current mirror factor $B$. The $B$ factor multiplies $f_T$ and $SR$, as shown in (5), and its maximum value is limited by its influence on the OTA stability. The designed OTA has $B$ equal to two.

2) Choosing the values for $(g_m/I_D)_1$ and $(g_m/I_D)_2$ is determined for each transistor from the experimental $(g_m/I_D)_1$ versus $(g_m/I_D)_2$ curves. Then, with the drain current value found in point 1, the $W/L$ of each transistor is found. The intended values of $(g_m/I_D)$ are chosen accordingly to their effect on the OTA performance. The dc gain is proportional to the product of $(g_m/I_D)_1$ and $(g_m/I_D)_2$ as shown in (4); $f_T$ is proportional to $g_m1$ as shown in (5), therefore, as the current is given, it is proportional to $(g_m/I_D)_1$. Hence $(g_m/I_D)_1$ and $(g_m/I_D)_2$ should be taken as high as possible. The maximum value we may choose is limited on one hand by the weak inversion maximum value of the technology (about 35 V\(^{-1}\)) in thin-film fully-depleted SOI MOS transistors and 25 V\(^{-1}\) in bulk MOS transistors) and on the other hand by the stability requirements because as we increase $(g_m/I_D)$, with fixed current, the transistor sizes and parasitic capacitances are increased and the phase margin is reduced. The value of the couple $(g_m/I_D)_1, (g_m/I_D)_2$ applied in the implemented design was the result of an exploration of the design space to obtain the combination that optimizes the trade-off between the dc gain and transition frequency for a given phase margin. The chosen values are $(g_m/I_D)_1$ equal to 28 and $(g_m/I_D)_2$ equal to 30, which correspond to an operating point in the moderate inversion region close to weak inversion.

The current mirror transistors are operated in strong inversion, with $(g_m/I_D)_2$ equal to eight, to guarantee good matching and noise properties.

The transistor lengths are determined by a trade-off between area and stability on one side and dc gain (due to the dependence of the Early voltage on the transistor length) on the other side. The chosen values are shown in Table I. Knowing the transistor lengths, all the transistor dimensions are determined, and the OTA frequency response can be predicted either by a SPICE or a symbolic analysis. When
TABLE II
CALCULATED, SIMULATED, AND MEASURED RESULTS OF THE MICROPW POWER SOI OTA CIRCUIT

<table>
<thead>
<tr>
<th></th>
<th>Synthesis Prog.</th>
<th>HSPICE</th>
<th>Measurements</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>($g_m/I_D)_1 (1/V)</td>
<td>28</td>
<td>29.4</td>
<td>28.3</td>
<td></td>
</tr>
<tr>
<td>($g_m/I_D)_2 (1/V)</td>
<td>30</td>
<td>31.6</td>
<td>30.5</td>
<td></td>
</tr>
<tr>
<td>$A_0$ (dB)</td>
<td>103.9</td>
<td>105.5</td>
<td>103</td>
<td>$@ C_L = 10\text{pF}$</td>
</tr>
<tr>
<td>$f_T$ (kHz)</td>
<td>324</td>
<td>336</td>
<td></td>
<td>$@ C_L = 10\text{pF}$</td>
</tr>
<tr>
<td>PM (deg.)</td>
<td>72.5</td>
<td>72</td>
<td></td>
<td>$@ C_L = 10\text{pF}$</td>
</tr>
<tr>
<td>$f_T$ (kHz)</td>
<td>261 *</td>
<td>270 *</td>
<td>271 *</td>
<td>$@ C_L = 12.3\text{pF}$</td>
</tr>
<tr>
<td>PM (deg.)</td>
<td>63.8 *</td>
<td>63 *</td>
<td>60 *</td>
<td>$@ C_L = 12.3\text{pF}$</td>
</tr>
<tr>
<td>SR (V/µA)</td>
<td>0.11</td>
<td>0.09</td>
<td>0.10</td>
<td>$@ C_L = 12.3\text{pF}$</td>
</tr>
<tr>
<td>Output Swing (Vpp)</td>
<td>2.02</td>
<td>2.2</td>
<td>1.93</td>
<td>$@ V_{DD} = 3\text{V}$</td>
</tr>
</tbody>
</table>

* Test version of the OTA with bonding PADS at the sources of the cascode transistors.

TABLE III
COMPARISON OF RESULTS OF $g_m/I_D$ BASED SYNTHESIS WITH CONVENTIONAL STRONG INVERSION AND WEAK INVERSION SYNTHESIS

<table>
<thead>
<tr>
<th></th>
<th>$g_m/I_D$ method</th>
<th>SI synthesis.</th>
<th>SI real</th>
<th>WI synthesis</th>
<th>WI real</th>
</tr>
</thead>
<tbody>
<tr>
<td>($g_m/I_D)_1 (1/V)</td>
<td>28</td>
<td>28</td>
<td>18.7</td>
<td>35</td>
<td>30</td>
</tr>
<tr>
<td>($g_m/I_D)_2 (1/V)</td>
<td>30</td>
<td>30</td>
<td>19.7</td>
<td>35</td>
<td>30.5</td>
</tr>
<tr>
<td>$A_0$ (dB)</td>
<td>103.9</td>
<td>103.9</td>
<td>96.7</td>
<td>107.1</td>
<td>104.6</td>
</tr>
<tr>
<td>$f_T$ (kHz)</td>
<td>324</td>
<td>351</td>
<td>236</td>
<td>395</td>
<td>344</td>
</tr>
<tr>
<td>PM (deg.)</td>
<td>72</td>
<td>84</td>
<td>86</td>
<td>64</td>
<td>68</td>
</tr>
<tr>
<td>W/L input pair</td>
<td>57.6</td>
<td>7.7</td>
<td>7.7</td>
<td>120.9</td>
<td>120.9</td>
</tr>
<tr>
<td>W/L cascode n</td>
<td>69.3</td>
<td>6.47</td>
<td>6.47</td>
<td>93.5</td>
<td>93.5</td>
</tr>
<tr>
<td>W/L cascode p</td>
<td>243.3</td>
<td>17.3</td>
<td>17.3</td>
<td>241.7</td>
<td>241.7</td>
</tr>
<tr>
<td>$\Sigma W/L$ ($\mu m^2$)</td>
<td>4900</td>
<td>1359</td>
<td>1359</td>
<td>6185</td>
<td>6185</td>
</tr>
</tbody>
</table>

a SPICE analysis is applied, a model valid in all regions of operation (like the one presented in [7]) is selected. If the phase margin ($PM$) is not acceptable, the values of $B, (g_m/I_D)$, or transistor lengths must be modified.

IV. EXPERIMENTAL RESULTS

The resulting design was realized in the 3-µm CMOS on SOI process of the Microelectronics Laboratory of the Université Catholique de Louvain.

The dc gain, transition frequency, and phase margin of the circuit prototype were measured. The transconductance of the input and cascode transistors were also determined in order to obtain the $g_m/I_D$ values.

Table II shows the excellent agreement observed between the performance figures calculated by the synthesis program, calculated by SPICE, and measured in the circuit prototype. The SPICE simulation used the HSPICE level 34 model which implements the model described in [7], with a set of parameters optimized to fit the SOI MOSFET characteristics.

In order to further demonstrate the interest and potential of the new methodology, we compare its results with the results obtained from the conventional design methods limited to the strong and weak inversion regions. These are shown in the columns “SI synthesis” and “WI synthesis” of Table III.

The strong inversion synthesis was based on the quadratic expression for the drain current in saturation as a function of the gate voltage. Extended to moderate inversion, this approximation is known to overestimate the $g_m/I_D$ related to a certain normalized current when compared to the real data. So that reciprocally, specified $g_m/I_D$ and $I_D$ are achieved by underestimated transistor sizes ($W/L$) and amplifier area (estimated as the “total” $W \cdot L$) as shown in Table III.

The weak inversion synthesis considered the exponential approximation for the drain current versus the gate voltage. This approximation predicts a value of $g_m/I_D$ equal to $35 (n = 1.1)$ independent of the current. Therefore, the $g_m/I_D$ value does not determine an $I_{D\text{sat}}$ value in this simplified model. The $I_{D\text{sat}}$ value must be chosen to guarantee weak inversion operation. The criterion applied was to choose $I_{D\text{sat}}$ ten times smaller than the value corresponding to the limit between the weak and strong inversion approximations equal to $2n \mu C_{ox}U_{D\text{sat}}^2$, which is a classical criterion to guarantee weak inversion operation [6]. This procedure leads to an overestimation of the transistor sizes required to achieve the specified $g_m/I_D$ and $I_D$ values.

The OTA performances achieved by the SI and WI synthesis are then estimated injecting the transistor sizes they provided into the OTA model which uses the real $g_m/I_D$ versus $I_D$
TABLE IV

<table>
<thead>
<tr>
<th></th>
<th>Bulk</th>
<th>SOI1</th>
<th>SOI2</th>
<th>SOI3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(g_m/I_D)1 (U/V)</td>
<td>19</td>
<td>28</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td>(g_m/I_D)2 (U/V)</td>
<td>17</td>
<td>30</td>
<td>27</td>
<td>17</td>
</tr>
<tr>
<td>A0 (dB)</td>
<td>95.6</td>
<td>103.9</td>
<td>102.7</td>
<td>95.6</td>
</tr>
<tr>
<td>I_tot (μA)</td>
<td>2.9</td>
<td>2</td>
<td>2</td>
<td>2.9</td>
</tr>
<tr>
<td>PM (°)</td>
<td>72.9</td>
<td>72.0</td>
<td>79.0</td>
<td>84.0</td>
</tr>
<tr>
<td>ΣW.L (μm²)</td>
<td>2628</td>
<td>4900</td>
<td>2881</td>
<td>1808</td>
</tr>
</tbody>
</table>

Fig. 4. Current consumption ratio and dc open loop gain difference between the bulk and SOI OTA versus transition frequency with phase margin of 60 degrees and C_L = 10 pF.

data. These estimations are shown in the columns “SI real” and “WI real” of Table III. The comparison shows that the strong inversion approach overestimates the transition frequency by about 50% and the gain by about 7 dB as the reduced transistor sizes correspond to larger normalized currents and hence smaller real g_m/I_D. The differences are lower if we consider the weak inversion approach because in the design under consideration the input pair and cascode transistors operate in moderate inversion close to the weak inversion region. Nevertheless, the transition frequency is overestimated by 15% and the estimated die area of the resulting design (estimated as the sum of the transistor active areas) is 25% larger than the area of the circuit obtained by the g_m/I_D approach.

V. ADVANTAGES OF THE SILICON-ON-INSULATOR TECHNOLOGY FOR LOW POWER ANALOG CIRCUITS

In this section we first compare the designed SOI OTA performance with the achievable performance of a similar bulk implementation. Then the current consumption and gain as a function of the transition frequency are compared for the considered OTA architecture in SOI and bulk technologies.

This section will demonstrate the usefulness of the proposed design methodology for the quick comparison of the performances of SOI versus bulk CMOS processes. We only need to introduce the technological data differences between both technologies, i.e., mainly the n factor, the “bottom-plate” nMOS and pMOS junction capacitances per unit area C_jn and C_jp, and the nMOS and pMOS sidewall capacitances per unit length C_jwren and C_jwrep. In our SOI technology (res. Bulk) n = 1.1 (1.5), C_jn = 0.06 fF/μm² (0.18 fF/μm²), C_jp = 0.06 fF/μm² (0.4 fF/μm²), C_jwren = 0.05 fF/μm (0.4 fF/μm), C_jwrep = 0.05 fF/μm (0.5 fF/μm). All other technology data are supposed the same for both technologies: the gate oxide thickness (t oxide = 30 nm), the effective mobility [μn = 500 e−m²/(V·s), μp = 190 e−m²/(V·s)], the drain and source region extensions (X_n = X_p = 8 μm) which multiplied by the transistor width give an estimation of the drain and source areas, the source—gate and drain—gate overlap (ov_n = ov_p = 0.15 μm) and the Early voltages (V_AN = V_Dp = 20 V @ 3 μm length). The current mirror factor B is equal to two in both implementations.

Table IV compares the performance of one bulk and three SOI implementations of the OTA. The four designs were optimized to achieve the same transition frequency, considering the same transistor lengths to simplify the comparison. SOI1 is the OTA designed in Section III with the aim to maximize the gain. This design has therefore the maximum values of (g_m/I_D)1 and (g_m/I_D)2 compatible with the required phase margin. The SOI2 design sought to trade-off a slight decrease in gain by a decrease in the OTA die area. We defined, for the sake of comparison, the objective of having approximately the same die area as the bulk implementation. The decrease in area is obtained with a slight decrease in (g_m/I_D)1 and (g_m/I_D)2.
The SOI3 design was aimed to have the same consumption and gain as the bulk implementation [i.e., the same values of \( \frac{g_m}{I_D} \) and \( \frac{g_m}{I_D} \)], in order to evaluate the reduction in die area that can be achieved for the same performances.

Fig. 5 shows the \( \frac{g_m}{I_D} \) of the differential pair transistors as a function of the transition frequency for both technologies, which explain the results of Fig. 4 following our design method. Fig. 5 also shows the current consumption for both technologies.

The results show that in order to obtain the same transition frequency and stability figures as SOI1, the bulk implementation requires an increase of 45% in current consumption, whereas the gain is about 8 dB lower. The lower \( \frac{g_m}{I_D} \) values achievable in bulk have indeed be compensated by an increase of \( I_D \) to maintain the same \( g_m \). The SOI1 OTA occupies, however, a larger die area than the bulk implementation due to the fact that the SOI lower parasitic capacitances allow to work closer to the weak inversion region (i.e., with bigger transistors for a given current) with acceptable phase margin. Nevertheless, the SOI2 OTA which has about the same area as the bulk OTA, still provides a much higher gain with a reduced power dissipation and an increased phase margin than the bulk OTA. The design SOI3 shows that if the same performances as in the bulk implementation are desired, the estimated die area is reduced by about 40%, without taking into account the area savings in SOI technology due to the absence of wells.

Fig. 4 shows the current consumption ratio and dc open loop gain difference between the bulk and SOI implementations of the OTA as a function of the specified transition frequency for the same phase margin of 60 degrees. As higher transition frequencies were sought, all the transistors were taken of minimum length (3 \( \mu m \)) in both SOI and bulk implementations, and the \( \frac{g_m}{I_D} \) of the current mirrors was taken equal to five.

The above examples show that the increased \( \frac{g_m}{I_D} \) values provided by the SOI technology, as well as the lower parasitic capacitances which allow to use higher values of \( \frac{g_m}{I_D} \) with acceptable phase margin lead to a reduction in current consumption of up to about 72% with a simultaneous increase in gain of up to 22 dB. On the other hand, the same performances as in bulk technology can be achieved with a much smaller circuit.

VI. Conclusion

The presented design methodology exploits the fact that the relationship between the ratio of the transconductance over dc drain current \( \frac{g_m}{I_D} \) and the normalized current \( \frac{I_D}{(W/L)} \) is a universal characteristic of all the transistors of the same type (nMOS or pMOS) belonging to the same process. This approach allows accurate sizing of analog circuits with a unified treatment from the strong to the weak inversion region. Particularly for low-power analog circuits, it is possible to take advantage of the moderate inversion region to optimize the power-speed compromise.

The proposed methodology was validated by the design of an SOI micropower OTA. The experimental results showed excellent correspondence with the predicted design data. When compared with the results achievable in bulk CMOS technologies, it appears that the increased \( \frac{g_m}{I_D} \) values and lower parasitic capacitances of fully depleted SOI devices allow a reduction of 72% of consumption with increased gain for the same bandwidth.

ACKNOWLEDGMENT

The authors wish to thank J.-P. Eggermont and B. Gentinne for their contributions in device characterization.

REFERENCES