On the design of very small transconductance OTAs with reduced input offset

Alfredo Arnaud  
DIE Universidad Católica  
8 de Octubre 2738  
Montevideo - Uruguay  
+598-2 4872717- ext.407  
email: aarnaud@ucu.edu.uy

Rafaela Fiorelli  
GME-IIE Univ.de la República  
Julio Herrera y Reissig 565/IIE  
Montevideo - Uruguay  
+598-2 7110974 - ext.123  
email: fiorelli@ling.edu.uy

Carlos Galup-Montoro  
LCI-EEL Univ.Fed. de Santa Catarina  
88040-900 Florianópolis - SC  
Brazil  
+55 48 331 9280  
email: carlos@eel.ufsc.br

ABSTRACT
In this paper it will be demonstrated, from the theory and measurements, that series-parallel (SP) mirrors allow building current copiers with copy factors of thousands, without degrading mismatch or noise performance. SP current-division will be then employed to design OTAs ranging from a few pS to a few nS, with up to 1V linear range, consuming in the order of 100nW, and with a reduced area. An integrated 3.3s time-constant integrator with up to 1V linear range, consuming in the order of 100nW, and with a reduced area. An integrated 3.3s time-constant integrator will also be presented. One-by-one several design non-idealities will be revised: linearity, offset, noise, leakages; as well as layout techniques. A final analysis concludes that SP-association of transistors allows to build very efficient transconductors, for demanding applications in the field of implantable electronics among others.

Categories and Subject Descriptors
B.7.1 [Integrated Circuits]: Types and Design Style – Input/output circuits, VLSI.

General Terms
Design.

Keywords
Analogue design, CMOS, low-power.

1. INTRODUCTION
In recent years there has been considerable research effort in the development of integrated transconductance amplifiers (OTAs), with very small transconductance and improved linear range due mainly to their application in biomedical circuits [1-6]. In a classical symmetrical OTA, voltage to current conversion is carried out in the input differential pair while the other transistors just copy the current to the output. Although bias current can be extremely low [7], thus obtaining an extremely low transconductance, a drawback in this case is the poor linearity as the input transistors operate in weak inversion. Several circuits were reported in the past to overcome the problem but the use of complex OTA architectures also increases noise, mismatch offset, and transistor area, and results in design trade-offs [1]. A rough classification of very low transconductance OTAs may include:
- Those circuits that modify the differential pair performing the voltage-to-current conversion in such a way to reduce the transconductance and increase its linear range (i.e.[4,8]).
- Those circuits that use current cancellation [5], or current division [9,10] to divide the OTA transconductance by a desired factor; but do not increase the input linear range.
- Circuits using voltage division techniques for the extension of the linear range [2,6].

Of course, techniques can be combined. Regarding their limitations: modified-differential pairs increase offset, noise, and are limited to a few nS in their transconductance; a further transconductance reduction requires the use of some kind of division scheme. Current cancellation and voltage division also show in general a large input offset and noise. Meanwhile, simple division of the output current of a differential pair by a high ratio has been widely considered an expensive technique in terms of area [2]. However, the use of series–parallel division of current in an OTA as in Fig.1 [9,10], allows the implementation of an area efficient current divider. In this paper we will examine in detail this circuit and SP current dividers in general. For the NMOS current mirrors in Fig.1, N unitary transistors $M_N$ are placed in-series and in-parallel to achieve an effective output transconductance $G_m$:

$$G_m = \frac{g_{m1}}{N^2}$$

(1)

$g_{m1}$ is the transconductance of the transistors $M_1$. To enhance linearity, a modified differential input pair can substitute $M_1$ in Fig.1, as shown in Fig.2 [8]. The effective transconductance $g_{m\_eff}$ of the pair in Fig.2 is calculated by small signal analysis, assuming that transistors $M_2$ operate in the linear region, each behaving as a resistor of value $2R$.

$$g_{m\_eff} = \frac{g_{m1}}{1 + n g_{m1}R}$$

(2)

$n$ is the slope factor [11], slightly greater than unity and weakly dependent on the gate voltage. The non-unitary current copiers of Fig.1 are not only area-efficient (their area is proportional to the square root of the copy factor), but are also mismatch-efficient, because they benefit from the improved matching of a large number of equal unitary transistors. With the appropriate placement, the designer can apply the most usual matching rules: common centroid geometry, and same surroundings; while using copy factors as large as thousands if for example $N=50$ or $N=100$ are selected. To preserve mismatch benefits while using moderate copy factors, the generic current mirror in Fig.3 can be used.
It is widely recognized that the performance of most analog or even digital MOS circuits is limited by random mismatch between transistors. Matching can be modelled by the random variations in geometric, process, and/or device parameters. The approach most employed by designers is to consider only variations in the threshold voltage and current factor. The composed transistors of a composite transistor will be inversely proportional to the equivalent channel length [13]. The copy factor $M$ in a current mirror is calculated as the ratio between the aspect ratios of the mirror of Fig.3:

$$\frac{I_{in}}{I_{Out}} = \frac{S_P}{R_Q} = M$$

where $P$, $R$, $Q$, and $S$ are the number of unitary transistors in series or in parallel. Classic current mirrors with a copy factor $M>>1$ -as in Fig.5(a)- are very sensitive to mismatch offset because at the output there is a single transistor $M_{in}$, with a reduced area that according to (3) presents high variations in the threshold voltage and current factor. The composed transistors of Fig.5(b) and Fig.5(c) can be used to implement a current mirror with also a copy factor $M=N^2$, but using the same number of unitary transistors at both input and output branches of the mirror. A better matching and a reduction in random offset are expected if usual layout matching rules are followed, because a large number of unitary transistors have been matched together. In this way, common centroid layout geometry is possible, even when matching transistors with a very different aspect ratio.

### 2.1 Mismatch calculation in a SP mirror

Even with a careful layout, fluctuations $\Delta \beta_{\text{off}}$ $\Delta V_{T_{\text{off}}}$ of each unitary transistor in Fig.3 produce an output current error term $\Delta I_{Out}$; then $I_{Out} = \left(\frac{I_{in}}{M}\right) + \Delta I_{Out}$. The designer requires a formula to estimate the SD of $I_{Out}$. In Fig.7, a composed transistor $M_N$ (with a transconductance $g_{\text{m}N}$) formed by a large number of series-stacked unit transistors $M_{in}$ is shown. $\Delta V_{T_{in}}, \Delta \beta_{\text{in}}$ fluctuations on unitary transistors affect the drain current $I_{Ds}$ and node voltages of each $M_{in}$. For each one is possible to write:

$$\Delta I_{Di} = -g_{m_{in}} \Delta V_{Ti} + \frac{I_{Ds}}{\beta_{\text{in}}} \Delta \beta_{\text{in}} - g_{m_{di}} \Delta V_{Si} + g_{m_{di}} \Delta V_{S(i-1)}$$

where $g_{m_{i}}, g_{m_{di}}, g_{m_{di}}$ are gate, source, and drain transconductances of $M_{in}$, respectively. (6) has been derived for a generic transistor but $\Delta I_{Di} = \Delta I_D$ is constant, because the transistors are series connected; also their channel charge densities at source and drain $Q_{Di} = Q_{Di}(i-1)$. Thus $g_{m_{di}(i-1)} = g_{m_{di}}$ [11]. Summing (6) for all the $P$ series transistors and can be utilized to obtain improved matching between devices. In Fig.4(a) two transistors $M_{SA}, M_{SB}$ are series connected; the equivalent ratio $\left(\frac{W}{L}\right)_{eq}$ of the composite transistor is [13]:

$$\left(\frac{W}{L}\right)_{eq} = \frac{(W/L)_S (W/L)_D}{(W/L)_S + (W/L)_D}$$

Eq.(4) may be extended to complex combinations of transistors to obtain different equivalent transistor geometries. For example, Fig.4(c) show measured drain current vs. drain voltage for two equivalent transistors: a single unitary sized $(W/L)_b = 4\mu m/10\mu m$, and a $10\times10$ array of the same transistor (Fig.4(b)). The two plots are similar, but note in the upper detail that the drain-source impedance $r_{ds}$ is much higher in the case of the composite transistor. As a rule of thumb, the output conductance of a composite transistor will be inversely proportional to the equivalent channel length [13].

2. SERIES-PARALLEL CURRENT DIVISION, MULTIPLICATION, AND MISMATCH.

Figure 1: PMOS-input symmetrical OTA with series-parallel current division to reduce transconductance without loss in linear range.

Figure 2: Active linearization of a differential pair to enhance linearity; $I_A-I_B = g_{m_{st}} (V_{in}^- - V_{in}^+)$.

Figure 3: Generic SP current mirror. Unitary transistors are identical $M_A = M_B = M_e$.

In (3) $A_{\text{eff}}, A_{\beta}$ are two technology parameters with typical values of $A_{\text{eff}}=13-30\text{mV}\cdot\mu m$ and $A_{\beta}=2-4\%/\mu m$ [12]. Series-parallel association of MOS transistors [13] is a useful circuit technique
Figure 5: Three M:1 current copiers. (a) Classic, (b) \( N = \sqrt{M} \) parallel transistors, copy to \( N \) series-stacked ones. (c) \( M \) parallel unitary transistors, copy to a \( N \times N \) array.

Figure 6: Calculated and measured \( I_{\text{out}} \), \( \sigma_{\text{out}}^2/I_{\text{out}}^2 \) in terms of the input current are shown for a 100:1 NMOS current mirror with the topologies of Fig.5(a) (\( M=100 \)) and Fig.5(c) (\( M=100, \ N=10 \)). \( \sigma_{\text{out}}^2 \) value was obtained from 10 samples of the circuit of the same batch.
\[ \sigma_{V_{in}}^2 = \frac{I_{in}^2}{g_mR} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{1}{R} \left( \sigma_{V_{r}}^2 \right) \]  

(9)

\[ V_{G} \text{ fluctuation is propagated to the output through } M_0. \text{ Also (8) is summed to calculate total } I_{out} \text{ SD in Fig.3} \]

\[ \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{1}{PQ} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{g_{mb}^2}{I_{out}} \left( \sigma_{V_{r}}^2 \right) \]  

(10)

\[ \left[ \frac{\sigma_{\beta}^2}{\beta^2} \right] , \left[ \frac{\sigma_{V_{r}}^2}{\beta^2} \right] , \text{ are the SD of the threshold voltage and the current factor for unitary transistors respectively.} \]

\section*{3. DESIGN OF SP VERY LOW TRANSCONDUCTORS}

Some other circuit properties should be studied. Linearity can be calculated in terms of the inversion level \( i_f \) [11] of the input differential pair (Fig.1). The expression is \( V_{Lin} = 3nq \left( \frac{\eta}{1+i_f} \right) \)

where \( V_{Lin} \) is the input linear range [10]. Linearity is further extended in the topology in Fig.2 [8].

\subsection*{3.1 Offset in the OTA}

The input pair \( M_1 \), PMOS current mirror transistors \( M_3 \), and SP current dividers, contribute to offset in the OTA of Fig.1. Summing their input referred contribution:

\[ \sigma_{V_{in}}^2 = 2 \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{I_{in}^2}{g_{ml} \cdot \Delta \Delta_{m} \cdot \Delta \Delta_{l}} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{4I_{in}}{N \cdot g_{ml} \cdot \Delta \Delta_{m} \cdot \Delta \Delta_{l}} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{2I_{in}^2}{g_{ml} \cdot \Delta \Delta_{m} \cdot \Delta \Delta_{l}} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) \]  

(11)

\[ \sigma_{V_{in}}^2 = \frac{2I_{in}^2}{g_{ml} \cdot \Delta \Delta_{m} \cdot \Delta \Delta_{l}} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{4I_{in}}{N \cdot g_{ml} \cdot \Delta \Delta_{m} \cdot \Delta \Delta_{l}} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) + \frac{2I_{in}^2}{g_{ml} \cdot \Delta \Delta_{m} \cdot \Delta \Delta_{l}} \left( \frac{\sigma_{\beta}^2}{\beta^2} \right) \]  

(12)

Expressions (11,12), can be extended to generic SP current division of Fig.3 changing \( N \) by \( \sqrt{PQ/R_S} \).

\subsection*{3.2 Noise Analysis}

SP mirror noise is calculated in Fig.3, with the aid of consistent thermal and flicker noise models [14]:

\[ S_{I_{out}} = 2nqB \left( \frac{g_{mb}}{I_{out}} \right) \left( \frac{1}{\mu R} + \frac{1}{\mu R} \right) \]  

(13)

\[ S_{I_{out}} = q^2 N_{out} \left( \frac{1}{\mu R} + \frac{1}{\mu R} \right) \]  

(14)

\[ v_{\text{input, th}}(f) \approx \frac{4nk_BT}{G_m} \eta \]  

(15)

\[ v_{\text{input, f}}(f) \approx \frac{2nk_BT}{N^3 C_{ox} \left( W/L \right)} \eta \left( \frac{2N_{out}}{N_{out}} + \frac{N_{out}}{N_{out}} \right) \frac{1}{f} \]  

(16)

The factor \( \eta = \left( \frac{1}{\sqrt{1+i_f+1}} \right) \) and \( \left( \frac{W/L}{(a+b+k)} \right) \) are the gate area of \( M_1, M_2, \) and \( M_3, \) respectively. Expressions (14) are very similar to that obtained for a simple symmetrical OTA, but a price is paid in noise for the linearization represented by the factor \( \eta \).

\subsection*{3.3 What is the minimum effective \( G_m \) possible with this technique?}

\( M_2, M_3 \) in Fig.1 may be biased with few pA or less; the only limitation being the sum of the leakage current at the source (drain) of each series transistor in \( M_{3(n+1)} \). In the target technology leakages in p-doped diffusions are much higher than those in an n-doped ones. For this reason, the presented OTAs include a PMOS differential pair, and NMOS SP current mirrors. From manufacturer’s data, estimated leakage current in a single source/drain was \( I_{out} \approx 30a \) for a \( 4 \mu m \times 2 \mu m \) n+ diffusion. The selected design criterion is that the leakages should be at least 10 times smaller than the bias current in the output branch.
\[ I_{\text{leak}} = N I_{\text{leak}} < I_{D1}/10N^2, \] which imposes a limit on the minimum achievable transconductance. A minimum 15pS OTA was estimated for \( N = 100 \), and \( (g_{m}/I_{D}) = 5 \) at the input pair, that is -although reasonable- an arbitrary worst case condition.

4. DESIGNED OTAS.
Several low, and very low transconductance OTAs named \( G_{\text{m1}} \) to \( G_{\text{m4}} \) were fabricated for test purpose. The design methodology is the one proposed in [10] while non idealities -noise, offset- are addressed using (11) to (14). \( G_{\text{m1}} \) is a 35pS OTA. It uses a \( g_{m1} = 170nS \) differential input pair in the configuration of Fig.1, with a \( N^2=4900 \) division factor. \( G_{\text{m2}} = 2.35nS \) OTA uses the differential pair of \( G_{\text{m1}} \) with a 9:1 division factor obtained with the topology of Fig.3 with \( P = 8, Q = 2, R = 1, S = 18 \). \( G_{\text{m3}}, G_{\text{m4}} \) are both 500mV linear range OTAs with a transconductance of 2.5nS, and 90pS respectively. The differential pair in Fig.2 with \( g_{m\text{-eff}} = 69nS \) is employed with 28:1, and 784:1 division factors.

In \( G_{\text{m1}} \), \( P, Q, R, S = 28, 1, 5, 5 \) and \( G_{\text{m4}} \) employs 28 parallel transistors copying to 28 series ones in the current divider. A careful layout plays a central role in obtaining a reduced mismatch. In all cases SP mirrors, as well as \( M_{1,3,4,5} \) are built with a large row of unitary transistors interconnected according to the desired topology. Unitary transistors of each composed one are alternated, and usual dummy structures at the row ends and same surroundings rules are respected. It should be pointed that SP OTAS allow a very efficient reutilization of layout blocks. In effect, widely different current mirrors and differential pairs are obtained just by changing SP connections at metal layer as indicated in Table I.

4.1 Measurement results.
Several measured characteristics of \( G_{\text{m1}} \) to \( G_{\text{m4}} \) are summarized in Table 2, while in Fig.8 the measured transfer functions of \( G_{\text{m3}} \) and \( G_{\text{m4}} \) are shown. It should be highlighted the reduced input offset of the OTAs, obtained with a moderate area and nano-power consumption. Ten circuit samples from the same batch were used of the OTAs, obtained with a moderate area and nano-power consumption. Ten circuit samples from the same batch were used.

A COMPARATIVE SURVEY
As pointed in the introduction, several very low transconductors and large time constant \( G_{m}\)-C filters have been reported. In [4], Sharpeskar et al. combine at the input gate degeneration, bulk driven transistors, and the so called bump transistor technique, to achieve a transconductor of 10nS with a linear range of \( \pm 1.7V \), around 20mV input offset and 1µW power consumption. In [1] several 10nS, sub-µW OTAs, using different input pair linearization techniques are compared; unfortunately no offset measurements are presented. These OTAs require the addition of some kind of division technique -like the one here presented- to achieve few Hz or sub-Hz range \( G_{m}\)-C filters. Techniques may include voltage attenuation [2,6], capacitor scaling [15], or

\[ G_{\text{m4}} \] was estimated for \( \alpha = \frac{D}{D_2} \).
current division/cancellation [3,10]. Table 3 puts together some previously reported large time constant integrators and filters. Although the comparative study is difficult because linearity, noise, and particularly offset, are not always measured in the same way, it is possible to conclude that the technique here presented is very efficient regarding mismatch, power consumption, and noise, without a significant overhead in silicon area.

6. CONCLUSIONS
A general expression was introduced, to estimate mismatch offset in series-parallel current division. Extremely large current multiplication (division) factors can be obtained by means of SP mirrors, without a significant loss in terms of area, offset, or noise. Series-parallel division of current in symmetrical OTAs mirrors, without a significant loss in terms of area, offset, or multiplication (division) factors can be obtained by means of SP linearity, noise, and input offset.

7. REFERENCES

Table 1: Several design characteristics of fabricated SP OTAs: input pair transconductance,division factor, unitary transistor size at the divider, area, and power consumption.

<table>
<thead>
<tr>
<th>OTA</th>
<th>Gm_pair</th>
<th>M (P-Q-R-S)</th>
<th>W/La</th>
<th>Area</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gm1</td>
<td>174</td>
<td>72</td>
<td>2.8</td>
<td>0.04</td>
<td>116</td>
</tr>
<tr>
<td>Gm2</td>
<td>174</td>
<td>70</td>
<td>1.7-2.0</td>
<td>0.09</td>
<td>113</td>
</tr>
<tr>
<td>Gm3</td>
<td>69</td>
<td>28</td>
<td>5-7</td>
<td>0.15</td>
<td>118</td>
</tr>
<tr>
<td>Gm4</td>
<td>69</td>
<td>784</td>
<td>28-1-28</td>
<td>0.15</td>
<td>113</td>
</tr>
</tbody>
</table>

Table 2: Several characteristics of OTAs: transconductance predicted-measured, measured linearity, predicted-measured input offset SD, maximum measured offset (10 samples), predicted–measured input noise in the band from .3 – 10Hz.

<table>
<thead>
<tr>
<th>OTA</th>
<th>Transc.</th>
<th>Vlin[mV]</th>
<th>Off.[mV]</th>
<th>Off_max [mV]</th>
<th>Noise [mVrms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gm1</td>
<td>2.4-2.6 nS</td>
<td>±160</td>
<td>8.0 – 4.4</td>
<td>8.3</td>
<td>48 – 89</td>
</tr>
<tr>
<td>Gm2</td>
<td>35-33pS</td>
<td>±160</td>
<td>5.4 – 2.1</td>
<td>4.0</td>
<td>193 – 160</td>
</tr>
<tr>
<td>Gm3</td>
<td>2.4-2.8 nS</td>
<td>±550</td>
<td>8.8 – 9.1</td>
<td>21</td>
<td>56 – 108</td>
</tr>
<tr>
<td>Gm4</td>
<td>89–100pS</td>
<td>±500</td>
<td>9.0 – 6.8</td>
<td>12</td>
<td>190 – na</td>
</tr>
</tbody>
</table>

Table 3: A comparative table of several reported division schemes used in OTAs, and large time constant Gm=C filters.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Ref.</th>
<th>Division Factor &amp; Time const.</th>
<th>Offset</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP Current Division</td>
<td>This work</td>
<td>70 to 4900 3.3s</td>
<td>2-9mV (SD)</td>
<td>Low noise, good linearity, nW power.</td>
</tr>
<tr>
<td>Voltage Division</td>
<td>[2]</td>
<td>Up to 10000 10s</td>
<td>Very large.</td>
<td>Poor linearity, small silicon area, nW power.</td>
</tr>
<tr>
<td>SP Current Division</td>
<td>[3]</td>
<td>2200 0.7s</td>
<td>130mV</td>
<td>Large offset probably due to non-symmetrical SP copy.</td>
</tr>
<tr>
<td>Current Div./Canc.</td>
<td>[5]</td>
<td>1000 100ms</td>
<td>40mV</td>
<td>Uses 10pF capacitors, random offset + 80mV systematic.</td>
</tr>
</tbody>
</table>