Design and Optimization of CMOS OTA with gm/Id Methodology using EKV model for RF Frequency Synthesizer Application.

Amine AYED,Hamadi GHARIANI,Mounir SAMET
Laboratoire d'Electronique et Technologies de l'Information
Ecole Nationale d'Ingénieurs de Sfax
Amine.Ayed@isecs.rnu.tn

Abstract_ **In this paper, we will focus exclusively in the obtention of optimum power designs, using the (*gm/ID*) methodology [1]. The introduction of a simple, gain bandwidth driven, automatic design algorithm for OTA is used as a starting point for the review of more advanced design methodology. This review leads to an automatic synthesis algorithm developed in MATLAB which systematically transits from high level specifications (total settling time) to the amplifier specifications (gain-bandwidth, slew rate) and then to transistor sizing. The design obtained complies with the high level specifications with minimum power consumption.**
Two main lines of study are followed here. The study of architectures for input and output stages that are suitable to be used on different environmental conditions, allow us to obtain an opamp cell that can be used in an ample spectrum of low-voltage, micropower applications.
The second line of study in analog design reuse focuses on the possibility of circuit performance tuning through the bias current, where preliminary results have already been obtained . The idea in this technique is to tune the power-speed trade off of the opamp cell using the bias current while keeping the performance in all other aspects.

I. INTRODUCTION

The architecture of OTA will be included in replica bias circuit for RF frequency synthesizer and particularly in ring VCO. A replica biasing circuit is used to adjust the gate bias of the PMOS load devices for a fixed swing at the output. The introduction of a simple, gain-bandwidth driven, automatic design algorithm for OTA is used as a starting point for the review of more advanced design methodology.
We begin with the review of the MOSFET model used in this work. The EKV model presents simple, single piece, continuous expressions and has many advantages regarding analog design. Specially, the fact that every equation is a function of the inversion level and a few physical-based parameters makes this model ideal to be used in automatic synthesis algorithms.

II. EKV FORMULATION OF MAJOR ANALOG MOS PARAMETERS IN BULK

The discussion of the adequacy of existing MOSFET models for reliable analog design is a relatively recent issue. The best known SPICE models are strong inversion-based models originally developed for digital application. They thus suffer from the troubles common to this kind of models concerning the discontinuities of the small-signal parameters, i.e. the current and charge derivatives, between the different regions of operation [2]. On the contrary, efficient analog design, and especially for low-voltage low-power (LVLP) applications of great interest, requires a modelling valid from weak to strong inversion and non-saturation to saturation conditions with smooth continuous transitions.
The charge-sheet model presented in [3] for bulk MOSFETs may fulfil such conditions. However it relies on the resolution of a highly non-linear implicit equation and is not computionaly efficient, nor widely available in SPICE. The charge-sheet formulation has not been extended towards a complete MOSFET standard SPICE model, including all the important analog device properties, such as noise, capacitance…, as well as an efficient and systematic parameter extraction procedure.
A much more convenient analog SPICE model known as the EKV model has been developed at the Ecole Polytechnique Fédérale de Lausanne (Suisse) [4]. Enz, Krummenacher and Vittoz indeed extended the initial Oguey-Czerveny's idea of a single empirical drain current equation continuous from weak to strong inversion, towards the derivation of a complete MOSFET model solving the problems stated above for analog simulations. In the present paper, we will first link the physical charge-sheet and empirical EKV MOS models.
In the case of analog circuit design, operational amplifiers (opamps) in particular, we are mostly concerned with the characteristics of MOS transistors in saturation and principally, the ratio of transconductance over drain current, the output conductance and the intrinsic gate capacitances.

III. THE (*gm/ID*) BASED METHODOLOGY FOR ANALOG DESIGN

The (*gm/ID*) based methodology allows an unified synthesis methodology in all regions of operation of the MOS transistor. It provides an alternative, taking advantage of the moderate inversion region, to obtain reasonable power-speed compromise [1]. This methodology has been widely used since its publication proving its advantages in analog circuits design [5, 6].

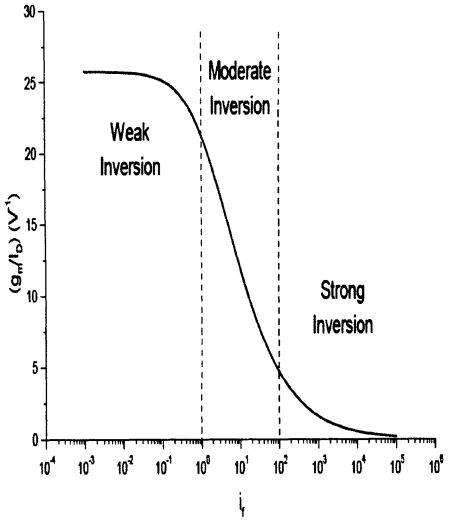

Fig.1. (*gm/ID*) ratio as a function of the inversion factor *if* for typical bulk-technology parameters.

The proposed methodology considers the relationship between the transconductance over drain current ratio (*gm/ID*) and the normalized drain current (*ID/(W/L)*) as a fundamental design tool. This choice of the (*gm/ID*) ratio is based in the following reasons

1. It gives an indication of the device operation region.
2. It is strongly related to the performance of analog circuits.
3. It provides a tool for calculating the transistor dimensions.

Finally, the ability to precisely obtain transistors dimension with this methodology lays in the fact that the (*gm/ID*) vs (*ID(W/L)*) characteristic is independent of transistor size, and therefore is a unique characteristic for all transistors of the same type in a given batch [1]. This "universal" (*gm/ID*) presented in equation (1) [4] shows that once a pair of values among (*gm/ID*), *gm* and *ID* are chosen, (*W/L*) ratio is unambiguously determined [1].

$$\frac{g_m}{I_D} = \frac{1}{nU_T} \cdot \frac{\left(1 - e^{\left(-\sqrt{IC}\right)}\right)}{\sqrt{IC}} \qquad (1)$$

with IC also called the inversion coefficient [4],

$$IC = \frac{I_D}{2n\mu C_{ox} \frac{W}{L}} U_T^{\,2}$$

and noting ID/(W/L) = I' the "adimensional" current normalized.

IV. SINGLE-STAGE OPERATIONAL TRANSCONDUCTANCE AMPLIFIER DESIGN

A well-know single-stage opamp(fig.2) is realized with a differential pair(M1-M2) as input stage and three current mirrors(M3-M8).

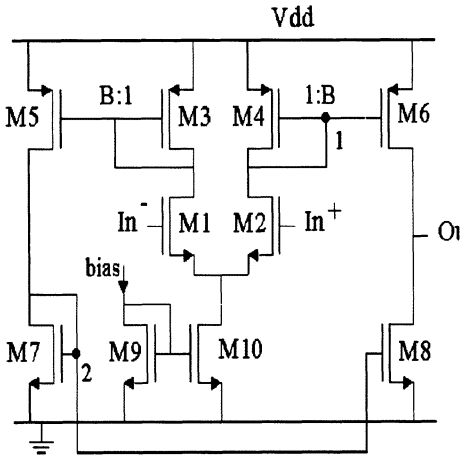

Fig.2. single-stage OTA architecture

The differential current issued from M1-M2 due to differential input voltage is mirrored towards the high impedance output node. Considering a symmetrical opamp, the DC open-loop gain and the gain-bandwith product for a mirror M4-M6 (M3M5) size ratio B are given by [7]

$$A_{v0,OTA} = \frac{Bg_{m1}}{(g_{d6} + g_{d8})} = Vea\left[\frac{g_m}{I_d}\right]_1 \qquad (2)$$

$$GBW_{OTA} = B\left[\frac{g_m}{I_d}\right]\frac{I_{d1}}{2\pi C_L} \qquad (3)$$

Where

$$B = \frac{(W/L)_6}{(W/L)_4} = \frac{(W/L)_5}{(W/L)_3} \qquad (4)$$

V_{ea} is the equivalent Early voltage, i.e., $V_{ea} = V_{ea6}\, V_{ea8}/(V_{ea6} + V_{ea8})$ due to parallel combination of M6 and M8 output conductances.

V_{ea} is also function of device length of M6 and M8 and of its drain voltage generally equal to the half power supply voltage. However, as a result of parasitic poles and zeros appearing at nodes 1 and 2, the opamp is not a first order system. But if the parasitic poles and zeros are sufficiently beyond the GBW as required to guarantee closed-loop stability, then the transition frequency is almost equal to the GBW.

A. Differential Input Pair Transistors M1-M2 Design

The design of the differential input pair is the same as for the amplifier basic block. The transition frequency for given C_L is specified. The device length and the equivalent Early voltage of the output node are imposed by the technology and the power supply. The choice of the g_m/I_d for the differential input pair M1-M2 determines the DC open-loop gain A_{vo} by (2), the drain current I_{d1} necessary to reach the transition frequency following equation (3) for a given ratio B and finally $(W/L)_1$. The problem then lies in the choices of $(g_m/I_d)_1$ and of ratio B.

Other architectures have a similar equation (3) for GBW and thus a similar design of the input pair.

B. Current Mirrors Design

The currents in the opamp branches are determined by the design of the differential input pair presented above and the ratio B of the p-type current mirror. The n and p mirrors are to be designed to repel the parasitic poles and zeros sufficiently beyond the transition frequency.

1) p-Type Current Mirror Design: The second pole of the opamp is determined by the total capacitance C_1 at the internal node 1 and the transconductance g_{m4} of M4 as follows:

$$\omega_{p1} = \frac{g_{m4}}{C_1} \qquad (5)$$

For stability and setting time considerations this second pole must be higher than f_T by a factor β_1

[7] $\omega_{p1} \geq \beta_1 (2\pi f_T)$ (6)

The capacitance C_1 is the sum of intrinsic, overlap, and drain-substrate capacitances of M2, M4, and M6 connected to node 1 (Fig.2). Considering that the only significant intrinsic capacitance in saturation is the capacitance between gate and source

C1=Cgs4+Cgs6+Cgso4+Cgso6+
+Cbd4+Cbd2+Cgdo2 (7)

Where
Cgs Gate-source intrinsic capacitance
Cgs0 Gate-source overlap capacitance
Cgd0 Gate-drain overlap capacitance
Cbd Drain-substrate capacitance

Expression (7) is valid provided the same device length is assumed for M4 and M6. The different capacitances can be estimated as follows:

Cgs4=2/3 C_{ox} W4 L4 Cgs04= C_{ox} ΔLW4
Cgd06=B C_{ox} W4 L4 Cgd02= C_{ox}ΔLW1
Cbd4= C_{ox} ξ W4 Cbd2= C_{ox} ξ W1

Where
C_{ox} Gate capacitance per unit area.
ΔL Length of the gate to drain or source overlap.
ξ Coefficient of the diffusion-substrate capacitance and is equal to t_{oxf} / t_{oxb} * L_d with
t_{oxf} gate oxide thickness,
t_{oxb} back gate oxide thickness, and
L_d minimal length of the drain diffusion area.

The choice of a $(g_m/I_d)_4$ value for M4 determines g_{m4} as the scaled drain current $I_{d4}(=I_{d1})$ is known. The $(W/L)_4$ is equal to the ratio between the drain current I_{d4} and the scaled drain current corresponding to this $(g_m/I_d)_4$ value.

Thus the position of the second pole can be directly computed by (5) and (6) since W4 is the only unknown. The $(g_m/I_d)_4$ value has to guarantee that the second pole is sufficiently beyond f_T , i.e., to verify condition (6). Moreover two other limits will restrict the possible range of $(g_m/I_d)_4$.

First, the device M4 must be operated in strong inversion for noise and mismatch considerations in current mirrors [7].

Secondly, $V_{dsat4-6}$ in inversely proportional to $(g_m/I_d)_4$ in strong inversion and consequently the output dynamic range decreases for lower $(g_m/I_d)_4$ [7].

The choice of g_m/I_d for the differential input pair and of the p mirror ratio (B) must allow to fin the value of $(g_m/I_d)_4$ satisfying conditions (6) and allowing a sufficiently wide output dynamic range .

For specified B, the compromise for $(g_m/I_d)_1$ is as follows.

A too low $(g_m/I_d)_1$ value provides a too high current consumption by (3), too low $(W/L)_1$ increasing offset and noise, and reduces the DC gain [7]. A too high $(g_m/I_d)_1$ value corresponds to too large $(W/L)_1$ and thus the part of capacitance C_1 relative to M1 may be too high to verify condition (6).

For specified $(g_m/I_d)_1$ the compromise for ratio B is as follows. A too low B value requires too high I_{d1} and $(W/L)_1$ from (6) and a too high value increases the C_1 capacitance. The typical value of this ratio is between 0.5 and 2[7].

It is worth nothing than for specified f_T the slew rate, i.e. SR=2 I_{d1}. B/C_L, only depends on the $(g_m/I_d)_1$ value in an inversely proportional way following (6) where I_{d1}. B is equal to the maximal

output current [7]. As on the other hand the DC gain is proportional to $(g_m/I_d)_1$, there is no universal value of $(g_m/I_d)_1$ value convenient to every application like micropower or high-speed circuits.

2) n-type Current Mirror Design: Due to the single-ended output the n-type current mirrors introduces a doublet consisting of a closely spaced pole and zero couple [7]. The pole p_2 is determined by the total capacitance C_2 at the internal node 2 and the transconductance g_{m7} of M7 and the zero z_1 is situated an octave beyond the pole

$$\omega_{p2} = \frac{g_{m7}}{C_2} \qquad \omega_{z1} = 2\omega_{p2} \qquad (8)$$

As for the second pole, the doublet must be sufficiently beyond the f_T by a factor β_2

$$\omega_{p2} \geq \beta_2 (2\pi f_T) \qquad (9)$$

The capacitance C_2 is the sun of intrinsic, overlap and drain-substrate capacitances of M5, M7 and M8 connected to node 2

$$C2=Cgs7+Cgs8+Cgso7+Cgso8+ +Cbd7+Cbd5+Cgdo5+Cgd08 \qquad (10)$$

The $(g_m/I_d)_7$ value of M7 determines the g_{m7}, the $(W/L)_7$ and thus the total capacitance C_2 and must verify the condition (9). This latter is easier to meet than (6). Indeed as the mobility of NMOS is higher than PMOS, for specified g_m/I_d and I_{d1}, thus same g_m, NMOS is narrower than a PMOS and consequently the NMOS capacitances are lower than PMOS ones and thus $\omega_{p2} \geq \omega_{p1}$

Reciprocally for $\omega_{p2} = \omega_{p1}$, $(g_m/I_d)_7$ will be higher than $(g_m/I_d)_4$ and V_{dsat8} will be lower than the V_{dsat6}.

B. Phase Margin

Supposing that the dominant pole produces a 90° phase loss at the transition frequency, the phase margin can be calculated by the following equation [7]:

$$\phi_M = 90° - \arctan\left(\frac{\omega_T}{\omega_{p1}}\right) - \arctan\left(\frac{\omega_T}{\omega_{p2}}\right) + \arctan\left(\frac{\omega_T}{\omega_{z1}}\right)$$

with $\omega_T = 2\pi f_T$

V. SYSTEMATIC DESIGN METHODOLOGY AND RESULTS

We now propose a design procedure based on the above results which follows the "gm/ID" design methodology for CMOS OTAs described previously. The approach will be illustrated by means of an example targeting a gain A_0 of 90 dB and a transition frequency f_T of 100 MHz, considering a load CL equal to 10 pF in a 0.35µm bulk CMOS technology. The sizing of M1 and M4 is based on the "gm/ID" methodology which proceeds along the lines described in figure 3 and is explained hereafter.

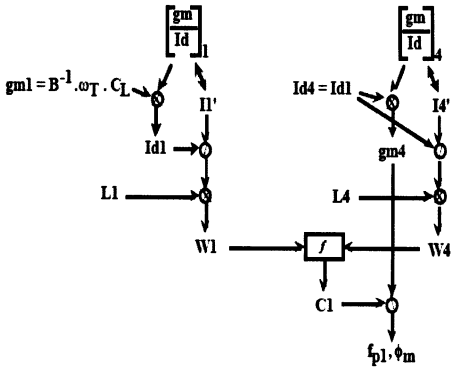

Fig.3. Design flow for the folded-cascode stage M1-M2

This approach is implanted with an automatic synthesis algorithm developed in MATLAB using symbolic equations such as lambertw function . A part of algorithm is shown hereafter.

```
clear all
%parameters declarations%
f1=0.5:0.5:40;
f4=0.5:0.5:40;
k1=1/(nn*UTn);
k4=1/(np*UTp);
for i=1:80
IC1(i)=lambertw(k1/((-k1+f1(i))^2));
   Ip1(i)=IC1(i)*(2*nn*Kn*(UTn^2));

IC4(i)=lambertw(k4/((-k4+f4(i))^2));
   Ip4(i)=IC4(i)*(2*np*Kp*(UTp^2));

ID1(i)=GBW*2*pi*CL/(B*f1(i));
ID4(i)=ID1(i);
gm1(i)=f1(i)*ID1(i);
gm4(i)=f4(i)*ID4(i);

W1(i)=ID1(i)*L1/Ip1(i);
W4(i)=ID4(i)*L4/Ip4(i);
```

Table 1 shows the Optimization results of systematic design methodology for gain-boosted regulated-cascode CMOS stages:

Parameters	Values
(W/L)1-2	31/2
(W/L)3-4	25/2
(W/L)5-6	51/2
(W/L)7-8	23/2
(gm/Id)1-2	22
(gm/Id)3-6	16.8
(gm/Id)7-8	15
Av0(dB)	85.3
fT(khz)	480
Margin phase(°)	88
CL(pF)	10

Table1. Optimization results

VI. SUMMARY

The 90 dB - 100 MHz bulk CMOS example nevertheless demonstrate that our design methodology may also converge towards solutions nearing the technology limits, in which intuitive or SPICE-based approaches would generally fail. Furthermore it is worth to point out that all the MATLAB calculations related to the synthesis procedure are achieved in short CPU times, on the order of minutes, so that they may be repeated a large number of times searching for the optimal design space or incorporating other specifications, e.g. the minimization of power consumption.

REFERENCES

[1] F. Silveira, D. Flandre, and P. G. Jespers, "A (*gm/ID*) based methodology for the design of cmos analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 9, pp. 1314–1319, Sept. 1996.

[2] Y.P. Tsividis and K. Suyama, *IEEE Journal of Solid-State Circuits* **SC-29**, 210 (1994).

[3] F. Van de Wiele, "A long-channel MOSFET model", Solid-State Electronics, 1979, **22**, pp. 991-997.

[4] C. Enz, F. Krummenacher and E.A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-woltage and low-current applications", Analog Integrated Circuit and Signal Processing, vol. 8, n°1, pp. 83-114, 1995.

[5] C. Galup-Montoro, M. Schneider, and A. Cunha, "A current-based MOSFET model for integrated circuit design," in *Low-Voltage / Low-Power Integrated Circuits and Systems: Low-Voltage Mixed-Signal Circuits*, E. Sanchez-Sinencio and A. Andreou, Eds. IEEE Press, ISBN 0-7803-3446-9, 1999, ch. 2, pp. 7–55.

[6] A. Cunha, M. Schneider, and C. Galup-Montoro, "An MOS transisotr model for analog circuit design," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 10, pp. 1510–1519, Oct. 1998.

[7] Jean-Paul Eggermont, Denis De Ceuster, Denis Flandre, " Design of SOI CMOS Operational Amplifiers for Applications up to 300°C » *IEEE Journal of Solid-State Circuits*, vol. 31, no. 2, February 1996.