A Design Methodology Using the Inversion Coefficient for Low-Voltage Low-Power CMOS Voltage References

Dalton Colombo[1], Christian Fayomi[2], Frederic Nabki[2], Luiz F. Ferreira[1], Gilson Wirth[1], and Sergio Bampi[1]

[1] Microelectronics Program (PGMICRO), UFRGS, Porto Alegre - RS, Brazil.
[2] Microelectronics Laboratory, Comp. Sc. Dept., UQAM, Montreal - QC, Canada.
e-mail: dmcolombo@inf.ufrgs.br

ABSTRACT

This paper presents an analog design methodology, which uses the selection of the inversion coefficient of MOS devices, to design low-voltage and low-power (LVLP) CMOS voltage references. The motivation of this work comes from the demand for analog design methods that optimize the sizing process of transistors working in subthreshold operation. The advantage of the presented method – compared to the traditional approaches for circuit design – is the reduction of design cycle time and the minimization of simulation iterations when the proposed equations are used. As a case study, a LVLP voltage reference based on subthreshold MOSFETs with a supply voltage of 0.7 V was designed in a 0.18-µm CMOS technology.

Index Terms: CMOS Voltage Reference, Analog Design Methodology, Inversion Coefficient, Low Voltage Design, Low Power Design

1. INTRODUCTION

The current trend towards low-voltage supply and low-power (LVLP) design is mainly caused by the growing demand for battery-operated portable equipment such as watches, smart phones and implantable devices (e.g., pacemakers and hearing aids). The operation of these circuits for long periods requires a significant amount of energy from the battery. However, battery size is often a limiting factor as its volume is limited for portability. In addition, the speed and high-integration density of transistors in single dies have increased heat dissipation to critical limits. Therefore, analog and mixed-signal circuits in the aforementioned systems must operate more and more under LVLP conditions.

The voltage reference is an essential component of several analog blocks such as data converters, voltage regulators and phase-locked loops. The reference must generate a precise output voltage that is ideally independent of process, supply, load and temperature variations. Currently, voltage references must be amenable to standard digital CMOS processes and thus operate with a supply bellow 1 V, while consuming few tens of nW to a few µWs.

Unfortunately, the design of such voltage references for commercial purposes becomes even more complicated due to the reduced time to market constraints. Accordingly, design methodologies have been proposed to reduce the cycle time of the analog design process, and enable the designer to explore different design options quickly while evaluating their tradeoffs [1-3]. In addition, several computational tools for automatic and optimized analog design have been developed [4-6].

An analog design methodology that provides good insight leading towards optimized design is the selection of the inversion coefficient (*IC*) of MOS transistors [7, 8]. The inversion coefficient is a numerical measure of the channel inversion, which depends on the applied bias voltage at the MOS terminals. In other words, the *IC* is a normalized number that is proportional to the quantity of free carriers in the channel region. The selection of the *IC* enables design within weak, moderate or strong inversion operation. Transistors operating in weak and moderate inversion are important for LVLP applications due to their low drain source saturation voltage (V_{DSAT}) and high transconductance efficiency (g_m/I_D). Using simple equations motivated by the EKV MOS model [9], the method in [7, 8] guides the designer in the manual selection of bias currents and transistor sizes, resulting in an optimized design.

Accordingly, this paper presents a transistor-level design methodology for low-voltage low-power CMOS voltage references that involves the selection

of the inversion coefficient. Using a spreadsheet-type design tool, drain bias current, inversion coefficient and channel length choices are mapped into channel width. The proposed method, used as initial design guidance, reduces the design time and minimizes the number of required simulations [10]. The circuit used as a study case, proposed originally in [11], was designed in a commercial 0.18-µm CMOS technology using BSIM3V3 models, and simulated with the HSPICE simulator.

This paper is organized in the following sequence. The inversion coefficient method is briefly explained in Sec-tion 2. Section 3 presents the design of the low-voltage and low-power CMOS reference. Simulation results and conclusions are presented in Sections 4 and 5, respectively.

2. INVERSION COEFFICIENT-BASED METHODOLOGY

As previously mentioned, the *IC* provides a numerical representation of the MOS inversion level. Weak inversion corresponds to $IC < 0.1$, while moderate inversion corresponds to $0.1 < IC < 10$. For $IC > 10$, MOS transistors are operating in strong inversion [7]. In weak inversion, the transtransport of carriers in the channel is dominated by diffusion; while in strong inversion, the prevailing transport mechanism is drift.

Table I shows Equations (1) – (5) that describes *IC* as a function of drain-source current, transistor aspect ratio ($S = W/L$) and the specific current I_0 (a technology and transistor-type dependent current) [7]. Note that parameters N_{SUB}, γ, μ_0, C_{OX} can be extracted from the BSIM3V3 model provided by the foundry. Parameter n_0 is the substrate factor which represents a loss of coupling efficiency between the gate and channel caused by the substrate or body, which acts as a back gate.

Table II shows the calculated values of I_0 and n_0, and also the typical values of the process parameters in the 0.18-µm CMOS technology used in this design.

Table III shows Equations (6)-(10) that describe channel width, gate area, transconductance efficiency (g_m/I_D), drain-source saturation voltage (V_{DSAT}) and effective gate-source voltage ($V_{EFF} = V_{GS} - V_T$) as function of inversion coefficient, respectively [7]. Equation (10) approaches Equation (11) for weak inversion (*i.e.* IC<< 1) and Equation (12) for strong inversion operation (*i.e.* IC >> 10). All these equations are valid only for saturation operation.

Equation (9) is very useful for the design of low-voltage circuits since it gives the boundary between the triode and saturation region in terms of the inversion level. This equation can be used to estimate the minimum supply required by voltage references to work properly. Note that in weak inversion, V_{DSAT} is nearly four times the thermal voltage, *i.e.* approximately 100 mV at 300 K.

For the LVLP voltage reference presented in this paper to work properly, its transistors must operate in either weak or moderate inversion, as will be shown in the next section. The traditional equations used to design transistors are given by (13) for strong inversion, and (14) for weak inversion operation for a drain-source voltage that is higher than 0.1 V. Parameter λ models the channel length modulation

Table I. Expressions for the Inversion Coefficient (IC).

Equations	
$IC = \frac{I_D}{I_0 \cdot S}$ (1)	
$I_0 = 2n_0 \cdot \mu_0 \cdot C_{OX} \cdot U_T^2$ (2)	
$n_0 = 1 + \frac{\gamma}{2\sqrt{\psi_0 + (V_{GS} - V_T)/n + V_{SB}}}$ (3)	
$\psi_0 \approx 2\phi_F + 4U_T$ (4)	
$\phi_F = U_T \cdot \ln(N_{SUB}/n_i)$ (5)	
Parameters	
IC = inversion coefficient	I_D = drain-source current, [A]
S = Transistor size aspect ratio	I_0 = technology current, [A]
n_0 = substrate factor	μ_0 = low-field mobility,[µA/V²]
C_{OX} = gate-oxide cap. [fF/µm²]	U_T = Thermal voltage [mV]
γ = body-effect factor, [V$^{1/2}$]	V_{SB} = source bulk voltage, [V]
ψ_0 = psi parameter, [V]	ψ_F = Fermi Potential, [V]
N_{SUB} = substrate doping concentration, [cm^{-3}]	n_i = silicon intrinsic carrier concentration, [cm^{-3}]

Table II. Device parameters at 300 K for the 0.18-µm process used

Parameters	nMOS	pMOS	Units
I_0	0.630	0.120	µA
n_0	1.234	1.211	-
μ_0	490.8	98.6	µA/V²
C_{OX}	7.79	7.53	fF/µm²
γ	0.471	0.421	V$^{1/2}$
ψ_0	0.987	0.973	V
N_{CH}	4E+17	3E+17	cm^{-3}
n_i	1.5E+10	1.5E+10	cm^{-3}

Table III. Expressions for W, W*L, g_m/I_D, V_{SAT} and V_{EFF} in saturation

$W = (L/IC) \cdot (I_D/I_0)$	(6)
$W \cdot L = (L^2/IC) \cdot (I_D/I_0)$	(7)
$g_m/I_D = 1/\left(n \cdot U_T \cdot \left(\sqrt{IC + 0.5\sqrt{IC}} + 1\right)\right)$	(8)
$V_{DSAT} = 2U_T \cdot \sqrt{IC + 0.25} + 3U_T$	(9)
$V_{EFF} = 2n \cdot U_T \cdot \ln\left(e^{\sqrt{IC}} - 1\right)$	(10)
$V_{EFF} = n \cdot U_T \cdot \ln(IC)$	(11)
$V_{EFF} = 2n \cdot U_T \cdot \sqrt{IC}$	(12)

phenomenon, while parameter n is often recognized as the slope factor when transistors are operating in weak inversion.

$$I_D = (1/2) \cdot \mu_0 \cdot C_{OX} \cdot S \cdot (V_{GS} - V_T)^2 \cdot (1 + \lambda \cdot V_{DS}) \quad (13)$$

$$I_D = 2 \cdot n \cdot \mu_0 \cdot C_{OX} \cdot S \cdot {U_T}^2 \cdot exp[(V_{GS} - V_T)/n \cdot U_T] \quad (14)$$

Nevertheless, as (6) provides a more conscious choice of the level of MOS inversion than the equations of the traditional approach, the proposed methodology uses it instead.

Figure 1 illustrates estimated values of V_{DSAT} and V_{EFF} for different values of IC, using (9) and (10). Using this figure, it is possible to make a comparison between different modes of operation: weak, moderate and strong inversion. Values of V_{DSAT} are around 100 mV for weak inversion and between 100 to 240 mV for moderate inversion, respectively. Such values are well suited to low voltage requirement, which represents the first advantage of operating within these regions. The second can be seen in Figure 2, which shows the simulated g_m/I_D as a function of V_{EFF}. One can see that transcondutance efficency is maximum in weak inversion. The efficiency drops to ~65% of its value in the center of moderate inversion, where V_{GS} - V_T is ~35 mV as annotated by point 'X' in the graph, and is lowest in strong inversion.

The shape of the g_m/I_D curve shown in Figure 2 is almost universal for MOS operation, and it is weakly dependent on channel length until velocity saturation effects reduce transcondutance efficiency [1, 7].

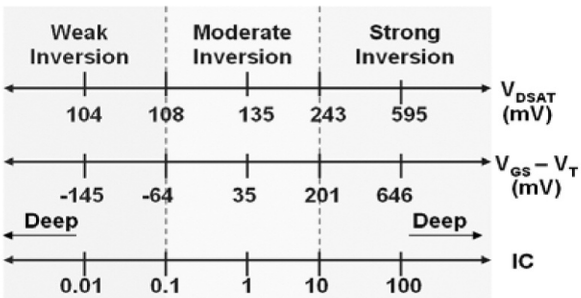

Figure 1. V_{DSAT} and (V_{GS}-V_T) as a function of IC

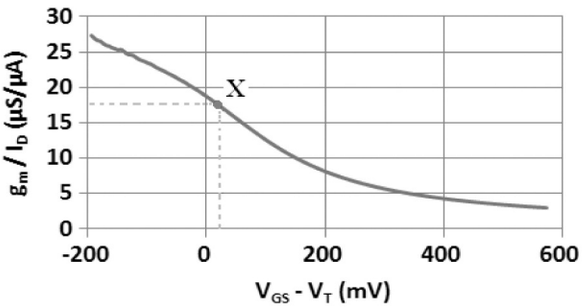

Figure 2. Simulated g_m/I_D as a function of V_{GS}-V_T

A. Related Work

As explained above, hand analysis expressions shown in Table III are based on the EKV MOS model. Another MOS model that presents the same basic physical variables of the EKV model is the ACM model [12]. This compact charge-based model is valid for all regions of operation, and presents equations with infinite order of continuity for the large and small-signal MOS parameters. The technology (or specific current) in the ACM model is four times smaller than that given by (2). The current version of the ACM model includes short channel effects, noise and process variability. Using equations derived from the ACM model, work in [2] proposes an analog design methodology useful for the design of CMOS amplifiers. The model authors in [13] present a graphical tool for transistor-level design using this model. The methodology proposed herein uses equations based on EKV model and it is focused on the design of subthreshold CMOS voltage references, which is different from what is discussed in the aforementioned works.

3. LOW VOLTAGE LOW POWER CMOS VOLTAGE REFERENCE

A widely used implementation for voltage references is the bandgap voltage reference (BGR). The traditional BGR generates an output voltage (V_{REF}), described by (15), through the addition of a base-emitter voltage (V_{BE}) to a properly scaled Proportional to Absolute Temperature (PTAT) voltage.

$$V_{REF} = V_{BE} + K_{PTAT} \cdot U_T \quad (15)$$

The constant K_{PTAT} is temperature-independent (to first order) gain factor needed to achieve the proper temperature compensation, as V_{BE} has a negative temperature coefficient. Its value is about 10, and is usually defined by the ratio of two resistances with the same temperature coefficients.

When the value of V_{REF} generated by (15) is close to V_{G0}, the bandgap energy of silicon divided by the electron charge (*i. e.* ~1.12 V), the conventional BGR architecture must be avoided in applications with supplies that are lower than 1 V. As a consequence of this limitation, other BGR topologies suitable for low voltage operation were proposed in [14-15]. In [14], a resistive divider is used to decrease V_{REF}, as given by (16), where ω is a ratio of resistances that can be adjusted to less than 1.

$$V_{REF} = \omega \cdot (V_{BE} + K_{PTAT} \cdot U_T) \quad (16)$$

Usually, the V_{BE} voltage required by (15) and (16) is implemented in CMOS by the means of parasitic

vertical or lateral bipolar transistors. However, some standard digital CMOS processes do not allow for parasitic BJTs, and if these devices are available, they are often not very well characterized and consequently electrical models are not very accurate [16].

An alternative solution for the implementation of temperature compensated voltage references is to replace the BJTs by MOSFETs operating in the subthreshold operation. In [17], V_{BE} is replaced by V_{GS} and the created output voltage is similar to that of (16), as given by:

$$V_{REF} = \omega \cdot (V_{GS} + K_{PTAT} \cdot U_T) \quad (17)$$

The gate-source voltage of a MOSFET, when biased with a current less than a certain technology-dependent value (I_X), decreases with temperature (T) in a quasi-linear fashion [18]. This behavior happens when the decrease in the threshold voltage caused by the increase of T, outweigh the effects caused by the decrease in mobility of the carriers. For bias currents higher than I_X, the opposite happens: the decrease of mobility caused by the increase of T, outweighs the effect of the reduction in the threshold voltage.

The temperature dependence of V_{GS} is given by (18) and (19) [17], where it is considered that the variation of n with temperature is small. These equations show the linear dependence between V_{GS} and T. Variables K_{T1} and V_{OFF} are BSIM3V3 parameters for temperature coefficient of V_T and offset voltage in subthreshold region at large W and L [19], respectively. For the technology used, K_{T1} and V_{OFF} are –0.282 and –0.120 V, respectively.

$$V_{GS}(T) \approx V_{GS}(T_0) + K_G \cdot [(T/T_0) - 1] \quad (18)$$

$$K_G \cong K_{T1} + V_{GS}(T_0) - V_T(T_0) - V_{OFF} \quad (19)$$

The value of V_{REF} generated by (17) is given by the threshold voltage extrapolated to zero Kelvin (V_{T0}) and a few process-dependent parameters, as shown in [20]. This result is analogous to the output of the bandgap voltage references. While BGRs have V_{REF} close to V_{GO}, CMOS references working in subthreshold operation have V_{REF} close to V_{T0}. Threshold voltages are, for instance, typically in the range of 600 and 400 mV, for 0.25-µm and 90-nm CMOS technologies, respectively.

As a consequence of having V_{REF} proportional to V_{TH0}, this type of voltage reference is able to operate with supply voltages that are lower than 1-V. However, it is well known that V_T is strongly dependent on substrate doping [21]. As the number of dopant atoms has a random distribution, the output voltage also presents a statistical fluctuation. However, if statistical data of process dispersion is available to the designer, it is possible to employ design techniques to mitigate the impact of variability. For example, if high accuracy performance is required by the voltage reference, a trimming circuit can be added in order to mitigate the impact of fabrication process effects [20].

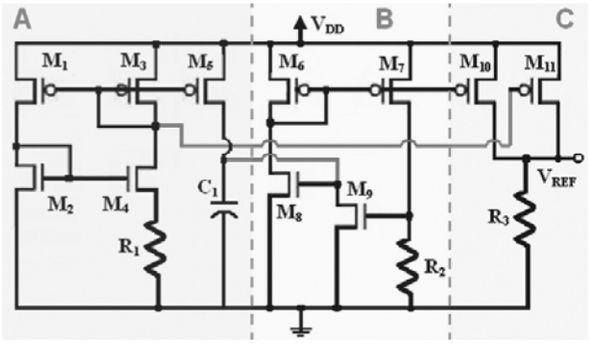

Figure 3. Simulated voltage reference used as a case study

The voltage reference used as a case study in this work was previously proposed in [11] and is shown in Figure 3. Its operation is as described by (17): a balanced summing of V_{GS} and U_T. By adjusting ω, it is possible to tune V_{REF} to higher or lower values than V_T.

The circuit is composed by tree sub circuits: a) a PTAT current generator, b) a V_{GS} current generator and c) a current adder. The original circuit in [11] works under 1-V of V_{DD}, but the aim of the approach proposed here is the nominal operation under 0.7 V. Moreover, the circuit in Figure 3 is designed to be robust against V_{DD} fluctuations in the range of 0.6 to 1.8 V.

The proposed design methodology, using equations based on the inversion coefficient, provides a step-by-step design process, for each sub-circuit of the voltage reference, which reduces the design time.

A. PTAT Current Generator

Sub-circuit A in Figure 3 is a classical supply independent current source, for which the current in R_1 (I_{R1}) is defined by the difference between the gate-source voltage of M_2 and of M_4 (ΔV_{GS}), and the resistance R_1. Using (14), one can show that ΔV_{GS} between M_2 and M_4 is described by (20), if the body effect is neglected and these devices are in weak inversion operation.

Using (11), Equation (20) can be expressed in terms of the IC, as given by (21). Current I_{R1}, described by (22), is proportional to U_T, which translates to a positive temperature coefficient.

$$\Delta V_{GS} = V_{GSM2} - V_{GSM4} = n \cdot U_T \cdot \ln(S_{M4}/S_{M2}) \quad (20)$$

$$\Delta V_{GS} = n \cdot U_T \cdot \ln(IC_{M2}/IC_{M4}) \quad (21)$$

$$I_{R1} = \Delta V_{GS}/R_1 \quad (22)$$

Making (IC_{M2}/ IC_{M4}) equal to 10, it is possible to provide a good layout matching between M_2 and M_4. When n and U_T (@ 300K) are respectively around 1.234 and 25.85 mV, the estimated value of ΔV_{GS} is 73.44 mV. Therefore, after the choice of I_{R1}, based on the power-consumption requirement, the resistance R_1 can be found. In this design, it is assumed that I_{R1} = 200 nA, which implies that R_1 = 367.22 kΩ.

As mentioned above, M_2 and M_4 must operate in weak inversion for (20) to be valid. Therefore, using (6) and choosing a value of *IC* lower than 0.1, the channels width of M_2 and M_4 can be estimated.

The dimensions of M_1, M_3 and M_5 were considered to be equal in order to provide good layout matching. Transistors in moderate inversion operation have less thermal noise and require less silicon area than those in weak inversion - for a given I_D [8]. Hence, M_1, M_3 and M_5 should operate in moderate inversion with the aim of reducing silicon area.

The moderate-values of V_{DSAT} of these transistors, when working in moderate inversion, are well suited to our supply voltage requirement, *i.e.* 0.6 V. Consequently, for M_1, M_3 and M_5, *IC* values between 0.1 and 10 were chosen.

Moreover, it is well known that the effects of mismatch and flicker noise on transistor performance are roughly given by (23) – (26) [22-23]. Parameters $\sigma^2(\Delta V_T)$, $\sigma^2(\Delta\mu * C_{OX} * W/L)$ and V_n^2 model the variance of the mismatch on threshold voltage , the variance of the mismatch on current gain; and the flicker noise voltage in a bandwidth of 1 Hz, respectively. In literature, the current gain is often called β. Variables AV_T, A_K and K_F are process-dependent constants; and f is the frequency. Parameter $\sigma(\Delta I_D/I_D)$ represents the square-root of the mismatch variance in the drain-source current, which includes ΔV_T and $\Delta\beta$ as the causes of variation. As can be seen in (25), the higher g_m/I_D, the higher the contribution of ΔV_T on the I_D mismatch. This is expected, as in weak inversion, the transconductance efficiency is maximized and thus a maximum mismatch is expected.

As a consequence of the above discussion, as the flicker noise and mismatch are approximately inversely proportional to the square root of transistor area, the channel lengths of M_1–M_5 were selected to be around 15 times the minimum value permitted by the technology.

Considering a channel length of 3 µm and a minimum gate area of nearly 50 µm², the maximum expected variance of V_T is equal to 1.8 mV. Constants A_{VT} for NMOS and PMOS transistors are 1.244x10^{-8} and 6.73x10^{-9}, respectively in the 0.18-µm CMOS technology used.

$$\sigma^2(\Delta V_T) = A_{VT}^2/\sqrt{W \cdot L} \quad (23)$$

Table IV. Transistor sizes, estimated and simulated values of g_m (sub-circuit A)

	Design Choices		Designed Sizes		
Device	I_D (nA)	IC	L (µm)	W (µm)	W*L (µm²)
M_2	200	0.06	3	15.87	47.6
M_4	200	0.006	3	158.7	476
$M_{1,3,5}$	200	0.25	3	19.97	59.9
Device	**Estimated g_m (µS)**	**Simulated g_{mS} (µS)**	**Error (g_m-g_{mS}) (%)**		
M_2	5.77	4.88	18.3		
M_4	6.13	5.47	12.2		
$M_{1,3,5}$	5.22	4.37	19.3		

$$(\sigma(\Delta\mu \cdot C_{OX} \cdot W/L)/(\mu \cdot C_{OX} \cdot W/L))^2 = (\sigma(\Delta\beta)/\beta)^2 = A_K^2/\sqrt{W \cdot L} \quad (24)$$

$$\left(\sigma(\Delta I_D/I_D)\right)^2 \cong (\sigma(\Delta\beta)/\beta)^2 + (g_m/I_D)^2 \cdot \sigma^2(\Delta V_T) \quad (25)$$

$$V_n^2 = K_F/(C_{OX} \cdot W \cdot L \cdot f) \quad (26)$$

After simulation, it was verified that the correct value of ΔV_{GS} is 86.67 mV; thus, the correct value of R_1 is 434.6 kΩ. The error on estimation of R_1 provided by (21) is nearly 15%. One of the causes of such error comes from the difference between V_T of M_2 and M_4 due to the body effect, which is neglected in this equation.

It is valid to add that the simulated temperature coefficient (TC) of ΔV_{GS} in the temperature range of -40 to 120 °C is ~0.203 mV/ °C.

Table IV presents the design choices (I_D, *IC* and *L*) for each transistor, and the estimated values of *W*, *W*L* and g_m using Equations (6) - (8). The simulated values of transconductance (g_{mS}) is also presented, in order to verify the accuracy of the proposed transistor sizing process compared with simulated results.

B. V_{GS} Current Generator

Sub-circuit B is responsible for the generation of a current with negative TC. Due to the feedback loop present in this circuit, I_{R2} is defined by (27), where the bias current of M_9 is a copy of I_{R1} (≈200 nA) through M_5.

$$I_{R2} = V_{GSM9}/R_2 \quad (27)$$

In order to reduce the power consumption and the required value of R_2 for a given current, M_9 must operate in weak inversion, and thus consequently presents a low V_{GS}. Using (11), (27) can be rewritten as a function of *IC*, as shown by (28). In the design, M_9 is defined as operating in the deep weak inversion with *IC* equal to 0.009, which corresponds to a V_{GS} of 176.3 mV.

Table V. Transistor sizes, estimated and simulated values of g_m (sub-circuit B)

	Design Choices		Designed Sizes		
Device	I_D (nA)	IC	L (µm)	W (µm)	W*L (µm²)
M6	203	0.25	3	20.31	60.9
M7	406	0.25	3	40.62	121.9
M8	203	0.25	3	3.87	11.6
M9	200	0.009	3	105.77	317.3
Device	**Estimated** g_m (µS)	**Simulated** g_{mS} (µS)	**Error** $(g_m\text{-}g_{mS})$ (%)		
M6	5.31	4.35	21.9		
M7	10.61	8.80	20.5		
M8	5.21	4.43	17.6		
M9	6.1	5.37	13.6		

$$I_{R2} = [n \cdot U_T \cdot \ln(IC_{M9}) + V_T]/R_2 \quad (28)$$

From the value of V_{GSM9} @200nA, it is possible to design for R_2 while considering the tradeoff between resistor area and power consumption. We have defined R_2 equal to R_1 for the purpose of providing good resistor matching. For R_2 = 434.6 kΩ, current I_{R2} is approximately 400 nA.

Transistors M_6 – M_8 should operate in moderate inversion in order to reduce silicon area. Hence, the *IC* for these devices is chosen between 0.1 and 1, similarly as was done in section A. The current of devices M_6 and M_8 is derived from M_9 and R_2, and is designed to be half of I_{R2} with the intention of reducing power consumption.

Table V presents the design choices (I_D, *IC* and *L*) for each transistor, and the estimated values of *W*, *W*L* and g_m using (6-8). The simulated value of g_m is also presented for comparison.

It is also important to estimate the temperature coefficient of V_{GSM9}, before designing sub-circuit C. Using (18) and (19), the TC of V_{GSM9} is estimated to be -1.152 mV/ °C; while the simulation shows a TC equal to -1.026 mV/ °C. The difference between the simulated and estimated values is 12.2%, and is predominantly due to the fact that Equations (18) and (19) consider the bias current of M_9 (I_{M5}) to be independent of temperature. As explained in section A however, the current provided by M_5 has a PTAT dependency. As such, if an ideal dc current source is used to bias M_9, the estimation provided by (18) and (19) differs by only 1.5% of the simulated results.

Nonetheless, the estimation provided by (18) and (19) gives acceptable values (*e. g.* less than 13% difference), because the relation of vgs_{M9} and I_{M5} in subthreshold operation is logarithmic.

C. Current Adder

Sub-circuit C is a current adder with resistor R_3 that converts current to voltage. The generated output voltage is described by (29). Using (22) and (27), Equation (29) can be rewritten as (30). Note that I_{M6} is half of I_{R2}.

$$V_{REF} = (I_{M10} + I_{M11}) \cdot R_3 = R_3 \cdot [(S_{M10}/S_{M6}) \cdot (V_{GSM9}/2 \cdot R_2) + (S_{M11}/S_{M3}) \cdot (\Delta V_{GS}/R_1)] \quad (29)$$

$$V_{REF} = \left[\begin{matrix}(S_{M10}/S_{M6}) \cdot (V_{GSM9}/2 \cdot R_2) + \\ (S_{M11}/S_{M3}) \cdot (\Delta V_{GS}/R_1)\end{matrix}\right] \cdot R_3 \quad (30)$$

As can be seen in (30), the temperature compensation of V_{REF} is provided by a proper sizing of the transistors. Note that R_3 does not affect the temperature compensation, and for this reason, V_{REF} can be adjusted to the desired level simply by adjusting R_3.

In order to provide a good layout matching, R_3 is defined to be equal to R_1. Moreover, (S_{M10}/S_{M6}) is set to be 1, with the aim of keeping low power consumption. Therefore, (30) can be rewritten as (31). Note that V_{REF} is a function of *IC* because V_{GSM9} and ΔV_{GS} are defined respectively by (11) and (21).

$$V_{REF} = (1/2) \cdot V_{GSM9} + (S_{M11}/S_{M3}) \cdot \Delta V_{GS} \quad (31)$$

To estimate the required value of (S_{M11}/S_{M3}) for temperature compensation, one can take the derivative with respect to temperature in (31). Imposing $(dV_{REF}/dt) = 0$, one can show that (S_{M11}/S_{M3}) can be found through (32). For the simulated values of TC_{VGSM9} and $TC_{\Delta VGS}$ which are -1.03 mV/ °C and 0.203 mV/ °C, respectively, (S_{M11}/S_{M3}) is found to be around 2.5; which corresponds to $S_{M11} = (50.57/3)$.

$$0 = (1/2) \cdot TC_{VGSM9} + (S_{M11}/S_{M3}) \cdot TC_{\Delta VGS} \quad (32)$$

In summary, the current mirror composed by M_{10} and M_{11} is sized to allow proper temperature compensation. As a consequence, the inversion level of these devices is set by TC_{VGSM9}, $TC_{\Delta VGS}$ and the inversion levels of M_3 and M_6. When M_3 and M_6 are operating in the beginning of moderate inversion, M_{10} and M_{11} are consequently operating in moderate inversion as well. Table VI summarizes the estimated values of *W*, *W*L* and g_m.

Table VI. Transistor sizes, estimated and simulated values of g_m (sub-circuit C)

	Design Choices		Designed Sizes		
Device	I_D (nA)	IC	L (µm)	W (µm)	W*L (µm²)
M_{10}	203.4	0.25	3	20.31	60.9
M_{11}	506.4	0.25	3	50.57	151.7
Device	**Estimated** g_m (µS)	**Simulated** g_{mS} (µS)	**Error** $(g_m\text{-}g_{mS})$ (%)		
M_{10}	5.31	4.41	20.4		
M_{11}	13.21	11.04	19.7		

Equations (1-12), (17-22) and (27-32) were implemented in a spreadsheet, which allows for quick design (and redesign). Finally, it is also worth adding that due to the high values of resistance needed to generate currents lower than 1 µA, the resistors would be made of high-resistance P-Poly material available in the technology.

4. SIMULATION RESULTS

After sizing all the transistors, the voltage reference was simulated and the key performance parameters obtained are presented in this section. The voltage reference variation caused by temperature variation (ΔV_{REF_TEMP}) was measured in the range of -40 to 120 °C. The output voltage variation caused by V_{DD} fluctuations (ΔV_{REF_VDD}) was measured considering a V_{DD} variation of 0.65 to 0.75 V. In addition, the output noise (ΔV_{REF_NOISE}) was integrated in a 500 kHz bandwidth, and the temperature coefficient, as described by (33), was also estimated.

$$TC = V_{REF}(T_0)^{-1} \cdot \left(\Delta V_{REF\_TEMP}/\Delta T\right) \qquad (33)$$

Figures 4 and 5 illustrate, respectively, V_{REF} as a function of temperature for the nominal R_3, and for three different values of R_3. As already explained, V_{REF} can be adjusted to the desired level by simply adjusting the value of R_3 without affecting the TC. The value of V_{REF} at 27 °C is 309.6 mV, while ΔV_{REF_TEMP} is equal to 3.2 mV – what represents a 1% of output variation in the temperature range.

Figure 6 shows the power supply rejection (PSR) performance of the voltage reference when a 5 pF load capacitance is connected at the output node. Its value at 10 kHz is 36.6 dB.

Figure 7 presents V_{REF} as a function of V_{DD}. The circuit works properly for V_{DD} down to 0.6 V, and presents an output variation of 2 mV for a ΔV_{DD} of 100 mV. This variation represents a line regulation performance of nearly 20 mV/V.

Figure 8 illustrates the frequency spectrum of the output noise voltage at node V_{REF}. As can be seen, the corner frequency is nearly 1 kHz. The integrated RMS output noise in a 500 kHz bandwidth is 277 µV. Simulation shows that flicker noise of M_2, thermal noise of R_1, channel thermal noise of M_3 and M_{11} correspond to: 10%, 28%, 35% and 6% of the output noise, respectively.

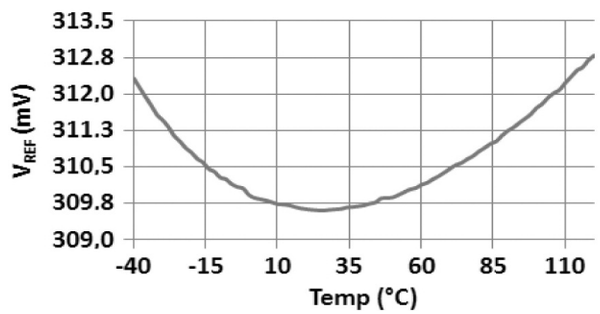

Figure 4. Simulated V_{REF} as a function of temperature

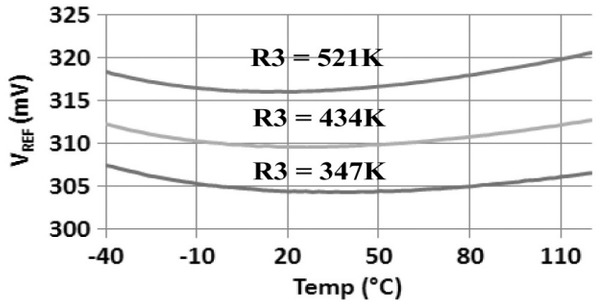

Figure 5. Simulated V_{REF} as a function of temperature for three different values of R_3

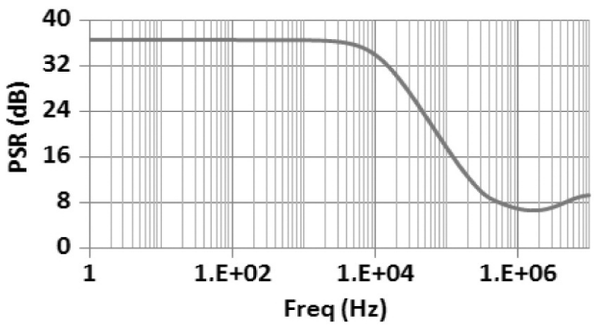

Figure 6. Simulated PSR

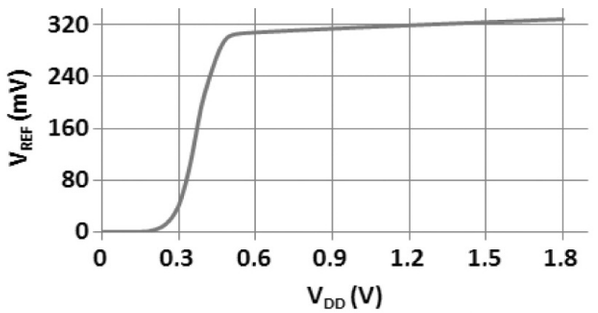

Figure 7. Simulated V_{REF} as a function of V_{DD}

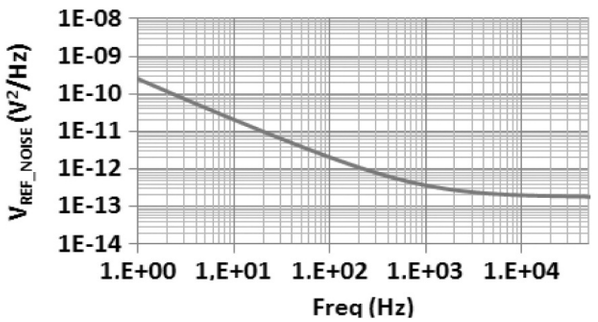

Figure 8. Frequency spectrum of the output noise

Table VII. Simulated Performance Parameters

V_{REF} (mV)	TC (ppm/°C)	ΔV_{REF_TEMP} (mV)	ΔT (°C)	ΔV_{REF_VDD} (mV)
309.6	130	3.2	160	1.98
Δ_{VREF_NOISE} (µV)	PSR (dB)	I_{supply} (µA)	VDD (V)	P (µW)
277	36.6	1.92	0.7	1.34

Table VIII. Simulation of V_{REF} using corner models

Model	# Corner Simulation								
	0	1	2	3	4	5	6	7	8
Nmos	T	F	S	F	S	F	F	S	S
Pmos	T	F	S	F	S	S	S	F	F
Res	T	S	F	F	S	S	F	S	F
V_{REF} (mV)	310	280	342	291	329	277	289	332	344

The current required by the circuit is 1.92 µA, and the power consumption is 1.34 µW. Lower power consumption can be achieved if higher values of resistances R_1 and R_2 are used at the cost of added area. Table VII presents all the results discussed so far.

A. Impact of Variability on V_{REF}

It is expected that process variations, such as mismatch and the tolerance of MOSFETs and resistors, will adversely affect the performance of the fabricated voltage reference. To mitigate these effects, trimming circuits can be included in the reference circuit design when the application requires high-accuracy. Such circuits are often a set of switches that permit the adjustment of the resistance or current in some branch of the circuit.

To estimate the impact of the process variation on the performance of the designed BGR, the circuit was simulated using corner and Monte Carlo models.

MOS transistors corner models used in simulations were: SS (slow-slow), FF (fast-fast), SF (slow-fast) and FS (fast-slow). The resistor models used were: SS and FF.

Figure 9 shows V_{REF} as a function of temperature using corner models. The minimum and maximum values of V_{REF}(27°C) found in these simulations are 277 mV and 344 mV, representing a variation of V_{REF} equal to 67 mV. Table VIII summarizes the simulation results of V_{REF} using corner models.

Figure 10 presents V_{REF} as a function of V_{DD} using corner models. In all simulation cases, the circuit operates properly with the minimum supply of 0.6 V.

Because mismatch is not taken into account in the former analysis, Monte Carlo simulation was also executed to have a complete estimation of fabrication process effects.

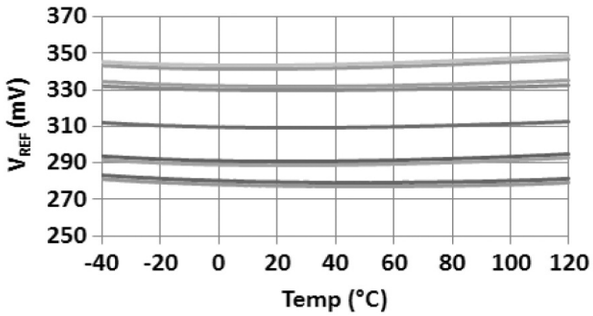

Figure 9. Simulated V_{REF} as a function of temperature using corner models

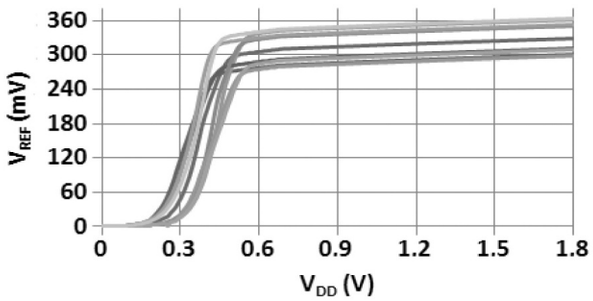

Figure 10. Simulated V_{REF} as a function of V_{DD} using corner models

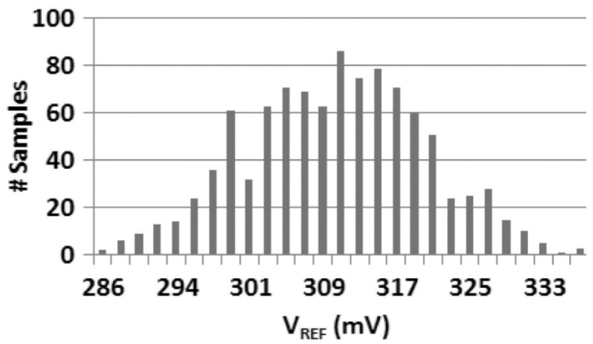

Figure 11. Monte Carlo simulation of V_{REF} at 27ºC

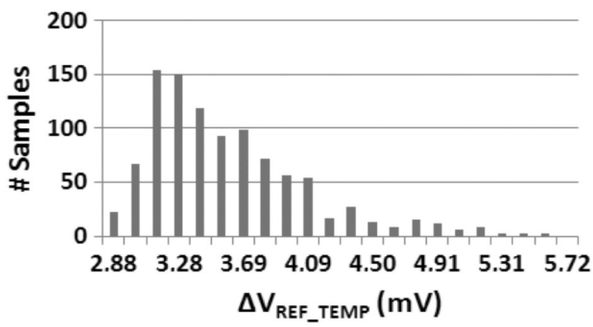

Figure 12. Monte Carlo analysis of $\_V_{REF\_TEMP}$

In the Monte Carlo analysis, the process and mismatch variations were included in a simulation considering one thousand samples. Figure 12 shows the histogram for V_{REF} at 27 °C. The mean value is

309.7 mV, while the sigma is equal to 9.6 mV. This represents for a 3-sigma requirement (99.73% of samples), a total variation of V_{REF} equal to ± 28.8 mV.

The temperature performance was also verified in the Monte Carlo analysis, as shown in Figure 12. The mean value of ΔV_{REF_TEMP} is 3.5 mV, while sigma is 0.51mV. In the worst case scenario, the output will vary nearly 5 mV in the temperature range of -40 to 120 °C.

The designed voltage reference is not able to supply current to the load. Therefore, in order to provide current to the application, the reference circuit must be connected to a buffer, that is, an operational amplifier in negative feedback configuration as illustrated in Figure 13. The amplifier must have high open-loop gain and present a frequency response as given by the voltage reference circuit.

In the non-inverting configuration shown in Figure 13, the output voltage (V_{OUT}) is given by (34). As V_{OUT} depends on the resistor ratio, which is insensitive to temperature variations in first approximation, it follows the temperature dependency of V_{REF}.

$$V_{OUT} = (1 + R_4/R_5) \cdot V_{REF} \quad (34)$$

Resistor R_5 (or R_4) can be designed to be adjustable through a set of switches that allow to increase or decrease its resistance. These switches can be simply transistors with low on-resistance. The trimming circuit composed by R_5 or R_4 is used to adjust V_{OUT} and therefore mitigate the variability caused by the fabrication process. For instance, in applications where a full scale variation of V_{REF} is equal to 67 mV – an estimation provided by corner simulations – is not acceptable, the above technique can be employed to ensure accuracy better than 10 mV [24]. In fact, the accuracy will be defined by the number of bits available for trimming.

It is worth noting that the amplifier should present low offset voltage (V_{OFFSET}), as this offset is directly transferred to the output. In addition, it is desirable to properly isolate the reference circuit and noisy traces, as the coupling noise is also transferred to V_{OUT}.

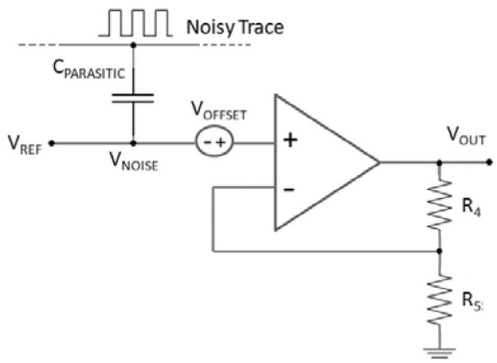

Figure 13. Operational amplifier in negative feedback configuration

B. Stability Analysis

Both sub-circuits of the designed voltage reference present feedback loops, and as a result there is a possibility for oscillation and instability.

Sub-circuit A presents a positive feedback loop through M_1 – M_4. However, if the loop gain is less than 1, the stability is guaranteed. In our design, the high value of R_1 guarantees a loop gain of 0.27. In addition, it is desired to design M_4 wider than M_2 [25]. As such, the sizing ratio S_{M4}/S_{M2} was designed to be equal to 10.

Regarding sub-circuit B, capacitor C_1 sets the dominant pole at the gate of M_8 [17]. It also sets the gain-bandwidth product of the feedback loop, ω_{GBW}, given by (35) [17].

$$\omega_{GBW} = [(g_{m9} \cdot g_{m8} \cdot R_2)/C_1] \cdot (S_{M7}/S_{M6}) \quad (35)$$

Therefore, capacitor C_1 should be properly sized in order to maintain a ω_{GBW} well below the value the other remaining poles, which are located at gate of M_6-M_7, and M_9 [17]. In this design, C_1 was set to 5 pF.

C. Supply ramp and start-up circuit

The current source implemented through sub-circuit A has two stable points of operation: (i) I_{R1} = 0 and (ii) $I_{R1} = \Delta V_{BE}/R_1$. Undesirable point (i) can happen if voltages at gate of M_1 and M_2 are near to V_{DD} and zero, respectively. If operation point (i) happens, M_5 will not supply current to M_9 and accordingly, I_{R2} will be equal to zero. In this situation, no voltage reference will be generated.

In order to guarantee the correct operation of sub-circuit A, a start-up circuit composed of M_{12} – M_{14} was added in the circuit as shown in Figure 14. The circuit injects current into device M_2, forcing its gate voltage to increase, and therefore establishing proper operation.

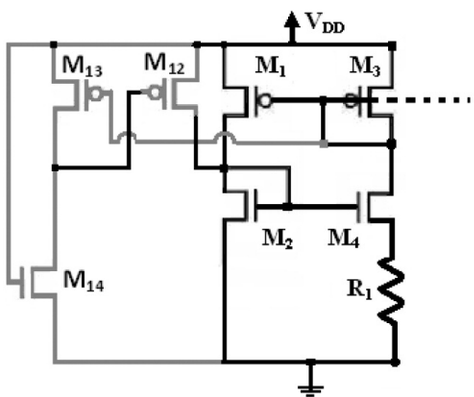

Figure 14. Start-up circuit

If the circuit has no current, the V_{SG} of M_1 will be zero, and consequently M_{13} will be turned off. As M_{14} has V_{GS} higher than V_T, it is always in conduction, which forces the V_{SG} of M_{12} to be V_{DD}. At this moment, M_{12} injects current in the drain of M_2, and thus, increases its gate-source voltage. This allows the circuit to converge to the correct operating point through the feedback formed by the current mirror M_1 and M_3. When the source-gate voltage of M_1 achieves its stable value, M_{13} turns on, which consequently forces V_{SG} of M_{12} to be small enough to turn it off, and thus the start-up circuit is disconnected from the voltage reference.

With the aim of verifying the correct operation point and also the stability of the circuit, it was simulated when applying a V_{DD} ramp. In this simulation, initial voltages conditions equal to 0.7 V and 0 V at the gates of M_1 and M_2, respectively were chosen (the worst case scenario). As can be seen in Figure 15, the reference with start-up circuit brings the reference to the proper operation point after 30 µs. Without the start-up circuit, the reference circuit fails to attain the desired operating point and V_{REF} remains at 0 V.

The transistor aspect ratios of M_{12}-M_{14} are (1µm/5µm), (80µm/0.3µm) and (0.3µm/20µm), respectively. Device M_{13} should have a large value of (W/L), while M_{14} should have a small (W/L) value in order to guarantee that the voltage at the drain of M_{13} is high enough to turn off M_{12} when the reference circuit achieves the correct operation point.

Finally, the methodology presented in this work is not only suitable to the circuit presented in Figure 3, but for any other voltage reference that must operate under low-voltage and low-power requirements using subthreshold MOSFETs. The advantage of the proposed methodology, if compared to the traditional approach – i.e. Equations (13-14) -, is the conscious choice of the MOS inversion level, which allows a quick and reasonably accurate estimation of transistor widths for a given set of design choices (e.g., current, channel length and IC). The estimations provided by Equations (1-12), (17-22) and (27-32) differ from the simulated results by 20%.

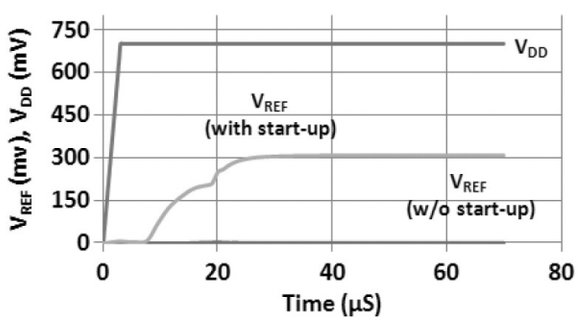

Figure 15. V_{REF} as a function of time

5. CONCLUSIONS

The methodology presented herein, when used as initial design guidance, reduces the design time of low-voltage and low-power voltage references based on subthreshold operation. Simulated results and estimations provided by the proposed equations are in agreement. The voltage reference used here as a case study operates under 0.7 V of supply while consuming only 1.34 µW. It is to be noted that the proposed methodology does not replace electrical simulations to evaluate circuit performance over process variations and layout parasitics, but leads to an optimization of the transistor sizing process in designs operating in weak, moderate or strong inversion.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support from the Emerging Leaders in the Americas Program (ELAP) provided by the government of Canada, Natural Sciences and Engineering Research Council of Canada (NSERC), ReSMIQ, and the Brazilian Funding Agencies CNPq and CAPES.

REFERENCES

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

[2]. R. L. Pinto, A.I.A. Cunha, M.C. Schneider, and C. Galup-Montoro, "An amplifier design methodology derived from a MOSFET current-based model". *In Proceedings of the 1998 International Symposium on Circuit and Systems*, 1998, pp. 301-304.

[3]. D. Stefanovi and M. Kayal, *Structured Analog CMOS Design – Analog Circuits and Signal Processing*. Springer. 2008, pp. 284.

[4]. B. Martin, "Automation comes to analog". *In IEEE Spectrum*. v. 38, no. 6, Jun., 2005. pp.70-75

[5]. A. L. Kundu. et al., "A Methodology for efficient design of analog circuits using an automated simulation based synthesis tool". *In Proceedings of the IEEE International Symposium on Circuits and Systems*. ISCAS'2008. pp. 732 – 735.

[6]. E. Yilmaz and G. Dündar, "Analog Layout Generator for CMOS Circuits". *IEEE Transactions on computer-aided design of integrated circuits and systems*, vol. 28, no. 1, January, 2009.

[7]. D. M. Binkely, "Tradeoffs and optimization in analog CMOS design*". In Proc of the IEEE 14th Int. Conf. Mixed Design of Integrated Circuits and Systems*. MIXDES'07. 47-60.

[8]. D. M. Binkley, *Analog CMOS Design, Tradeoffs and Optimization*. John Wiley and Sons. 2008. pp.632

[9]. Enz, C. Krummenacher, F. and Vittoz, E. A. "An analytical MOS transistor Model valid in all regions of operation and dedicated to low-voltage and low-current applications". *Journal Analog Integrated Circuits and Signal Processing*, vol. 8, 1995, pp.83-114.

[10]. D. M. Colombo, G. Wirth and C. Fayomi, "Design Methodology using Inversion Coefficient for Low-Voltage Low Power CMOS Voltage Reference". *In Proceedings of the 23rd Symposium on Integrated circuits and system design. SBCCI'2010. 2010. pp.43-48.*

[11]. W. Jianping, et al., "A Novel Low-Voltage Low-Power CMOS Voltage Reference Based on Subthreshold MOSFETs". *In Proceedings of the 6th IEEE International Conference on ASIC*. ASICON 2005. pp. 369-373.

[12]. A. I. A. Cunha, M. C. Schneider and C. Galup-Montoro, "An MOS Transistor Model for Analog Circuit Design". *IEEE Journal of Solid-State Circuits*. vol. 33, no. 10, Oct. 1998.

[13]. P. Giacomelli, M. C. Schneider and C. Galup-Montoro, "MOSVIEW: A Graphical Tool for MOS Analog Design". In *Proceedings of the 2003 IEEE International Conference on Microelectronics Systems Education*. 2003, pp.45 - 46.

[14]. Banba, H. et al., "A CMOS Bandgap Reference Circuit with sub-1-V operation", *IEEE Journal of Solid-State Circuits* (May 1999). vol. 34, no. 5, pp. 670-674.

[15]. C. J. Fayomi, G. Wirth, H. Achigui and A. Matsuzawa, "Sub 1 V CMOS bandgap references design techniques: a survey". *Journal Analog Integrated Circuits and Signal Processing,* vol. 62, no. 2, 2010, pp. 141-157.

[16]. Y. Dai, et al., "Threshold voltage based CMOS voltage reference". *In IEEE Proceedings of Circuits, Devices and Systems*. vol. 151, no. 1, 2004, pp. 58-62.

[17]. G. Giustolisi, et al., "A Low-Voltage Low Power Voltage Reference Based on Subthreshold MOSFETs". *IEEE Journal of Solid-State Circuits*, vol. 38, no.1, 2003, pp. 151-154.

[18]. I. M. Filanovsky and A. Allam, "Mutual compensation of mobility and threshold voltage temperature effects with application in CMOS Circuits". *IEEE Transactions on Circuits and Systems*, July, 2001. vol. 48, pp. 876-884.

[19]. W. Liu, Jin, X., Xi, X., et al. BSIM3v3.3 MOSFET Model – User's Manual. Retrieved June 02, 2010 from http://www-device.eecs.berkeley.edu/~bsim3.

[20]. L. H. C. Ferreira, T. C. Pimenta and R. L. Moreno, "A CMOS threshold voltage reference source for very-low-voltage applications". *Microelectronics Journal*, vol. 39, no. 12, pp.1867-1873, 2008.

[21]. Y. Tsividis. *Operation and Modeling of the MOS Transistor*. Oxford University Press, New York, USA. 2º edition. pp. 640. 1999.

[22]. B. Razavi, "*Design of Analog CMOS Integrated Circuits*". McGraw Hill Higher Education, first edition, pp. 704, 2003.

[23]. P. R. Kinget, "Device Mismatch and Tradeoffs in the Design of Analog Circuits". *IEEE Journal of Solid State Circuits*. vol. 40. no.6, 2005, pp.1212 – 1224.

[24]. G. Rincon-Mora, "Voltage references – From Diodes to Precision High-Order Bandgap Circuits". Wiley - IEEE Press. pp. 192. 2001.

[25]. A. S. Sedra and K. C. Smith, "Microelectronics Circuit". Oxford University Press. Fourth edition. pp. 1360. 1997.