The compact all-region MOSFET model: theory and applications

Carlos Galup and Marcio Schneider
Federal University of Santa Catarina (UFSC)
Electrical and Electronics Engineering Department
Florianópolis, SC, Brazil 88040-900
carlosgalup@gmail.com

***Abstract*— A compact presentation of the basic theory of the MOS transistor is given. Instead of the usual approach of furnishing separate analytical formulas for the strong and weak inversion regions of the MOS transistor, we provide simple formulas which are valid in all operating regions, including moderate inversion. We review ultra-low-power circuits operating near the threshold condition that allow the automatic extraction of the specific current IS and the threshold voltage VT of MOS transistors, which are the most fundamental parameters for technology characterization, circuit design, and testing.**

Keywords— all-region MOSFET model, inversion coefficient, inversion level, MOSFET model, parameter extraction

I. INTRODUCTION

Since the 1970s the most popular compact MOSFET models for circuit design have been the strong and the weak inversion models. Interpolation models, such as EKV [1], allow the design in the transition (moderate inversion) region. Regional and interpolation models are no longer acceptable for circuit design in advanced, low voltage technologies, where the moderate inversion region is increasingly important [2]. For this reason, the accurate all-region MOSFET model [3]-[6], strongly based on the MOS transistor theory, will be summarized in this paper. The very simple analytical expression for the transconductance-to-current ratio (g_m/I_D), valid in all the operating regions, is derived in Section II. Based on the g_m/I_D model, the complete dc model is obtained in Section III. Finally, as a fundamental application, the design of circuit for the extraction of both the specific current and the threshold voltage is presented in Section IV.

II. THE g_m/I_D MODEL FOR THE SATURATED TRANSISTOR

The weak inversion (WI) model of the saturated MOS transistor shown below is similar to the expression for the collector current of a bipolar junction transistor (BJT).

$$I_{WI} = I_S e^{V_{GS}/n\phi_t} \Rightarrow I_{WI} = n\phi_t g_m \qquad (1)$$

where ϕ_t is the thermal voltage and n is the slope factor ranging form from 1.1 to 1.5 for bulk technologies. The transconductance g_m is proportional to the transistor current and independent of the geometry of the transistor, as for the BJT.

On the other hand, the transconductance g_m for the strong inversion (SI) quadratic model of the MOSFET in saturation depends not only on the drain current, but also on the geometric ratio W/L, the mobility μ and the oxide capacitance per unit area C'_{ox} .

$$I_{SI} = \frac{1}{2n}\left(\mu C'_{ox}\frac{W}{L}\right)\left(V_{GS}-V_T\right)^2 \Rightarrow I_{SI} = \frac{ng_m^2}{2\mu C'_{ox}\left(W/L\right)} \qquad (2)$$

It is important to remark that the WI current is a linear function of g_m, while the SI current is a quadratic one. Thus, if we add the two expressions, as shown in (3), for $g_m \to 0$ I_{SI} is negligible, while for $g_m \to \infty$ I_{WI} is negligible. Consequently, (3) can be regarded as an interpolation function of the WI and SI asymptotic behaviors of the drain current, but it is more than that. Although introduced heuristically, (3) is one of the main equations of a physics-based all-region MOSFET model [6]. Equation (3) is instrumental for MOSFET amplifier design.

$$I_D = I_{WI} + I_{SI} = ng_m\phi_t\left[1+\frac{g_m}{2\mu C'_{ox}\phi_t\left(W/L\right)}\right] \qquad (3)$$

For a given g_m we have the choice of the operation region depending on the value of W/L. Figure 1 summarizes the importance of (3). Let us assume for the time being that the channel length is fixed. If we reduce (W/L) in order to decrease the area, we pay for this reduction with an increase in current consumption. Equivalently, if we try to save power by reducing the excess current, more silicon area is spent.

Equation (3) can be rewritten as shown below

$$\frac{I_D}{ng_m\phi_t} = 1+\frac{ng_m\phi_t}{4\mu nC'_{ox}\frac{\phi_t^2}{2}\left(W/L\right)} = 1+\frac{ng_m\phi_t i_f}{4I_D} \qquad (4)$$

$$I_S = \mu nC'_{ox}\frac{\phi_t^2}{2}\left(W/L\right) \qquad i_f = \frac{I_D}{I_S} \qquad (5)$$

where I_s is called the specific current and i_f is the forward inversion level or inversion coefficient. Values of i_f greater

This work was supported in part by CNPq and CAPES (research agencies of the Brazilian government)

978-1-5386-4859-9/18/$31.00 ©2018 IEEE

than 100 characterize strong inversion, and the transistor operates in weak inversion up to i_f=1.

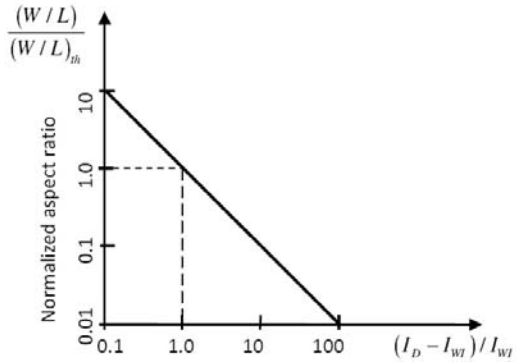

Fig.1 Relation between the normalized aspect ratio and the normalized excess current in a MOSFET amplifier design.

Intermediate values of i_f, from 1 to 100, indicate moderate inversion. Equation (4) is a quadratic equation for g_m/I_D, whose solution is shown below

$$\frac{ng_m\phi_t}{I_D} = \frac{2}{1+\sqrt{1+i_f}} \tag{6}$$

The remarkable simple expression (6) allows MOSFET design in all inversion regions using the inversion level as the key variable.

III. THE I-V MOSFET MODEL

A. I-V Model in Saturation

An important difference between BJTs and MOSFETs is that for the former we have one transconductance while for the latter we must consider at least two transconductances since MOSFETs are four-terminal devices. The gate and the source transconductances are defined as [1]

$$g_m = \frac{\partial I_D}{\partial V_G} \qquad g_{ms} = -\frac{\partial I_D}{\partial V_S} \tag{7}$$

The body effect reduces the gate transconductance with respect to the source transconductance. In fact, the variation in the source voltage contributes directly to the inversion charge variation, whereas the variation in the gate voltage contributes to variations in both the inversion and depletion charges. Thus, the two transconductances are related by [1]

$$g_m = \frac{g_{ms}}{n} \tag{8}$$

The source voltage for which the channel is at threshold is called the pinch-off voltage V_P, which can be approximated by $V_P = (V_{GB} - V_{T0})/n$, where V_{TO} is the threshold voltage at V_{SB} = 0 [1] and, as in (8), the denominator n models the body effect.

As shown in the Appendix, integrating (6) with respect to V_S, we get the all-region dc model (9) for the saturated MOSFET. Threshold occurs for i_f=3. In effect, for i_f=3 $V_P = V_S$.

$$V_P - V_S = \phi_t\left[\sqrt{1+i_f} - 2 + \ln\left(\sqrt{1+i_f} - 1\right)\right] \tag{9}$$

The semi-log plot in Fig. 2 shows the common source characteristics of a saturated MOSFET. The approximate straight line in the low-current region is associated with the exponential WI regime and the curve knee with moderate inversion.

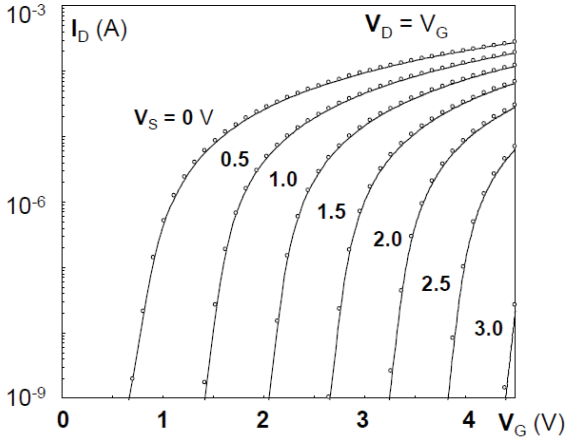

Fig. 2 Common source characteristics of a saturated nMOS transistor: (-) simulated results using (9), and (o) measured data [6].

B. I-V Model in all Regions

Owing to the symmetry of the MOS transistor, the drain current in the triode region can be written as the difference between the forward component $I(V_G,V_S)$ and the reverse component $I(V_G,V_D)$ [1], as shown in Fig. 3 and given below.

$$I_D = I(V_G,V_S) - I(V_G,V_D) = I_F - I_R = I_S(i_f - i_r) \tag{10}$$

where i_r is the reverse inversion level. In saturation $i_r << i_f$ and $I_D \approx I_F = I_S i_f$.

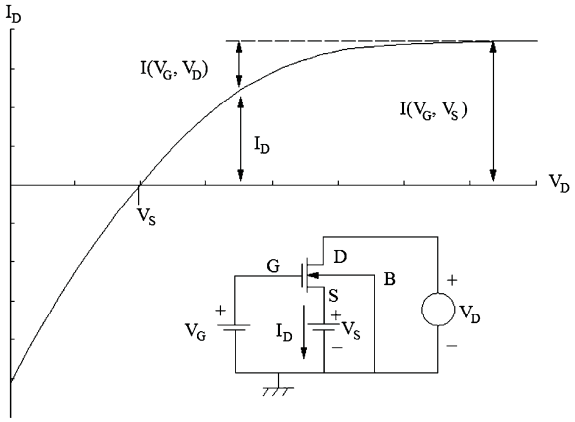

Fig. 3 Output characteristics of a long-channel NMOS transistor at constant V_S and V_G. Voltages are referred to the substrate.

Due to transistor symmetry, the relation between drain voltage and reverse inversion level is obtained substituting V_S by V_D and i_f by i_r in (9). Thus, from (9) we easily obtain expression

(11) for V_{DS} in terms of the inversion levels at source i_f and drain i_r.

$$\frac{V_{DS}}{\phi_t} = \sqrt{1+i_f} - \sqrt{1+i_r} + \ln\left(\frac{\sqrt{1+i_f}-1}{\sqrt{1+i_r}-1}\right) \qquad (11)$$

In Fig. 4 we compare the normalized output characteristics in WI (a), MI (b) and SI (c-f) and the curves obtained using (11).

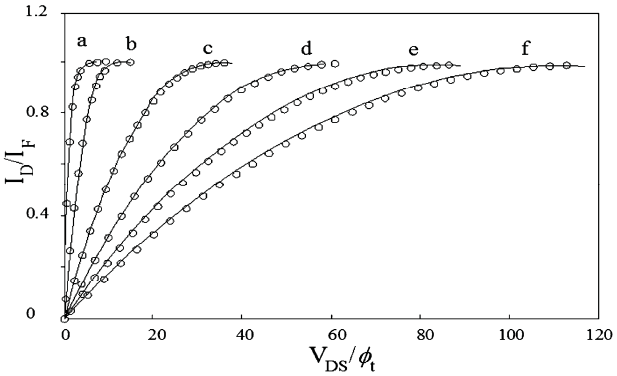

Fig. 4 Normalized output characteristics of an NMOS transistor.
a) i_f= 4.5x 10^{-2}; (b) i_f= 65; (c) i_f= 9.5x10^2 ; (d) i_f= 3.1x 10^3 ; (e) i_f= 6.8x 10^3; (f) i_f= 1.2x 10^4 [6].

IV SPECIFIC CURRENT EXTRACTORS

The self cascode MOSFET (SCM) of Fig. 5 is the core of several self-biased current sources [7]-[10]. In the (SCM) of Fig. 5, M1 is in triode region and M2 is in the saturation region. Thus

$$I_D = I_{S2} i_{f2} = I_{S1}(i_{f1} - i_{r1}) \qquad (12)$$

Since $i_{r1} = i_{f2}$ we can calculate i_{f1} in terms of i_{f2} as

$$i_{f1} = \left(1+\frac{I_{S2}}{I_{S1}}\right) i_{f2} = \left(1+\frac{S_2}{S_1}\right) i_{f2} = \left(1+\frac{W_2 L_1}{W_1 L_2}\right) i_{f2} = \alpha i_{f2} \qquad (13)$$

where α depends only on geometric ratios. Applying (11) to transistor M_1 we get

$$\frac{V_{REF}}{\phi_t} = \sqrt{1+\alpha i_{f2}} - \sqrt{1+i_{f2}} + \ln\left(\frac{\sqrt{1+\alpha i_{f2}}-1}{\sqrt{1+i_{f2}}-1}\right) \qquad (14)$$

If M1 is operates in WI ($i_{f2} << 1$), (14) reduces to

$$\frac{V_{REF}}{\phi_t} \to \ln \alpha \qquad (15)$$

Thus, V_{REF} is proportional to absolute temperature (PTAT) with a slope dependent on the ratio of the aspect ratios of the SCM transistors only [7]- [9].

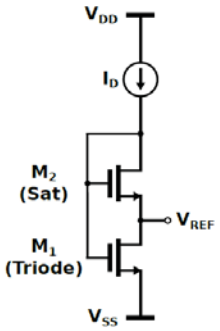

Fig. 5 Self cascode MOSFET (SCM)

In the self-biased current source of Fig. 6 [10] two SCMs operate at the same current, due to the pMOS current mirror, and the operational amplifier connected in closed loop forces the intermediate voltages of both SCMs to be the same. V_{off} represents the sum of the opamp input offset voltage and the effect of the threshold voltage mismatch of the SCM transistors. If V_{off} = 0 and M_1 and M_2 operate in WI, it follows that

$$\ln \alpha_1 = \sqrt{1+\alpha_3 i_{f4}} - \sqrt{1+i_{f4}} + \ln\left(\frac{\sqrt{1+\alpha_3 i_{f4}}-1}{\sqrt{1+i_{f4}}-1}\right) \qquad (16)$$

where $\alpha_1 = 1+\frac{S_2}{S_1}$ $\alpha_3 = 1+\frac{S_4}{S_3}$

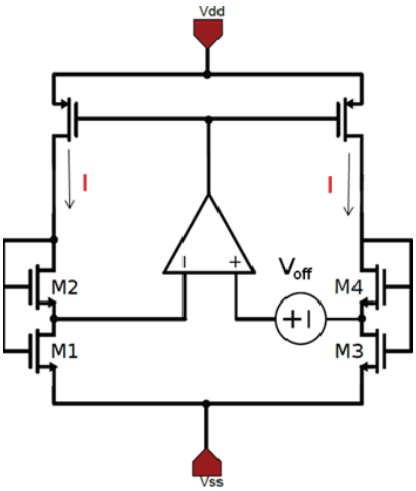

Fig. 6 Modular self-biased specific current extractor [10]

Thus, i_{f4} (the inversion level of M_4) is constant and dependent only on the geometrical ratios α_1 and α_3, and the output reference current $I = I_{S4} i_{f4}$ is proportional to the specific current of M_4. Thus, the circuit can be regarded as a specific current extractor, useful to bias transistors at constant inversion levels, independent of the temperature and technology [7]. By operating all transistors in either weak or moderate inversion, the specific current extractor can be designed as an ultra-low-power circuit [9].

If i_{f4} = 3, the gate voltage of M4 equals V_{T0}; thus, the circuit in Fig. 6 operates also as a threshold voltage extractor.
Figure 7 shows a transistor-only implementation of the current source of Fig. 6. Here, the intermediate branch (M5-M6-M8), which substitutes for the operational amplifier of Fig. 6, imposes the equality of the potentials Vx and Vy,. In effect, if M6 is N times wider than M2, its source voltage equals V_x. On the other hand, if M3 is M times wider than M5, its drain voltage equals V_y. Thus, the series connection of M5 and M6 forces Vx =Vy [7].

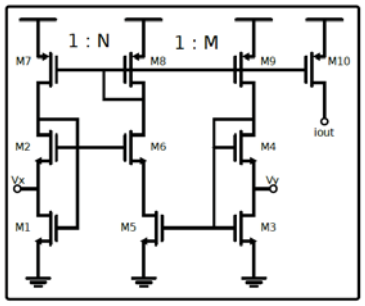

Fig. 7 Specific self-biased current extractor [7]

Since the potentials at the intermediate nodes of the SCMs in Fig. 7 are no longer useful, transistors M1 and M2 can be merged into the single transistor M1, as well as M3 and M4 into M3, as shown in Fig. 8.
To conclude, it is important to remark that many other applications of the compact all-region model are presented in references [6] and [11].

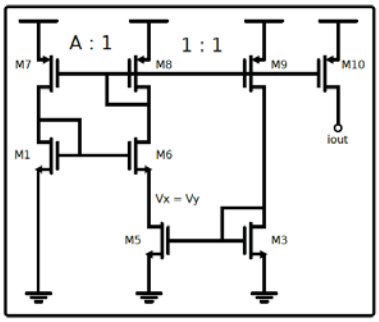

Fig. 8 Specific self-biased current extractor [8]

CONCLUSIONS

Starting from the classical weak and strong inversion models, a compact all-region expression for MOSFET sizing has been obtained, and from this law, the derivation of the accurate MOSFET model valid in weak, moderate and strong inversion has been presented. The use of the forward and reverse inversion levels of the transistor as key variables has been emphasized, allowing the design of circuits with transistors operating in all operating regions (WI, MI, SI), as well as in the triode and saturation regions. As an example, the analysis of a self-biased current source which allows the extraction of the specific current, as well as the threshold voltage, has been presented.

APPENDIX

From (6), (7) and (10), we obtain

$$g_{ms} = -\frac{dI_D}{dV_S} = \frac{2I_S}{\phi_t}\left(\sqrt{1+i_f}-1\right) \qquad \text{(A1)}$$

Integrating (A1)

$$V_P - V_S = \phi_t\left[\sqrt{1+i_f}+\ln\left(\sqrt{1+i_f}-1\right)-\sqrt{1+i_p}-\ln\left(\sqrt{1+i_p}-1\right)\right] \qquad \text{(A2)}$$

It is easy to verify that the choice of i_p = 3 gives the cleanest definition of pinch-off.

REFERENCES

[1] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low voltage and low-current applications," *J. Analog Integr. Circuits Signal Process*., vol. 8, pp. 83-114, July 1995.

[2] J. Watts, C. McAndrew, C. Enz, C. Galup-Montoro, G. Gildenblat, C. Hu, R. van Langevelde, M. Miura-Mattausch, R. Rios, C.-T. Sah, "Advanced compact models for MOSFETs," *Proc. WCM 2005*, pp. 3-12.

[3] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro, "An explicit physical model for the long-channel MOS transistor including small signal parameters," *Solid-State Electronics*, vol. 38, no 11, pp. 1945- 1952, Nov. 1995.

[4] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro, "An MOS transistor model for analog circuit design", *IEEE J. Solid-State Circuits*, vol. 33, no. 10, pp. 1510-1519, Oct. 1998.

[5] C. Galup-Montoro and M. C. Schneider, *MOSFET Modeling for Circuit Analysis and Design*, World Scientific, Singapore, 2007.

[6] M. C. Schneider and C. Galup-Montoro, *CMOS Analog Design Using All-Region MOSFET Modeling*, Cambridge University Press, 2010.

[7] P. Heim, S. Schultz, and M. A. Jabri, "Technology-independent biasing technique for CMOS analogue micropower implementations of neural networks," *Proceedings of the 4th IEEE International Workshop on Cellular Neural Networks and their Applications (CNNA-95)*, pp. 9-12, 1995.

[8] H. J. Oguey and D. Aebischer, "CMOS current reference without resistance," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 7, pp. 1132-1135, July 1997.

[9] E. M. Camacho-Galeano, C. Galup-Montoro and M. C. Schneider, "A 2-nW 1.1-V Self-Biased Current Reference in CMOS Technology", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 2, pp. 61-65, February 2005.

[10] A. F. Fernando Ordoñez Hurtado, "Design methodology of a modular CMOS ultra-low Power self-biased current source", M.S. Thesis, Federal Santa Catarina 2017, available at http://www.lci.ufsc.br/work_master.htlm.

[11] Paul Jespers, *The gm/ID Methodology, a sizing tool for low-voltage analog CMOS Circuits*, Springer 2010.