

Gate current: Modeling, ΔL extraction and impact on RF performance

R. van Langevelde, A.J. Scholten, R. Duffy*, F.N. Cubaynes*, M.J. Knitel and D.B.M. Klaassen

Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

Phone: +31-40-2742418; Fax: +31-40-2744113; E-mail: ronald.van.langevelde@philips.com

* Philips Research Leuven, Kapeldreef 75, B-3001 Leuven, Belgium

Abstract

In this paper a new physical gate leakage model is introduced, which is both accurate and simple. It only uses 5 parameters, making parameter extraction straightforward. As a result the model can be used to extract effective length for modern CMOS technologies. The influence of gate current on the RF performance is studied.

Introduction

With CMOS technology downscaling the increasing gate leakage current seriously starts to affect circuit design. Nevertheless, the compact modeling of gate current is a relatively unexplored field [1-3]. In this paper a new physics-based gate leakage model is introduced, which is at least as accurate as other models [3, 4], but much simpler. It only uses 5 parameters, making parameter extraction easy and straightforward. Furthermore, the model validity range is tested for thin oxides. Next, we show how the gate current can be used to extract ΔL .

Up till now the study of gate current impact on circuit behaviour has been limited to digital considerations, e.g. stand-by power dissipation and charge storage [5]. In this paper we study, for the first time, the impact of gate current on RF performance.

Gate current model

The new gate current model is shown in Figs. 1 and 2. It is part of a new compact MOS model, MOS Model 11, which is based on the explicit formulation of surface potential ψ_s [6, 7]. In the inversion region the intrinsic gate current I_{GC} consists of electrons (NMOS) or holes (PMOS) tunneling from the inversion layer. The current continuity equation ($\partial I_{DS}/\partial x = -W \cdot J_{GC}$) is solved to calculate I_{GC} , under the assumption that I_{GC} only induces a small perturbation of the potential distribution along the channel ($\partial I_{DS}/\partial x \approx 0$). Using eq. (1) and (2), I_{GC} can be solved from eq. (3). Under the same small-perturbation assumption the partitioning of I_{GC} can be simplified to eqs. (4) and (5). In accumulation the intrinsic gate current consists of electrons tunneling from gate to bulk in NMOS or vice-versa in PMOS [3]. Since no channel current is flowing, the solution of I_{GB} is straightforward. The overlap regions are considered as MOS-structures with different flat-band voltage and body factor. The overlap currents consist of electrons (NMOS) or holes (PMOS) tunneling from the accumulation layer in the overlap regions to the gate or vice-versa. Including all the above components, the model gives an accurate

$$\text{Used variables: } Q_{inv} = -C_{ox} \cdot (V_{ox} - k_0 \cdot \sqrt{\psi_s}) \equiv -C_{ox} \cdot V_{inv}$$

$$V_{ox} = V_{GB} - V_{FB} - \psi_p - \psi_s, \quad \psi_p = \left(\sqrt{V_{GB} - V_{FB} - \psi_s + \frac{kp^2}{4}} - \frac{kp}{2} \right)^2$$

$$\text{Channel current: } I_{DS} = -\mu \cdot W \cdot \left[Q_{inv} \cdot \frac{\partial \psi_s}{\partial x} - \frac{kT}{q} \cdot \frac{\partial Q_{inv}}{\partial x} \right] \quad (1)$$

Direct-tunnelling probability [8]:

$$P_{DT}(V_{ox}) \propto \frac{V_{ox}}{I_{ox}} \cdot e^{-\frac{B \cdot I_{ox} \cdot [1 - (1 - V_{ox}/\chi_B)^{3/2}]}{V_{ox}}} \approx \frac{V_{ox}}{I_{ox}} \cdot e^{-\frac{3}{2} \cdot \frac{B \cdot I_{ox}}{\chi_B}} \cdot e^{\frac{3}{8} \cdot \frac{B \cdot I_{ox}}{\chi_B^2} \cdot V_{ox}}$$

Gate-to-channel current in inversion ($V_{GB} > V_{FB}$) [8]:

$$J_{GC} = -Q_{inv} \cdot P_{DT}(V_{ox}) \approx \frac{I_{GINV}}{W \cdot L} \cdot V_{ox} \cdot V_{inv} \cdot e^{B^* \cdot V_{ox}} \quad (2)$$

$$I_{GC} = W \cdot \int_0^L J_{GC} \cdot dx \quad (3)$$

$$\text{Partition: } I_{GS} = W \cdot \int_0^L (1 - x/L) \cdot J_{GC} \cdot dx \quad (4), \quad I_{GD} = I_{GC} - I_{GS} \quad (5)$$

Gate-to-bulk current in accumulation ($V_{GB} < V_{FB}$):

$$J_{GB} = -C_{ox} \cdot V_{ox} \cdot P_{DT}(-V_{ox}), \quad I_{GB} \approx I_{GACC} \cdot V_{ox}^2 \cdot e^{-B^* \cdot V_{ox}} \quad (6)$$

Gate/source or gate/drain overlap current (where $X = S$ or D):

$$V_{oxov} = V_{GX} - V_{FBov} - \psi_{pov} - \psi_{sov}$$

$$\psi_{pov} = \left(\sqrt{V_{GX} - V_{FBov} - \psi_{sov} + \frac{kp^2}{4}} - \frac{kp}{2} \right)^2$$

$$I_{GOV} \approx I_{GOV} \cdot V_{oxov}^2 \cdot (e^{B^* \cdot V_{oxov}} - e^{-B^* \cdot V_{oxov}}) \quad (7)$$

Figure 1: Equations for gate current model. V_{FB} and V_{FBov} denote flat-band voltage of the intrinsic resp. overlap regions, and k_0 and kp denote body factor of the bulk silicon resp. gate polysilicon. B is a physical constant, and χ_B is the oxide potential barrier (both differ for electrons and holes). Parameters I_{GINV} and I_{GACC} scale with gate area ($\propto W \cdot L$), and I_{GOV} scales with width ($\propto W$).

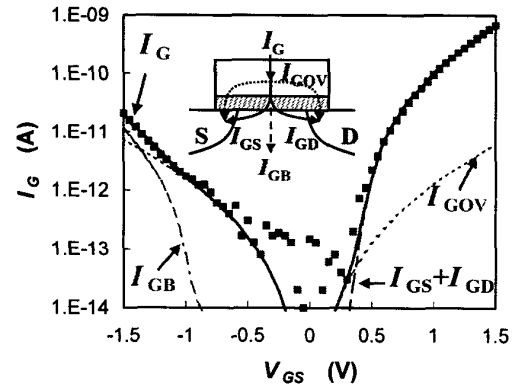


Figure 2: Gate current as a function of gate bias, symbols are measurements and lines are modelled results. All current components, intrinsic (I_{GC} and I_{GB}) and extrinsic (I_{GOV}), as described in Fig. 1 are shown. (NMOS, $W/L = 10\mu\text{m}/0.6\mu\text{m}$, $t_{ox} = 2.2\text{nm}$, $V_{DS} = 0.0\text{V}$)

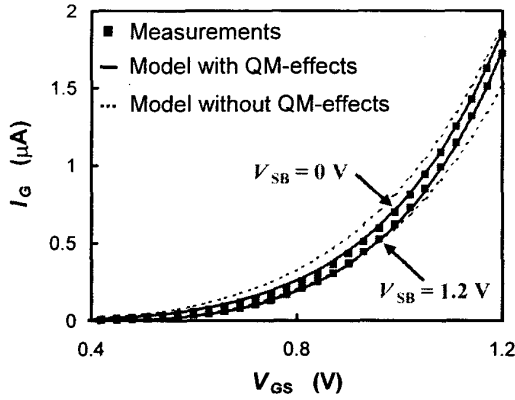


Figure 3: Gate current as a function of V_{GS} for two values of V_{SB} . Quantum-mechanical lowering of potential barrier χ_B is taken into account. (NMOS, $W/L = 10\mu\text{m}/10\mu\text{m}$, $t_{ox} = 1.7\text{nm}$, $V_{DS} = 50\text{mV}$)

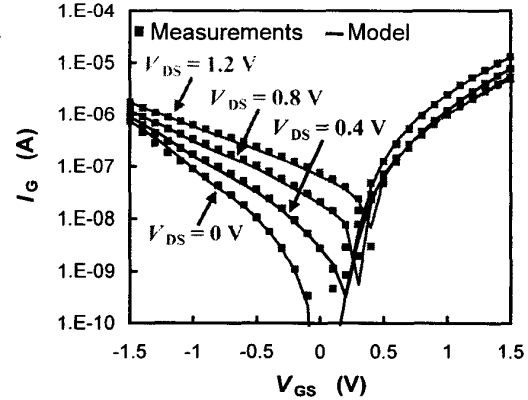


Figure 5: Gate current as a function of gate bias for various values of drain bias. (NMOS, $W/L = 10\mu\text{m}/0.6\mu\text{m}$, $t_{ox} = 1.4\text{nm}$)

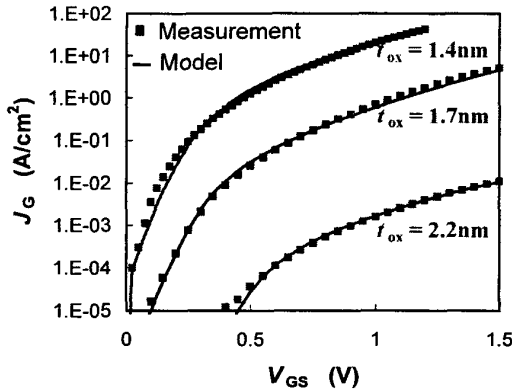


Figure 4: Gate current density as a function of gate bias for different values of oxide thickness. (NMOS, $V_{DS} = 0.0\text{V}$)

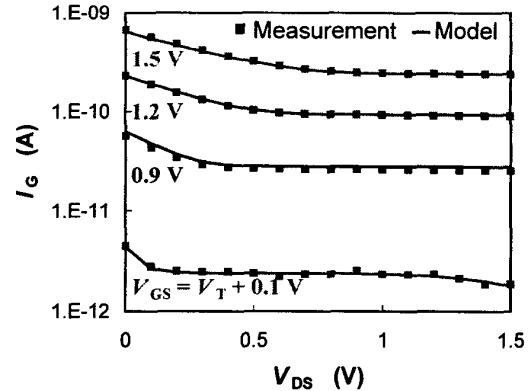


Figure 6: Gate current as a function of drain bias for various values of gate bias. (NMOS, $W/L = 10\mu\text{m}/10\mu\text{m}$, $t_{ox} = 2.2\text{nm}$)

description of I_G (Fig. 2), using only 5 adjustable parameters: I_{GINV} , I_{GACC} , I_{GOV} , B_n^* (for electrons) and B_p^* (for holes).

Carriers at the oxide interface are confined to a narrow potential well, resulting in energy quantization, which affects the surface potential [9] and the effective oxide potential barrier [8]. An effective potential barrier $\chi_B - \Delta\chi_B$ has been introduced (adding no parameters), where $\Delta\chi_B \propto E_{eff}^{2/3}$ and E_{eff} is the effective field. It results in an accurate description of the quantum effects at various V_{SB} (Fig. 3). Similar accuracy is obtained for various t_{ox} (Fig. 4), and for the V_{DS} dependence (Figs. 5 and 6). The I_{GS}/I_{GD} partition cannot be verified using measurements. In Fig. 7, it is therefore verified by breaking down the MOSFET into $N = 10$ equal segments, each described by the above model, similar to [10]. The partition, which follows naturally from the segmentation model, is accurately reproduced by (4) and (5) adding no parameters.

Model validation for thin oxides

The perturbation assumption may not hold true for future technologies where J_{GC} will be large. In Fig. 8, this is in-

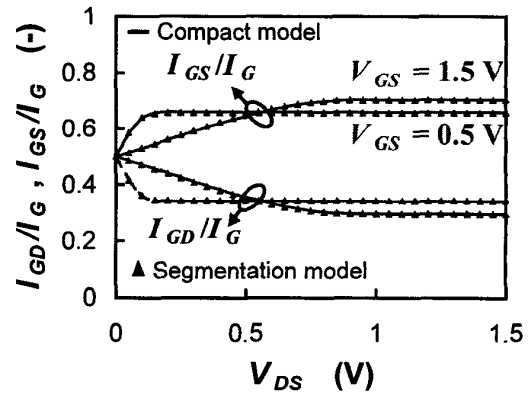


Figure 7: Modelled partition of I_{GS} and I_{GD} current components as a function of drain bias are verified using a segmentation model ($N = 10$). (NMOS, $W/L = 10\mu\text{m}/0.6\mu\text{m}$, $t_{ox} = 2\text{nm}$)

vestigated by comparing the compact model with the segmentation model using parameters scaled according to the ITRS roadmap [11] down to 50nm CMOS for pure SiO_2 gate oxides (worst case). Clearly, the perturbation assumption is valid for

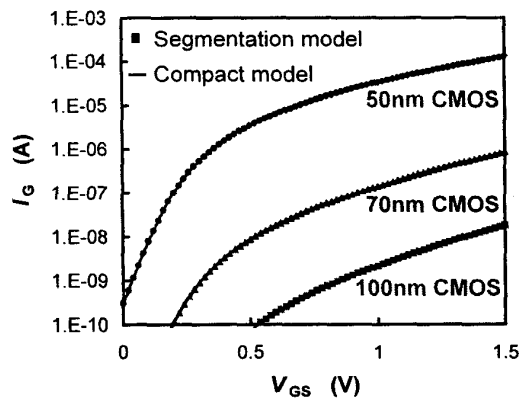


Figure 8: The small perturbation assumption is verified using a $N = 10$ segmentation model for minimum channel length devices in future technologies, following from ITRS-specifications.

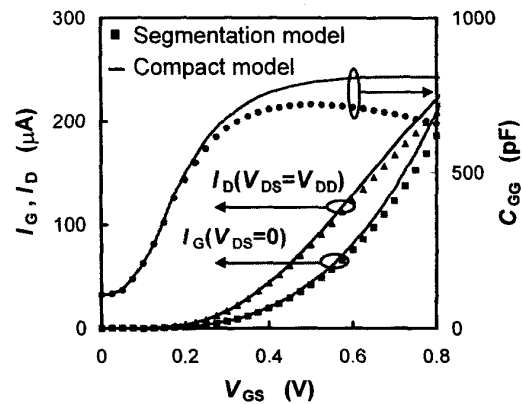


Figure 10: The influence of gate current on the potential distribution, see Fig. 9, results in different electrical behavior: gate capacitance, drain current as well as gate current decrease as compared to the perturbation theory.

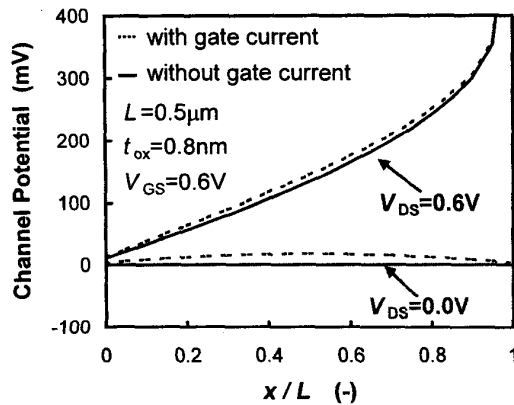


Figure 9: The predicted influence of gate current on the channel potential distribution for an $L = 0.5 \mu\text{m}$ device in 50nm technology. The influence of I_G is no longer negligible, which results in different behavior, see Fig. 10.

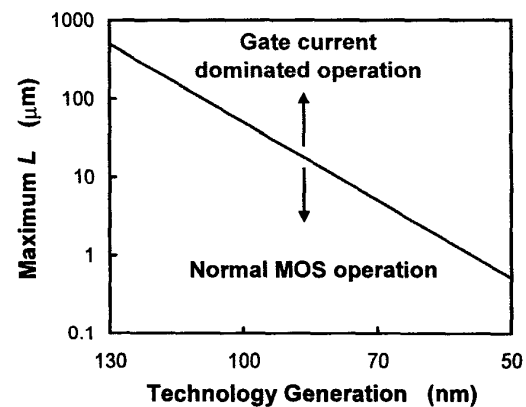


Figure 11: The maximum value of channel length below which the transistor still functions as a MOSFET (i.e. $I_{DS} \gg I_G$) for future CMOS technologies scaled according to the ITRS roadmap [11].

short-channel devices down to the 50nm node. However, for longer channel lengths this may not be the case, see Fig. 9, resulting in a decrease in gate capacitance, I_G and I_D , see Fig. 10. The overall MOSFET behavior is dominated by I_G , and our model is no longer accurate. As a rule of thumb the perturbation assumption breaks down when I_G is in the same order of magnitude as the on-current. The maximum channel length below which the transistor still operates as a MOSFET, is shown for future generations in Fig. 11.

Extraction of ΔL

A realistic value of $\Delta L \equiv L_{\text{drawn}} - L$ is a prerequisite for compact modeling, since it affects I - V , C - V as well as noise characteristics. In addition, a ΔL which is consistently used in both DC and AC is required for statistical simulation [12]. For modern devices with pocket implants, however, the traditional extraction method of ΔL using gain factor β is not accurate anymore [13, 14]. An alternative method using C_{GB} measurements looks promising [14, 15], but the increased I_G in fu-

ture devices interferes with accurate capacitance measurements. However, since both $I_{G\text{INV}}$ and $I_{G\text{ACC}}$ scale with gate area, see Fig. 1, either one can be used to extract ΔL (or ΔW). The accumulation behavior is insensitive to V_T and poly-depletion favoring $I_{G\text{ACC}}$. Plotting the extracted $I_{G\text{ACC}}$ as a function of mask length, ΔL can be determined (Fig. 12). The obtained ΔL corresponds with the one obtained from C_{GB} .

Impact on RF performance

In order to investigate the impact of I_G on RF performance we have simulated the cut-off frequency f_T , the maximum oscillation frequency f_{max} , the input impedance, the minimum noise figure (taking into account shot noise in the gate leakage), linearity and the matching behavior for future short-channel devices. It was found that f_T , f_{max} as well as the linearity figure are not affected by I_G down to 50nm. The input impedance, however, changes significantly with increasing I_G , see Fig. 13, resulting in an increase of the input resistance, which may facilitate impedance matching for RF design. This increased in-

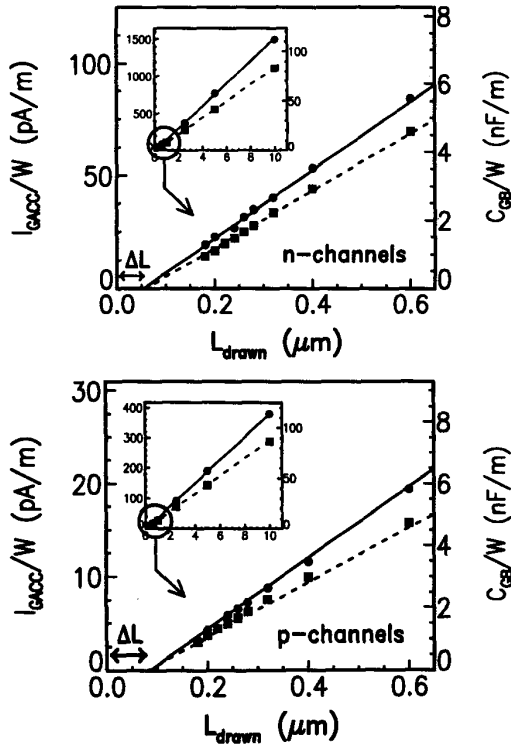


Figure 12: Extraction of ΔL using capacitance C_{GB} in accumulation: squares are measured results ($|V_{GS}| = 1.8V$, $V_{DS} = 0V$) and dashed line is scaling rule. The same value of ΔL can be extracted by optimizing gate current model for different mask lengths and using the results for extracted parameter I_{GACC} : bullets are extracted results and solid line is scaling rule. ($0.18\mu m$ CMOS, $t_{ox} = 3.2nm$)

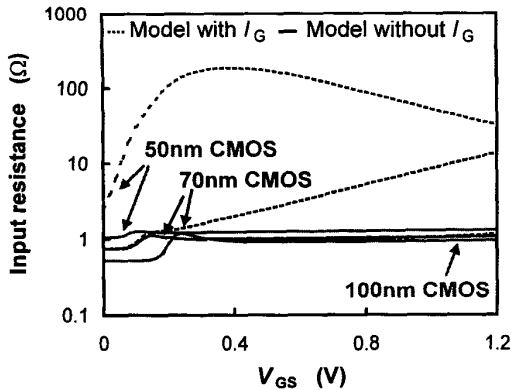


Figure 13: Predicted input resistance, $Re(1/Y_{GG})$, as a function of gate bias for short-channel devices in future CMOS technologies as calculated from segmentation model. The influence of gate current is shown. (NMOS, $W = 60 \times 3\mu m$, $V_{DS} = V_{DD}$, $f = 2GHz$)

put resistance furthermore results in an increase in minimum noise figure, which is significant at $f = 2GHz$ and decreases with increasing frequency. The impact of I_G on matching is studied by looking at the influence of t_{ox} variations on the on-

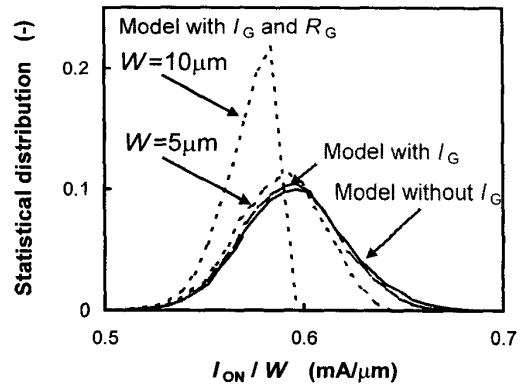


Figure 14: Statistical analysis of the on-current of a 50nm n -MOS transistor assuming only oxide thickness variations ($\sigma = 5\%$). The influence of gate current (solid lines) and polysilicon gate resistance ($R_{poly} = 10\Omega/\square$) for two channel widths (dashed lines) are investigated.

current, see Fig. 14. The spread in I_{on} only changes slightly due to I_G for the 50nm device. However, gate resistance R_G may lead to an effective gate bias drop resulting in a decrease of I_D and I_G [16], thus affecting matching behavior considerably, see Fig. 14. This can be prevented by minimizing R_G using multi-finger layouts [17].

Conclusions

A new model for gate current has been developed suitable for compact MOS models. It includes physical partitioning and quantum-mechanical quantization effects. Using only 5 parameters, the model gives an accurate description over the whole operation region, and scales well with geometry and oxide thickness. The model can furthermore be used to extract ΔL for modern CMOS technologies where conventional extraction methods fail. It has been found that, except for the noise figure, most of the RF performance of short-channel devices will be unaffected by the increased gate current with CMOS downscaling. Furthermore, if care is taken to minimize gate resistance, the matching properties will not change significantly.

References

- [1] P. O'Sullivan *et al.*, *ESSDERC'99*, p.488
- [2] C.-H. Choi *et al.*, *IEDM'99*, p.735
- [3] K.M. Cao *et al.*, *IEDM'00*, p.815
- [4] www-device.eecs.berkeley.edu/~bsim3
- [5] P.J. Wright *et al.*, *IEEE Trans. El. Dev.*, 37, p.1884, 1990
- [6] www.semiconductors.philips.com/Philips_Models
- [7] R. van Langevelde *et al.*, *ESSDERC'01*, p.81
- [8] Z.A. Weinberg, *J. Appl. Phys.*, 53, p.5052, 1982
- [9] R. Rios *et al.*, *IEDM'95*, pp.937
- [10] A.J. Scholten *et al.*, *IEDM'99*, p.163
- [11] www.itrs.net/1999_SIA_Roadmap
- [12] M. van Dort *et al.*, *IEDM'95*, p.941
- [13] H. van Meer *et al.*, *IEEE El. Dev. Lett.*, 21, p.133, 2000
- [14] A.J. Scholten *et al.*, *ESSDERC'01*, p.311
- [15] T.S. Hsieh *et al.*, *ICMTS'01*, p.263
- [16] M. Koh *et al.*, *IEDM'98*, p.919
- [17] L.F. Tiemeijer *et al.*, *IEDM'01*