Nat.Lab. Unclassified Report 2003/00239

Date of issue: April 2003

Physical Background of MOS Model 11

Level 1101

R. van Langevelde, A.J. Scholten and D.B.M. Klaassen

Authors' address data: R. van Langevelde WAY41; Ronald.van.Langevelde@philips.com
A.J. Scholten WAY41; Andries.Scholten@philips.com
D.B.M. Klaassen WAY41; D.B.M.Klaassen@philips.com

Unclassified Report: 2003/00239

Title: Physical Background of MOS Model 11
Level 1101

Author(s): R. van Langevelde, A.J. Scholten and D.B.M. Klaassen

Part of project: Compact modelling

Customer: Philips Semiconductors

Keywords: MOS Model 11, compact modelling, MOSFET, CMOS, circuit simulation, integrated circuits

Abstract: A new compact model for MOS transistors has been developed, MOS Model 11 (MM11), the successor to MOS Model 9. MM11 not only gives an accurate description of charges and currents and their first-order derivatives (transconductance, conductance, capacitances), but also of their higher-order derivatives. In other words, it gives an accurate description of MOSFET distortion behaviour, and as such MM11 is suitable for digital, analog as well as RF circuit design.
MOS Model 11 is a symmetrical, surface-potential-based model. It includes an accurate description of all physical effects important for modern and future CMOS technologies, such as, e.g., gate tunnelling current, gate-induced drain leakage, influence of pocket implants, poly-depletion, quantum-mechanical effects and bias-dependent overlap capacitances.
Including all of the above, MOS Model 11 achieves superior accuracy as compared to MOS Model 9 without an increase in the number of parameters and with only a slight increase in simulation time.

The goal of this report is to present the complete physical background of MM11, Level 1101, including a description of steady-state currents, charges and noise sources. This report gives a more in-depth description and derivation of the model equations as used in the report NL-UR 2002/802.

List of Symbols

This list is not a complete list of the symbols used in this report, it excludes those symbols which are only used locally in a particular chapter. In the following, subscripts i and j denote one of the MOSFET terminals, i.e., drain (D), gate (G), source (S) or bulk (B).

Symbol	Unit	Description
A_{GIDL}	AV^{-3}	Gain factor for gate-induced drain leakage current, see Section 4.2
B_{GIDL}	V	Probability factor for gate-induced drain leakage current, see Section 4.2
B_{acc}	V	Probability factor for gate tunnelling current in accumulation, see Section 5
B_{inv}	V	Probability factor for gate tunnelling current in inversion, see Section 5
C_{GIDL}	−	Factor for the lateral field dependence of the gate-induced drain leakage current, see Section 4.2
C_{GDO}	F	Oxide capacitance for the gate–drain overlap, see Section 6.2
C_{GSO}	F	Oxide capacitance for the gate–source overlap, see Section 6.2
C_{OX}	F	Total gate oxide capacitance, $C_{\mathrm{OX}} = C_{\mathrm{ox}} \cdot W \cdot L$
C_{b}	$\mathrm{F/m}^2$	Linearisation factor of Q_{b}, $C_{\mathrm{b}} = -\partial Q_{\mathrm{b}}/\partial \psi_{\mathrm{s}}\vert_{\psi_{\mathrm{s}}=\bar{\psi}}$
C_{g}	$\mathrm{F/m}^2$	Linearisation factor of Q_{g}, $C_{\mathrm{g}} = -\partial Q_{\mathrm{g}}/\partial \psi_{\mathrm{s}}\big\vert_{\psi_{\mathrm{s}}=\bar{\psi}}$
C_{ij}	F	Transcapacitance between node i and j, see eq. (6.5)
C_{inv}	$\mathrm{F/m}^2$	Linearisation factor of Q_{inv}, $C_{\mathrm{inv}} = -\partial Q_{\mathrm{inv}}/\partial \psi_{\mathrm{s}}\vert_{\psi_{\mathrm{s}}=\bar{\psi}}$
C_{ox}	$\mathrm{F/m}^2$	Gate oxide capacitance per unit area, $C_{\mathrm{ox}} = \epsilon_{\mathrm{ox}}/t_{\mathrm{ox}}$
$\mathcal{E}_{\mathrm{C}}$	eV	Energy level of lower edge of conduction band
$\mathcal{E}_{\mathrm{F}}$	eV	Fermi energy level
$\mathcal{E}_{\mathrm{V}}$	eV	Energy level of upper edge of valence band
$\mathcal{E}_{\mathrm{g}}$	eV	Energy bandgap between $\mathcal{E}_{\mathrm{C}}$ and $\mathcal{E}_{\mathrm{V}}$ in silicon, $\mathcal{E}_{\mathrm{g}} = 1.12\mathrm{eV}$ at $T = 300\mathrm{K}$
E_{Si}	V/m	Transversal electric field at $\mathrm{Si/SiO_2}$-interface, $E_{\mathrm{Si}} = Q_{\mathrm{g}}/\epsilon_{\mathrm{Si}}$
E_{eff}	V/m	Effective transversal electric field, see Section 3.3.1
E_{ox}	V/m	Transversal electric field in gate oxide, $E_{\mathrm{ox}} = Q_{\mathrm{g}}/\epsilon_{\mathrm{ox}}$
E_{x} , E_{y}	V/m	Lateral ($E_{\mathrm{x}} = -\partial\psi/\partial x$) and transversal ($E_{\mathrm{y}} = -\partial\psi/\partial y$) electric field
G_{R}	−	Expression for series resistance, see Section 3.4
G_{Th}	−	Expression for self-heating, see Section 3.6.3
G_{mob}	−	Expression for mobility reduction, see Section 3.3.1
G_{vsat}	−	Expression for velocity saturation, see Section 3.3.2
G_{tot}	−	Expression for mobility reduction, velocity saturation, series resistance, channel length modulation and self-heating, see Sections 3.3-3.6
$G_{\Delta\mathrm{L}}$	−	Expression for channel length modulation, see Section 3.6.1
I_{GINV}	AV^{-2}	Gain factor for intrinsic gate tunnelling current in inversion, see Section 5.1
I_{GACC}	AV^{-2}	Gain factor for intrinsic gate tunnelling current in accumulation, see Section 5.2
I_{GOV}	AV^{-2}	Gain factor for source/drain overlap gate tunnelling current, see Section 5.3

Symbol	Unit	Description
$I_{G_{ov}}$	A	Gate current in gate-source (I_{Gov_0}) or gate-drain (I_{Gov_L}) overlap region, see Section 5.3
I_{drift} , I_{diff}	A	Drift or diffusion component of I_{DS}, $I_{DS} = I_{drift} + I_{diff}$
I_i	A	Current flowing into node i
I_{ij}	A	Current flowing from node i to node j
L	m	Effective channel length
L_{mask}	m	Mask or drawn channel length
N_{FA}	$V^{-1}m^{-4}$	First coefficient of the flicker noise, see Section 7.1
N_{FB}	$V^{-1}m^{-2}$	Second coefficient of the flicker noise, see Section 7.1
N_{FC}	V^{-1}	Third coefficient of the flicker noise, see Section 7.1
N_A	m^{-3}	Impurity concentration in bulk
N_{OV}	m^{-3}	Effective impurity concentration in gate-overlapped drain or source extension
N_P	m^{-3}	Impurity concentration in polysilicon gate
N_T	J	Coefficient of the thermal noise, $N_T = 4 \cdot k_B \cdot T$
Q_b	C/m^2	Bulk charge density formed by accumulation or depletion layer
$\bar{Q}_b$	C/m^2	Average bulk charge density, $\bar{Q}_b = Q_b(\psi_s = \bar{\psi})$
Q_g	C/m^2	Gate charge density
$\bar{Q}_g$	C/m^2	Average gate charge density, $\bar{Q}_g = Q_g(\psi_s = \bar{\psi})$
Q_{inv}	C/m^2	Inversion-layer charge density
Q_{inv_0}, Q_{inv_L}	C/m^2	Inversion-layer charge density at source or drain side
$\bar{Q}_{inv}$	C/m^2	Average inversion-layer charge density, $\bar{Q}_{inv} = Q_{inv}(\psi_s = \bar{\psi})$
Q^*_{inv}	C/m^2	Effective inversion-layer charge density, $Q^*_{inv} = Q_{inv} + \phi_T \cdot C_{inv}$
$\bar{Q}^*_{inv}$	C/m^2	Average effective inversion-layer charge density, $\bar{Q}^*_{inv} = Q^*_{inv}(\psi_s = \bar{\psi})$
Q_j	C	Total nodal charge of node j
Q_{ov}	C/m^2	Charge density in gate-overlapped source/drain extension
Q_{ov_0}, Q_{ov_L}	C/m^2	Charge density in gate-overlapped source or drain extension
R_S	Ω	Series resistance
R_{S_0}	Ω	Bias-independent part of series resistance, see Section 3.4
S_{I_D}, S_{I_G}	A^2/Hz	Noise spectral density of drain (S_{I_D}) and gate (S_{I_G}) current
T	K	Temperature in channel region in MOS transistor
V	V	Electron quasi-Fermi potential
$V_{DS_{sat}}$	V	Drain-source saturation voltage, see Appendix H
$V_{DS_{sat\infty}}$	V	Drain-source saturation voltage for long-channel device, $V_{DS_{sat\infty}} = \psi_{sat} - \phi_B - V_{SB}$
V_{DS_x}	V	Ohmic/saturation smoothing function, see Appendix A.2
V_{FB}	V	Flat-band voltage

Symbol	Unit	Description
V_{FBov}	V	Flat-band voltage of gate-overlapped source/drain extension
V^*_{GB}	V	Effective gate-bulk bias, $V^*_{\mathrm{GB}} = V_{\mathrm{GB}} - V_{\mathrm{FB}}$
V_{GX}	V	Gate-source ($X = S$) or gate-drain ($X = D$) voltage
V_{P}	V	Characteristic voltage of channel length modulation, see Section 3.6.1
V_{T}	V	Threshold voltage
V_{ij}	V	Voltage between terminal i and j
V_{ov}	V	Oxide voltage in gate-overlapped source/drain extension, $V_{\mathrm{ov}} = -Q_{\mathrm{ov}}/C_{\mathrm{ox}}$
$V_{\mathrm{ov_0}}, V_{\mathrm{ov_L}}$	V	Oxide voltage in gate-overlapped source or drain extension
V_{ox}	V	Oxide voltage in intrinsic channel region, $V_{\mathrm{ox}} = Q_{\mathrm{g}}/C_{\mathrm{ox}}$
W	m	Effective channel width
W_{mask}	m	Mask or drawn channel width
a_1	–	Factor of the weak–avalanche current, see Section 4.1
a_2	V	Exponent of the weak–avalanche current, see Section 4.1
a_3	–	Factor of the drain–source voltage above which weak–avalanche occurs, see Section 4.1
f	Hz	Frequency
g	A·m/V	Channel conductance, $g = \mu \cdot W \cdot Q_{\mathrm{inv}}$
g_{ds}	A/V	Output conductance, $g_{\mathrm{ds}} = \partial I_{\mathrm{D}}/\partial V_{\mathrm{DS}}$
g_{m}	A/V	Transconductance, $g_{\mathrm{m}} = \partial I_{\mathrm{D}}/\partial V_{\mathrm{GS}}$
g_{mb}	A/V	Substrate transconductance, $g_{\mathrm{mb}} = \partial I_{\mathrm{D}}/\partial V_{\mathrm{BS}}$
$\hbar$	J·s	Reduced Planck constant, $\hbar = 1.05458 \cdot 10^{-34}$J·s
k_{B}	J/K	Boltzmann constant, $k_{\mathrm{B}} = 1.3806226 \cdot 10^{-23}$J/K
k_0	$\mathrm{V}^{1/2}$	Body-effect factor of bulk, $k_0 = \sqrt{2 \cdot q \cdot \epsilon_{\mathrm{Si}} \cdot N_{\mathrm{A}}}/C_{\mathrm{ox}}$
k_{ov}	$\mathrm{V}^{1/2}$	Body-effect factor of source/drain extension, $k_{\mathrm{ov}} = \sqrt{2 \cdot q \cdot \epsilon_{\mathrm{Si}} \cdot N_{\mathrm{OV}}}/C_{\mathrm{ox}}$
k_{P}	$\mathrm{V}^{1/2}$	Body-effect factor of polysilicon gate, $k_{\mathrm{P}} = \sqrt{2 \cdot q \cdot \epsilon_{\mathrm{Si}} \cdot N_{\mathrm{P}}}/C_{\mathrm{ox}}$
m_0	–	Parameter for (short-channel) subthreshold slope, $m_0 = m_{\mathrm{S}} - 1$
m_{S}	–	Parameter for short-channel subthreshold slope, $m_{\mathrm{S}} = 1 + m_0$
n	m^{-3}	Free electron concentration, $n \approx N_{\mathrm{A}} \cdot \exp((\psi - \phi_{\mathrm{B}} - V)/\phi_{\mathrm{T}})$
n_{i}	m^{-3}	Intrinsic carrier concentration, $n_{\mathrm{i}} = 1.45 \cdot 10^{+16}\mathrm{m}^{-3}$ at $T = 300$K
p	m^{-3}	Free hole concentration, $p \approx N_{\mathrm{A}} \cdot \exp(-\psi/\phi_{\mathrm{T}})$
q	C	Elementary unit charge, $q = 1.6021918 \cdot 10^{-19}$C
t_{ox}	m	Gate oxide thickness, see Fig. 2.1
v	m/s	Carrier velocity
v_{sat}	m/s	Saturation velocity
x , y , z	m	Lateral, transversal and orthogonal coordinates, see Fig. 2.1

Symbol	Unit	Description
Δ_{acc}	V	Approximate function for impact of holes in the inversion region, see eq. (2.14)
ΔQ_{inv}	C/m^2	Inversion-layer charge density difference, $\Delta Q_{inv} = Q_{inv_L} - Q_{s_0}$
$\Delta V_{G_{dibl}}$	V	Increase in effective gate bias due to DIBL, see Section 3.2
$\Delta V_{G_{sf}}$	V	Increase in effective gate bias due to static feedback, see Section 3.6.2
$\Delta\psi$	V	Surface potential difference, $\Delta\psi = \psi_{s_L} - \psi_{s_0}$
α	–	Factor of channel length modulation, see Section 3.6.1
β	A/V^2	Gain factor, $\beta = \mu_0 \cdot C_{ox} \cdot W/L$
ϵ_{Si}	F/m	Dielectric permittivity of silicon, $\epsilon_{Si} = 1.0447720 \cdot 10^{-10}$F/m
ϵ_{ox}	F/m	Dielectric permittivity of SiO_2, $\epsilon_{ox} = 3.4531438 \cdot 10^{-11}$F/m
η_{mob}	–	Effective field parameter for dependence on Q_{inv} and Q_{dep}, see Section 3.3.1
θ_R	1/V	Coefficient of series resistance, $\theta_R = 2 \cdot \beta \cdot R_{S_0}$
θ_{R1}	V	Numerator of gate bias dependent part of R_S, see Section 3.4
θ_{R2}	V	Denominator of gate bias dependent part of R_S, see Section 3.4
θ_{Th}	V^{-3}	Coefficient of self-heating, see Section 3.6.3
θ_{ph}	V^{-1}	Coefficient of mobility reduction due to phonon scattering, see Section 3.3.1
θ_{sr}	V^{-1}	Coefficient of mobility reduction due to surface roughness scattering, see Section 3.3.1
θ_{sat}	1/V	Velocity saturation parameter, $\theta_{sat} = \mu_0/(v_{sat} \cdot L)$
μ	m^2/V·s	Carrier mobility
μ_0	m^2/V·s	Low-field bulk mobility
ν	–	Exponent of field dependence of mobility model, see Section 3.3.1
ϕ_B	V	Surface potential at onset of strong inversion, $\phi_B = 2 \cdot \phi_F$
ϕ_F	V	Intrinsic Fermi potential, $\phi_F = \phi_T \cdot \ln(N_A/n_i)$
ϕ_T	V	Thermal voltage, $\phi_T = k_B \cdot T/q$
ψ	V	Electrostatic potential with respect to neutral bulk
$\bar{\psi}$	V	Average surface potential, $\bar{\psi} = (\psi_{s_L} + \psi_{s_0})/2$
ψ_p	V	Surface potential in polysilicon gate, see Fig. 2.3
ψ_s	V	Surface potential, see Fig. 2.3
ψ_{s_0} , ψ_{s_L}	V	Surface potential at source or drain side
$\psi_{s_{ov}}$	V	Surface potential in gate overlap region, see Fig. 2.6
ψ_{sat}	V	Surface potential in weak inversion, see eq. (3.13)
σ_{dibl}	V$^{-1/2}$	Drain-induced barrier-lowering parameter, see Section 3.2
σ_{sf}	V$^{-1/2}$	Static-feedback parameter, see Section 3.6.2
ω	rad/s	Angular frequency, $\omega = 2 \cdot \pi \cdot f$

Contents

1 Introduction

Since its introduction in the 1960s, MOS transistor technology has been subject to a ceaseless decrease in transistor dimensions and resulting progress in performance. The decrease in transistor dimensions has allowed for an exponential increase with time in the number of components per chip and in operation speed. Furthermore, the downscaling of MOS technology has led to the on-going integration of many different functions, both digital and analogue, on a single chip. Nowadays, the realization of very complex and high-speed circuits in CMOS technology is the order of the day.
The growth in circuit complexity has made indispensable the use of computer-aided simulation tools which allow the circuit designer to predict and optimize circuit behaviour before the circuit is realized in silicon, so-called *circuit simulators*. Circuit simulators contain mathematical models for the quantitative description of the terminal behaviour of the circuit elements as a function of bias conditions, temperature, and device geometry. The above models are often referred to as *compact models*, and they are a critical link in the translation of CMOS process properties into IC performance. Evidently, the benefit of a circuit simulator depends heavily on the accuracy of the compact models used, on the other hand the models should be as simple as possible in order to limit circuit simulation time. Continuity, accuracy, scalability, and simulation performance are basic requirements for such a compact MOSFET model.
Generally three different types of compact models can be distinguished: physics-based models, empirical models and table look-up models. In the first type, the model equations are analytical expressions which have been derived directly from device physics. In the second type, the relations are of a curve fitting nature, and in the last type, the characteristics are reconstructed via tables of measured data. The latter two types offer little physical insight in contrast to the physics-based model, which makes use of physical significant parameters that allow for statistical modelling and that obey well-defined geometrical and temperature scaling rules. As a consequence, most compact MOSFET models in existing circuit simulators are physics-based models. These models include MOS Model 9 (MM9) [1], the BSIM4 model [2] and the EKV model [3].
Within Philips, MM9 has been the compact model of choice since the early 1990s. However, with the continuous downscaling of CMOS technologies, the demands on compact MOS models have become more and more stringent. Modern CMOS technologies are not only suitable for digital and analogue but also for RF applications. A compact model should thus be accurate for RF as well as for digital and analogue design. In addition, new physical effects come into prominence with technology downscaling. The model should accurately describe all the important physical effects of modern and future technologies. Unfortunately all existing models (including MM9) fail to satisfy the above demands, and as a result, a new compact MOS model has been developed over the past years, called MOS Model 11 (MM11) [4]-[17], the successor to MM9. The new MM11 fulfills all of the demands for advanced compact MOS models as discussed in the following Sections.

1.1 V_{T}-Based versus ψ_{s}-Based Models

In MOS modelling, the transition region between subthreshold and superthreshold behaviour, often referred to as the *moderate inversion region*, has traditionally received little attention. As the supply voltage is scaled down to lower values with CMOS downscaling, however, this region becomes an increasingly larger fraction of the overall logic swing in digital circuits. Furthermore, in analogue and RF design, MOSFETs are typically biased in this operation region around threshold. An accurate and physical description of the moderate inversion region is thus essential.
Conventional models such as, e.g., BSIM4 and MM9 are based on the formulation of threshold voltage V_{T} [1, 2, 18]. These so-called V_{T}-based models have been developed using regional approximations which are joined together by suitable smoothing functions over the moderate inversion region,

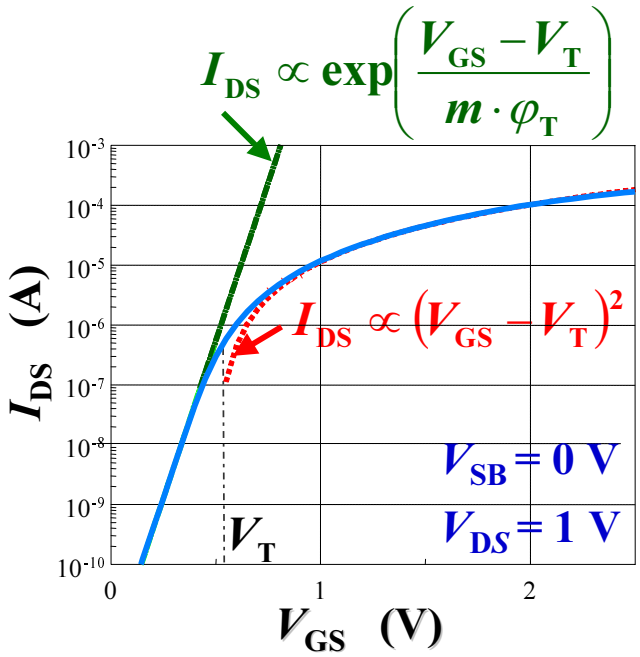


Figure 1.1: V_T-based models such as BSIM4 and MM9 make use of regional approximations for the drain-source channel current I_{DS} in the subthreshold ($V_{GS} < V_T$) and superthreshold region ($V_{GS} > V_T$) (dashed lines). These approximations break down around threshold (i.e., $V_{GS} = V_T$). In order to circumvent this problem, the approximations are joined together by suitable smoothing functions over the transition region. These mathematical smoothing functions have no physical basis.

see Fig. 1.1. With the increasing importance of the moderate inversion region, this type of model relies more on the mathematical smoothing functions and less on the physics-based approximations in the subthreshold and the superthreshold region. This has led to the current trend of increasing model complexity and an increasing number of parameters. Some newer models are not based on V_T formulations but on inversion charge Q_{inv} formulations [3, 19, 20], so-called Q_{inv}-based models. These models, nevertheless, still use a more or less empirical description of the moderate inversion region.

Over the years, in an attempt to increase the physical content, especially in the moderate inversion region, the focus has gradually shifted from V_T-based (and Q_{inv}-based) models to charge-sheet models based on the formulation of surface potential ψ_s [21]-[31]. These so-called ψ_s-based models are inherently single-piece and give a physics-based and accurate description in all operation regions. Unfortunately, these models require an iterative solution of surface potential, which is generally considered to be computationally expensive. This is undesirable for circuit simulation, and as such it is a drawback of ψ_s-based models. Nowadays, nevertheless, it is felt that the extra computation time is worth the return in accuracy. New models in the public domain, still under development, such as the HiSIM model [22]-[25] and the SP model [27]-[31] are ψ_s-based models.

MM11 is based on ψ_s formulations. In addition, MM11 makes use of a suitable explicit approximation of ψ_s which reduces computation time and still preserves accuracy [9].

1.2 Demands for Analogue and RF Design

Analogue and RF design typically require an accurate description of not only currents and capacitances but also of the small-signal behaviour, the noise behaviour and the distortion behaviour. Espe-

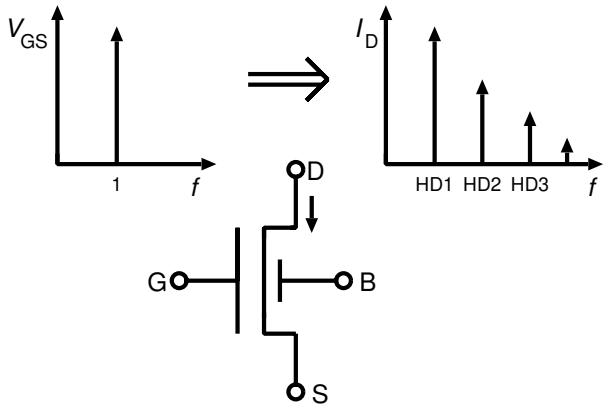
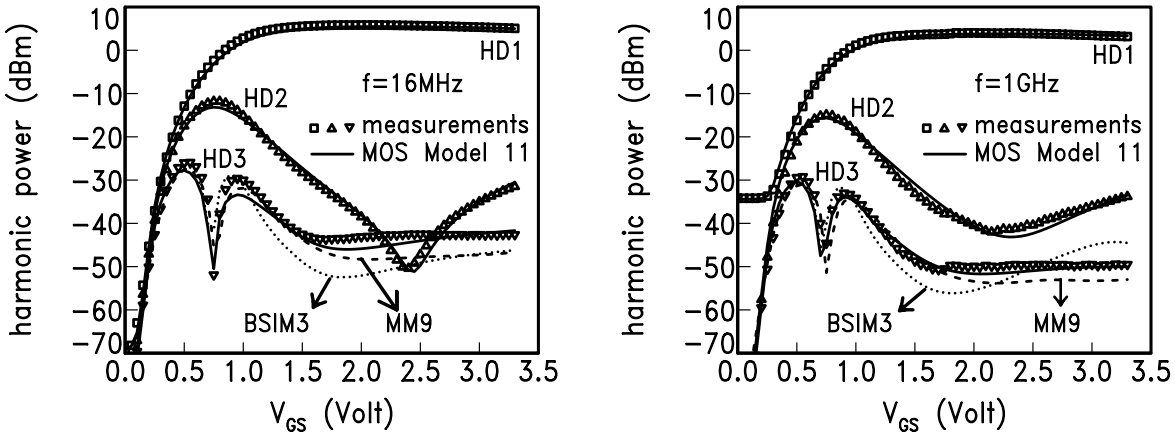


Figure 1.2: In a typical amplifier structure, a sinusoidal input signal V_{GS} results in a distorted output signal I_D containing the ground harmonic (HD1) as well as undesired higher-order harmonics (HD2, HD3,. . .) due to the non-linear dependence of I_D on V_{GS}.

cially the latter has received little attention, and is thus a special point of interest. Another point of interest for analogue and RF design is the drain-source symmetry of the MOS model.

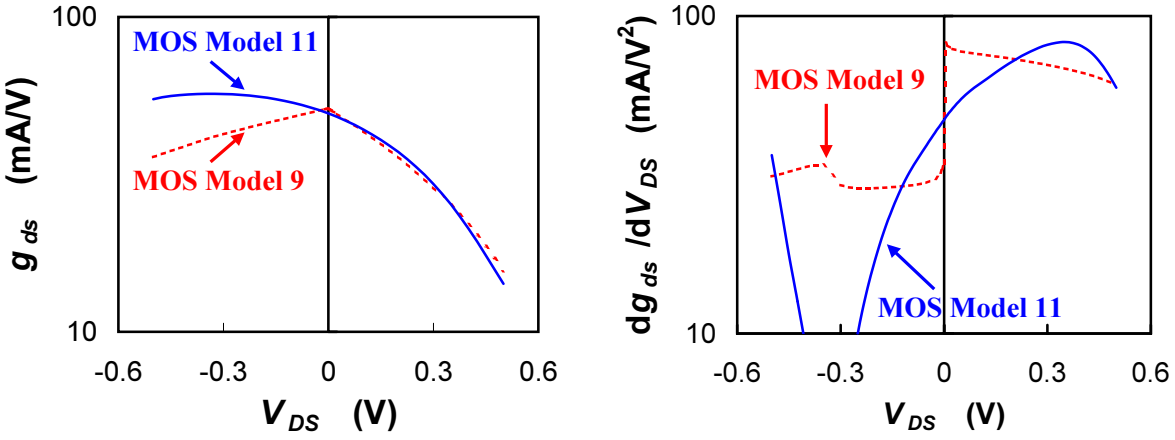
Distortion Modelling: The distortion behaviour of a MOSFET in an amplifier structure is illustrated in Fig. 1.2. In circuits using balanced topologies, even-order harmonics can be reduced by about 40dB. As a result, the 3rd-order harmonic forms a lower limit for the total distortion [32]. A compact MOS model should thus accurately describe the drain current and its higher-order derivatives (up to at least 3rd-order), MM11 has especially been developed for this purpose. As a result, compared to other compact models, MM11 contains improved expressions for mobility reduction [5], velocity saturation and various conductance effects [6, 7]. This improved model gives an accurate description of distortion at both low and high frequencies [10], see Fig. 1.3.



(a) (b)

Figure 1.3: Measured and modelled harmonic distortion as a function of dc gate bias V_{GS} of a $16 \times 10/0.35\mu$m n-type MOSFET in 0.35μm technology at (a) low frequency ($f = 16$MHz), and (b) high frequency ($f = 1$GHz). ($V_{DS} = 3.3$V, $V_{SB} = 0$V and $P_{in} = -5$dBm)

Drain-Source Symmetry: For applications where the MOSFET is used as a gate-controlled resistor [32], the model should be symmetrical with respect to source and drain at zero drain-source bias



(a) (b)

Figure 1.4: The use of a non-symmetric model description with respect to drain and source at $V_{DS} = 0$, such as used in conventional models such as MM9 and BSIM4, will result in (a) a kink in the conductance g_{ds} (i.e., $\partial I_{DS}/\partial V_{DS}$) and (b) discontinuities in the higher-order derivatives at $V_{DS} = 0$. In MM11, care has been taken to preserve symmetry.

(i.e., $V_{DS} = 0$) in order not to give erroneous results. In most MOS models, the expression for channel current I_{DS} has been derived for $V_{DS} \geq 0$, and source and drain are simply interchanged internally when $V_{DS} < 0$. In a typical MOSFET structure, however, the choice of source and drain is arbitrary, in other words, the device is symmetric with respect to source and drain when $V_{DS} = 0$. This should be reflected in the channel current expression, i.e., the same expression multiplied by -1 should be obtained if drain and source are interchanged.

Typical V_T-based models are intrinsically asymmetrical since the threshold voltage is defined at the source side only. This asymmetry leads to discontinuities in the higher-order derivatives of channel current at $V_{DS} = 0$, see Fig. 1.4, which is undesirable in applications where the MOSFET is used as a gate-bias controlled resistor. Care has to be taken in the derivation of the model expressions to preserve symmetry [6, 26], this has been done in MM11 [6, 7], see Fig. 1.4.

With respect to drain-source symmetry, it should be pointed out here that models are often mistakingly classified as being source-referenced or bulk-referenced [33], where the former is supposed to be aymmetrical and the latter symmetrical. Source-referenced models make use of V_{GS}, V_{DS} and V_{SB} as independent variables, whereas bulk-referenced models make use of V_{GB}, V_{DB} and V_{SB}. It will be clear that any bulk-referenced model can be converted to a source-referenced model by replacing V_{GB} and V_{DB} by $V_{GS} + V_{SB}$ and $V_{DS} + V_{SB}$, respectively. In other words, all advantages of a body-referenced model (such as, e.g., drain-source symmetry) can be carried over to a corresponding source-referenced model or vice-versa. As a consequence, it is misleading to talk about source-referenced and bulk-referenced models, and it makes more sense to talk about symmetrical and asymmetrical models.

1.3 Important Physical Effects:

The electrical behaviour of realistic MOSFETs deviates considerably from the ideal MOSFET behaviour owing to the influence of various physical effects. These effects have to be accurately taken into account in a compact MOS model. Traditionally, most models incorporate physical effects such as mobility reduction, bias-dependent series resistances, velocity saturation, drain-induced barrier lowering, static feedback, channel length modulation, self-heating and impact ionization. As modern

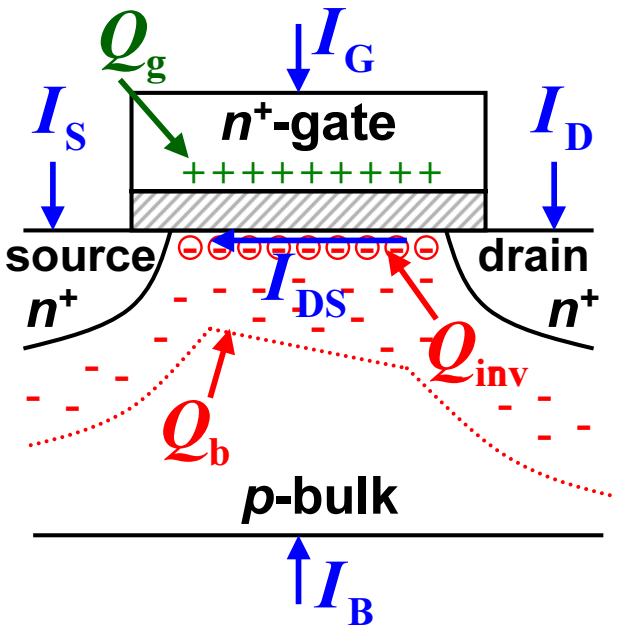


Figure 1.5: Basic structure of an n-type MOSFET, where all the most important electrical quantities are indicated: the nodal currents (I_D, I_S, I_G and I_B), the channel current (I_{DS}) and various charge densities (Q_{inv}, Q_b and Q_g).

CMOS process technology scales down to sub-100nm dimensions, however, certain physical effects, which did not affect circuit design before, become important. With the downscaling of gate oxide thickness, the gate poly-depletion effect becomes increasingly important. In addition, the gate tunnelling current is no longer negligible, and quantum-mechanical quantisation effects become important. Furthermore, the use of pocket implants in order to reduce short-channel effects does affect the electrical behaviour of MOSFETs. For short-channel devices, the input capacitance is dominated by the overlap capacitances, which are bias-dependent. All of the above effects should be taken into account in an advanced compact model, they have been implemented in MM11 [12, 13, 14].

1.4 Outline of Report

In a MOS transistor, we can distinguish three regions: i) the intrinsic channel region, ii) the gate/source and gate/drain overlap regions, and iii) the drain/bulk and source/bulk junction regions. It should be pointed out that MM11 only provides a model for the intrinsic channel region and the (gate/source and gate/drain) overlap regions. Junction charges, junction leakage currents and interconnect capacitances are not included. They are covered by separate models, which are not treated in this report. This report aims at giving a comprehensive physical derivation of the model equations used in MOS Model 11, Level 1101 [17]. These equations make use of electrical or miniset parameters, which are valid for a device with a specific geometry and a specific temperature. Since most of these parameters scale with geometry and with temperature, the process as a whole is characterised by an enlarged set of scaling parameters, the maxiset. These scaling parameters can be linked to the electrical parameters by physics-based geometrical and temperature scaling rules. In theory, these scaling rules can be based on physical relations. In practice, however, it is found that the use of semi-empirical scaling rules result in a more accurate and robust description. As a result, in this report, we will only focus on the derivation of the model equations based on miniset parameters. The geometrical and temperature scaling is left out of consideration here.
The MM11 equations give a description of all transistor-action related quantities: nodal currents,

nodal charges and noise-power spectral densities. Before relations can be derived for these quantities, we need to derive expressions for surface potential and important electrical quantities such as charge densities, see Fig. 1.5. This is done in Chapter 2. Next, we focus our attention on the steady-state (or dc) behaviour of MOSFETs. For ideal MOSFET operation, both bulk current and gate current are considered zero, and as a result, the drain-source channel current determines the MOSFET behaviour, see Fig. 1.5. The channel current and all the phenomena that affect it, are discussed in Chapter 3. For realistic devices, however, the bulk and gate current are not zero, and the modelling of these currents is treated in Chapter 4 and 5, respectively. In Chapter 6, we turn our attention towards the dynamic behaviour of MOSFETs, which is described in terms of quasi-static nodal charges. Finally, we derive the expressions for the noise-power spectral densities in Chapter 7.

The treatment of MOS device physics in this report is rather concise in some parts, the interested reader is therefore referred to [34]-[36] for a more in-depth discussion. Furthermore, it will be clear that outlining the physical background of MM11 asks for numerous derivations. In order to keep the report readable, the more elaborate derivations are given in separate Appendices.

2 MOSFET Basics

As the name Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) suggests, the MOS transistor consists of a semiconductor substrate on which a thin layer of insulating oxide (SiO_2) of thickness t_{ox} is grown. A conducting layer (metal or heavily doped silicided polysilicon) called the gate electrode is deposited on top of the oxide. The basic structure of an n-type MOSFET[1] is shown in Fig. 2.1. The p-type doped silicon region, commonly referred to as the bulk or substrate, is contacted via the bulk contact. Two heavily n-type doped regions of depth X_j, called the source and the drain, are formed in the substrate on either side of the gate. The gate overlaps slightly with the source and drain regions. The region between the source and drain junctions is called the channel region, which has a length L (in the x-direction) and a width W (in the z-direction). Due to the manufacturing tolerances the mask length/width differs slightly from the final gate polysilicon length/width, furthermore the lateral under diffusion of the source and drain junctions also has its effect on the actual channel length/width. As a result, both L and W may differ from the actual mask dimensions, L_{mask} and W_{mask}, and the gate overlaps both the source and drain extensions, the so-called *overlap regions*.

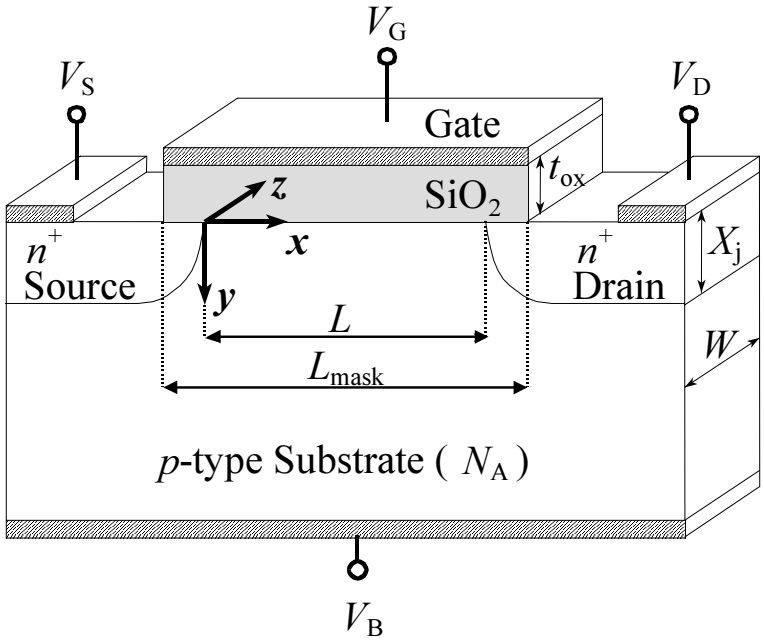


Figure 2.1: The basic structure of an n-type MOS transistor.

When a voltage V_{GB} is applied between the gate and the bulk, the band structure near the Si-SiO_2 interface is changed. For the moment, we assume source and drain are grounded ($V_{SB} = V_{DB} = 0$), in this case three different situations can be distinguished (in the channel region): *accumulation*, *depletion* or *weak inversion* and *strong inversion*, as is shown in Fig. 2.2. Note that the gate overlap regions behave differently than the channel region.

For negative or very low gate voltage values, holes are attracted to the surface and a thin layer with a positive charge, a so-called *accumulation layer*, is formed. With increasing gate voltage, the band bending becomes less, until at a distinct gate voltage value no band bending occurs. This is called the flat-band voltage V_{FB}.
Beyond this point, the band bending is opposite to the accumulation condition, a negative charge is being built up. In effect, the positive voltage at the gate will repel the holes from the silicon surface

[1] With the appropriate change of signs for charge and potential, the following also holds for p-type devices which are fabricated with an n-type doped substrate or a p-type doped substrate with an implanted n-well.

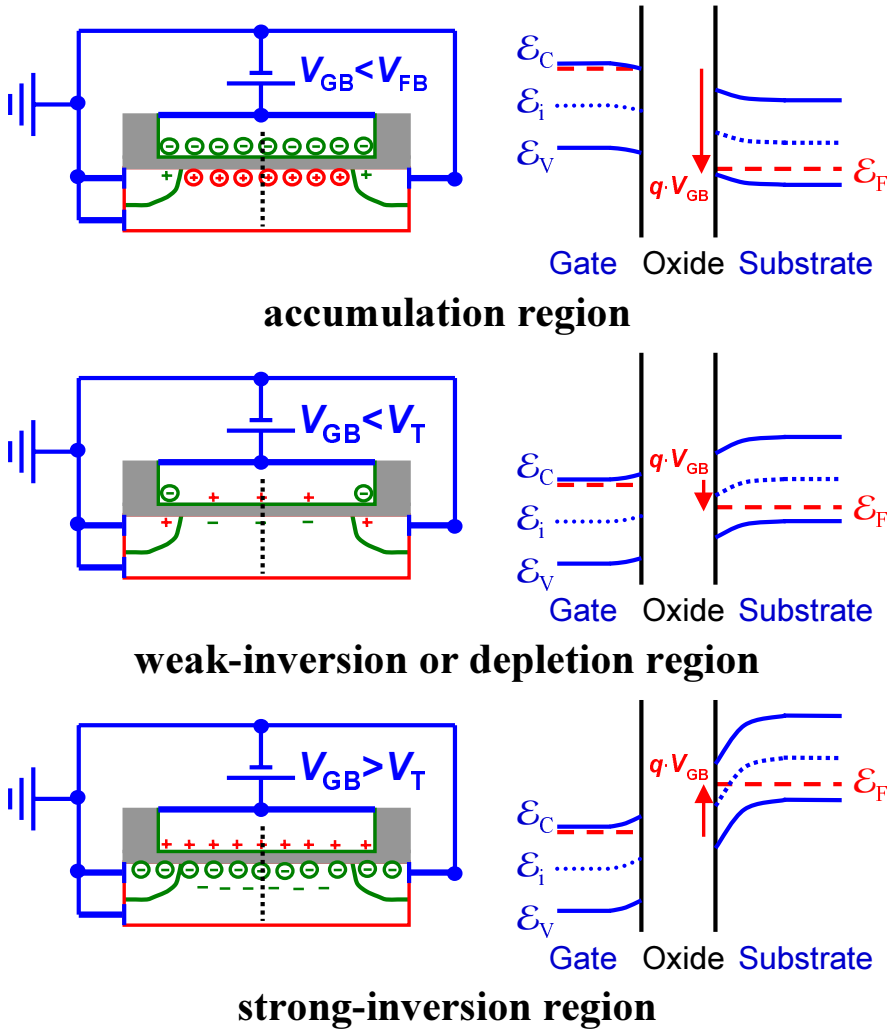


Figure 2.2: The charge distribution and the corresponding energy-band diagram along the indicated dashed line (in y-direction) in an n-MOSFET for the three operation conditions: *accumulation*, *depletion* or *weak inversion* and *strong inversion*.

and thus expose the negatively charged (immobile) acceptor ions. This charge is called the depletion charge. Since holes are depleted at the surface, it is referred to as the depletion condition.
When V_{GB} is increased still more, the downward band bending becomes stronger. In fact, the band bending may cause the midgap energy $\mathcal{E}_i$ to cross over the constant Fermi level $\mathcal{E}_F$, see Fig. 2.2. In this case, the surface behaves like an n-type material as opposed to the original p-type material, hence the name inversion region. A conducting layer with a negative (mobile) charge Q_{inv}, a so-called *inversion layer*, is formed. This layer shields the underlying silicon from the gate potential, and as a result the band bending does not extend deeper into the silicon after the (strong) inversion layer has been formed.
The inversion charge can be contacted via the source and drain region, and a current, the so-called *channel current*, will flow through the inverted area when a potential difference V_{DS} is applied between drain and source. Since the inversion charge depends heavily on the gate potential, the gate can be used to control the current through the channel.

In this chapter the so-called surface potential, on which MM11 is based, will be introduced and treated in depth, and expressions based on surface potential formulations will be derived for the charge densities and other important quantities such as electric fields. Since MM11 not only takes into

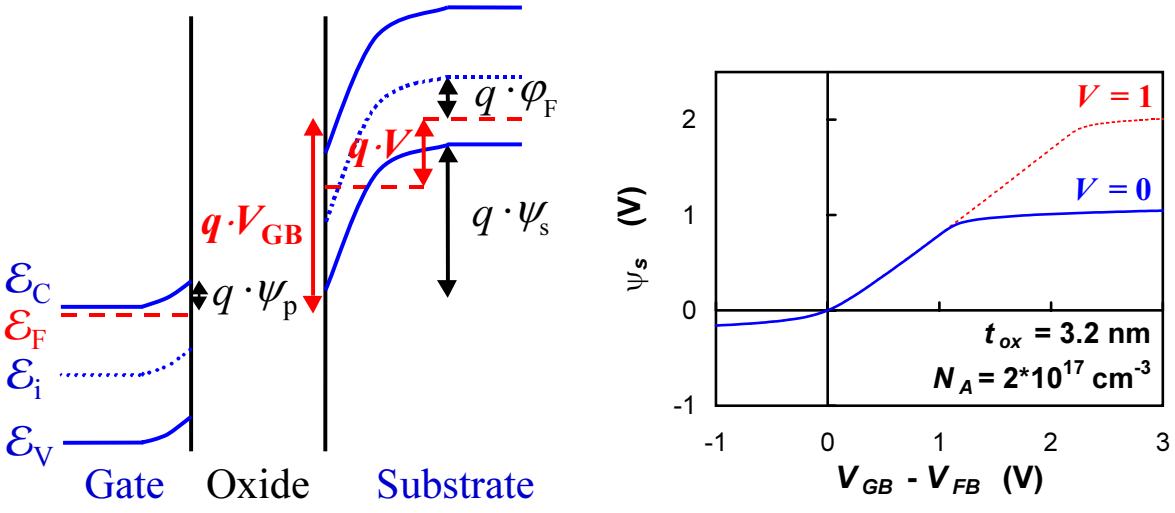


Figure 2.3: (a) The energy-band diagram (in transversal direction) of an n-MOSFET for $V_{GB} > V_{FB}$, where ψ_s is the surface potential, ψ_p is the potential drop in the gate due to the poly-depletion effect, V is the quasi-Fermi potential and ϕ_F is the intrinsic Fermi-potential ($\phi_B = 2 \cdot \phi_F$). (b) The surface potential as a function of gate bias for different values of quasi-Fermi potential V (as calculated from (2.7) with $m_S = 1$).

account the intrinsic MOS region (where the channel is formed) but also the extrinsic MOS regions (i.e., the gate-source and the gate-drain overlap regions), this Chapter is split up in two Sections. First the surface potential and the charge densities in the intrinsic region are treated in Section 2.1, and next the same is done for the extrinsic regions in Section 2.2.

2.1 Intrinsic Region

In this Section we will discuss the formulation of surface potential in a typical MOS structure (Section 2.1.1), followed by a description of the various important charge densities and electric fields (Section 2.1.2).

2.1.1 Surface Potential

In order to be able to calculate electrical quantities such as currents, charges and noise, we first start off with the definition of surface potential. The surface potential ψ_s is defined as the electrostatic potential at the gate oxide/substrate interface with respect to the neutral bulk (due to band bending, see Fig. 2.3 (a)). For the moment, it is assumed that the gate is ideal and no depletion occurs in the polysilicon (i.e., $\psi_p = 0$ in Fig. 2.3 (a)). In this case, the surface potential ψ_s can be calculated using the following derivation.
In the p-type substrate, a space charge $\rho(x, y)$ is present [35]:

$$\rho(x, y) = q \cdot [p(x, y) - n(x, y) - N_A] \tag{2.1}$$

where N_A is the net acceptor doping concentration. The electron and hole density, n and p, are given by Maxwell-Boltzmann statistics:

$$\begin{aligned} n(x,y) &\approx N_A \cdot \exp\left(\frac{\psi(x,y)-V(x)-2\cdot\phi_F}{\phi_T}\right) \\ p(x,y) &\approx N_A \cdot \exp\left(-\frac{\psi(x,y)}{\phi_T}\right) \end{aligned} \tag{2.2}$$

where ψ is the electrostatic potential with respect to the neutral bulk and ϕ_T ($= k_B \cdot T/q$) is the thermal voltage. The intrinsic Fermi potential ϕ_F (as shown in Fig. 2.3 (a)) is defined by $\phi_T \cdot \ln(N_A/n_i)$, and $V(x)$ denotes the electron quasi-Fermi potential, which ranges from V_{SB} at the source side ($x = 0$) to V_{DB} at the drain side ($x = L$). The Poisson equation for the electrostatic potential ψ is written as:

$$\nabla^2\psi = -\frac{\rho(x,y)}{\epsilon_{Si}} \tag{2.3}$$

In order to obtain an approximate analytical solution of (2.3), usually the assumption is made that $\partial^2\psi/\partial x^2 \ll \partial^2\psi/\partial y^2$. This is called the *gradual channel approximation*, which is valid for long-channel devices. The Poisson equation can now be rewritten as:

$$\frac{\partial^2\psi}{\partial y^2} \approx \frac{q\cdot N_A}{\epsilon_{Si}} \cdot \left[1 - \exp\left(-\frac{\psi}{\phi_T}\right) + \exp\left(\frac{\psi - V - \phi_B}{\phi_T}\right)\right] \tag{2.4}$$

where $\phi_B = 2\cdot\phi_F$. As boundary conditions both ψ and $\partial\psi/\partial y$ are taken to be equal to zero deep in the neutral bulk. Using $\partial^2\psi/\partial y^2 = 1/2 \cdot \partial(\partial\psi/\partial y)^2/\partial y$, the total charge Q_s per unit area in the semiconductor can be obtained from Gauss' law:

$$\begin{aligned} Q_s &= \epsilon_{Si} \cdot \left.\frac{\partial\psi}{\partial y}\right|_{y=0} \\ &= \pm k_0 \cdot C_{ox} \cdot \sqrt{\psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right] + \phi_T \cdot \exp\left(-\frac{V+\phi_B}{m_S\cdot\phi_T}\right) \cdot \left[\exp\left(\frac{\psi_s}{m_S\cdot\phi_T}\right) - 1\right]} \end{aligned} \tag{2.5}$$

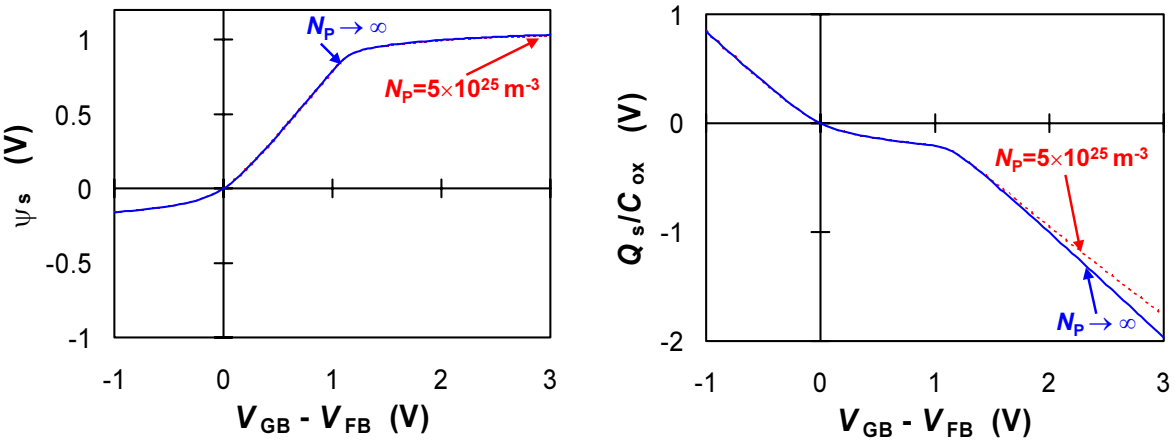
where C_{ox} is the gate oxide capacitance per unit area given by ϵ_{ox}/t_{ox}, and k_0 is the body factor given by $\sqrt{2\cdot q\cdot\epsilon_{Si}\cdot N_A}/C_{ox}$. In the above equation, a new parameter m_S has been introduced in order to take into account short-channel effects as will be discussed in Section 3.2. For the time being, m_S is taken to be its theoretical value of 1. The charge density Q_s is negative for $V_{GB} > V_{FB}$ (i.e., inversion) and positive for $V_{GB} < V_{FB}$ (i.e., accumulation).
Applying Gauss' theorem at the oxide interface, the above charge can also be related to the applied gate bias:

$$Q_s = -C_{ox} \cdot \left(V^*_{GB} - \psi_s\right) \tag{2.6}$$

where the effective gate-bulk bias V^*_{GB} is given by $V_{GB} - V_{FB}$. Equating (2.5) and (2.6), an implicit relation for $\psi_s(V_{GB}, V)$ is found:

$$\left(\frac{V^*_{GB} - \psi_s}{k_0}\right)^2 = \psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right] + \phi_T \cdot \exp\left(-\frac{V+\phi_B}{m_S\cdot\phi_T}\right) \cdot \left[\exp\left(\frac{\psi_s}{m_S\cdot\phi_T}\right) - 1\right] \tag{2.7}$$

The surface potential ψ_s cannot be solved analytically from the above relation, and thus it has to be solved iteratively. In Fig. 2.3 (b), the surface potential ψ_s is given as a function of applied gate bias for different values of quasi-Fermi potential V. Three distinct regions of operation can be observed, the *accumulation region* (i.e., $\psi_s < 0$ and $V^*_{GB} < 0$), the *weak-inversion or depletion region* (i.e., $0 < \psi_s < \phi_B + V$ and $0 < V^*_{GB} < \phi_B + V + k_0 \cdot \sqrt{\phi_B + V}$) and the *strong-inversion region* (i.e., $\psi_s > \phi_B + V$ and $V^*_{GB} > \phi_B + V + k_0 \cdot \sqrt{\phi_B + V}$).



(a) (b)

Figure 2.4: (a) The surface potential ψ_s (as given by (2.10)) and (b) the total charge density Q_s (as given by (2.8)) as a function of gate bias V^*_{GB} for an ideal gate, i.e., $N_P \to \infty$ (solid line), and for a polysilicon gate with $N_P = 5 \times 10^{25}\text{m}^{-3}$ (dashed line). (n-MOS, $V = 0$, $N_A = 2 \times 10^{23}\text{m}^{-3}$, $t_{ox} = 3.2\text{nm}$ and $m_S = 1$)

Poly-Depletion Effect: In practice, the polysilicon gate is not an ideal conductor. The use of a polysilicon gate in modern technologies results in an unwanted effect, the so-called *poly-depletion effect* [37, 38]. For high values of normal electric field, a depletion layer is not only formed in the silicon substrate but in the polysilicon gate as well, resulting in a potential drop ψ_p across the polysilicon depletion layer (i.e., the electrostatic potential at the gate/gate-oxide interface with respect to the neutral gate, see Fig. 2.3 (a)). As a consequence, the above-derived relations become inaccurate. The poly-depletion effect particularly affects the MOS C-V-characteristics [38], and has to be taken into account. According to Appendix B.1, the total charge density Q_s becomes:

$$Q_s = -C_{ox} \cdot \left(V^*_{GB} - \psi_s - \psi_p\right) \tag{2.8}$$

where the potential ψ_p is given by:

$$\psi_p = \begin{cases} 0 & \text{for: } V^*_{GB} \le 0 \\ \left(\sqrt{V^*_{GB} - \psi_s + {k_P}^2/4} - k_P/2\right)^2 & \text{for: } V^*_{GB} > 0 \end{cases} \tag{2.9}$$

where k_p is the gate body factor given by $\sqrt{2 \cdot q \cdot \epsilon_{Si} \cdot N_P}/C_{ox}$ and N_P is the net donor doping concentration in the polysilicon gate. Including poly-depletion, the implicit relation for surface potential (2.7) can be rewritten as:

$$\left(\frac{V^*_{GB} - \psi_s - \psi_p}{k_0}\right)^2 = \psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right] \tag{2.10}$$

$$+ \phi_T \cdot \exp\left(-\frac{V + \phi_B}{m_S \cdot \phi_T}\right) \cdot \left[\exp\left(\frac{\psi_s}{m_S \cdot \phi_T}\right) - 1\right]$$

The influence of poly-depletion can be seen in Fig. 2.4, where the surface potential ψ_s and the total charge density Q_s are shown for an ideal metal gate (i.e., $N_P \to \infty$ and consequently $\psi_p = 0$) and

for a typical poly-silicon gate. It is clear that, although it hardly influences the surface potential, poly-depletion reduces the charge density in the strong inversion region and, as a result, it will affect the currents and the capacitances in the MOSFET.

The implicit relation (2.10) for surface potential can only be solved iteratively, which is generally considered to be computationally expensive. This is undesirable for VLSI circuit simulation, and as such it is a drawback of ψ_s-based models. Nevertheless, these days it is felt that the extra computation time is worth the return in accuracy. Moreover, it is possible to reduce the computation time by using a suitable approximation of ψ_s [9, 29]. In MM11, an approximate explicit solution of surface potential is used, see Appendix C, resulting in no increase of computation time as compared to MM9.

Quantum-Mechanical Effects: For modern CMOS technologies, the use of a thinner gate oxide t_{ox} and higher channel doping N_A leads to very high normal electric fields at the Si-SiO_2 interface. As a result, quantum-mechanical effects start to play a role [39, 40], and the above calculation of surface potential no longer holds [37, 41, 42]. Consequently, the above-derived relation (2.10) is no longer accurate and a modified relation has to be used, see Appendix D. Generally speaking, quantum-mechanical effects result in an effective increase in intrinsic Fermi potential ϕ_F and in an effective (bias-dependent) increase in oxide thickness. Since the former can be taken into account by adjusting certain model parameters (such as ϕ_B), only the latter effect is taken into account in MM11. See Appendix D for a more in-depth discussion of quantum-mechanical effects.

2.1.2 Charge Densities

In general, electrical quantities such as currents, charges and noise can be written in terms of the surface potential at the source side ψ_{s_0} ($x = 0$) and at the drain side ψ_{s_L} ($x = L$), which can both be calculated from (2.10) using $V = V_{SB}$ and $V = V_{DB}$, respectively. In order to calculate electrical quantities, a distinction has to be made between the mobile electron charge supplied by the source and drain, i.e., the inversion layer charge density Q_{inv}, and the immobile charge modulated by the bulk, i.e., the bulk charge density Q_b. In order to calculate Q_{inv} we need to integrate the electron density n along the y-direction. Unfortunately this does not result in an analytical expression, and an approximation is thus needed, see Appendix E. Since the inversion layer in general is much less thick than the depletion layer, it can be assumed that its thickness can be neglected. This concept is commonly referred to as the *charge sheet approximation* [44]. Under this assumption, no potential is dropped across the inversion layer and, according to (E.5), Q_b can simply be written as:

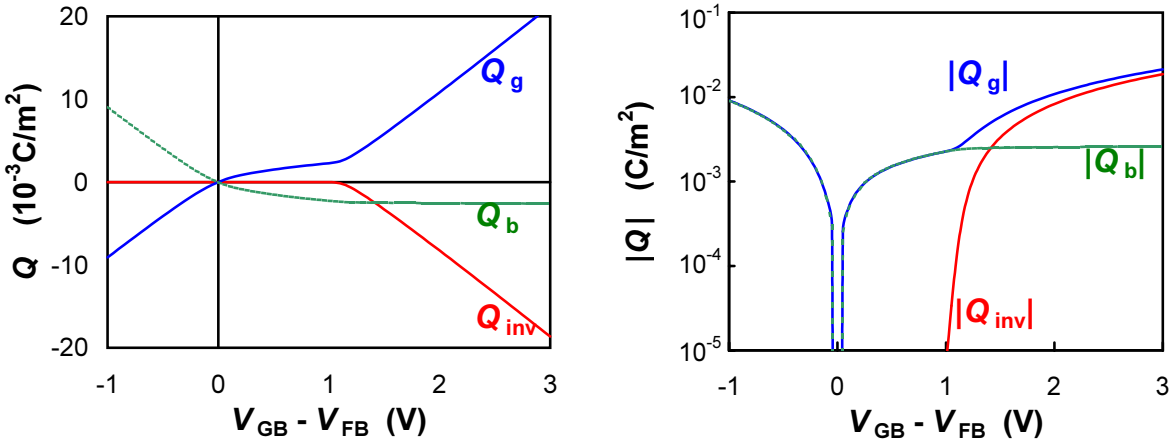
$$|Q_b| = C_{ox} \cdot k_0 \cdot \sqrt{\psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right]} \tag{2.11}$$

In the accumulation region ($V^*_{GB} < 0$), Q_b is positive and equal to the accumulation charge density, and in the inversion region ($V^*_{GB} > 0$), it is negative and equal to the depletion charge density. The inversion layer charge density Q_{inv} becomes:

$$Q_{inv} = Q_s - Q_b = -C_{ox} \cdot \left(V^*_{GB} - \psi_s - \psi_p\right) - Q_b \tag{2.12}$$

Since the inversion layer charge is negligible in the accumulation region and the influence of holes is marginal in the inversion region, Q_{inv} can be approximated by:

$$Q_{inv} = \begin{cases} 0 & \text{for: } V^*_{GB} \le 0 \\ -C_{ox} \cdot \left[\, V^*_{GB} - \psi_s - \psi_p - k_0 \cdot \sqrt{\psi_s + \Delta_{acc}} \,\right] & \text{for: } V^*_{GB} > 0 \end{cases} \tag{2.13}$$



(a) (b)

Figure 2.5: The inversion-layer charge density $Q_{\rm inv}$, the bulk charge density $Q_{\rm b}$ and the gate charge density $Q_{\rm g}$ plotted as a function of gate bias $V^*_{\rm GB}$ on a (a) linear and (b) logarithmic scale. (n-MOS, $V = 0$, $N_{\rm A} = 2 \times 10^{23}{\rm m}^{-3}$, $t_{\rm ox} = 3.2{\rm nm}$ and $m_{\rm S} = 1$)

where $\Delta_{\rm acc}$ is a simple function of $V^*_{\rm GB}$, which approximates the influence of holes in the inversion region, see Appendix C:

$$\Delta_{\rm acc} = \phi_{\rm T} \cdot \left[\exp\left(-{\rm Acc} \cdot \frac{V^*_{\rm GB}}{\phi_{\rm T}} \right) - 1 \right] \tag{2.14}$$

Here, Acc is defined as:

$${\rm Acc} = \left. \frac{\partial \psi_{\rm s}}{\partial V_{\rm GB}} \right|_{V^*_{\rm GB}=0} = \frac{1}{1 + k_0 \,/\, \sqrt{2 \cdot \phi_{\rm T}}} \tag{2.15}$$

Using (2.5) and the above approximation $\Delta_{\rm acc}$, we can also write for $Q_{\rm inv}$ (for $V^*_{\rm GB} > 0$):

$$Q_{\rm inv} = -k_0 \cdot C_{\rm ox} \cdot \left[\sqrt{\psi_{\rm s} + \Delta_{\rm acc} + \phi_{\rm T} \cdot e^{-\frac{V+\phi_{\rm B}}{m_{\rm S}\cdot\phi_{\rm T}}} \cdot \left[e^{\frac{\psi_{\rm s}}{m_{\rm S}\cdot\phi_{\rm T}}} - 1 \right]} - \sqrt{\psi_{\rm s} + \Delta_{\rm acc}} \right] \tag{2.16}$$

The gate charge density $Q_{\rm g}$ is simply given by $-Q_{\rm inv} - Q_{\rm b}$:

$$Q_{\rm g} = C_{\rm ox} \cdot \left(V^*_{\rm GB} - \psi_{\rm s} - \psi_{\rm p} \right) = \begin{cases} C_{\rm ox} \cdot \left(V^*_{\rm GB} - \psi_{\rm s} \right) & \text{for: } V^*_{\rm GB} \le 0 \\ C_{\rm ox} \cdot \left[\frac{2\cdot\left(V^*_{\rm GB} - \psi_{\rm s}\right)}{1+\sqrt{1+4/k_{\rm P}^2\cdot\left(V^*_{\rm GB} - \psi_{\rm s}\right)}} \right] & \text{for: } V^*_{\rm GB} > 0 \end{cases} \tag{2.17}$$

The charge densities $Q_{\rm inv}$, $Q_{\rm b}$ and $Q_{\rm g}$ are shown in Fig. 2.5 as a function of gate bias $V^*_{\rm GB}$. In accumulation, the accumulation layer and consequently both $|Q_{\rm g}|$ and $|Q_{\rm b}|$ as well increase approximately linearly with $V^*_{\rm GB}$. In weak inversion, the depletion layer increases only weakly with $V^*_{\rm GB}$, and as a result both $Q_{\rm g}$ and $Q_{\rm b}$ are only weakly dependent on $V^*_{\rm GB}$. On the other hand, $Q_{\rm inv}$ is very small, but increases exponentially with $V^*_{\rm GB}$. Finally, in strong inversion, the depletion layer is shielded off from the gate by the inversion layer, and consequently $Q_{\rm b}$ is approximately independent of $V^*_{\rm GB}$, whereas

$|Q_{\mathrm{inv}}|$ increases almost linearly with V^*_{GB}.

Care has to be taken to preserve drain-source symmetry in the model expressions [6, 26], see Section 1.2. In order to preserve symmetry, the linearisation of quantities such as currents and charges should be done with respect to the average surface potential $\bar{\psi}$ instead of the source or drain potential [6, 7, 30]. The average surface potential $\bar{\psi}$ is defined as:

$$\bar{\psi} = \frac{\psi_{\mathrm{s_L}} + \psi_{\mathrm{s_0}}}{2} \tag{2.18}$$

In addition, we can define the surface potential difference $\Delta\psi$:

$$\Delta\psi = \psi_{\mathrm{s_L}} - \psi_{\mathrm{s_0}} \tag{2.19}$$

A first-order Taylor polynomial of (2.13) around $\psi_{\mathrm{s}} = \bar{\psi}$ results in (for $V^*_{\mathrm{GB}} > 0$):

$$Q_{\mathrm{inv}} \approx \bar{Q}_{\mathrm{inv}} - C_{\mathrm{inv}} \cdot \left(\psi_{\mathrm{s}} - \bar{\psi}\right) \tag{2.20}$$

where the average inversion-layer charge density $\bar{Q}_{\mathrm{inv}}$ is given by:

$$\bar{Q}_{\mathrm{inv}} = Q_{\mathrm{inv}}\left(\psi_{\mathrm{s}} = \bar{\psi}\right) = -C_{\mathrm{ox}} \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right)}} - k_0 \cdot \sqrt{\bar{\psi} + \Delta_{\mathrm{acc}}}\right] \tag{2.21}$$

and

$$C_{\mathrm{inv}} = -\left.\frac{\partial Q_{\mathrm{inv}}}{\partial \psi_{\mathrm{s}}}\right|_{\psi_{\mathrm{s}}=\bar{\psi}} = -C_{\mathrm{ox}} \cdot \left(\frac{1}{\sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right)}} + \frac{k_0}{2 \cdot \sqrt{\bar{\psi} + \Delta_{\mathrm{acc}}}}\right) \tag{2.22}$$

In addition, we can define the inversion-layer charge difference ΔQ_{inv}:

$$\Delta Q_{\mathrm{inv}} = Q_{\mathrm{s_L}} - Q_{\mathrm{s_0}} = -C_{\mathrm{inv}} \cdot \Delta\psi \tag{2.23}$$

The gate charge density Q_{g} can be approximated in a similar way, taking a first-order Taylor polynomial of (2.17) around $\psi_{\mathrm{s}} = \bar{\psi}$:

$$Q_{\mathrm{g}} \approx \bar{Q}_{\mathrm{g}} - C_{\mathrm{g}} \cdot \left(\psi_{\mathrm{s}} - \bar{\psi}\right) \tag{2.24}$$

where:

$$\bar{Q}_{\mathrm{g}} = Q_{\mathrm{g}}\left(\psi_{\mathrm{s}} = \bar{\psi}\right) = \begin{cases} C_{\mathrm{ox}} \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right) & \text{for: } V^*_{\mathrm{GB}} \le 0 \\ C_{\mathrm{ox}} \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right)}}\right] & \text{for: } V^*_{\mathrm{GB}} > 0 \end{cases} \tag{2.25}$$

and

$$C_{\mathrm{g}} = -\left.\frac{\partial Q_{\mathrm{g}}}{\partial \psi_{\mathrm{s}}}\right|_{\psi_{\mathrm{s}}=\bar{\psi}} = \begin{cases} C_{\mathrm{ox}} & \text{for: } V^*_{\mathrm{GB}} \le 0 \\ C_{\mathrm{ox}} \cdot \left[\frac{1}{\sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \bar{\psi}\right)}}\right] & \text{for: } V^*_{\mathrm{GB}} > 0 \end{cases} \tag{2.26}$$

The bulk charge density Q_b is now simply given by:

$$Q_b = -Q_g - Q_{inv} \approx \bar{Q}_b - C_b \cdot \left(\psi_s - \bar{\psi}\right) \tag{2.27}$$

where:

$$\bar{Q}_b = Q_b\left(\psi_s = \bar{\psi}\right) = -\bar{Q}_g - \bar{Q}_{inv} \tag{2.28}$$

and

$$C_b = -\left.\frac{\partial Q_b}{\partial \psi_s}\right|_{\psi_s=\bar{\psi}} = -C_g - C_{inv} \tag{2.29}$$

In addition to the above charge densities, we will define here some other important quantities such as the transversal electric field in the gate oxide E_{ox}:

$$E_{ox} = \frac{Q_g}{\epsilon_{ox}} \tag{2.30}$$

and the transversal electric field at the Si/SiO_2-interface E_{Si}:

$$E_{Si} = \frac{Q_g}{\epsilon_{Si}} \tag{2.31}$$

which are important in the calculation of gate current.
From a physical point of view, it makes sense to use charge densities and electrical fields in the derivation of expressions. In the MM11 equations [17], however, for convenience's sake most variables are written in terms of voltages. The above-mentioned quantities can be translated into MM11 (auxiliary) variables. The average surface potential $\bar{\psi}$ in (2.18) is equal to $\bar{\psi}_{inv}$ in MM11. The inversion layer charge density Q_{inv} in (2.13) corresponds to the effective gate drive V_{G_T} in MM11:

$$V_{G_T} = -\frac{Q_{inv}}{C_{ox}} \tag{2.32}$$

and equivalently:

$$\bar{V}_{G_T} = -\frac{\bar{Q}_{inv}}{C_{ox}} \tag{2.33}$$

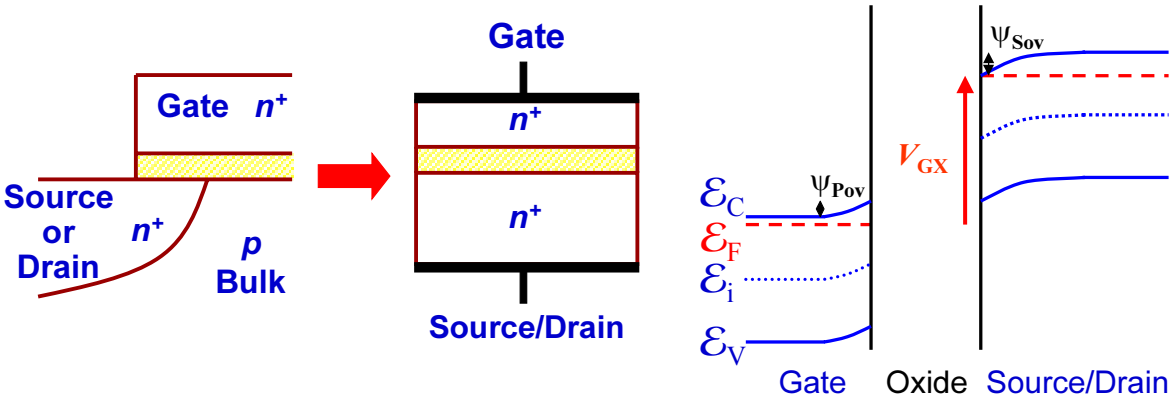
The average gate charge density $\bar{Q}_g$ in (2.25) translates into V_{ox} in MM11:

$$V_{ox} = \frac{\bar{Q}_g}{C_{ox}} \tag{2.34}$$

And finally, the quantities C_{inv} and C_g are translated into ξ and ξ_{ox}, respectively:

$$\xi = -\phi_T \cdot \frac{C_{inv}}{C_{ox}} \tag{2.35}$$

$$\xi_{ox} = \phi_T \cdot \frac{C_g}{C_{ox}} \tag{2.36}$$



(a) (b)

Figure 2.6: (a) The gate/source or gate/drain overlap region in an n-MOSFET can be approximately treated as an n^+-gate/oxide/n^+-bulk MOS capacitance where the source or drain, respectively, acts as a bulk. (b) The corresponding energy-band diagram (in transversal direction) of the gate overlap region for $V_{GX} > V_{FBov}$, where $\psi_{s_{ov}}$ is the surface potential and $\psi_{p_{ov}}$ is the potential drop in the gate due to the poly-depletion effect.

2.2 Extrinsic Region

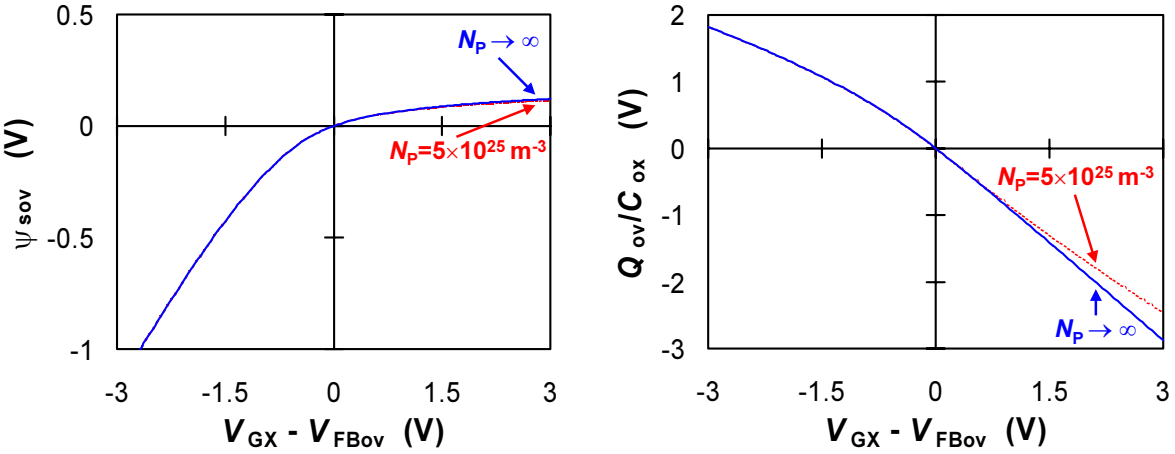
In this Section, we will discuss the formulation of surface potential in the gate/source and gate/drain overlap regions (Section 2.2.1), followed by a description of the charge densities and other important quantities (Section 2.2.2). In the following general derivation, we denote the source or drain terminal by X. In order to keep the derivations manageable, the gate overlap region is treated as an n^+-gate/oxide/n^+-bulk MOS capacitance where the source or drain acts as a bulk, see Fig. 2.6. Although the impurity doping concentration in the n^+-source/drain extension region is non-uniform in both lateral and transversal direction, it is assumed that an effective constant donor doping concentration N_{OV} can be defined for this structure, resulting in an effective flat-band voltage V_{FBov} and body factor k_{ov} ($= \sqrt{2 \cdot q \cdot \epsilon_{Si} \cdot N_{OV}}/C_{ox}$). Since the source/drain extension is highly n-type doped, $V_{FBov} \approx 0$ and $k_{ov} \gg k_0$.
For $V_{GX} > V_{FBov}$, a negatively charged accumulation layer is formed in the overlapped n^+-extension and a positively charged depletion layer is formed in the overlapping gate, whereas, for $V_{GS} < V_{FBov}$, a positively charged depletion layer is formed in the overlapped n^+-extension and a negatively charged accumulation layer is formed in the overlapping gate.

2.2.1 Surface Potential

According to Appendix B.2, assuming that only accumulation and depletion occur in the overlapped n^+-region[2], the implicit expression for the surface potential $\psi_{s_{ov}}$ in the source region is given by:

$$\left(\frac{V_{GX} - V_{FBov} - \psi_{p_{ov}} - \psi_{s_{ov}}}{k_{ov}}\right)^2 = -\psi_{s_{ov}} + \phi_T \cdot \left[\exp\left(\frac{\psi_{s_{ov}}}{\phi_T}\right) - 1\right] \tag{2.37}$$

[2]Since the source/drain extension has a very high doping concentration, an inversion layer in the overlap region will only be formed at very negative gate bias values V_{GX}. This effect has been neglected, see Appendix B.2.



(a) (b)

Figure 2.7: (a) The surface potential $\psi_{s_{ov}}$ (as given by (2.37)) and (b) the total charge density Q_{ov} (as given by (2.38)) as a function of gate bias $V_{GX} - V_{FBov}$ for an ideal gate, i.e., $N_P \to \infty$ (solid line), and for a polysilicon gate with $N_P = 5 \times 10^{25}\text{m}^{-3}$ (dashed line). (n-MOS, $V = 0$, $N_{OV} = 1 \times 10^{25}\text{m}^{-3}$, $t_{ox} = 3.2\text{nm}$ and $m_S = 1$)

where the potential drop in the polysilicon gate material due to the poly-depletion effect $\psi_{p_{ov}}$ is given by:

$$\psi_{p_{ov}} = \begin{cases} 0 & \text{for: } V_{GX} \le V_{FBov} \\ \left(\sqrt{V_{GX} - V_{FBov} - \psi_{s_{ov}} + k_P{}^2/4} - k_P/2\right)^2 & \text{for: } V_{GX} > V_{FBov} \end{cases}$$

The surface potential in the overlap region $\psi_{s_{ov}}$ is plotted as a function of gate bias in Fig. 2.7 (a). It is clear that, as is the case in the intrinsic region (see 2.4 (a)), poly-depletion hardly affects $\psi_{s_{ov}}$. Again the surface potential $\psi_{s_{ov}}$ can be explicitly approximated along the same lines as described in Appendix C.

2.2.2 Charge Densities

The gate overlap regions may contribute significantly to the total gate current, the gate-induced drain leakage, and the total gate capacitance. In order to be able to calculate these electrical quantities, we need an expression for the charge density in the overlapped source/drain extension Q_{ov}, given by (see Appendix B.2):

$$\begin{aligned} Q_{ov} &= -C_{ox} \cdot \left(V_{GX} - V_{FBov} - \psi_{p_{ov}} - \psi_{s_{ov}}\right) \qquad (2.38) \\ &= \begin{cases} -C_{ox} \cdot \left(V_{GX} - V_{FBov} - \psi_{s_{ov}}\right) & \text{for: } V_{GX} \le V_{FBov} \\ -C_{ox} \cdot \left[\frac{2 \cdot \left(V_{GX} - V_{FBov} - \psi_{s_{ov}}\right)}{1 + \sqrt{1 + 4/k_P{}^2 \cdot \left(V_{GX} - V_{FBov} - \psi_{s_{ov}}\right)}}\right] & \text{for: } V_{GX} > V_{FBov} \end{cases} \end{aligned}$$

and for the voltage across the oxide V_{ov} (from gate to source/drain), simply given by:

$$V_{ov} = -Q_{ov}/C_{ox} \qquad (2.39)$$

Note that V_{ov} given above corrsponds to the same variable used in MM11.
The charge density Q_{ov} can be seen as a function of gate bias in Fig. 2.7 (b). It is clear that, although it hardly influences the surface potential, poly-depletion reduces the charge density in the accumulation region and as a result it will affect the overlap capacitance in the MOSFET.

3 Channel Current Modelling

In this Chapter, a general expression for the channel current I_{DS} based on surface-potential formulations will be developed. Starting with an expression for the ideal drain current in Section 3.1, the attention will be subsequently focussed on the subthreshold region (subthreshold slope and drain-induced barrier lowering) and on the modelling of physical effects such as mobility reduction, velocity saturation, series resistance, and conductance effects (channel length modulation, static feedback and self-heating) in Sections 3.2 through 3.6. Finally, the impact of pocket implants on channel current will be briefly discussed in Section 3.7.

3.1 Ideal Channel Current

For the calculation of the current that flows from drain to source, the so-called channel current I_{DS}, it is assumed that the hole current as well as recombination/generation can be neglected. In the ideal case, it is furthermore assumed that there is a current flow in the x-direction only. In other words, the bulk current I_B and gate current I_G are zero. The channel current can simply be written as [43, 44], see Appendix E.1:

$$I_{DS} = -\mu \cdot W \cdot Q_{inv} \cdot \frac{dV}{dx} \tag{3.1}$$

where μ is the carrier mobility. Based on the charge sheet approximation, making some further assumptions (see Appendix E.1) the channel current can be approximated by [44, 45, 46]:

$$I_{DS} = I_{drift} + I_{diff} \tag{3.2}$$

where I_{drift} and I_{diff} denote a drift and a diffusion current, respectively:

$$I_{drift} = -\mu \cdot W \cdot Q_{inv} \cdot \frac{\partial \psi_s}{\partial x} \tag{3.3}$$

$$I_{diff} = \mu \cdot W \cdot \phi_T \cdot \frac{\partial Q_{inv}}{\partial x} \tag{3.4}$$

From the above equations it is clear that $\partial V / \partial \psi_s$ can be written as:

$$\frac{\partial V}{\partial \psi_s} = 1 - \frac{\phi_T}{Q_{inv}} \cdot \frac{\partial Q_{inv}}{\partial \psi_s} = \frac{Q^*_{inv}}{Q_{inv}} \tag{3.5}$$

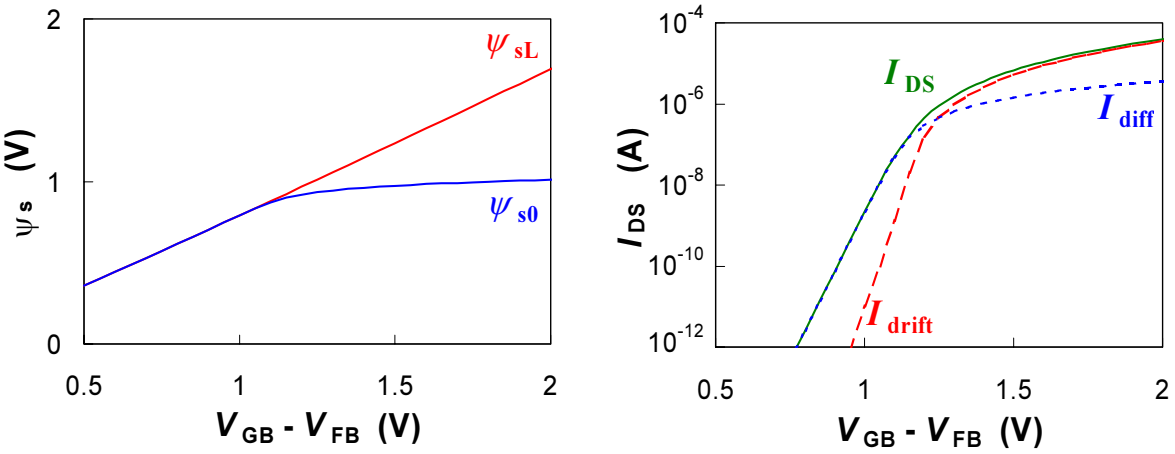
where the variable Q^*_{inv} is defined as:

$$Q^*_{inv} = Q_{inv} - \phi_T \cdot \frac{\partial Q_{inv}}{\partial \psi_s} = Q_{inv} + \phi_T \cdot C_{inv} = \bar{Q}^*_{inv} - C_{inv} \cdot (\psi_s - \bar{\psi}) \tag{3.6}$$

Here, $\bar{Q}^*_{inv}$ corresponds to $\bar{V}^*_{G_T}$ in the MM11 expressions ($\bar{V}^*_{G_T} = -\bar{Q}^*_{inv}/C_{ox}$), and it is simply given by:

$$\bar{Q}^*_{inv} = \bar{Q}_{inv} + \phi_T \cdot C_{inv} \tag{3.7}$$

Since the mobile charge density Q_{inv} is only non-zero in the inversion region, we will concentrate on the condition $V^*_{GB} > 0$ for the calculation of I_{DS}. Considering that I_{DS} is constant along the channel



(a) (b)

Figure 3.1: (a) The calculated surface potential at the source (ψ_{s_0}) and the drain (ψ_{s_L}) side, and (b) the corresponding channel current I_{DS} ($= I_{drift} + I_{diff}$) as a function of effective gate bias V^*_{GB}. (n-MOS, $\beta = 140\mu A/V^2$, $N_A = 2 \times 10^{23} m^{-3}$, $N_P \to \infty$, $t_{ox} = 3.2nm$, $V_{DB} = 1V$ and $V_{SB} = 0V$)

(i.e., $\partial I_{DS}/\partial x = 0$), and integrating (3.1) from the source end of the channel ($x = 0$) to the drain end ($x = L$), using (3.5), we can calculate the channel current:

$$I_{DS} = -\frac{W}{L} \cdot \int_{\psi_{s_0}}^{\psi_{s_L}} \mu \cdot Q^*_{inv} \cdot d\psi_s \tag{3.8}$$

Assuming for the moment that the mobility μ is constant along the channel (i.e., $\mu = \mu_0$), the above integral can be solved, resulting in:

$$I_{DS} = -\mu_0 \cdot \frac{W}{L} \cdot \bar{Q}^*_{inv} \cdot \Delta\psi = -\beta \cdot \frac{\bar{Q}^*_{inv}}{C_{ox}} \cdot \Delta\psi \tag{3.9}$$

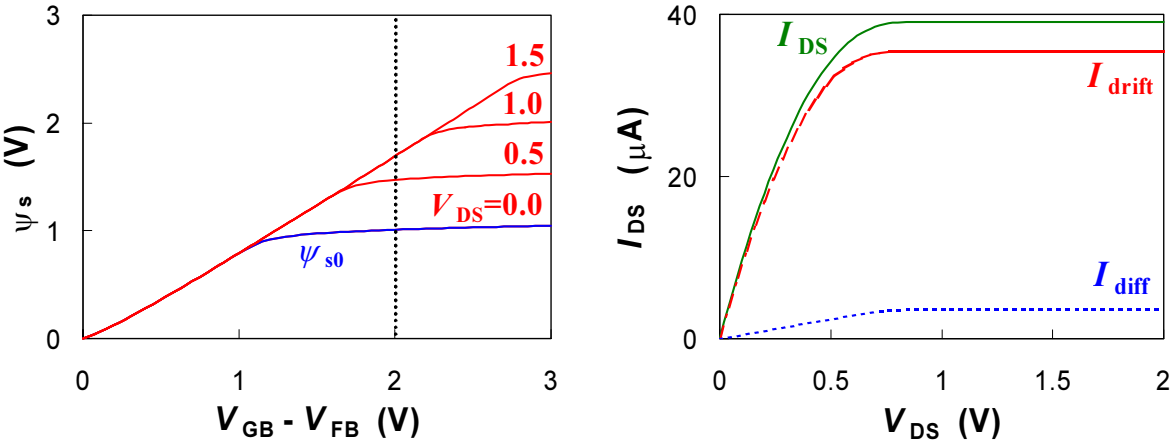
where β is the so-called gain factor given by $\mu_0 \cdot C_{ox} \cdot W/L$. Using (3.7), we can again make a distinction between a drift component:

$$I_{drift} = -\beta \cdot \frac{\bar{Q}_{inv}}{C_{ox}} \cdot \Delta\psi \tag{3.10}$$

and a diffusion component:

$$I_{diff} = -\beta \cdot \phi_T \cdot \frac{C_{inv}}{C_{ox}} \cdot \Delta\psi = \beta \cdot \phi_T \cdot \frac{\Delta Q_{inv}}{C_{ox}} \tag{3.11}$$

The channel current and its components have been plotted as a function of gate bias in Fig. 3.1. At low gate bias, both source and drain are in weak inversion. In this bias region, the so-called *subthreshold region*, the channel current is mainly due to diffusion. On the other hand, at high gate bias, the source is in strong inversion, and the drift component is dominant. This bias region is commonly referred to as the *superthreshold region*. In the transition region between subthreshold and superthreshold both drift and diffusion are of importance, this region is often referred to as *the moderate inversion region* [34].
The channel current and its components have been plotted as a function of drain bias in Fig. 3.2.



(a) (b)

Figure 3.2: (a) The calculated surface potential at the source (ψ_{s_0}) and the drain (ψ_{s_L}) side as a function of effective gate bias V^*_{GB} for various values of drain-source bias V_{DS}, and (b) the corresponding channel current I_{DS} ($= I_{drift} + I_{diff}$) as a function of drain-source bias V_{DS}. (n-MOS, $\beta = 140\mu A/V^2$, $N_A = 2 \times 10^{23} m^{-3}$, $t_{ox} = 3.2nm$, $V^*_{GB} = 2V$ and $V_{SB} = 0V$)

Here the MOSFET is biased in the superthreshold region, and consequently the source is in strong inversion. At low drain-source bias V_{DS}, the drain is in strong inversion as well. As a result, $\Delta\psi$ is approximately equal to V_{DS} and the channel current I_{DS} increases almost linearly with V_{DS}. In this case, the MOSFET acts as a (gate-bias dependent) resistor between drain and source, and as a consequence, this bias region is referred to as the *ohmic* or *linear region*. At high V_{DS}, the drain is in weak inversion. The surface potential at the drain ψ_{s_L} is pinned to a value independent of V_{DS}, and as a consequence, I_{DS} becomes independent of V_{DS} as well. In other words I_{DS} saturates for V_{DS} above a certain saturation voltage $V_{DS_{sat\infty}}$, hence the name *saturation region* for this bias region. Saturation occurs when the channel at the drain side is pinched off, i.e., the inversion-layer charge density at the drain Q_{inv_L} approaches zero[3]. In the superthreshold region, the saturation voltage $V_{DS_{sat\infty}}$ can be calculated by approximating ψ_{s_L} by $\phi_B + V_{SB} + V_{DS_{sat\infty}}$ at the onset of saturation. As mentioned above, in the saturation region ψ_{s_L} is pinched at its weak-inversion value ψ_{sat}, see Appendix C, and as a result we can simply write:

$$V_{DS_{sat\infty}} = \psi_{sat} - \phi_B - V_{SB} \tag{3.12}$$

where ψ_{sat} is given by (C.10), repeated here for completeness' sake:

$$\psi_{sat} = \left(\frac{\sqrt{P_D \cdot \left(V^*_{GB} + \Delta_{acc}\right) + {k_0}^2/4} - k_0/2}{P_D} \right)^2 - \Delta_{acc} \tag{3.13}$$

where $P_D = 1 + (k_0/k_P)^2$. Note that ψ_{sat} is independent of V_{DS}.

The description of channel current I_{DS} derived in this Section is only valid for an ideal long-channel MOS transistor. In order to get an accurate description for practical devices, several physical effects have to be taken into account: short-channel effects in subthreshold such as drain-induced barrier

[3] In reality, however, as Q_{inv_L} approaches zero, the lateral electric field $-\partial\psi_s/\partial x$ increases to very high values resulting in a breakdown of the gradual channel approximation and the occurrence of velocity saturation, see Section 3.3.2

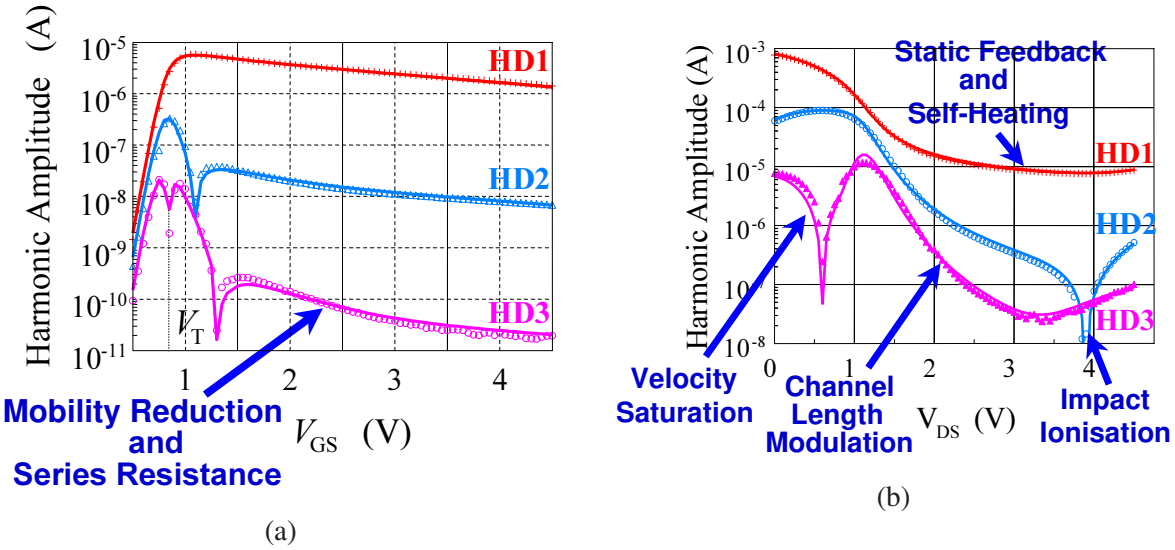


Figure 3.3: Measured (symbols) and modelled (lines) harmonic distortion in drain current I_D (see Fig. 1.2) for a sinusoidal input signal ($v_{in} = \hat{v} \cdot \sin(2 \cdot \pi \cdot f)$ with $f = 1$kHz) applied to (a) gate terminal ($\hat{v} = 50$mV) as a function of dc gate bias V_{GS} (n-MOS, $W/L = 10\mu$m$/1\mu$m, $V_{DS} = 0.1$V and $V_{SB} = 0.0$V) and (b) drain terminal ($\hat{v} = 0.5$V) as a function of dc drain bias V_{DS} (n-MOS, $W/L = 10\mu$m$/1\mu$m, $V_{GS} = 2.5$V and $V_{SB} = 0.0$V). The modelled results are obtained with MM11. The various physical effects that determine the distortion behaviour in different regions of operation are indicated.

lowering (DIBL), mobility effects such as mobility reduction and velocity saturation, the influence of series-resistance and conductance effects, such as channel length modulation, static feedback and self-heating. In MOS Model 11, special attention has been paid to the implementation of the above effects so that an accurate description of distortion behaviour is obtained, see Fig. 3.3[4]. In addition, the implementation requires special attention with regard to drain-source symmetry as well. In the following Sections, the above-mentioned physical effects will be discussed.

3.2 Subthreshold Current

The diffusion current I_{diff}, given by (3.11), is dominant in the subthreshold region, where both ψ_{s_0} and ψ_{s_L} are approximately equal to ψ_{sat}, see Fig. 3.1. This implies that even a very small error in the values of ψ_{s_0} and ψ_{s_L} will result in a large relative error in the difference $\Delta\psi$ on which (3.11) relies. The above difficulty can be circumvented by rewriting Q_{inv} in (2.16) to:

$$Q_{inv} = \frac{-k_0 \cdot C_{ox} \cdot \phi_T \cdot \exp\left(-\frac{V+\phi_B}{m_S \cdot \phi_T}\right) \cdot \left[\exp\left(\frac{\psi_s}{m_S \cdot \phi_T}\right) - 1\right]}{\sqrt{\psi_s + \Delta_{acc} + \phi_T \cdot \exp\left(-\frac{V+\phi_B}{m_S \cdot \phi_T}\right) \cdot \left[\exp\left(\frac{\psi_s}{m_S \cdot \phi_T}\right) - 1\right]} + \sqrt{\psi_s + \Delta_{acc}}} \tag{3.14}$$

which corresponds to variable V_{inv} in the MM11 expressions ($V_{inv} = -Q_{inv}/C_{ox}$). The diffusion component I_{diff} can now simply be written as:

$$I_{diff} = \beta \cdot \phi_T \cdot \frac{Q_{inv_L} - Q_{inv_0}}{C_{ox}} \tag{3.15}$$

[4]The phenomenon of impact ionisation will be treated in Section 4.1.

so that simply assuming both ψ_{s_0} and ψ_{s_L} are equal to ψ_{sat} in the subthreshold region leads to accurate results. An approximate expression of the subthreshold current reads:

$$I_{DS} \approx \frac{\beta \cdot {\phi_T}^2 \cdot k_0}{2 \cdot \sqrt{\psi_{sat} + \Delta_{acc}}} \cdot \exp\left(\frac{\psi_{sat} - V_{SB} + \phi_B}{m_S \cdot \phi_T}\right) \cdot \left[1 - \exp\left(-\frac{V_{DS}}{m_S \cdot \phi_T}\right)\right] \tag{3.16}$$

which clearly shows that I_{DS} has an exponential dependence on ψ_{sat} and consequently on V_{GB}, and that the influence of drain bias V_{DS} is only noticeable at very low values of V_{DS}. A value V_{limit} above which the subthreshold I_{DS} hardly changes with V_{DS} can be defined, the saturation voltage. Its choice is a bit arbitrary but V_{limit} has been chosen to be equal to $4 \cdot \phi_T$.
In short-channel devices, the gradual channel approximation no longer holds, i.e., $\partial^2\psi/\partial x^2$ is no longer negligible w.r.t. $\partial^2\psi/\partial y^2$. As a result the subthreshold behaviour starts to deviate from the ideal long-channel behaviour, this is ascribed to so-called *short-channel effects*. These effects result in an increase in the so-called *subthreshold swing* and an increase in the channel current I_{DS} with V_{DS}, which cannot be explained by the long-channel expression (3.16). The latter effect is commonly referred to as the drain-induced barrier lowering (DIBL) effect.

Subthreshold Swing: An important measure in subthreshold is the subthreshold swing S_{swing}, defined as the change in gate bias V_{GB} required to reduce the subthreshold I_{DS} by one decade [36]:

$$S_{swing} = \frac{\partial V_{GB}}{\partial \log(I_{DS})} \approx \ln(10) \cdot \frac{I_{DS}}{g_m} \tag{3.17}$$

The subthreshold swing gives a measure for how much one can decrease I_{DS} by decreasing the gate bias. Using (3.16) and assuming the change in $1/\sqrt{\psi_{sat} + \Delta_{acc}}$ with V_{GB} is negligible with respect to the exponential dependence of I_{DS} on V_{GB}, we can write for S_{swing}:

$$S_{swing} = \ln(10) \cdot \phi_T \cdot m_S \cdot \frac{\partial V_{GB}}{\partial \psi_{sat}} \tag{3.18}$$

In the ideal case, $m_S = 1$ and the subthreshold swing is determined by $\partial\psi_{sat}/\partial V_{GB}$, see Fig. 3.4. As can be seen, the subthreshold swing is bias dependent and always higher than $\ln(10) \cdot \phi_T \approx$ 60mV/decade (at room temperature). In short-channel MOSFETs, however, S_{swing} has been found to considerably exceed the ideal S_{swing}, which is due to the breakdown of the gradual channel approximation. An analytical calculation of the subthreshold behaviour in the short-channel case is complicated, but, as (3.18) indicates, an increase in S_{swing} can be empirically incorporated by using parameter m_S ($m_S > 1$), see Fig. 3.4.

Drain-Induced Barrier Lowering: For short-channel devices in subthreshold, an increase in drain current with drain-source bias V_{DS} is empirically observed that cannot be explained by the long-channel current expression (3.16). This effect is called drain-induced barrier lowering (DIBL) [47]. It is often explained by the fact that, as V_{DS} increases, the drain depletion region moves closer to the source depletion region, resulting in a significant field penetration from the drain into the source. Due to this field penetration, the potential barrier at the source is lowered, resulting in an increased injection of electrons by the source, giving rise to an increased drain current. The effect is schematically shown in Fig. 3.5.
In conventional V_T-based models, DIBL is often modelled by a linear decrease in threshold voltage $\Delta V_{T_{dibl}}$ with V_{DS}. In ψ_s-based models, however, it is more convenient to replace the gate bias V_G by an effective gate bias $V_G + \Delta V_{G_{dibl}}$, where $\Delta V_{G_{dibl}}$ is positive and equal to $\Delta V_{T_{dibl}}$.
Different models have been developed to calculate the DIBL effect [48]-[50]; these models concentrate on the calculation of the surface potential minimum and translating this in an effective decrease

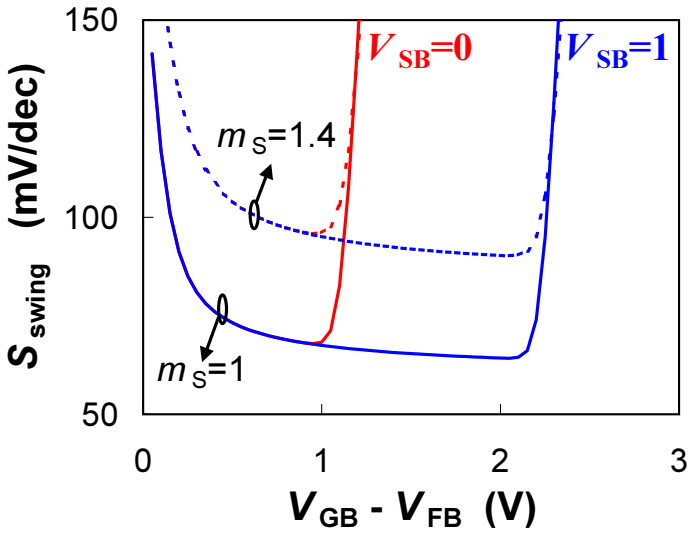


Figure 3.4: The simulated subthreshold swing S_{swing} (as given by (3.17)) as a function of gate bias for the ideal (long-channel) case ($m_S = 1.0$) and for the non-ideal (short-channel) case ($m_S = 1.4$). (n-MOS, $\beta = 140\mu\text{A/V}^2$, $N_A = 2\times10^{23}\text{m}^{-3}$, $N_P \to \infty$, $t_{ox} = 3.2\text{nm}$ and $V_{DS} = 1\text{V}$)

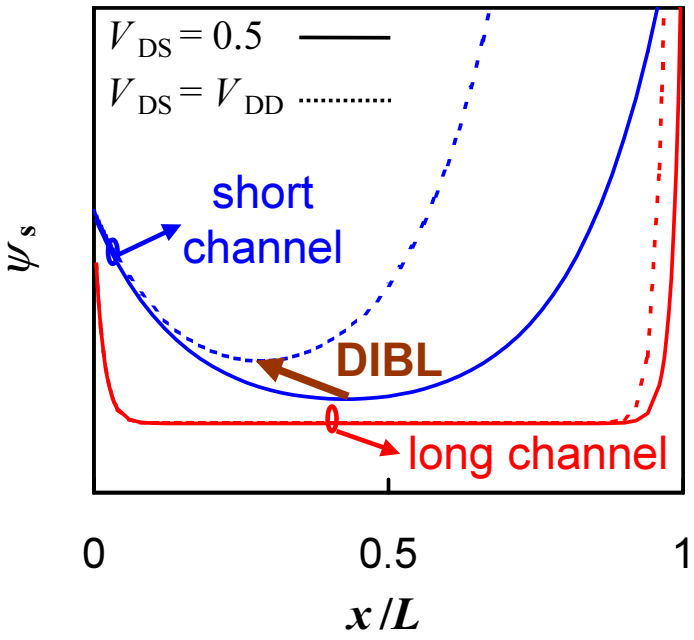


Figure 3.5: Surface potential ψ_s as a function of position along the channel for a MOSFET biased in the subthreshold region. Only channel length L and drain bias V_{DS} are varied. The effect of drain-induced barrier lowering (DIBL) is indicated. DIBL results in an increase of minimum ψ_s with V_{DS} at constant V_{GS}, it is important for short-channel transistors.

of threshold voltage. A pseudo two-dimensional analysis [48] predicts an exponential dependence of $\Delta V_{T_{dibl}}$ on channel length, whereas the voltage doping transformation [49, 50] predicts a $1/L^2$-dependence. Especially the latter model appears to give accurate results. Here, the substrate doping concentration N_A is replaced by an effective concentration N_{eff} [49]:

$$N_{eff} = N_A - \frac{2 \cdot \epsilon_{Si} \cdot V_{DS}^*}{q \cdot L^2} \tag{3.19}$$

where

$$\begin{aligned} V_{DS}^* &= V_{DS} + 2 \cdot (V_{bi} - \phi_B) + 2 \cdot \sqrt{(V_{bi} - \phi_B) \cdot (V_{DS} + V_{bi} - \phi_B)} \\ &\approx 2 \cdot V_{DS} + 4 \cdot (V_{bi} - \phi_B) \end{aligned} \tag{3.20}$$

Here, V_{bi} is the built-in potential between the n^+-type source/drain extension and the p-type substrate, given by $\phi_T \cdot \ln\left(N_A \cdot N_J / {n_i}^2\right)$, and N_J is the source/drain-extension doping concentration. Note that V_{DS}^* consists of a V_{DS}-dependent part and a V_{DS}-independent part. Introducing (3.19) in the description of k_0, we can calculate the change in threshold voltage ΔV_T [49, 50]:

$$\Delta V_T = k_0 \cdot \sqrt{V_{SB} + \phi_B} \cdot \left(1 - \sqrt{\frac{N_{eff}}{N_A}}\right) \approx k_0 \cdot \sqrt{V_{SB} + \phi_B} \cdot \frac{\epsilon_{Si} \cdot V_{DS}^*}{q \cdot N_A \cdot L^2} \tag{3.21}$$

The threshold voltage difference consists of a V_{DS}-independent part and a V_{DS}-dependent part. The former gives a channel length dependence of threshold voltage, and typically results in a sharp decrease in threshold voltage for short-channel MOSFETs, the so-called V_T roll-off. In MM11, this has simply been taken into account in the length scaling of body factor k_0 and intrinsic potential ϕ_B. The V_{DS}-dependent part of the above ΔV_T is DIBL, and we can write for $\Delta V_{G_{dibl}}$:

$$\Delta V_{G_{dibl}} = \Delta V_{T_{dibl}} = k_0 \cdot \sqrt{V_{SB} + \phi_B} \cdot \frac{2 \cdot \epsilon_{Si} \cdot V_{DS}}{q \cdot N_A \cdot L^2} = \sigma_{dibl} \cdot \sqrt{V_{SB} + \phi_B} \cdot V_{DS} \tag{3.22}$$

where σ_{dibl} is a channel length dependent empirical parameter. Note that, at higher back bias values V_{SB}, the influence of the DIBL effect becomes more important.

It should be noted that in a compact model the denotations S and D are interchanged for $V_{DS} < 0$ V, and as a result $\Delta V_{G_{dibl}}$ given by (3.22) is effectively dependent on the absolute value $|V_{DS}|$. In other words, when using (3.22), symmetry with respect to source and drain at $V_{DS} = 0$ V is not preserved. Realizing that the influence of DIBL is only noticeable at high drain bias, we can replace V_{DS} in (3.22) by an effective drain-source bias $V_{DS_{eff}}$:

$$V_{DS_{eff}} = \frac{{V_{DS}}^4}{\left({V_{limit}}^2 + {V_{DS}}^2\right)^{3/2}} \tag{3.23}$$

The above expression has been chosen in such a way that $\partial^i \Delta V_{G_{dibl}} / \partial {V_{DS}}^i = 0$ at $V_{DS} = 0$ V up to $i = 3$. As a result, it ensures continuity of I_{DS} at $V_{DS} = 0$ V up to the third-order derivative.

3.3 Mobility Effects

Up till now, the expressions for the channel current have been derived under the assumption that the carrier mobility in the inversion layer is constant. In reality, however, this is not true. Carriers in the

channel undergo increased scattering with increasing fields when they move under the influence of the normal electric field E_y (i.e., $-\partial\psi/\partial y$) and the lateral electric field E_x (i.e., $-\partial\psi/\partial x$) due to the gate bias V_{GS} and the drain bias V_{DS}, respectively. The dependence of mobility on normal electric field is often referred to as *mobility reduction*, whereas the dependence on lateral electric field is often referred to as *velocity saturation*.

3.3.1 Mobility Reduction

Although the mobility reduction model of MM11 has already been extensively treated in [4]-[7], in this section, for completeness' sake, a brief overview of the model derivation underlying the mobility reduction model will be given. For a more extensive and general overview of inversion layer mobility, the reader is referred to the large bulk of literature published on this matter [51]-[70].
In a MOSFET structure, the normal electric field restricts the channel to a sheet layer in which two-dimensional confinement effects and scattering cause the mobility to depend on bias conditions. It has been found that the inversion layer mobility follows a universal curve independent of the substrate bias V_{SB}, the substrate impurity concentration N_A or the gate oxide thickness t_{ox} when plotted as a function of the effective normal electric field E_{eff}, defined by [54]-[62]:

$$E_{eff} = -\frac{Q_b + \eta \cdot Q_{inv}}{\epsilon_{Si}} \tag{3.24}$$

In most papers for an inversion layer on a (100) oriented surface, η is taken to be $1/2$ for electrons and $1/3$ for holes. Generally, however, η is dependent on device process technology (e.g., doping profile, threshold voltage implant) [54], temperature [58] and surface orientation [62].
With the scaling down of MOS transistors towards the requisite deep submicron dimensions, which involves the use of thinner gate oxide and higher substrate doping, the thickness of the inversion layer becomes in the order of a few Å, which is smaller than the De Broglie wavelength of the carriers [40]. As a result, channel mobility has to be treated quantum-mechanically, [52, 55, 60]. Quantum-mechanical self-consistent calculations show that energy subbands of electrons and holes are formed in the different energy valleys. The spacing of these subbands increases with increasing normal electric field, see Appendix D. Actual modelling of scattering processes in the inversion layer is very complex due to the quantum-mechanical nature of these processes and the fact that in most cases more than one subband is filled. Therefore for mobility, a simplified semi-empirical approach has been adopted. As has been done elsewhere [51]-[70], mobility can be described by considering three mechanisms which dominate the scattering of charge carriers in the inversion layer at the Si-SiO_2 interface, see Fig. 3.6:

Coulomb Scattering (μ_C): Charged centres near the Si-SiO_2 interface can be of the same charge type as the mobile inversion charge leading to Coulomb repulsion. This results in scattering, which is important for lightly inverted surfaces, high surface-charge densities or substrate doping concentrations, and less important for heavily inverted surfaces because of carrier screening. Empirically, it was found that Coulomb scattering limited mobility μ_C is given by [61, 70]:

$$\mu_C \propto -\frac{Q_{inv}}{N_A} \tag{3.25}$$

It is obvious that Coulomb scattering is not governed by a universal relation on E_{eff}, therefore at low Q_{inv}-values the universal law does not hold. The above type of scattering is mostly of influence in the subthreshold region. Here, I_{DS} is dominated by the exponential dependence of Q_{inv} on gate bias, and Coulomb scattering is therefore neglected for the purpose of this work.

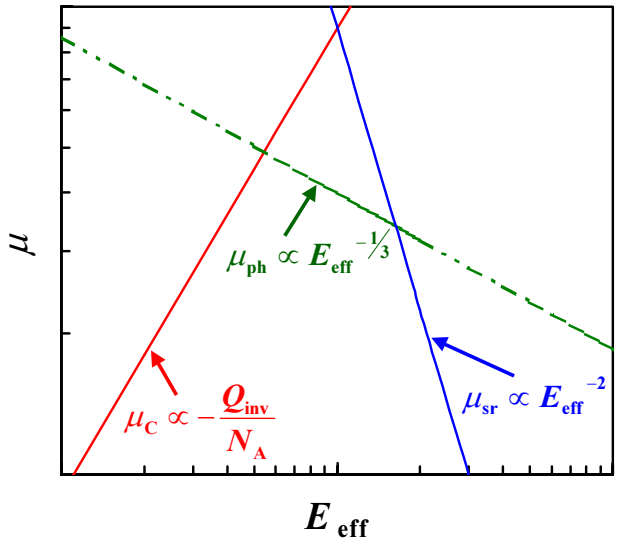


Figure 3.6: Electron mobility μ as a function of effective normal electric field E_{eff} on a double logarithmic scale. Three scattering mechanisms which determine mobility are indicated: Coulomb scattering (μ_{C}), phonon scattering (μ_{ph}) and surface roughness scattering (μ_{sr}).

Phonon Scattering (μ_{ph}): Surface phonons from the quantum vibrations of the crystal lattice scatter the mobile charge carriers. Under the assumption that carriers in the inversion layer only occupy the lowest subband, the mobility determined by acoustic phonon scattering is given by [60, 62]:

$$\mu_{\text{ph}} \propto \left(- \frac{11/32 \cdot Q_{\text{inv}} + Q_{\text{b}}}{\epsilon_{\text{Si}}} \right)^{-1/3} \tag{3.26}$$

Experimentally, it was found for both holes and electrons that:

$$\mu_{\text{ph}} \propto {E_{\text{eff}}}^{-1/3} \tag{3.27}$$

For electrons where $\eta = 1/2$ in (3.24), expression (3.27) deviates slightly from (3.26), which is ascribed to the fact that electrons occupy several subbands at intermediate values of E_{eff} [60]. The assumption of single subband occupation in this case does not hold. The lowest hole subband compared to the lowest electron subband quickly reaches a high occupancy at lower field strength.

Surface Roughness Scattering (μ_{sr}): The interface between the silicon crystal and the gate oxide is not atomically smooth. The above interface roughness scatters the mobile charge carriers. This type of scattering is especially important under strong inversion conditions, because the strength of the interaction is governed by the distance of the carriers from the surface; the closer the carriers are to the surface, the stronger the scattering due to surface roughness will be. For electrons, it was found experimentally that the mobility limited by surface roughness scattering μ_{sr} has the following dependence on effective field:

$$\mu_{\text{sr}} \propto {E_{\text{eff}}}^{-2} \tag{3.28}$$

For holes, on the other hand, it was found experimentally that:

$$\mu_{\text{sr}} \propto {E_{\text{eff}}}^{-1} \tag{3.29}$$

The difference between (3.28) and (3.29) for electrons and holes, respectively, is often ascribed to the fact that, at high transverse fields, holes tend to congregate further away from the interface than electrons do. The larger average distance leads to a reduced influence of the interface roughness and thus to less surface roughness scattering for holes.

In theory the above-described mechanisms can be incorporated into one channel mobility (μ_{eff}), using Matthiessen's rule:

$$\frac{1}{\mu_{\text{eff}}} = \frac{1}{\mu_0} + \frac{1}{\mu_{\text{sr}}} + \frac{1}{\mu_{\text{ph}}} \tag{3.30}$$

where μ_0 is the carrier mobility limited by ionized impurity scattering and acoustic phonon scattering in the bulk material. Equation (3.30) leads to:

$$\mu_{\text{eff}} = \frac{\mu_0}{1 + \left(\theta^*_{\text{ph}} \cdot E_{\text{eff}}\right)^{1/3} + \left(\theta^*_{\text{sr}} \cdot E_{\text{eff}}\right)^{\text{n}}} = \frac{\mu_0}{G_{\text{mob}}} \tag{3.31}$$

where θ^*_{ph} and θ^*_{sr} are empirical parameters, and $\text{n} = 2$ for electrons and $\text{n} = 1$ for holes. Strictly speaking, Matthiessen's rule is not valid, because it tacitly assumes that the momentum relaxation times due to the different scattering mechanisms have the same energy dependence. In order to correctly account for the various scattering sources, a weighted statistical averaging of the relaxation times should be performed. In [4]-[7] it has been found experimentally, based on 3rd-order distortion measurements in the linear region, that a more accurate description of mobility reduction is given by:

$$\mu_{\text{eff}} = \frac{\mu_0}{G_{\text{mob}}} = \frac{\mu_0}{1 + \left[\left(\theta^*_{\text{ph}} \cdot E_{\text{eff}}\right)^{\nu/3} + \left(\theta^*_{\text{sr}} \cdot E_{\text{eff}}\right)^{2\cdot\nu}\right]^{1/\nu}} \quad \text{for electrons,} \tag{3.32}$$

and:

$$\mu_{\text{eff}} = \frac{\mu_0}{G_{\text{mob}}} = \frac{\mu_0}{\left[1 + \left(\theta^*_{\text{ph}} \cdot E_{\text{eff}}\right)^{\nu/3} + \left(\theta^*_{\text{sr}} \cdot E_{\text{eff}}\right)^{\nu}\right]^{1/\nu}} \quad \text{for holes.} \tag{3.33}$$

In the above two equations, $\nu = 2$ at room temperature. Note that both (3.32) and (3.33) are variations on the Matthiessen's rule based expression (3.31), which goes to say that they both reproduce the ${E_{\text{eff}}}^{-1/3}$ dependence at low values of E_{eff} and the ${E_{\text{eff}}}^{-\text{n}}$ at high values of E_{eff}.

For an accurate implementation of the mobility reduction effect in the channel current expression, mobility μ in (3.10) and (3.11) has to be replaced by μ_{eff} as given by (3.32) or (3.33) for n-MOS and p-MOS, respectively. However, since the effective electric field E_{eff} is dependent on the inversion layer charge density Q_{inv} and bulk charge density Q_{b}, both E_{eff} and μ_{eff} are not constant along the channel. As a result, the integrals in (3.10) and (3.11) cannot be solved analytically. For reasons of simplicity, E_{eff} is assumed to be constant along the channel at the average surface potential $\psi_{\text{s}} = \bar{\psi}$, reducing (3.24) to:

$$E_{\text{eff}} = -\frac{\left(\bar{Q}_{\text{b}} + \eta \cdot \bar{Q}_{\text{inv}}\right)}{\epsilon_{\text{Si}}} \tag{3.34}$$

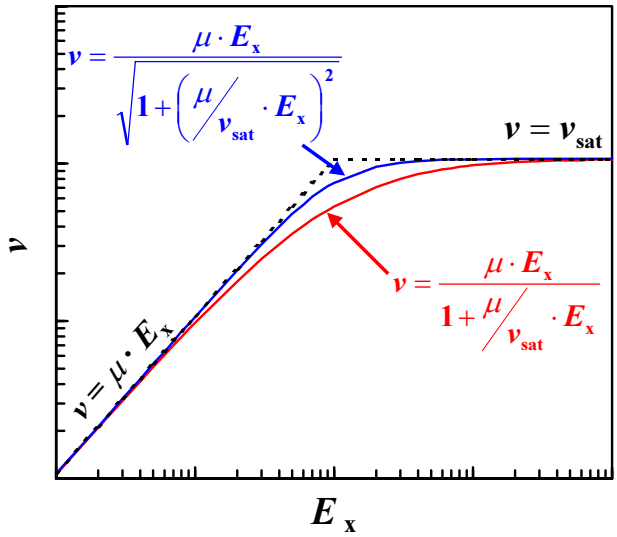


Figure 3.7: Electron drift velocity v versus lateral electric field E_x on a double logarithmic scale. The velocity is proportional to E_x at low E_x and approaches a constant value v_{sat} at high E_x. This type of behaviour is generally referred to as velocity saturation. The empirical relations (3.36) and (3.39) for electron velocity saturation are also shown.

The above method ensures preservation of drain-source symmetry at $V_{DS} = 0$. In both (3.10) and (3.11), we can now replace the constant mobility μ_0 (incorporated in gain factor β) by the effective mobility μ_{eff}:

$$I_{DS} = -\frac{\beta}{G_{mob}} \cdot \frac{\bar{Q}^*_{inv}}{C_{ox}} \cdot \Delta\psi \tag{3.35}$$

where G_{mob} is given by (3.32) and (3.33) in the case of n-type and p-type MOSFETs, respectively. In order to obtain the MM11 expression for mobility reduction, we need to replace the effective field E_{eff} in (3.34) by an effective voltage $V_{eff} = \epsilon_{Si} \cdot \eta \cdot E_{eff}/C_{ox}$. In addition, we can write for the MM11 mobility-parameters $\eta_{mob} = 1/\eta$, $\theta_{sr} = \theta^*_{sr}/\eta \cdot C_{ox}/\epsilon_{Si}$ and $\theta_{ph} = \theta^*_{ph}/\eta \cdot C_{ox}/\epsilon_{Si}$.

3.3.2 Velocity Saturation

With an increase in lateral electric field E_x, carriers gain sufficient energy to be scattered by optical phonons, resulting in a decrease of mobility μ_{eff} and eventually resulting in the saturation of drift velocity v. This limiting high-field drift velocity is referred to as the *saturation velocity* v_{sat}. Ideally, the carrier velocity can be described by $\mu_{eff} \cdot E_x$ at low values of E_x and by v_{sat} at high values of E_x, see Fig. 3.7. In most compact MOS models, the following empirical relation for velocity saturation is used [71]:

$$v = \frac{\mu_{eff} \cdot \frac{\partial \psi_s}{\partial x}}{1 + \frac{\mu_{eff}}{v_{sat}} \cdot \left|\frac{\partial \psi_s}{\partial x}\right|} \tag{3.36}$$

The use of this relation, however, does not result in a symmetrical description w.r.t. source and drain (at $V_{DS} = 0$), and it furthermore does not give an accurate description of MOSFET distortion behaviour [6, 7]. For a more accurate and symmetrical description of velocity saturation, an adjusted

form of the Scharfetter-Gummel expression [72, 73] can be used [6, 7]:

$$v = \frac{\mu_{\text{eff}} \cdot \frac{\partial \psi_s}{\partial x}}{\sqrt{1 + \left[\left(\frac{\mu_{\text{eff}}}{v_c} \cdot \frac{\partial \psi_s}{\partial x} \right)^2 \Big/ \sqrt{G^2 + \left(\frac{\mu_{\text{eff}}}{v_c} \cdot \frac{\partial \psi_s}{\partial x} \right)^2} \right] + \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot \frac{\partial \psi_s}{\partial x} \right)^2}} \tag{3.37}$$

where v_c is a parameter corresponding to the velocity of the longitudinal acoustic phonons and G is a fitting parameter. Equation (3.37) is valid for both electrons and holes; the second term in the denominator describes the hole velocity behaviour (i.e., $\mu_{\text{eff}}/v_{\text{sat}} \approx 0$ for holes):

$$v = \frac{\mu_{\text{eff}} \cdot \frac{\partial \psi_s}{\partial x}}{\sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_c} \cdot \frac{\partial \psi_s}{\partial x} \right)^2 \Big/ \sqrt{G^2 + \left(\frac{\mu_{\text{eff}}}{v_c} \cdot \frac{\partial \psi_s}{\partial x} \right)^2}}} \quad \text{for holes} \tag{3.38}$$

whereas the third term describes the electron velocity behaviour (i.e., $\mu_{\text{eff}}/v_c \approx 0$ for electrons):

$$v = \frac{\mu_{\text{eff}} \cdot \frac{\partial \psi_s}{\partial x}}{\sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot \frac{\partial \psi_s}{\partial x} \right)^2}} \quad \text{for electrons} \tag{3.39}$$

For the moment, let us concentrate on the influence of velocity saturation on I_{DS} in n-MOSFETs. Using (3.39), we can rewrite eqs. (3.2)-(3.4) to:

$$I_{\text{DS}} = -\frac{\mu_{\text{eff}} \cdot W \cdot Q^*_{\text{inv}}}{\sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot \frac{\partial \psi_s}{\partial x} \right)^2}} \cdot \frac{\partial \psi_s}{\partial x} \tag{3.40}$$

The above differential equation can be solved analytically, which results in an implicit relation for I_{DS}, see Appendix F.1. Here, however, we will use a simplified yet accurate method to solve (3.40). Integrating from source to drain, we can rewrite (3.40) to:

$$I_{\text{DS}} \cdot \int_0^L \sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot \frac{\partial \psi_s}{\partial x} \right)^2} \cdot \mathrm{d}x = -\mu_{\text{eff}} \cdot W \cdot \int_{\psi_{s_0}}^{\psi_{s_L}} Q^*_{\text{inv}} \cdot \mathrm{d}\psi_s \tag{3.41}$$

where the integral in the right-hand side has already been solved in Section 3.1. The impact of velocity saturation on I_{DS} can now be described by:

$$I_{\text{DS}} = -\frac{\beta}{G_{\text{vsat}}} \cdot \frac{\bar{Q}^*_{\text{inv}}}{C_{\text{ox}}} \cdot \Delta\psi \tag{3.42}$$

where:

$$G_{\text{vsat}} = \frac{G_{\text{mob}}}{L} \cdot \int_0^L \sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot \frac{\partial \psi_s}{\partial x} \right)^2} \cdot \mathrm{d}x \tag{3.43}$$

A first-order approximation of the integral in (3.43) can be obtained by assuming that the lateral electric field $-\partial\psi_s/\partial x$ is constant along the channel and equal to $\Delta\psi/L$ [6, 7, 26]. In this case, G_{vsat} is simply written as:

$$G_{\text{vsat}} = \sqrt{{G_{\text{mob}}}^2 + (\theta_{\text{sat}} \cdot \Delta\psi)^2} \tag{3.44}$$

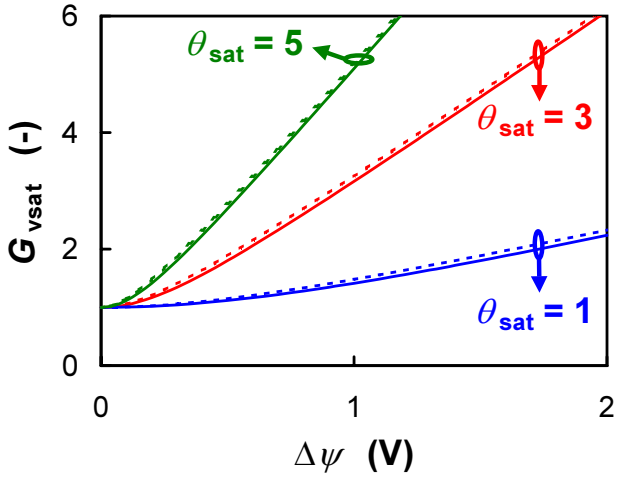


Figure 3.8: Velocity saturation expression G_{vsat} as a function of potential drop $\Delta\psi$ for various values of θ_{sat}. Solid lines and dashed lines indicate results given by (3.44) and (3.45), respectively. For the sake of implicity, the influence of G_{mob} has been neglected, i.e., $G_{\mathrm{mob}} = 1$.

where parameter θ_{sat} is equal to $\mu_0/(v_{\mathrm{sat}} \cdot L)$ in theory, but is considered as an empirical parameter for practical puproses. The above description of velocity saturation has been found to give an accurate description of drain-induced third-order harmonic distortion [6, 7]. Still, a slightly better approximation of (3.43) is obtained by assuming that the lateral electric field $-\partial\psi_{\mathrm{s}}/\partial x$ increases linearly along the channel from 0 at the source to $2 \cdot \Delta\psi/L$ at the drain (i.e., $-\partial\psi_{\mathrm{s}}/\partial x = 2 \cdot \Delta\psi/L^2 \cdot x$). In this case, G_{vsat} is given by:

$$G_{\mathrm{vsat}} = \frac{G_{\mathrm{mob}}}{2} \cdot \left[\sqrt{1+\Gamma^2} + \frac{\ln\left(\Gamma + \sqrt{1+\Gamma^2}\right)}{\Gamma} \right] \tag{3.45}$$

where:

$$\Gamma = \frac{2 \cdot \theta_{\mathrm{sat}} \cdot \Delta\psi}{G_{\mathrm{mob}}} \tag{3.46}$$

In Fig. 3.8, the difference between (3.45) and (3.44) is shown. It is clear that the difference is only marginal, nevertheless, the use of (3.45) results in a slightly better description of transconductance at high drain bias. In addition, it has been found empirically that a further improvement in accuracy can be obtained by replacing (3.46) by:

$$\Gamma = \frac{2 \cdot \theta_{\mathrm{sat}} \cdot \Delta\psi}{\sqrt{G_{\mathrm{mob}}}} \tag{3.47}$$

which results in a slightly different gate bias dependency. In MM11, the channel current expression including velocity saturation is given by (3.42), where the electron velocity saturation expression is given by (3.45) and (3.47).

For p-type MOSFETs, we have to use (3.38) instead of (3.39). In this case, the integration along the channel is less straightforward. For simplicity's sake, we approximate (3.38) by:

$$v = \frac{\mu_{\text{eff}} \cdot \frac{\partial \psi_s}{\partial x}}{\sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_c} \cdot \frac{\partial \psi_s}{\partial x}\right)^2} \Big/ \sqrt{G^2 + \left(\frac{\mu_0}{v_c} \cdot \frac{\Delta\psi}{L}\right)^2}} \quad \text{for holes} \tag{3.48}$$

and as a consequence, for holes we can simply use (3.45) where:

$$\Gamma = \frac{2 \cdot \theta_{\text{sat}} \cdot \Delta\psi}{\sqrt{G_{\text{mob}}} \cdot \left[G^2 + {\theta_{\text{sat}}}^2 \cdot \Delta\psi^2\right]^{\frac{1}{4}}} \tag{3.49}$$

Here, parameter θ_{sat} is equal to $\mu_0/(v_c \cdot L)$ in theory, but is considered as an empirical parameter for practical purposes. Parameter G has been found to be of minor influence, and is taken equal to 1.
In other words, the n-type MOSFET equations can simply be used for a p-type MOSFET when θ_{sat} is replaced by $\theta_{\text{sat}}/(1 + {\theta_{\text{sat}}}^2 \cdot \Delta\psi^2)^{1/4}$.

In very short-channel devices, the channel length may become comparable to the mean free path of the carriers, the carrier transport may become quasi-ballistic and as a result, carrier velocity may locally reach a value larger than the saturation velocity v_{sat}. This effect is called *velocity overshoot* [74]. In this case, the maximum velocity v_{sat} becomes dependent on the device structure, particularly on the channel length L [75]. Consequently, the channel length dependence of parameter θ_{sat} will differ from the theoretical $1/L$-dependence. As a result, a more robust length-scaling relation is used for θ_{sat} in MM11 [17].
A more physical way of modelling the quasi-ballistic transport would be to use a scattering matrix based model [76]. However, this seems to be important only for very short channel lengths ($L <$ 10nm) [76], and has therefor so far been left out of account in MM11.

3.4 Series Resistance

The source and drain junction portion of a MOSFET are parasitic components. These n^+-p-junction region form a resistive and a capacitive element. Since source and drain are an essential part of the device, these parasitic elements cannot be eliminated. The capacitive element is taken into account in the junction diode model (JUNCAP), and as a consequence, attention will be focused on the resistive element. In general, the resistive element of the source/drain junction is assumed to be caused by a contact resistance, the sheet resistance of the heavily doped source/drain diffusion region and a spreading resistance due to current crowding. In this case, the element can be modelled by a constant resistance, which is inversely proportional with channel width W, although the above proportionality may no longer hold for narrow-width MOSFETs [77].
As the drain current may reach large values, especially for short-channel devices, the voltage drop across the source and drain series resistance is no longer negligible and has to be taken into account. The effect of the source and drain resistance R_S and R_D, respectively, on the channel current can be taken into account by using the equivalent circuit given in Fig. 3.9. From a viewpoint of computational efficiency, however, this is not a good solution, since two additional nodes per transistor have to be used in this case for circuit simulation. As a consequence, it is much more practical to include these resistances directly into the MOSFET device equations without adding additional nodes. Under

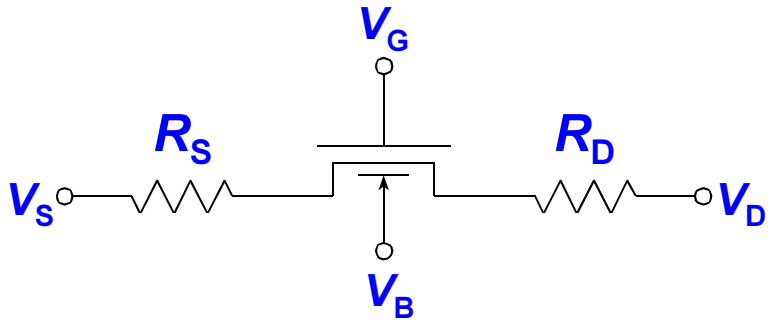


Figure 3.9: Equivalent circuit for a MOS transistor taking into account source (R_S) and drain (R_D) series resistance.

the assumption that source and drain resistance are equal (i.e., $R_S = R_D$), the general drain-source channel current expression becomes, see Appendix G:

$$I_{DS} = -\frac{\beta}{G_{tot}} \cdot \frac{\bar{Q}^*_{inv}}{C_{ox}} \cdot \Delta\psi \tag{3.50}$$

where:

$$G_{tot} = \frac{G_{vsat} + G_R}{2} \cdot \left[1 + \sqrt{1 - \frac{4 \cdot G_R / G_{vsat}}{(G_{vsat} + G_R)^2} \cdot \left({G_{vsat}}^2 - {G_{mob}}^2\right)}\right] \tag{3.51}$$

and:

$$G_R = -2 \cdot \beta \cdot \bar{Q}^*_{inv} / C_{ox} \cdot R_S \tag{3.52}$$

In addition, neglecting the influence of the diffusion component (i.e., $\bar{Q}^*_{inv} \approx \bar{Q}_{inv}$), we can write:

$$G_R = -\theta_R \cdot \bar{Q}_{inv} \,/ C_{ox} \tag{3.53}$$

where θ_R is a model parameter, theoretically given by $2 \cdot \beta \cdot R_S$. The impact of source/drain resistance is found in a reduction of the transconductance g_m and the device current driving capability.

In most compact MOSFET models, the series resistance is considered to be bias independent. In drain engineered devices such as Lightly Doped Drain (LDD) MOSFETs, however, the source and drain series resistance depends heavily on drain and gate bias[5] even for moderately doped junctions [80, 81]. For a correct description of distortion behaviour, this bias dependence has to be implemented in the drain current expression [5]. As can be seen in Fig. 3.10, the series resistance in the drain/source region consists of four components: a contact resistance R_{co}, a sheet resistance R_{sh}, a spreading resistance R_{sp} due to current crowding in the vicinity of the channel end, and an accumulation layer resistance R_{acc} [81]. In modern MOSFETs, both the sheet resistance R_{sh} and the spreading resistance R_{sp} are quite small due to the use of silicide. The accumulation layer resistance R_{acc} is due to overlap of the polysilicon gate on the drain/source region, where an accumulation layer is formed. The accumulation layer charge density Q_{ov} increases with increasing gate voltage, this results in a gate voltage dependent resistance R_{ac}. The contact resistance, sheet resistance and spreading resistance are independent of terminal voltages and can be written in terms of the device parameters; they are,

[5] Sometimes both series resistance and channel length are considered to be bias dependent [78], where the latter is the result of the channel broadening effect [79]. Since it is not possible to distinguish the series resistance from this broadening effect, here channel length is assumed to be constant and all bias dependency is included in the series resistance.

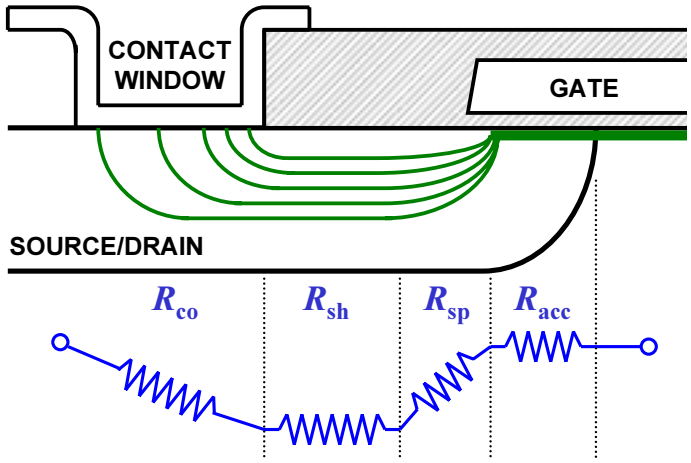


Figure 3.10: Schematic diagram showing current pattern in the source/drain region of a MOS transistor and the representative resistance components, consisting of a contact resistance R_{co}, a sheet resistance R_{sh}, a spreading resistance R_{sp} due to current crowding in the vicinity of the channel end, and an accumulation layer resistance R_{acc} [81].

furthermore, inversely proportional to channel width W. The complete gate-voltage dependent source resistance R_S can roughly be given by [82]:

$$R_S = R_{S_0} \cdot \left(1 + \frac{\theta_{R1}}{\theta_{R2} - \bar{Q}_{inv}/C_{ox}}\right) \tag{3.54}$$

where R_{S_0} is the bias-independent part of the series resistance inversely proportional to W, and parameters θ_{R1} and θ_{R2} determine the bias-dependent part. Using the above expression, we can rewrite (3.52) into:

$$G_R = -\theta_R \cdot \left(1 + \frac{\theta_{R1}}{\theta_{R2} - \bar{Q}_{inv}/C_{ox}}\right) \cdot \frac{\bar{Q}_{inv}}{C_{ox}} \tag{3.55}$$

where θ_R is now equal to $2 \cdot \beta \cdot R_{S_0}$. Finally, the MM11 expression for series resistance can be obtained by replacing $-\bar{Q}_{inv}/C_{ox}$ by $\bar{V}_{G_T}$.

3.5 Saturation Voltage

The calculation of surface potential so far has been based on ideal MOSFET operation, in which case the channel at the drain end pinches off when $\psi_{s_L} = \psi_{sat}$, in other words when the inversion-layer charge density at the drain Q_{inv_L} approaches zero[6]. Above threshold, this corresponds to an ideal drain-source saturation voltage $V_{DS_{sat\infty}} = \psi_{sat} - \phi_B - V_{SB}$. This so-called pinch-off behaviour is automatically included in a ψ_s-based model. For short-channel devices, however, the description of pinch-off is no longer realistic, owing to the fact that carriers reach velocity saturation even before the pinch-off condition is fulfilled. Furthermore, part of the drain-source voltage is dropped across the series resistances at both drain and source side, which has its effect on saturation voltage as well. In the non-ideal case, the saturation voltage $V_{DS_{sat}}$ may thus deviate significantly from the ideal case

[6] In reality, however, as Q_{inv_L} approaches zero, the lateral electric field $-\partial\psi_s/\partial x$ increases to very high values resulting in a breakdown of the gradual channel approximation and the occurence of velocity saturation, even for an infinitely long device.

$V_{DS_{sat\infty}}$. The saturation voltage $V_{DS_{sat}}$ is derived in Appendix H and is given by (H.17).
For the calculation of surface potential ψ_{s_L}, the difference in saturation voltage (i.e., $V_{DS_{sat}} - V_{DS_{sat\infty}}$) can be taken into account by clamping the drain-bulk voltage V_{DB} at an effective drain-bulk voltage $V_{DS_x} + V_{SB}$ [26]. Here, V_{DS_x} is a function which changes smoothly from V_{DS} in the ohmic region (i.e., for $V_{DS} < V_{DS_{sat}}$) to $V_{DS_{sat}}$ in the saturation region (i.e., for $V_{DS} > V_{DS_{sat}}$). The smooth transition from ohmic to saturation region can be obtained by using the smoothing function V_{DS_x} as defined in [26], see Appendix A.2:

$$V_{DS_x} = \frac{V_{DS} \cdot V_{DS_{sat}}}{\left[{V_{DS}}^{2 \cdot m} + {V_{DS_{sat}}}^{2 \cdot m} \right]^{\frac{1}{2 \cdot m}}} \tag{3.56}$$

where m is an empirical fitting parameter which determines the smoothness of the transition. In practice, m is taken equal to 8 for a long-channel device (i.e., $L = 10\mu$m), equal to 2 for a minimum channel length device (i.e., $L = L_{min}$), and $1/m$ is taken to scale linearly with channel length L. The surface potential at the drain end ψ_{s_L} can now simply be evaluated from the implicit relation (2.10) by replacing V by $V_{DS_x} + V_{SB}$.

3.6 Effects on Conductance

As was mentioned in Section 3.5, the transistor is in the saturation mode for a drain voltage V_{DS} larger than $V_{DS_{sat}}$. Here, the drain end of the channel is weakly inverted as opposed to the source end, resulting in a channel current I_{DS} that ideally is independent of drain voltage and a conductance g_{ds} ($= \partial I_{DS}/\partial V_{DS}$) that is equal to zero. As the inversion layer charge density Q_{inv_L} at the drain end becomes very small, the lateral electric field E_x increases to very high values owing to the continuity of current flow. This results in a saturation of carrier velocity and a breakdown of the gradual channel approximation at the drain end. Hence, two-dimensional effects such as *channel length modulation* and *static feedback* become apparent at the drain end, and as a result the channel current becomes dependent on drain voltage and the conductance g_{ds} becomes non-zero. Another phenomenon that results in a non-zero conductance in saturation, is *self-heating*.
In typical MOSFET amplifier structures, the MOS transistors are biased in saturation, and the amplification is determined by the ratio of transconductance over conductance g_m/g_{ds}. An accurate modelling of conductance is thus essential.

3.6.1 Channel Length Modulation

When V_{DS} is increased beyond $V_{DS_{sat}}$, the velocity saturation point moves towards the source, causing effectively that the channel length L is shortened by a length ΔL. This movement is referred to as channel length modulation[7], and it effectively causes the conductance to be non-zero in the saturation region. The so-called *electrical channel length* L_{elec} is given by:

$$L_{elec} = L \cdot \left(1 - \frac{\Delta L}{L}\right) = L \cdot G_{\Delta L} \tag{3.57}$$

Replacing L by L_{elec} in (3.50) and (3.52), we can rewrite G_{tot} so that it includes channel length modulation:

$$G_{tot} = \frac{G_{vsat} \cdot G_{\Delta L} + G_R}{2} \cdot \left[1 + \sqrt{1 - \frac{4 \cdot G_R/G_{vsat}}{(G_{vsat} \cdot G_{\Delta L} + G_R)^2} \cdot \left({G_{vsat}}^2 - {G_{mob}}^2\right)}\right] \tag{3.58}$$

[7] As a matter of fact, the definition of channel length modulation and saturation voltage $V_{DS_{sat}}$ is a consequence of the use of the gradual channel approximation, i.e., a one-dimensional solution of a two-dimensional system.

An exact calculation of ΔL is needed, but this is not easily accomplished, since it requires a two-dimensional solution of Poisson's equation for the channel saturation region at the drain end. This equation can only be solved using numerical techniques. Several methods have been proposed to circumvent this problem, but nowadays the pseudo two-dimensional model [83] has been widely accepted, since it has been shown to describe the channel length modulation accurately.
The pseudo two-dimensional analysis is based on the application of Gauss' law to a specific area at the drain end of the channel. Making some assumptions for this area the 2-D Poisson equation reduces to a one-dimensional differential equation, which can be solved analytically. The analysis is worked out in more detail in Appendix I. The channel length modulation can now be written as:

$$G_{\Delta L} = 1 - \frac{\Delta L}{L} = 1 - \alpha \cdot \ln\left[\frac{V_{DS} - V_{DS_{sat}} + \sqrt{\left(V_{DS} - V_{DS_{sat}}\right)^2 + {V_P}^2}}{V_P}\right] \tag{3.59}$$

where, in view of the numerous simplifications made to obtain this equation, α and V_P are considered as empirical parameters, and α is inversely proportional with channel length L. As a result, α will increase with decreasing channel length L, and the impact of channel length modulation will become more important. A smooth transition from the ohmic region (where $G_{\Delta L} = 1$) to the saturation region can be obtained by replacing $V_{DS_{sat}}$ in (3.59) by the smoothing function V_{DS_x}:

$$G_{\Delta L} = 1 - \frac{\Delta L}{L} = 1 - \alpha \cdot \ln\left[\frac{V_{DS} - V_{DS_x} + \sqrt{\left(V_{DS} - V_{DS_x}\right)^2 + {V_P}^2}}{V_P}\right] \tag{3.60}$$

This is the expression for channel length modulation as used in MM11.

3.6.2 Static Feedback

When the average distance between the conducting drain and the channel becomes small, an increase of the drain bias beyond saturation induces some excess mobile charge in the inversion layer. This electrostatic coupling between drain and channel region is often referred to as static feedback [84]-[87]. The increase of mobile charge can be estimated analytically by assuming that the field lines originating from the drain follow a cylindrical path [85]. This results in a linear increase of Q_{inv} with drain voltage V_{DS}, which is supported by two-dimensional device simulations [7]. This behaviour cannot be explained by equation (2.13). However, abandoning the gradual channel approximation and using again a pseudo two-dimensional analysis, the inversion layer charge Q_{inv} can be approximated by [88]:

$$Q_{inv}(x) = -C_{ox} \cdot \left[V^*_{GB} - \psi_s - \psi_p - k_0 \cdot \sqrt{\psi_s + \Delta_{acc}} + \frac{\kappa \cdot \epsilon_{Si}}{C_{ox}} \cdot X_{dep} \cdot \frac{\partial^2 \psi_s}{\partial x^2}\right] \tag{3.61}$$

where X_{dep} is the depletion layer width and κ is a parameter determined by the distribution of $\partial^2\psi_s/\partial x^2$ over y. As the gradient of the electric field along the channel $\partial^2\psi_s/\partial x^2$ is no longer negligible, its inclusion in (3.61) leads to an increase of inversion-layer charge density ΔQ_{sf} with increasing drain voltage V_{DS}. The channel current should be calculated from (3.1) introducing (3.61). Unfortunately, the resulting expression cannot be solved analytically, and as a result it is difficult to give a quantitative description of the static feedback effect.
It has been found, though, that the above increase ΔQ_{sf} can be modelled by an increase in effective gate bias $\Delta V_{G_{sf}}$ along the same lines as was done for the DIBL effect. The effective gate bias

shift $\Delta V_{G_{sf}}$ is determined by $\Delta Q_{sf}/C_{ox}$ averaged along the channel. Empirically, it has been found that $\Delta V_{G_{sf}}$ is given by [6, 7]:

$$\Delta V_{G_{sf}} = \sigma_{sf} \cdot \sqrt{V_{DS_{sat}}} \cdot V_{DS} \approx \sigma_{sf} \cdot \sqrt{V_{DS_{sat\infty}}} \cdot V_{DS} \tag{3.62}$$

where, in theory, σ_{sf} depends on oxide thickness, drain doping profile and substrate doping, but, in practice, it is considered as an empirical parameter. Experimentally it has been found that σ_{sf} is inversely proportional to the channel length L [86, 87]. Note that the effective gate bias shift $\Delta V_{G_{sf}}$ increases with gate bias.

Again, as in the case of the DIBL effect, the straightforward implementation of (3.62) would result in the breakdown of the model symmetry w.r.t. drain and source (at $V_{DS} = 0$ V). In order to preserve symmetry, V_{DS} can be replaced by the effective drain-source bias $V_{DS_{eff}}$ given by (3.23). To some extent, the static feedback phenomenon is similar to the DIBL effect, compare (3.22) and (3.62). Nevertheless, physically it is different since mobile carriers are induced instead of modifying the fixed depletion charge, therefore DIBL only occurs in the subthreshold region. Both effects can be combined in a single gate-bias shift ΔV_G, which is determined by a transition from DIBL in weak inversion to static feedback in strong inversion:

$$\Delta V_G = \begin{cases} \Delta V_{G_{dibl}} & \text{for:} \quad \Delta V_{G_{dibl}} > \Delta V_{G_{sf}} \\ \Delta V_{G_{sf}} & \text{for:} \quad \Delta V_{G_{dibl}} \leq \Delta V_{G_{sf}} \end{cases} \tag{3.63}$$

The above expression can be made C_∞-continuous by using hyp-smoothing functions, see Appendix A. DIBL and static feedback can now simply be taken into account by replacing the gate bias V_G in all model equations by the effective gate bias $V_G + \Delta V_G$.

3.6.3 Self-Heating

Although self-heating is more common in SOI devices, where conductance may even become negative due to this effect, it may also significantly affect the channel current I_{DS} in bulk submicron MOSFETs at high gate bias. The effective working temperature T in the inversion layer increases linearly with the dissipated power P_{dis} [89]-[91]:

$$T = T_{amb} + R_{Th} \cdot P_{dis} \tag{3.64}$$

where T_{amb} is the ambient temperature, and R_{Th} is the thermal resistance of the device, dependent on the device geometry[8]. The dissipated power P_{dis} is given by:

$$P_{dis} = I_{DS} \cdot V_{DS} \tag{3.65}$$

From (3.64), it is clear that an increase in drain bias results in an increase in intrinsic temperature, which naturally influences all kinds of electrical parameters. The parameter which is mostly affected, is the carrier mobility μ_0, while all other electrical parameters change much less with temperature. The thermal dependence of mobility is given by [36]:

$$\mu_0(T) = \mu_0(T_{amb}) \cdot \left(\frac{T}{T_{amb}} \right)^{-\eta_\beta} \tag{3.66}$$

[8] In theory, the thermal resistance R_{Th} decreases with increasing channel area $W \cdot L$

where η_β is different for electrons and holes and has a value between 1.5 and 2.0. Now, inserting (3.64) into (3.66) and making a first-order Taylor-series, we get:

$$\mu_0(V_{\rm DS}) \approx \mu_0(T_{\rm amb}) \cdot \left(1 - \frac{\eta_\beta \cdot R_{\rm Th} \cdot I_{\rm DS} \cdot V_{\rm DS}}{T_{\rm amb}}\right) \tag{3.67}$$

This equation can be incorporated into the channel current expression (3.50), using (3.58):

$$I_{\rm DS} = -\frac{\mu_0(T_{\rm amb}) \cdot W/L \cdot \bar{Q}^*_{\rm inv} \cdot \Delta\psi \cdot \left(1 - \frac{\eta_\beta \cdot R_{\rm Th} \cdot I_{\rm DS} \cdot V_{\rm DS}}{T_{\rm amb}}\right)}{\frac{G_{\rm vsat} \cdot G_{\Delta L} + G_{\rm R}}{2} \cdot \left[1 + \sqrt{1 - \frac{4 \cdot G_{\rm R}/G_{\rm vsat}}{(G_{\rm vsat} \cdot G_{\Delta L} + G_{\rm R})^2} \cdot \left({G_{\rm vsat}}^2 - {G_{\rm mob}}^2\right)}\right]} \tag{3.68}$$

which can be rewritten to a simple expression for $I_{\rm DS}$:

$$I_{\rm DS} = -\frac{\beta}{G_{\rm tot}} \cdot \frac{\bar{Q}^*_{\rm inv}}{C_{\rm ox}} \cdot \Delta\psi \tag{3.69}$$

Here, $G_{\rm tot}$ has been redefined so that it includes self-heating:

$$G_{\rm tot} = G_{\rm Th} + \frac{G_{\rm vsat} \cdot G_{\Delta L} + G_{\rm R}}{2} \cdot \left[1 + \sqrt{1 - \frac{4 \cdot G_{\rm R}/G_{\rm vsat}}{(G_{\rm vsat} \cdot G_{\Delta L} + G_{\rm R})^2} \cdot \left({G_{\rm vsat}}^2 - {G_{\rm mob}}^2\right)}\right] \tag{3.70}$$

The function $G_{\rm Th}$ describes the self-heating effect, and is given by:

$$G_{\rm Th} = -\theta_{\rm Th} \cdot \frac{\bar{Q}^*_{\rm inv}}{C_{\rm ox}} \cdot \Delta\psi \cdot V_{\rm DS} \tag{3.71}$$

where $\theta_{\rm Th}$ is theoretically given by $\mu_0 \cdot C_{\rm ox} \cdot W \cdot \eta_\beta \cdot R_{\rm Th}/T_{\rm amb}$ but is more practically considered as a geometry-dependent empirical parameter. From (3.69)-(3.71), it is clear that, like the effect of series resistance, self-heating results in a reduction of transconductance, particularly at high drain and gate bias. Finally, approximating $\bar{Q}^*_{\rm inv}$ by $\bar{Q}_{\rm inv}$ and replacing $-\bar{Q}_{\rm inv}/C_{\rm ox}$ by variable $\bar{V}_{\rm G_T}$, we obtain the expression for $G_{\rm Th}$ as used in MM11.

It should be mentioned here that the model for self-heating introduced above is quasi-static and is therefore only valid for frequencies below $1/\tau_{\rm Th}$. Here, $\tau_{\rm Th}$ is the thermal relaxation time of the device, which is equal to $(2 \cdot \pi \cdot R_{\rm Th} \cdot C_{\rm Th})^{-1}$, where $C_{\rm Th}$ is the thermal capacitance. The thermal relaxation time $\tau_{\rm Th}$ is independent of device geometry, and is in the order of 10 to 100 ns. The above self-heating model will thus be valid for frequencies up to approximately 10 to 100 MHz.
In most circuit simulators, thermal effects can be described by an external equivalent network consisting of the thermal resistance $R_{\rm Th}$ and the thermal capacitance $C_{\rm Th}$ in parallel. This network is driven by the input $P_{\rm dis}$, which is calculated using the MOSFET model, and generates the output ΔT, which is subsequently used in recalculating the channel current. In this way, the frequency dependency of self-heating can be taken into account.

3.7 Impact of Pocket Implants

In present-day CMOS technologies, pocket implants are widely used to reduce short-channel effects such as threshold-voltage roll-off and punch-through [92], see Fig. 3.11. These pocket implants are regions of high doping concentration of the same type as the channel near the source and drain extensions. As a result, the impurity doping concentration at the Si/SiO_2-interface is no longer uniform

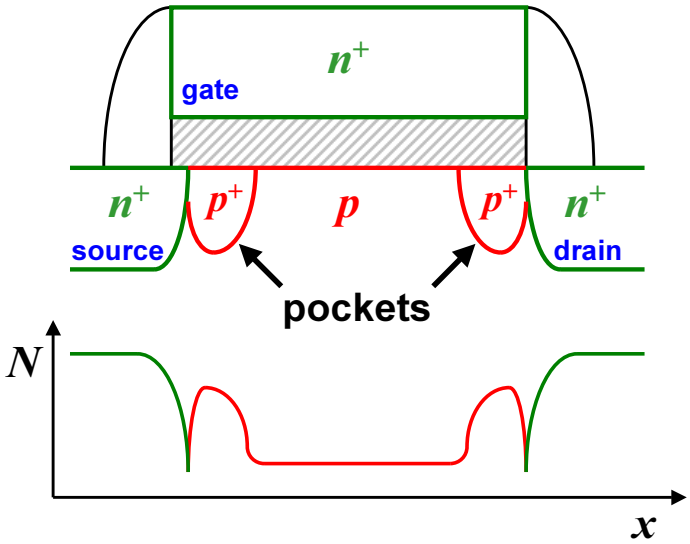


Figure 3.11: In modern MOSFETs, extra doping is implanted in the channel near the source and drain extensions in order to reduce short-channel effects. These implants result in regions of high doping concentration of the same type as the channel, commonly referred to as *pocket implants*. The impurity doping concentration $N(x)$ (at the Si/SiO_2-interface) is no longer uniform along the channel.

along the channel. Furthermore, as the channel length is decreased, the pockets occupy a relatively larger part of the channel and consequently their impact increases. In other words, the average channel doping $\bar{N}_A$ ($= \int_0^L N_A(x) \cdot dx/L$) increases with decreasing channel length L. Since several physical quantities such as charge densities, electrical fields and carrier mobility are dependent on the channel doping, the use of pocket implants may affect the electrical behaviour of MOSFETs considerably.
Pocket implants have been found to affect the threshold voltage [93, 94, 95], the output conductance [96, 97] and the mobility [98]. For a physical implementation of the impact of pocket implants, we should solve (3.1) using a position-dependent channel doping $N_A(x)$. The resulting equation is, however, not analytically solvable. In practice, it has been found that the model based on uniform channel doping (as derived in this Chapter) still gives an accurate description for a device with a specific geometry. This basically implies that the use of pocket implants merely affects the length scaling of certain electrical parameters. The impact of pocket implants on threshold voltage can therefore be taken into account in the length scaling of the body factor k_0 and the intrinsic potential ϕ_B. Furthermore, the impact on output conductance can be taken into account in the length scaling of the DIBL parameter σ_{dibl}, the static feedback parameter σ_{sf} and the channel length modulation parameter α. Finally, the impact on mobility μ_0 can be taken into account in the length scaling of the gain factor β [13]. As most conventional extraction methods of effective channel length L are based on the assumption that μ_0 is channel length independent, the latter unfortunately means that these methods are no longer valid for pocket-implanted devices [98]-[100]. New methods for effective channel length extraction, either based on capacitance measurements [13, 101] or on gate current measurements [14, 102], have to be used.

4 Bulk Current Modelling

Up till here, it has been assumed that the bulk current in a MOSFET is equal to zero. In the MOSFET channel and overlap regions, however, a bulk current may be generated between drain and bulk or between source and bulk due to the so-called effects of *impact ionisation* and *gate-induced drain leakage* (GIDL). These effects will be discussed in this Chapter. In addition to the above effects, a bulk current may be generated between gate and bulk due to tunnelling through the gate oxde. This effect, however, will be covered in Chapter 5.
The drain-bulk and source/bulk n^+/p-junctions act as diodes, and as a result they will also contribute to the bulk current. As mentioned before, the steady-state and capacitive behaviour of these junctions is not taken into account in MOS Model 11, but is taken into account in a separate junction diode model JUNCAP instead. The bulk current due to junction leakage is thus left out of consideration in this Chapter.

4.1 Impact Ionisation

For a MOSFET biased in saturation, the electric field E_x at the drain side may reach very high values. In this case, electrons travelling through the channel from source to drain are accelerated and gain so much energy that they can create extra electron-hole pairs by exciting electrons from the valence band into the conduction band, see Fig. 4.1. The latter is commonly referred to as impact ionisation. In this way, an avalanche of free carriers may arise and the initial flux of carriers is multiplied, until, possibly, complete breakdown occurs. In a typical MOSFET, however, only low-level avalanche multiplication or *weak-avalanche* occurs. Impact ionisation may result in a significant bulk current I_B [103], assuming that all the generated holes are collected by the bulk terminal[9]. As a consequence, the drain current I_D is no longer equal to the channel current I_{DS}:

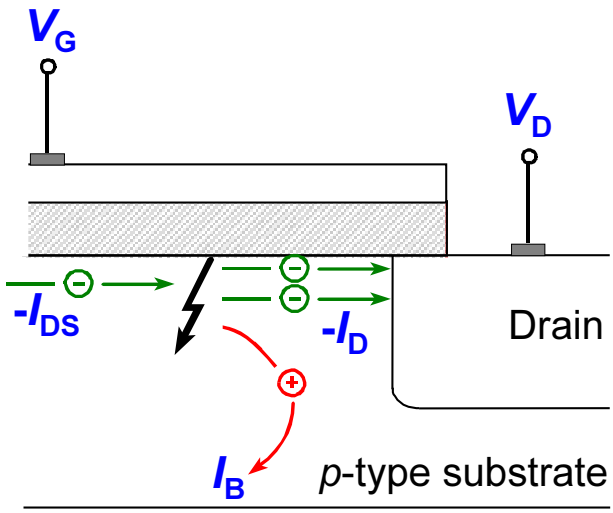


Figure 4.1: Schematic illustration of impact ionisation in an n-MOS transistor. Electrons travelling through the channel from source to drain are accelerated and gain so much energy that they can create extra electron-hole pairs by exciting electrons from the valence band into the conduction band. The generated electrons and holes are collected by the drain and bulk terminal, respectively.

[9] In practice, a large part of the generated hole current is collected by the bulk terminal and a small part recombines with electrons which are supplied by the source terminal [7]. The latter part is neglected in this report.

$$I_{\mathrm{D}} = I_{\mathrm{DS}} + I_{\mathrm{B}} \tag{4.1}$$

The bulk current is formed by an avalanche current I_{avl}. For low-level avalanche multiplication, the avalanche current can be expressed as [36]:

$$I_{\mathrm{avl}} = I_{\mathrm{DS}} \cdot \int_0^L \alpha_{\mathrm{n}} \cdot \mathrm{d}x \tag{4.2}$$

where I_{DS} is given by the general channel current expression (3.69), and α_{n} is the electron impact ionisation coefficient per unit length. Since this coefficient is higher for electrons (α_{n}) than for holes (α_{p}), the effect of bulk current is more severe in n-channel than in p-channel MOSFETs. The impact ionisation coefficient is a strong function of the lateral electric field E_{x} [104]:

$$\alpha_{\mathrm{n}} = A_{\mathrm{i}} \cdot \exp\left(-\frac{B_{\mathrm{i}}}{E_{\mathrm{x}}}\right) \tag{4.3}$$

where A_{i} and B_{i} are called the impact ionisation constants. Due to the exponential dependence of α_{n} on electric field, the impact ionisation will dominate at the position where the electric field is maximum, i.e., at the drain end. Using a pseudo two-dimensional calculation, as was done for channel length modulation in Appendix I, the avalanche current can be approximated by [36]:

$$I_{\mathrm{avl}} \approx a_1 \cdot I_{\mathrm{DS}} \cdot \exp\left(-\frac{b}{E_{\mathrm{x_L}}}\right) \tag{4.4}$$

where a_1 and b are parameters and $E_{\mathrm{x_L}}$ is the peak value of the lateral drain field. In Appendix I, an expression for $E_{\mathrm{x_L}}$ is found that is approximately equal to $(V_{\mathrm{DS}} - V_{\mathrm{DS_{sat}}})/l_{\mathrm{c}}$, where l_{c} is a characteristic length. Equation (4.4) can be rewritten to:

$$I_{\mathrm{avl}} \approx a_1 \cdot I_{\mathrm{DS}} \cdot \exp\left(-\frac{a_2}{V_{\mathrm{DS}} - V_{\mathrm{DS_{sat}}}}\right) \tag{4.5}$$

where $a_2 = b \cdot l_{\mathrm{c}}$. The above approximation, however, is somewhat oversimplified and leads to substantial errors. In practice, this deficiency can be corrected by modifying the voltage dependence of $E_{\mathrm{x_L}}$ with an experimental term $(V_{\mathrm{DS}} - a_3 \cdot V_{\mathrm{DS_{sat}}})/l_{\mathrm{c}}$, in which a_3 is a technology dependent fitting parameter. The avalanche current now becomes:

$$I_{avl} = \begin{cases} 0 & \text{for:} \quad V_{DS} \le a_3 \cdot V_{\mathrm{DS_{sat}}} \\ a_1 \cdot I_{\mathrm{DS}} \cdot \exp\left(-\frac{a_2}{V_{\mathrm{DS}} - a_3 \cdot V_{\mathrm{DS_{sat}}}}\right) & \text{for:} \quad V_{\mathrm{DS}} > a_3 \cdot V_{\mathrm{DS_{sat}}} \end{cases} \tag{4.6}$$

The use of a smoothing function is superfluous here, because I_{avl} and its higher-order derivatives go to zero for $V_{\mathrm{DS}} \to a_3 \cdot V_{\mathrm{DS_{sat}}}$. The results of equation (4.6) are close to experimental data as can be seen in Fig. 4.2.

Over the years, the constant downward scaling of device dimensions in CMOS technologies has resulted in a reduction of gate oxide thickness t_{ox}, a reduction of junction depth X_{j} and an enhancement of channel implant N_{A}. Up till about 1998, CMOS technologies were scaled down according to the so-called *constant-voltage scaling* strategy, which meant that the supply voltage did not scale with each technology generation [105]. For this strategy, the peak electric field increases with downscaling of minimum channel length at comparable drain bias, and as a result the impact ionisation effect

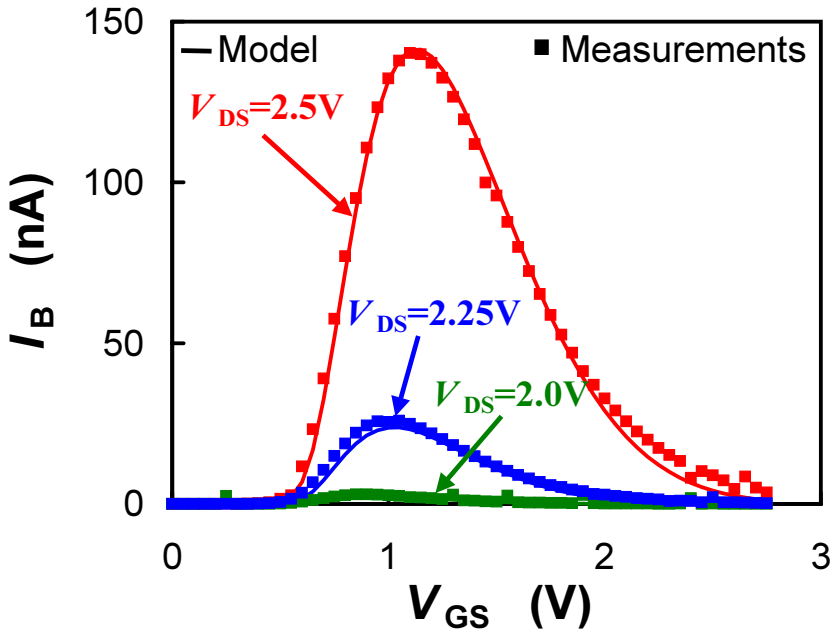


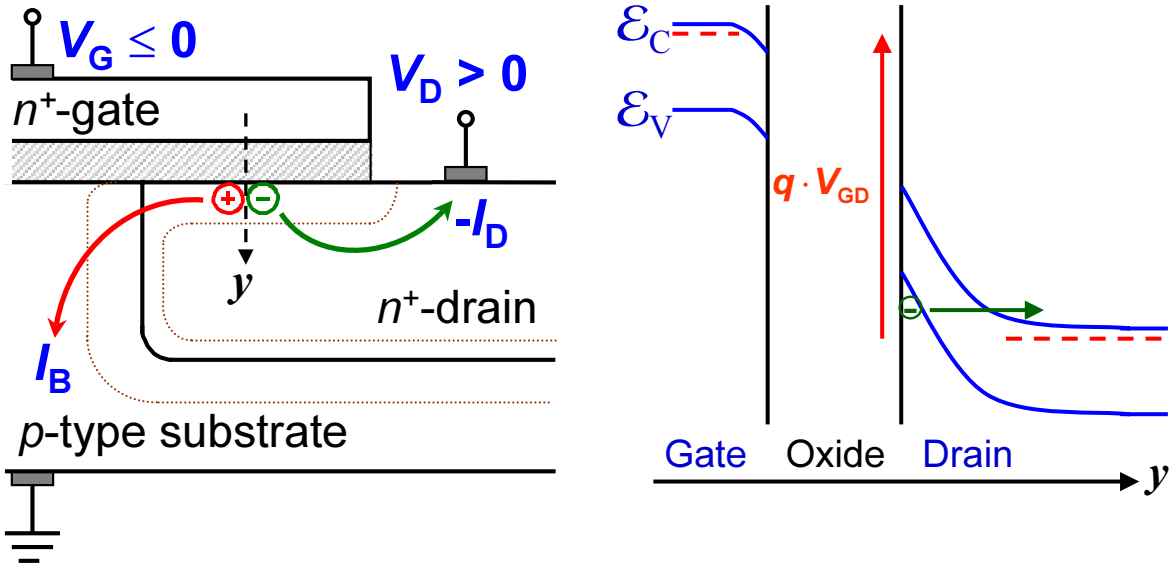
Figure 4.2: Measured (*symbols*) and modelled (*lines*) bulk current I_B as a function of gate bias V_{GS} for different values of drain bias V_{DS}, using (4.6). (n-MOS, $W/L = 10\mu\text{m}/0.5\mu\text{m}$, $t_{ox} = 5.0\text{nm}$ and $V_{SB} = 0.0\text{V}$)

becomes more and more important with each technology generation.
Since about 1998, however, CMOS technologies have been and continue to be scaled down according to the so-called *constant-field scaling* strategy [105]. In this strategy, the supply voltage is also scaled down with each technology generation, so that the lateral and transversal electric fields in the devices remain approximately the same. For modern CMOS technologies, the supply voltage is so low that carriers can hardly obtain energy enough to be able to ionise electron-holes pairs[10]. As a result, the effect of impact ionisation will no longer play an important role in modern and future CMOS technologies, and other effects such as junction leakage or gate-induced drain leakage will determine the bulk current.

4.2 Gate-Induced Drain Leakage

When the MOSFET is in off-state, a significant leakage current flowing from drain to bulk can be detected at a drain voltage much lower than the breakdown voltage [106]. This drain leakage current is caused by the gate-induced high electric field in the gate-to-drain overlap region, and as a result it has been named gate-induced drain leakage (GIDL). Since off-state leakage in MOSFETs is one of the major issues for retention time degradation in dynamic random access memories (or DRAMs), GIDL has been subject to intensive research [106]-[115].
The physical explanation of the GIDL phenomena is depicted in Fig. 4.3. For negative gate-drain bias V_{GD}, a depletion region is formed underneath the gate-to-drain overlap region and a high transversal field is created in the depletion region. Electron-hole pairs are generated by the tunnelling of valence band electrons into the conduction band (as indicated in Fig. 4.3 (b)) and collected by the drain and bulk separately. The above-mentioned tunnelling can either occur via band-to-band tunnelling (BBT) or via trap-assisted tunneling (TAT). Band-to-band tunnelling has a stronger dependence on electric field than trap-assisted tunneling, and hence, it is generally assumed that the band-to-band tunnelling

[10] In a 0.12μm CMOS technology, for example, the supply voltage is 1.2 V, whereas the energy bandgap $\mathcal{E}_g$ is approximately equal to 1.12 eV.



(a) (b)

Figure 4.3: (a) Cross-section of gate-drain overlap region, and (b) corresponding energy-band diagram along the y-direction (as indicated in (a)). Electron-hole pairs are generated by tunnelling of valence band electrons into the conduction band resulting in a leakage current between drain and bulk. This effect is called gate-induced drain leakage (GIDL), although the same effect can also occur at the source side (gate-induced source leakage or GISL).

mechanism is dominant. Nevertheless, trap-assisted tunneling may become dominant for specific conditions, such as low field and low temperatures [111, 112, 114]. Since the drain leakage current is typically dominated by other phenomena (e.g., subthreshold current or junction leakage) for these specific conditions, the influence of trap-assisted tunnelling is neglected in MOS Model 11.
According to Appendix J, the band-to-band tunnelling current density J_{BBT} in the overlap region can be approximated by:

$$J_{BBT} \propto {E_{tov_L}}^2 \cdot \exp\left(-\frac{B^{\#}_{GIDL}}{E_{tov_L}}\right) \tag{4.7}$$

where $B^{\#}_{GIDL}$ is considered as an empirical parameter, theoretically proportional to ${\mathcal{E}_g}^{3/2}$. The maximum electric field E_{tov_L} at the Si/SiO_2-interface in the overlapped drain extension consists of a (dominant) transversal component (equal to $C_{ox} \cdot V_{ov_L}/\epsilon_{Si}$) and a lateral component empirically proportional to V_{DB}. The maximum electric field E_{tov_L} can be written as:

$$E_{tov_L} = \frac{C_{ox}}{\epsilon_{Si}} \cdot \sqrt{{V_{ov_L}}^2 + (C_{GIDL} \cdot V_{DB})^2} \tag{4.8}$$

where C_{GIDL} is an empirical parameter. In most cases, the transversal component will be much larger than the lateral component, consequently C_{GIDL} will be very small and E_{tov_L} is approximately equal to the transversal field $C_{ox} \cdot V_{ov_L}/\epsilon_{Si}$. The total gate-induced drain leakage current I_{GIDL} between drain and bulk can be calculated by integrating J_{BBT} over the total area of the overlapped drain extension. Assuming that the maximum field E_{tov_L} is constant over this area, we can simply write:

$$I_{GIDL} \propto W \cdot \Delta L_{ov} \cdot {E_{tov_L}}^2 \cdot \exp\left(-\frac{B^{\#}_{GIDL}}{E_{tov_L}}\right) \tag{4.9}$$

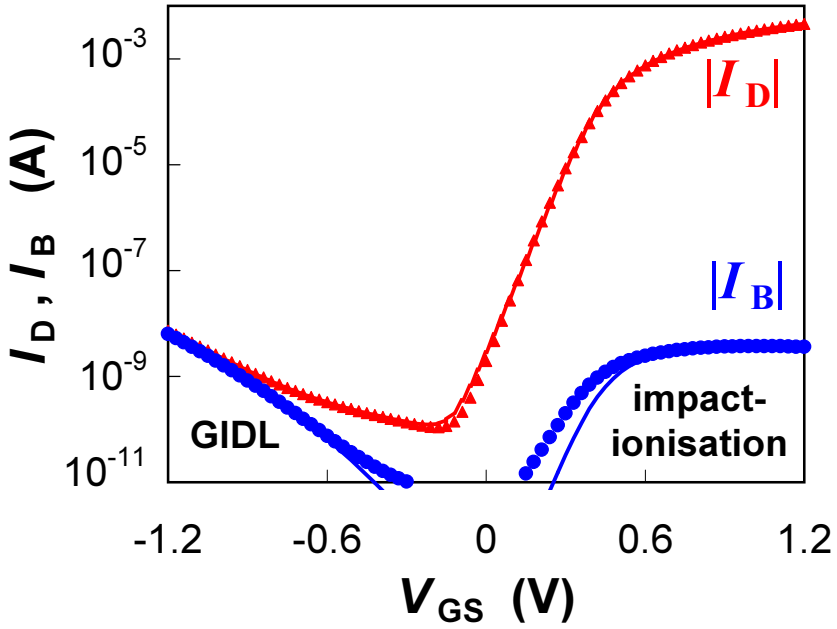


Figure 4.4: Measured (symbols) and simulated (lines) drain and bulk current as a function of V_{GS} at high drain bias ($V_{DS} = 1.2$V). GIDL becomes dominant for negative values of V_{GS}, it is accurately described by (4.11). (n-MOS, $W/L = 10\mu$m$/0.13\mu$m, $t_{ox} = 2.0$nm and $V_{SB} = 0.0$V)

where ΔL_{ov} is the length of the (gate-source or gate-drain) overlap region.
In the derivation of I_{GIDL}, it has been assumed that all electron-hole pairs generated by tunnelling are collected by the drain and bulk separately. This implies that even at zero drain-bulk bias ($V_{DB} = 0$), a current will be flowing between drain and bulk (i.e., $I_{GIDL} \neq 0$), which is not physical. In order to have a more physical description of I_{GIDL} that goes to zero for $V_{DB} = 0$, all tunnelling components in the overlapped drain-bulk junction should be taken into account, i.e., not only the transversal component but also the lateral component. This would, however, lead to a very complex expression. In order to ensure that $I_{GIDL} = 0$ for $V_{DB} = 0$ and that I_{GIDL} changes sign when V_{DB} changes sign, the following simple empirical expression is used instead of (4.9):

$$I_{GIDL} \propto W \cdot \Delta L_{ov} \cdot V_{DB} \cdot {E_{tov_L}}^2 \cdot \exp\left(-\frac{B^{\#}_{GIDL}}{E_{tov_L}}\right) \tag{4.10}$$

The empirical addition of V_{DB} in (4.10) w.r.t. (4.9) is allowed as the bias dependency of I_{GIDL} is all but determined by the exponential term $\exp(-B^{\#}_{GIDL}/E_{tov_L})$. Finally, we can write the MM11-equation:

$$I_{GIDL} = A_{GIDL} \cdot V_{DB} \cdot {V_{tov_L}}^2 \cdot \exp\left(-\frac{B_{GIDL}}{V_{tov_L}}\right) \tag{4.11}$$

where A_{GIDL} is an empirical parameter proportional to $W \cdot \Delta L_{ov} \cdot C_{ox}/\epsilon_{Si}$, B_{GIDL} is equal to $\epsilon_{Si} \cdot B^{\#}_{GIDL}/C_{ox}$ and V_{tov_L} is given by:

$$V_{tov_L} = \sqrt{{V_{ov_L}}^2 + (C_{GIDL} \cdot V_{DB})^2} \tag{4.12}$$

Equation (4.11) has been found to accurately describe the gate-induced drain leakage behaviour of MOSFETs, see Fig. 4.4.
In the above derivation, we have focussed on the gate-induced drain leakage. The same phenomenon,

however, can also occur at the source side, in which case it is referred to as gate-induced source leakage (GISL). The electric field in the overlapped source region is typically not as high as the field in the drain region, and as a result, GISL will not really impact the source leakage. Nonetheless, it is still important to include GISL in the model in order to preserve drain-source symmetry. Analogous to the derivation of GIDL, we can write the GISL current I_{GISL} between source and bulk as:

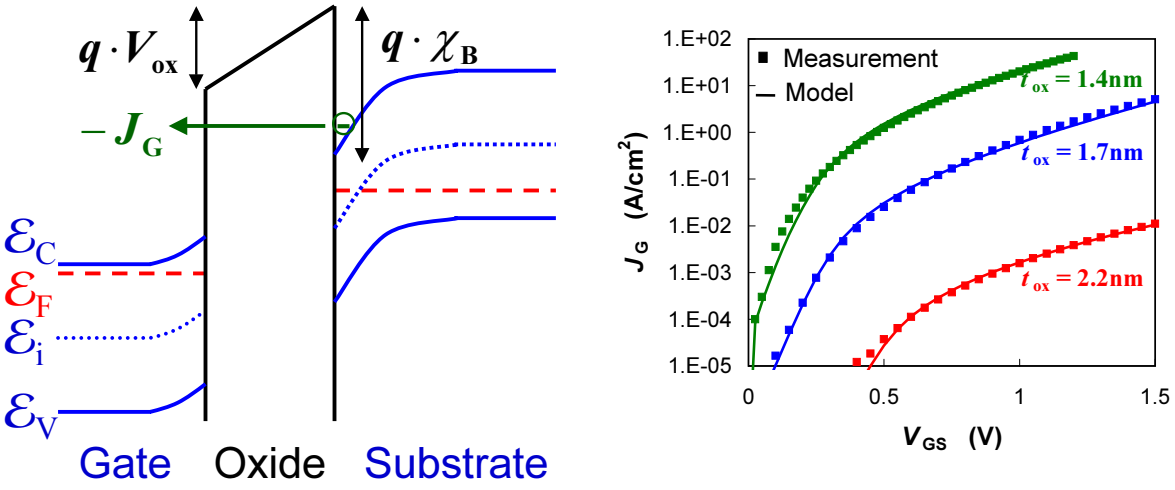
$$I_{\mathrm{GISL}} = A_{\mathrm{GIDL}} \cdot V_{\mathrm{SB}} \cdot {V_{\mathrm{tov_0}}}^2 \cdot \exp\left(-\frac{B_{\mathrm{GIDL}}}{V_{\mathrm{tov_0}}}\right) \tag{4.13}$$

where $V_{\mathrm{tov_0}}$ is given by:

$$V_{\mathrm{tov_0}} = \sqrt{{V_{\mathrm{ov_0}}}^2 + (C_{\mathrm{GIDL}} \cdot V_{\mathrm{SB}})^2} \tag{4.14}$$

5 Gate Current Modelling

Up till here, it has been assumed that the gate current in a MOSFET is equal to zero. From a classical point of view, this assumption holds true, since carriers in the inversion layer cannot cross the potential barrier of the gate oxide, see Fig. 5.1 (a). From a quantum-mechanical point of view, however, carriers may tunnel through the potential barrier resulting in a non-zero gate current density J_G. The probability of tunnelling increases exponentially with decreasing oxide thickness t_{ox}, resulting in an exponentially increasing J_G, see Fig. 5.1 (b). With CMOS technology scaling, t_{ox} is constantly scaled down, and consequently gate current can no longer be neglected for modern and future CMOS technologies as it may start to affect circuit performance [116, 117].



(a) (b)

Figure 5.1: (a) The energy-band diagram of an n-MOS structure in inversion ($V^*_{GB} > 0$), where χ_B is the oxide potential barrier, i.e., the difference between the conduction-band levels in SiO_2 and Si at the interface. Electrons in the inversion layer may tunnel to the gate resulting in a non-zero gate current density J_G. (b) The gate current density J_G as a function of gate bias V_{GS} for different values of oxide thickness t_{ox}; J_G increases exponentially with decreasing t_{ox} (n-MOS, $V_{DS} = V_{SB} = 0$ V).

Over the years, the gate tunnelling current has been subject to extensive research [118]-[131]. Nevertheless, the compact modelling of gate current still is a relatively unexplored field [132]-[135], although it is receiving increasingly more attention [136, 137]. In MM11, a new physics-based gate leakage model is used [14], which is at least as accurate as other models [2, 134], but much simpler. This model will be treated in this Chapter.
According to Appendix K.1, the gate current density J_G can generally be written as:

$$J_G = q \cdot N \cdot P_{tunnel}\,(V_{ox}, \chi_B, A, B) \tag{5.1}$$

where N is the number of mobile carriers per unit area and P_{tunnel} is the transmission probability of a carrier tunnelling:

$$P_{tunnel}\,(V_{ox}, \chi_B, A, B) = A \cdot V_{ox} \cdot \exp\left(-\frac{B}{V_{ox}} \cdot \left[1 - \left(1 - \frac{V_{ox}}{\chi_B}\right)^{3/2}\right]\right) \tag{5.2}$$

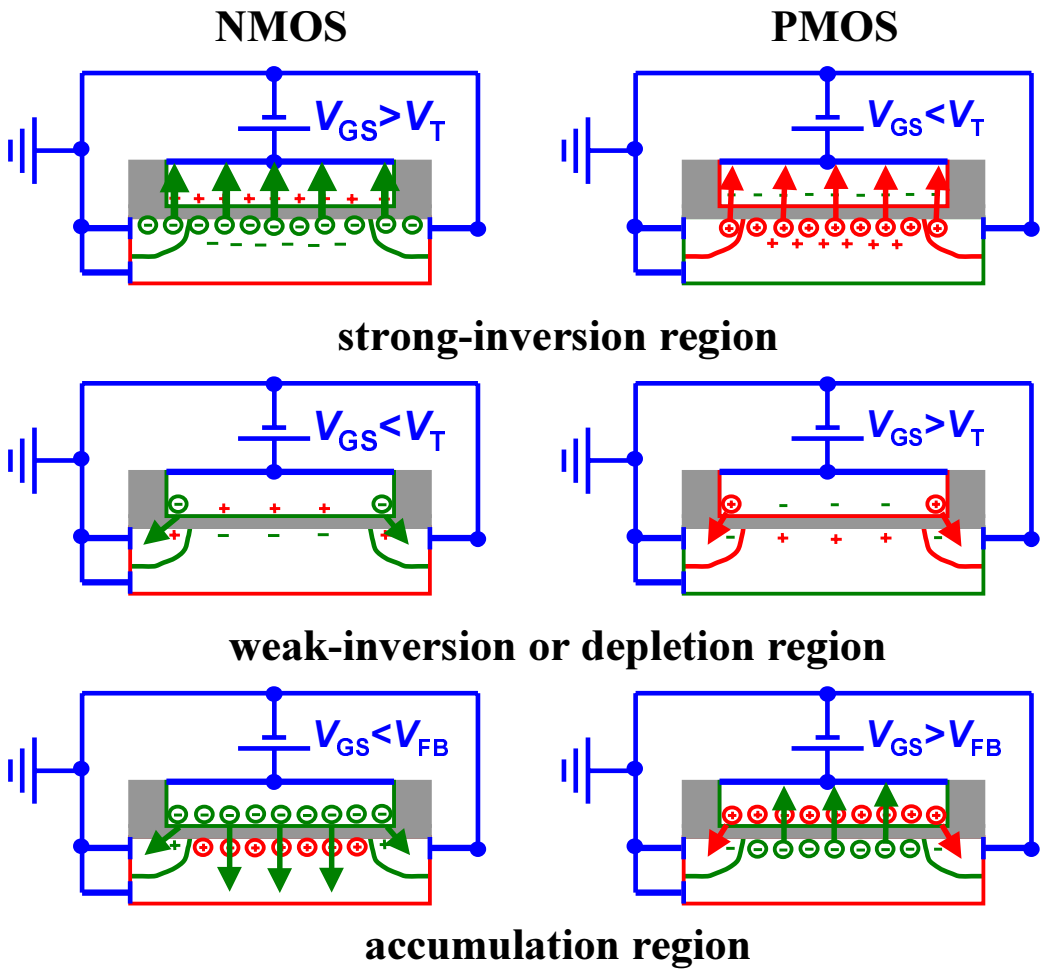


Figure 5.2: The gate tunnelling components in the different operation regions. The intrinsic (channel) region behaves differently from the gate overlap regions. In the intrinsic region, electrons (n-MOS) or holes (p-MOS) tunnel from the channel to the gate in inversion, whereas in accumulation electrons tunnel from the gate to the substrate (n-MOS) or vice-versa (p-MOS). In the overlap regions, electrons (n-MOS) or holes (p-MOS) tunnel from the source/drain-extension to the gate or vice-versa, depending on the bias conditions. Overall three gate current components can be distinguished: the gate-to-channel current, the gate-to-bulk current and the gate overlap current.

Here V_{ox} is the oxide voltage ($= Q_g/C_{ox}$). In addition, the prefactor A and the probability factor B are physical constants given by (K.8) and (K.6), respectively, but in view of the approximations made A and B are treated as empirical parameters. In Fig. 5.1 (b) it can be seen that (5.1) gives an accurate description of J_G for different values of t_{ox}.

In a typical MOSFET structure we can distinguish different gate current components, see Fig. 5.2 and Tab. 5.1. In general, three main gate current components can be distinguished: the gate-to-channel I_{GC}, the gate-to-bulk I_{GB} and the gate overlap component $I_{G_{ov}}$, see Fig. 5.3. The above-mentioned components will be subsequently treated in the following Sections.

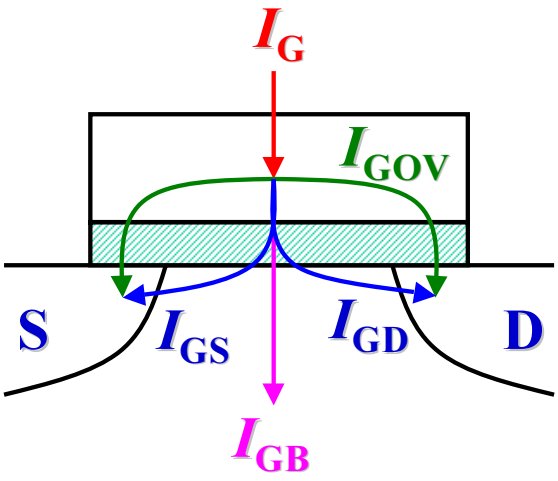
5.1 Gate-to-Channel Current

According to Appendix K.1, the gate-to-channel current density J_{GC} can be written as:

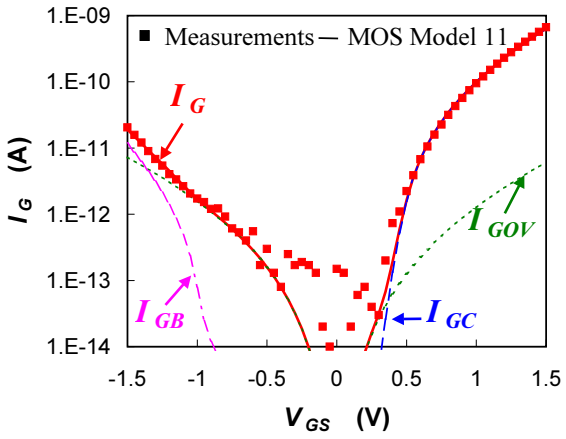
$$J_{GC} = -Q_{inv} \cdot P_{inv} \tag{5.3}$$

Table 5.1: The type of carriers that contribute to the gate tunnelling components as indicated in Fig. 5.2. The type of carriers determine the value of oxide energy barrier χ_B that has to be used ($\chi_{B_N} = 3.1$V for electrons, $\chi_{B_P} = 4.5$V for holes). Furthermore, the probability factor B is different for electrons and holes. In the last row, the direction of gate current is indicated.

Type	Intrinsic MOSFET		Overlap Regions
	Accumulation	Inversion	
n-MOS	electrons	electrons	electrons
p-MOS	electrons	holes	holes
	I_{GB}	I_{GS} / I_{GD}	I_{GS} / I_{GD}



(a)



(b)

Figure 5.3: (a) The different gate current components in a MOSFET. One can distinguish the intrinsic components, i.e., the gate-to-channel current I_{GC} ($= I_{GS} + I_{GD}$) and the gate-to-bulk current I_{GB}, and the extrinsic components, i.e., the gate/source and gate/drain overlap components $I_{G_{ov}}$. (b) Measured and modelled gate current as a function of gate bias V_{GS} at $V_{DS} = V_{SB} = 0$ V, the different gate current components are also shown. n-MOS, $W/L = 10/0.6\mu$m, $t_{ox} = 2$nm.

where the tunnelling probability P_{inv} is equal to $P_{tunnel}(V_{ox}, \chi_B, A_{inv}, B_{inv})$, A_{inv} is the probability prefactor and B_{inv} is the probability factor for carriers in the inversion layer tunnelling to the gate. In order to calculate the total gate-to-channel current I_{GC}, the current continuity equation has to be solved:

$$\frac{\partial I_{DS}(x)}{\partial x} = -W \cdot J_{GC}(x) \tag{5.4}$$

where I_{DS} is given by (3.1) and is no longer constant along the channel. Differential equation (5.4) cannot be solved analytically, and as a consequence it needs to be approximated. The current continuity equation is solved under the assumption that J_{GC} only induces a small perturbation of the potential distribution along the channel (i.e., $\partial I_{DS}/\partial x \approx 0$). Note that this assumption implies that I_{DS} is (approximately) constant along the channel and all equations derived in Chapter 3 are still valid. The

total gate-to-channel current I_{GC} is obtained by integrating J_{GC} along the channel:

$$I_{GC} = W \cdot \int_0^L J_{GC} \cdot \mathrm{d}x \tag{5.5}$$

As a further approximation the exponential term in the tunnelling probability P_{inv} is linearised:

$$P_{inv} = A_{inv} \cdot V_{ox} \cdot \exp\left(-\frac{B_{inv}}{V_{ox}} \cdot \left[1 - \left(1 - \frac{V_{ox}}{\chi_B}\right)^{\frac{3}{2}}\right]\right) \approx A^{\#}_{inv} \cdot V_{ox} \cdot \exp\left(B^{\#}_{inv} \cdot V_{ox}\right) \tag{5.6}$$

where:

$$A^{\#}_{inv} = A_{inv} \cdot \exp\left(-\frac{3}{2} \cdot \frac{B_{inv}}{\chi_B}\right) \tag{5.7}$$

$$B^{\#}_{inv} = \frac{3}{8} \cdot \frac{B_{inv}}{{\chi_B}^2} \tag{5.8}$$

Neglecting the influence of velocity saturation, $\mathrm{d}x$ in (5.5) can be replaced by (L.2). Furthermore defining $\phi = \psi_s - \bar{\psi}$ and $\mathrm{d}\phi = \mathrm{d}\psi_s$, eq. (5.5) can be rewritten in:

$$I_{GC} = -\frac{A^{\#}_{inv} \cdot W \cdot L}{\bar{Q}^{*}_{inv} \cdot \Delta\psi} \cdot \int_{-\Delta\psi/2}^{\Delta\psi/2} Q_{inv} \cdot Q^{*}_{inv} \cdot V_{ox} \cdot \exp\left(B^{\#}_{inv} \cdot V_{ox}\right) \cdot \mathrm{d}\phi \tag{5.9}$$

where Q_{inv} and Q^{*}_{inv} are given by (2.20) and (3.6), respectively, and the oxide voltage V_{ox} in the channel region is simply given by:

$$V_{ox} = \frac{Q_g}{C_{ox}} = \frac{\bar{Q}_g - C_g \cdot \phi}{C_{ox}} = \bar{V}_{ox} - \partial V_{ox} \cdot \phi \tag{5.10}$$

It is straightforward to solve the integral (5.9), although it results in a lengthy equation. In order to simplify the result, we take a third-order Taylor polynomial around $\Delta\psi = 0$, which results in:

$$I_{GC} = -A^{\#}_{inv} \cdot W \cdot L \cdot \bar{Q}_{inv} \cdot \bar{V}_{ox} \cdot \exp\left(B^{\#}_{inv} \cdot \bar{V}_{ox}\right) \cdot P_{GC} \tag{5.11}$$

where:

$$P_{GC} = 1 + \frac{{r_B}^2 + 2 \cdot r_B \cdot (r + r^{*} + r_{ox}) + 2 \cdot (r \cdot r^{*} + r \cdot r_{ox} + r^{*} \cdot r_{ox})}{24} \cdot \Delta\psi^2 \tag{5.12}$$

Here r, r^{*}, r_{ox} and r_B are dimensionless variables defined by:

$$r = C_{inv}/\bar{Q}_{inv} \tag{5.13}$$

$$r^{*} = C_{inv}/\bar{Q}^{*}_{inv} \tag{5.14}$$

$$r_{ox} = \partial V_{ox}/\bar{V}_{ox} \tag{5.15}$$

$$r_B = B^{\#}_{inv} \cdot \partial V_{ox} \tag{5.16}$$

In the strong inversion region where I_{GC} is important, equation (5.12) can be further simplified by assuming $r^{*} \approx r$, which results in:

$$P_{GC} = 1 + \frac{{r_B}^2 + 2 \cdot r_B \cdot (2 \cdot r^{*} + r_{ox}) + 2 \cdot \left({r^{*}}^2 + 2 \cdot r^{*} \cdot r_{ox}\right)}{24} \cdot \Delta\psi^2 \tag{5.17}$$

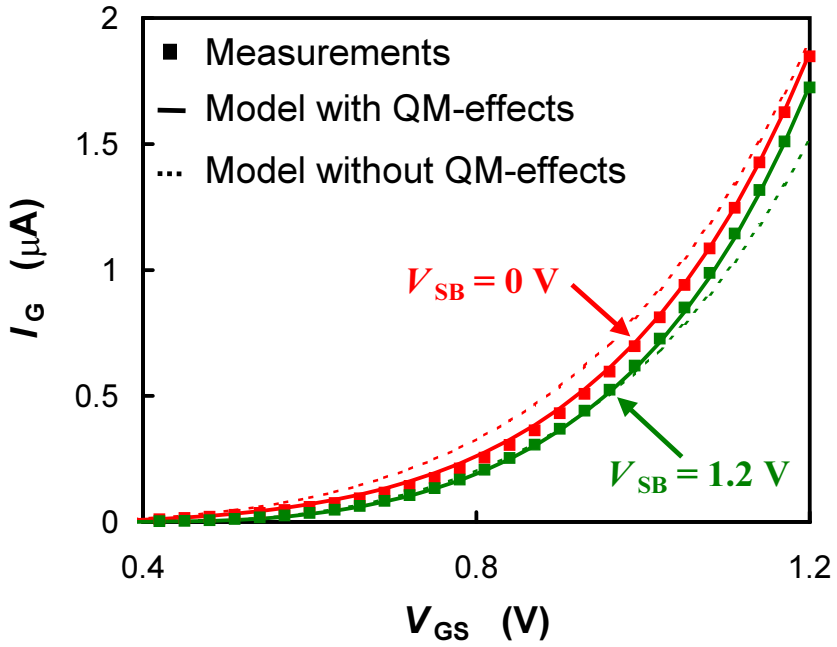


Figure 5.4: Gate current as a function of V_{GS} for two values of V_{SB}. Quantum-mechanical lowering of potential barrier χ_B is taken into account. (n-MOS, $W/L = 10\mu m/10\mu m$, $t_{ox} = 1.7nm$, $V_{DS} = 50mV$)

Finally, we can revert the term $A^{\#}_{inv} \cdot \exp(B^{\#}_{inv} \cdot \bar{V}_{ox})$ in (5.11) back to the original tunnel probability $P_{tunnel}\left(\bar{V}_{ox}, \chi_B, A_{inv}, B_{inv}\right)$, resulting in:

$$I_{GC} = -I_{GINV} \cdot \frac{\bar{Q}_{inv}}{C_{ox}} \cdot \bar{V}_{ox} \cdot P_{GC} \cdot \exp\left(-\frac{B_{inv}}{\bar{V}_{ox}} \cdot \left[1 - \left(1 - \frac{\bar{V}_{ox}}{\chi_B}\right)^{\frac{3}{2}}\right]\right) \tag{5.18}$$

where parameter I_{GINV} is theoretically equal to $A_{inv} \cdot W \cdot L \cdot C_{ox}$, but is used as an empirical parameter. The gate-to-channel current I_{GC} can be seen in Fig. 5.3 (b) as a function of gate bias for a typical n-MOS transistor at $V_{DS} = 0$.

Carriers at the oxide interface are confined to a narrow potential well, quantum-mechanically resulting in a splitting of the conduction energy band into discrete subbands and in a displacement of the inversion-layer carrier distribution from the interface, see Appendix D. This affects the effective oxide potential barrier [118]. Owing to these quantum-mechanical effects, the electrons in the inversion layer are not situated at the bottom of the conduction band but in an energy level which effectively lies $\Delta\chi_B$ above the conduction band. Assuming that only the lowest energy subband is occupied by electrons and using (D.10), the value of $\Delta\chi_B$ can be given by:

$$\Delta\chi_B = QM_\psi \cdot \left(\frac{\epsilon_{Si} \cdot \bar{E}_{eff}}{C_{ox}}\right)^{2/3} \tag{5.19}$$

where QM_ψ is equal to $QM \cdot {C_{ox}}^{2/3}$, and QM is a physical constant ($QM_N = 5.951993\ Vm^{4/3}/C^{2/3}$ for electrons and $QM_P = 7.448711\ Vm^{4/3}/C^{2/3}$ for holes). As a result, the effective oxide potential barrier $\chi_{B_{eff}}$ is lowered by an amount of $\Delta\chi_B$:

$$\chi_{B_{eff}} = \chi_B - \Delta\chi_B \tag{5.20}$$

In order to take this quantum-mechanical barrier lowering into account, B_{inv} in the above equations has to be replaced by:

$$B_{\mathrm{eff}} = B_{\mathrm{inv}} \cdot \left(\frac{\chi_{\mathrm{B_{eff}}}}{\chi_{\mathrm{B}}} \right)^{3/2} \tag{5.21}$$

The resulting model gives in an accurate description of the quantum effects at various V_{SB}, see Fig. 5.4.

Source/Drain Partitioning: The carriers tunnelling from the inversion layer to the gate contributing to I_{GC} are supplied by both source (I_{GS}) and drain (I_{GD}). The partitioning of I_{GC} into I_{GS} and I_{GD} can simply be given by, see Appendix K.2:

$$I_{\mathrm{GD}} = W \cdot \int_0^L \frac{x}{L} \cdot J_{\mathrm{GC}} \cdot \mathrm{d}x \tag{5.22}$$

$$I_{\mathrm{GS}} = W \cdot \int_0^L \left(1 - \frac{x}{L}\right) \cdot J_{\mathrm{GC}} \cdot \mathrm{d}x = I_{\mathrm{GC}} - I_{\mathrm{GD}} \tag{5.23}$$

Neglecting the influence of velocity saturation, x and $\mathrm{d}x$ in (5.22) can be replaced by (L.4) and (L.2), respectively. Using linearisation (5.6), we obtain:

$$I_{\mathrm{GD}} = \frac{A^{\#}_{\mathrm{inv}} \cdot W \cdot L}{2 \cdot \bar{Q}^{*}_{\mathrm{inv}} \cdot \Delta\psi} \cdot \int_{-\Delta\psi/2}^{\Delta\psi/2} \frac{{Q^{*}_{\mathrm{inv}}}^2 - {Q^{*}_{\mathrm{inv_0}}}^2}{C_{\mathrm{inv}} \cdot \bar{Q}^{*}_{\mathrm{inv}} \cdot \Delta\psi} \cdot Q_{\mathrm{inv}} \cdot Q^{*}_{\mathrm{inv}} \cdot V_{\mathrm{ox}} \cdot \exp\left(B^{\#}_{\mathrm{inv}} \cdot V_{\mathrm{ox}}\right) \cdot \mathrm{d}\phi \tag{5.24}$$

The straightforward solution of this integral results in a lengthy equation. In order to simplify the result, we take a third-order Taylor polynomial around $\Delta\psi = 0$, which results in:

$$I_{\mathrm{GD}} = -A^{\#}_{\mathrm{inv}} \cdot W \cdot L \cdot \bar{Q}_{\mathrm{inv}} \cdot \bar{V}_{\mathrm{ox}} \cdot \exp\left(B^{\#}_{\mathrm{inv}} \cdot \bar{V}_{\mathrm{ox}}\right) \cdot \left(\frac{P_{\mathrm{GC}}}{2} - P_{\mathrm{GD}}\right) \tag{5.25}$$

where:

$$\begin{aligned} P_{\mathrm{GD}} = {} & [r_{\mathrm{B}} + r + r_{\mathrm{ox}}] \cdot \frac{\Delta\psi}{12} + \left[{r_{\mathrm{B}}}^3 + {r_{\mathrm{B}}}^2 \cdot \left(3 \cdot r + 2 \cdot r^{*} + 3 \cdot r_{\mathrm{ox}}\right)\right. \\ & + 2 \cdot r_{\mathrm{B}} \cdot \left(2 \cdot r \cdot r^{*} + 3 \cdot r \cdot r_{\mathrm{ox}} + 2 \cdot r^{*} \cdot r_{\mathrm{ox}} - {r^{*}}^2\right) \\ & \left. + 4 \cdot r \cdot r^{*} \cdot r_{\mathrm{ox}} - 2 \cdot r \cdot {r^{*}}^2 - 2 \cdot {r^{*}}^2 \cdot r_{\mathrm{ox}}\right] \cdot \frac{\Delta\psi^3}{480} \end{aligned} \tag{5.26}$$

Equation (5.26) can be further simplified by assuming $r^{*} \approx r$, which results in:

$$\begin{aligned} P_{\mathrm{GD}} = {} & \left[r_{\mathrm{B}} + r^{*} + r_{\mathrm{ox}}\right] \cdot \frac{\Delta\psi}{12} + \left[{r_{\mathrm{B}}}^3 + {r_{\mathrm{B}}}^2 \cdot \left(5 \cdot r^{*} + 3 \cdot r_{\mathrm{ox}}\right)\right. \\ & \left. + 2 \cdot r_{\mathrm{B}} \cdot \left({r^{*}}^2 + 5 \cdot r^{*} \cdot r_{\mathrm{ox}}\right) + 2 \cdot {r^{*}}^2 \cdot r_{\mathrm{ox}} - 2 \cdot {r^{*}}^3\right] \cdot \frac{\Delta\psi^3}{480} \end{aligned} \tag{5.27}$$

Finally, we can revert the term $A^{\#}_{\mathrm{inv}} \cdot \exp(B^{\#}_{\mathrm{inv}} \cdot \bar{V}_{\mathrm{ox}})$ in (5.25) back to the original tunnel probability $P_{\mathrm{tunnel}}\left(\bar{V}_{\mathrm{ox}}, \chi_{\mathrm{B}}, A_{\mathrm{inv}}, B_{\mathrm{eff}}\right)$, resulting in:

$$I_{\mathrm{GD}} = -I_{\mathrm{GINV}} \cdot \frac{\bar{Q}_{\mathrm{inv}}}{C_{\mathrm{ox}}} \cdot \bar{V}_{\mathrm{ox}} \cdot \left(\frac{P_{\mathrm{GC}}}{2} - P_{\mathrm{GD}}\right) \cdot \exp\left(-\frac{B_{\mathrm{eff}}}{\bar{V}_{\mathrm{ox}}} \cdot \left[1 - \left(1 - \frac{\bar{V}_{\mathrm{ox}}}{\chi_{\mathrm{B}}}\right)^{\frac{3}{2}}\right]\right) \tag{5.28}$$

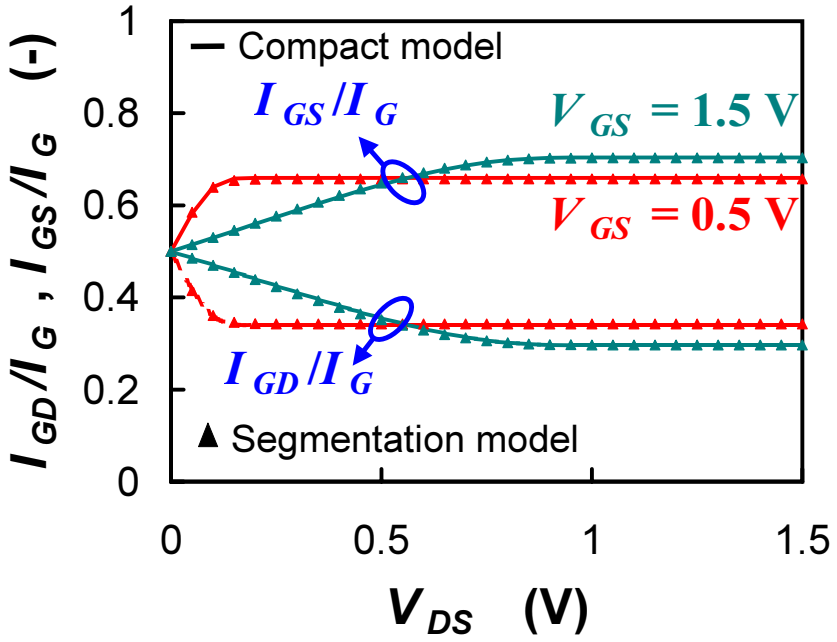


Figure 5.5: Modelled partition of I_{GS} and I_{GD} current components as a function of drain bias are verified using a segmentation model ($N = 10$). (n-MOS, $W/L = 10\mu m/0.6\mu m$, $t_{ox} = 2nm$)

The resulting I_{GS}/I_{GD} partition cannot be verified using measurements. In Fig. 5.5, it is therefore verified by breaking down the MOSFET into $N = 10$ equal segments, each described by the above model, similar to [138]. The partition, which follows naturally from the segmentation model, is accurately reproduced by (5.23) and (5.25) adding no parameters.

In the above derivations, only electrons tunnelling from the inversion layer to the gate have been taken into account. In reality, electrons in the gate will also tunnel to the inversion layer, resulting in a small tunneling component. The two tunnelling components have opposite signs and cancel out when zero bias is applied (i.e., $V_{GS} = V_{DS} = V_{SB} = 0$). The above-derived expressions for I_{GC}, I_{GD} and I_{GS}, however, do not become zero for this zero bias condition[11]. This is not physical. For a more physical description, the electrons tunnelling from gate to channel have to be taken into account, which results in a complex expression. A simple method which forces I_{GC} to zero for zero bias condition (and hardly affects the accuracy at other bias conditions), is to subtract a bias-independent value equal to I_{GC} at $V_{GS} = V_{DS} = V_{SB} = 0$ from I_{GC}. An even simpler method is used in MM11. Realising that $\bar{V}_{ox}$ is approximately equal to $V_{GS} - V_{DS_x}/2$ in the inversion region, this method replaces $\bar{V}_{ox}$ by $V_{GS} - V_{DS_x}/2$ in the prefactor of the expressions (5.18) and (5.28) for I_{GC} and I_{GD}, respectively. Furthermore, to arrive at the MM11-expressions, we need to replace r^*, r_{ox} and r_B by the MM11-variables ξ^*, ∂V_{ox} and B^*_{inv}, respectively. Finally, in order to obtain an accurate source/drain partitioning in subthreshold, we replace the term $-\bar{Q}_{inv} \cdot r^* \cdot \Delta\psi/C_{ox}$ ($\approx -\bar{Q}_{inv} \cdot r \cdot \Delta\psi/C_{ox} = \Delta Q_{inv}/C_{ox}$) in (5.28) by the MM11-variable $V_{inv_0} - V_{inv_L}$.

5.2 Gate-to-Bulk Current

For an n-type MOS transistor operating in accumulation, an accumulation layer of holes is formed in the p-type substrate and an accumulation layer of electrons is formed in the n^+-type polysilicon gate.

[11] Expressions (5.18) and (5.28) for I_{GC} and I_{GD}, respectively, only become zero for $V^*_{GB} = 0$.

Since the oxide energy barrier for electrons ($\chi_{B_N} = 3.1$V) is considerably lower than that for holes ($\chi_{B_P} = 4.5$V), the gate current will mainly consist of electrons tunnelling from the gate to the bulk silicon, where they are swept to the bulk terminal. In this case the (intrinsic) gate current density J_{GB} can be written as:

$$J_{GB} = \begin{cases} 0 & \text{for: } V_{GB} > V_{FB} \\ -Q_b \cdot P_{tunnel}\,(-V_{ox}, \chi_B, A_{acc}, B_{acc}) & \text{for: } V_{GB} \leq V_B \end{cases} \tag{5.29}$$

where A_{acc} is the probability prefactor and B_{acc} is the probability factor for carriers in the accumulation layer tunnelling to the gate. For n-type MOSFETs, both the gate-to-channel and the gate-to-bulk currents consist of electrons tunnelling, and as a result $B_{inv} = B_{acc}$. For p-type MOSFETs, however, this is not the case, see Tab. 5.1, and consequently $B_{inv} \neq B_{acc}$. The total gate-to-bulk current can be obtained by integrating J_{GB} along the channel:

$$I_{GB} = W \cdot \int_0^L J_{GB} \cdot \mathrm{d}x \tag{5.30}$$

Since no channel current is flowing, the solution of I_{GB} is straightforward. In accumulation, $\partial\psi_s/\partial x \approx 0$ and $\psi_{s_0} = \psi_{s_L}$, and as a result we can simply write (for $V^*_{GB} < 0$):

$$I_{GB} = I_{GACC} \cdot \frac{\bar{Q}_b}{C_{ox}} \cdot \bar{V}_{ox} \cdot \exp\left(\frac{B_{acc}}{\bar{V}_{ox}} \cdot \left[1 - \left(1 + \frac{\bar{V}_{ox}}{\chi_B}\right)^{3/2}\right]\right) \tag{5.31}$$

where parameter I_{GACC} is theoretically equal to $A_{acc} \cdot W \cdot L \cdot C_{ox}$, but is used as an empirical parameter. Again, the influence of quantum-mechanical quantization can be taken into account by making use of an effective oxide barrier lowering. However, in order to limit calculation time, quantum-mechanical oxide barrier lowering is neglected in this case. The gate-to-bulk current I_{GB} can be seen in Fig. 5.3 (b) as a function of gate bias for a typical n-MOS transistor at $V_{DS} = 0$.
In the above derivations, only electrons tunnelling from the gate to the channel region have been taken into account. In reality, electrons in the channel region will also tunnel to the gate, resulting in a small tunneling component. The two tunnelling components have opposite signs and cancel out when zero bias is applied (i.e., $V_{GS} = V_{DS} = V_{SB} = 0$). The above-derived expressions for I_{GB}, however, does not become zero for this zero bias condition[12]. This is not physical. In order to force I_{GB} to zero at $V_{GS} = V_{DS} = V_{SB} = 0$, in MM11, we simply replace $\bar{V}_{ox}$ by V_{GB} in the prefactor of expression (5.31) for I_{GB}.

5.3 Gate Overlap Current

Apart from the intrinsic components I_{GC} and I_{GB}, considerable gate current can be generated in the gate/source- and gate/drain-overlap regions. The overlap regions are considered as two-terminal MOS-structures with different flat-band voltage $V_{FB_{ov}}$ and body factor k_{ov}, see Section 2.2. In the following general discussion, we denote the source or drain by X. For $V_{GX} > V_{FBov}$ a negatively charged accumulation layer is formed in the overlapped n^+-source/drain extension and a positively charged depletion layer is formed in the overlapping gate. In this case the overlap gate current will mostly consist of electrons tunnelling from the source/drain accumulation layer to the gate, it is simply given by:

$$I_{G_{ov}} = W \cdot \Delta L_{ov} \cdot Q_{ov} \cdot P_{tunnel}\,(V_{ov}, \chi_B, A_{inv}, B_{inv}) \tag{5.32}$$

[12] Expression (5.31) for I_{GB} becomes zero for $V^*_{GB} = 0$.

where Q_{ov} and V_{ov} are given by (2.38) and (2.39), respectively, and ΔL_{ov} is the length of the gate/source or gate/drain overlap region.
For $V_{GX} < V_{FBov}$, the situation is reversed, a positively charged depletion layer is formed in the overlapped n^+-source/drain extension and a negatively charged accumulation layer is formed in the overlapping gate. In this case, the overlap gate current will mostly consist of electrons tunnelling from the gate accumulation layer to the source/drain, it is given by:

$$I_{G_{ov}} = W \cdot \Delta L_{ov} \cdot Q_{ov} \cdot P_{tunnel}\,(-V_{ov},\, \chi_B,\, A_{inv},\, B_{inv}) \tag{5.33}$$

Adding up (5.32) and (5.33), we obtain for the total gate overlap current:

$$I_{G_{ov}} = I_{GOV} \cdot {V_{ov}}^2 \cdot \left\{ \exp\left(-\frac{B_{inv}}{V_{ov}} \cdot \left[1 - \left(1 - \frac{V_{ov}}{\chi_B} \right)^{3/2} \right] \right) \right. \tag{5.34}$$

$$\left. - \exp\left(\frac{B_{inv}}{V_{ov}} \cdot \left[1 - \left(1 + \frac{V_{ov}}{\chi_B} \right)^{3/2} \right] \right) \right\}$$

where the parameter I_{GOV} is theoretically equal to $A_{inv} \cdot W \cdot \Delta L_{ov} \cdot C_{ox}$, but is used as an empirical parameter. Physically $I_{G_{ov}}$ should be zero for $V_{GX} = 0$, this is, however, not the case. Along the same lines as in Sections 5.1 and 5.2, we simply replace one of the V_{ov}'s in the prefactor of (5.34) by V_{GX} In order to force $I_{G_{ov}}$ to zero for $V_{GX} = 0$. The resulting expression is the MM11 expression. In Fig. 5.3 (b), the gate overlap current $I_{G_{ov}}$ is shown as a function of gate bias for a typical n-MOS transistor at $V_{DS} = 0$ (i.e. $I_{Gov_L} = I_{Gov_0}$).

5.4 Total Gate Current

Including all the above components, the model gives an accurate description of I_G over the whole operation region, see Figs. 5.1 (b) and 5.3 (b), using only 5 adjustable parameters: I_{GINV}, I_{GACC}, I_{GOV}, B_{inv} and B_{acc}. Similar accuracy is obtained for the V_{DS} dependence, see Figs. 5.6 and 5.7.

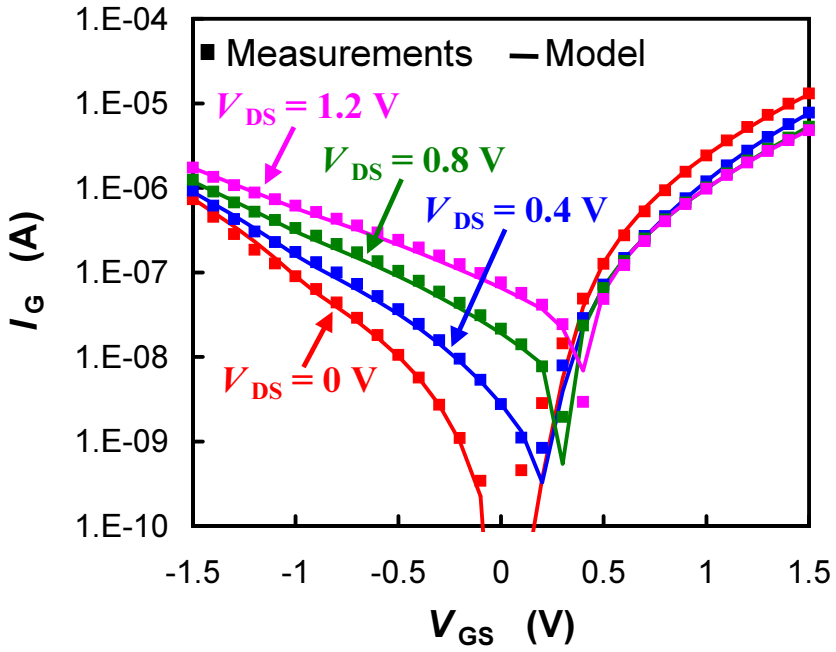


Figure 5.6: Gate current as a function of gate bias for various values of drain bias (n-MOS, $W/L = 10\mu\text{m}/0.6\mu\text{m}$, $t_{ox} = 1.4\text{nm}$).

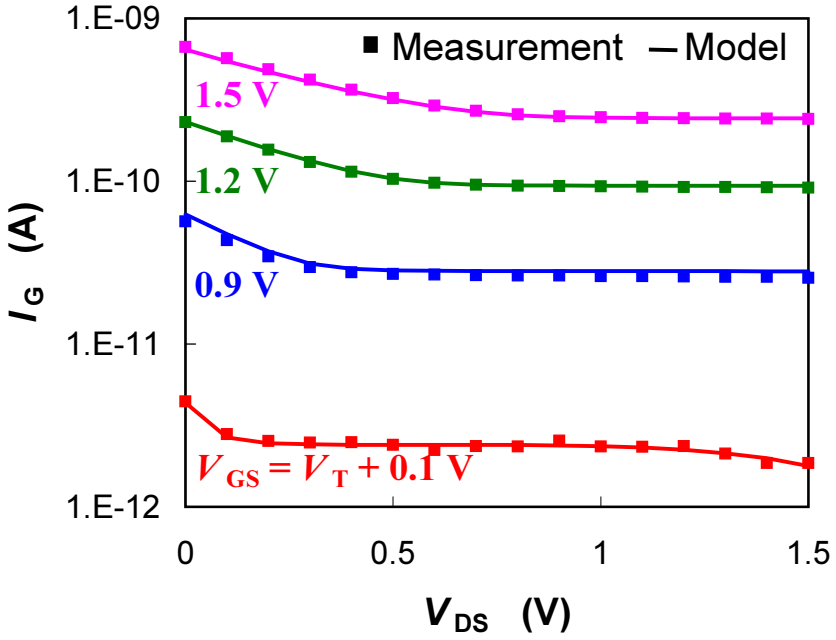


Figure 5.7: Gate current as a function of drain bias for various values of gate bias (n-MOS, $W/L = 10\mu\text{m}/10\mu\text{m}$, $t_{ox} = 2.2\text{nm}$).

6 Charge modelling

The dynamic behaviour of a MOSFET is not only determined by the time dependence of the steady-state currents, but also by the time dependence of the various charges in the transistor. These charges give rise to a capacitive behaviour of the MOSFET. In a typical MOS structure, we can distinguish intrinsic and extrinsic charges. The latter are due to the gate/source and gate/drain overlap regions. The drain/source junctions also contribute to the capacitive behaviour of the MOSFET, but this is not taken into account here. As mentioned before, it is described by a separate junction diode model.

6.1 Intrinsic Charges

In the intrinsic MOS transistor, charges can be attributed to the four terminals, see Fig. 6.1.

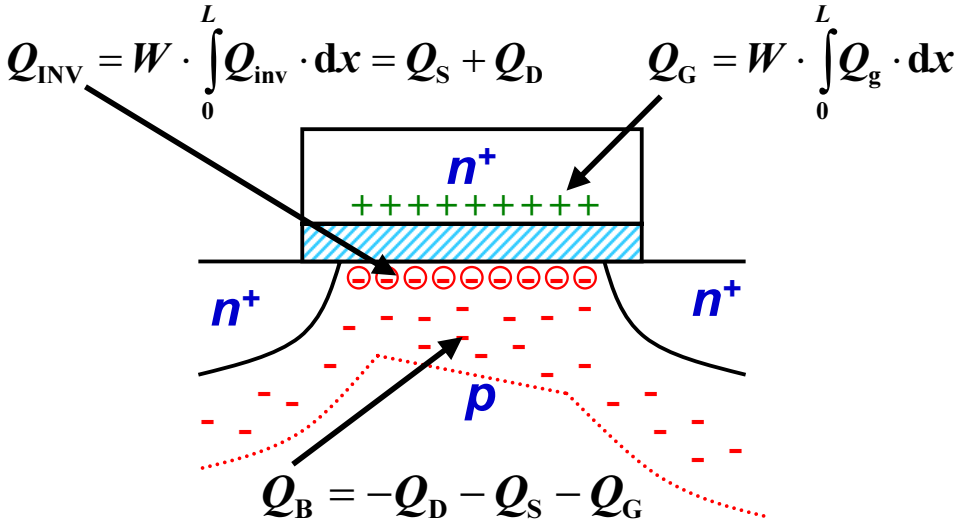


Figure 6.1: Cross-section of an n-MOSFET structure (biased in the inversion region). In a quasi-static approximation, charges can be attributed to the four terminals as indicated, where the total inversion-layer charge Q_{INV} is partitioned in a source (Q_S) and drain (Q_D) charge.

The total gate charge Q_G can be calculated by integrating Q_g along the channel:

$$Q_G = W \cdot \int_0^L Q_g \cdot dx \tag{6.1}$$

The total inversion-layer charge is split up in a source Q_S and a drain Q_D charge, which can be calculated using the Ward-Dutton charge partitioning scheme [139]:

$$Q_S = W \cdot \int_0^L \left(1 - \frac{x}{L}\right) \cdot Q_{inv} \cdot dx \tag{6.2}$$

$$Q_D = W \cdot \int_0^L \frac{x}{L} \cdot Q_{inv} \cdot dx \tag{6.3}$$

This partitioning scheme results in a bias-dependent or dynamic charge partitioning as opposed to a fixed partitioning as used in, e.g., BSIM4. Since charge neutrality holds for the complete transistor,

the total bulk charge Q_B is simply given by:

$$Q_B = W \cdot \int_0^L Q_b \cdot dx = -Q_S - Q_D - Q_G \tag{6.4}$$

Using the 4 above-mentioned charges, we can define 16 capacitances C_{ij}:

$$C_{ij} = \begin{cases} \frac{\partial Q_i}{\partial V_j} & \text{for: } i = j \\ -\frac{\partial Q_i}{\partial V_j} & \text{for: } i \neq j \end{cases} \tag{6.5}$$

where i and j denote the terminal S, D, G or B[13]. Owing to charge neutrality and to the fact that the 4 terminal voltages can be reduced to 3 voltage differences without losing information, the 16 above capacitances consist of 9 independent capacitances and 7 capacitances which can be written in terms of the 9 independent capacitances (e.g., $C_{GG} = C_{DG} + C_{SG} + C_{BG} = C_{GD} + C_{GS} + C_{GB}$).
It should be noted here that in general $C_{ij} \neq C_{ji}$, in other words the capacitances are non-reciprocal. This non-reciprocity is a direct consequence of the charge conversation law and the non-linear multiple terminal nature of the MOS device. Only at $V_{DS} = 0$, as the device is passive, the capacitances are reciprocal (i.e., $C_{ij} = C_{ji}$) and symmetrical w.r.t. source and drain (i.e., $C_{iD} = C_{iS}$ or $C_{Dj} = C_{Sj}$). It can be shown, however, that owing to the use of the charge sheet approximation, the capacitances are not exactly reciprocal at $V_{DS} = 0$, see Appendix E.2. The latter effect is nonetheless negiligible when compared to the non-reciprocity (at $V_{DS} = 0$) that occurs in V_T-based models.

The total charges as defined by (6.1)-(6.4) have to be calculated. In the following derivation, we will neglect the influence of velocity saturation, which is allowed as pointed out in Appendix F.2. Starting with the source and drain charge, equations (6.2) and (6.3) can be solved using (2.20). Using (L.2) and (L.4), see Appendix L, one can rewrite (6.3) to:

$$Q_D = \frac{W \cdot L}{2} \cdot \int_{Q^*_{inv_0}}^{Q^*_{inv_L}} \left(Q^*_{inv} - \phi_T \cdot C_{inv}\right) \cdot \frac{\left(Q^{*\,2}_{inv} - Q^{*\,2}_{inv_0}\right) \cdot Q^*_{inv}}{\left(\bar{Q}^*_{inv} \cdot \Delta Q^*_{inv}\right)^2} \cdot dQ^*_{inv} \tag{6.6}$$

which yields:

$$Q_D = \frac{W \cdot L}{2} \cdot \left(\frac{3 \cdot Q^{*\,3}_{inv_L} + 6 \cdot Q^*_{inv_0} \cdot Q^{*\,2}_{inv_L} + 4 \cdot Q^{*\,2}_{inv_0} \cdot Q^*_{inv_L} + 2 \cdot Q^{*\,3}_{inv_0}}{15 \cdot \bar{Q}^*_{inv} \cdot \bar{Q}^*_{inv}} - \phi_T \cdot C_{inv} \right) \tag{6.7}$$

Applying the same approach, one can calculate Q_S using (6.2):

$$Q_S = \frac{W \cdot L}{2} \cdot \left(\frac{2 \cdot Q^{*\,3}_{inv_L} + 4 \cdot Q^*_{inv_0} \cdot Q^{*\,2}_{inv_L} + 6 \cdot Q^{*\,2}_{inv_0} \cdot Q^*_{inv_L} + 2 \cdot Q^{*\,3}_{inv_0}}{15 \cdot \bar{Q}^*_{inv} \cdot \bar{Q}^*_{inv}} - \phi_T \cdot C_{inv} \right) \tag{6.8}$$

Bearing in mind that $Q^*_{inv_L} = \bar{Q}^*_{inv} + \Delta Q^*_{inv}/2$ and $Q^*_{inv_L} = \bar{Q}^*_{inv} - \Delta Q^*_{inv}/2$, and defining:

$$F_j = -\frac{1}{2} \cdot \frac{\Delta Q^*_{inv}}{\bar{Q}^*_{inv}} \tag{6.9}$$

[13] For certain bias conditions and certain combinations of i and j, the capacitance defined in (6.5) may be negative. Therefore, and to avoid confusion, the partial derivative C_{ij} is sometimes referred to as transcapacitance or capacitive coefficient.

it can be shown that (6.7) and (6.8) reduce to:

$$Q_{\mathrm{D}} = \frac{1}{2} \cdot \frac{C_{\mathrm{OX}}}{C_{\mathrm{ox}}} \cdot \left[\bar{Q}^*_{\mathrm{inv}} - \frac{\Delta Q^*_{\mathrm{inv}}}{6} \cdot \left(F_{\mathrm{j}} + \frac{F_{\mathrm{j}}^2}{5} - 1 \right) - \phi_{\mathrm{T}} \cdot C_{\mathrm{inv}} \right] \tag{6.10}$$

$$Q_{\mathrm{S}} = \frac{1}{2} \cdot \frac{C_{\mathrm{OX}}}{C_{\mathrm{ox}}} \cdot \left[\bar{Q}^*_{\mathrm{inv}} - \frac{\Delta Q^*_{\mathrm{inv}}}{6} \cdot \left(F_{\mathrm{j}} - \frac{F_{\mathrm{j}}^2}{5} + 1 \right) - \phi_{\mathrm{T}} \cdot C_{\mathrm{inv}} \right] \tag{6.11}$$

where C_{OX} is the total oxide capacitance given by $C_{\mathrm{ox}} \cdot W \cdot L$. The above two equations form the basis for the Q_{D} and Q_{S} equations used in MM11.
In order to calculate Q_{G} using (6.1), we can use (2.24). Furthermore, using (L.2) and bearing in mind that $\partial Q^*_{\mathrm{inv}} / \Delta Q^*_{\mathrm{inv}} = \partial \psi_{\mathrm{s}} / \Delta \psi$, one can rewrite (6.1) to:

$$Q_{\mathrm{G}} = W \cdot L \cdot \int_{\psi_{\mathrm{s}_0}}^{\psi_{\mathrm{s}_{\mathrm{L}}}} \left[\bar{Q}_{\mathrm{g}} - C_{\mathrm{g}} \cdot \left(\psi_{\mathrm{s}} - \bar{\psi} \right) \right] \cdot \frac{\left[\bar{Q}^*_{\mathrm{inv}} - C_{\mathrm{inv}} \cdot \left(\psi_{\mathrm{s}} - \bar{\psi} \right) \right]}{\bar{Q}^*_{\mathrm{inv}}} \cdot \frac{\mathrm{d}\psi_{\mathrm{s}}}{\Delta \psi} \tag{6.12}$$

which yields:

$$Q_{\mathrm{G}} = \frac{C_{\mathrm{OX}}}{C_{\mathrm{ox}}} \cdot \left(\bar{Q}_{\mathrm{g}} + \frac{1}{12} \cdot \frac{C_{\mathrm{g}} \cdot C_{\mathrm{inv}}}{\bar{Q}^*_{\mathrm{inv}}} \cdot \Delta \psi^2 \right) = \frac{C_{\mathrm{OX}}}{C_{\mathrm{ox}}} \cdot \left(\bar{Q}_{\mathrm{g}} - \frac{C_{\mathrm{g}}}{C_{\mathrm{inv}}} \cdot \frac{\Delta Q^*_{\mathrm{inv}}}{6} \cdot F_{\mathrm{j}} \right) \tag{6.13}$$

The total bulk charge Q_{B} is simply calculated from (6.4):

$$Q_{\mathrm{B}} = -\frac{C_{\mathrm{OX}}}{C_{\mathrm{ox}}} \cdot \left[\bar{Q}_{\mathrm{inv}} + \bar{Q}_{\mathrm{g}} - \frac{\Delta Q^*_{\mathrm{inv}}}{6} \cdot F_{\mathrm{j}} \cdot \left(1 + \frac{C_{\mathrm{g}}}{C_{\mathrm{inv}}} \right) \right] \tag{6.14}$$

The presence of source and drain resistance also affects the above charge model. In order to take the effect of series resistance into account, as pointed out in Appendix G.2, we simply replace $\Delta Q^*_{\mathrm{inv}}$ in (6.9)-(6.14) by:

$$\Delta Q^*_{\mathrm{inv}} \approx \frac{\Delta Q_{\mathrm{inv}}}{1 + \frac{G_{\mathrm{R}}}{G_{\mathrm{tot}}}} \tag{6.15}$$

Note that the above-defined $\Delta Q^*_{\mathrm{inv}}$ corresponds to the variable $\Delta V_{\mathrm{G_T}}$ ($= \Delta Q^*_{\mathrm{inv}} / C_{\mathrm{ox}}$) as used in MM11.

Following the above approach, an accurate description of the accumulation region and the poly-depletion effect are automatically included in the charge model, see Fig. 6.2. Note, however, that an electrical oxide thickness has to be used which is larger than the physical oxide thickness. This is due to the negligence of quantum-mechanical effects.

Quantum-mechanically, the inversion/accumulation charge concentration is not maximum at the Si-SiO_2-interface (as it would be in the classical case), but reaches a maximum at a distance Δy from the interface [39]. This quantum-mechanical effect can be taken into account by an effective oxide thickness $t_{\mathrm{ox_{eff}}}$, which is bias-dependent, see Appendix D:

$$t_{\mathrm{ox_{eff}}} = t_{\mathrm{ox}} \cdot \left[1 + QM_{\mathrm{t_{ox}}} \cdot \left(\frac{C_{\mathrm{ox}}}{\epsilon_{\mathrm{Si}} \cdot E_{\mathrm{eff}}} \right)^{1/3} \right] \tag{6.16}$$

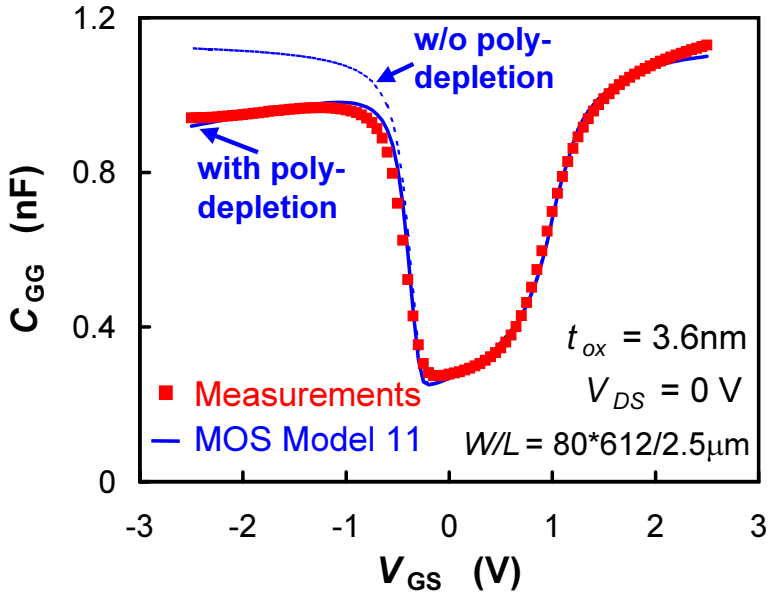


Figure 6.2: Input capacitance C_{GG} as a function of gate bias for a long-channel p-MOSFET. The influence of the poly-depletion effect is also shown. The used electrical t_{ox} is 3.6nm, which differs from the physical t_{ox} of 3.2nm due to quantum-mechanical effects.

where $QM_{t_{ox}}$ is equal to $2/5 \cdot QM \cdot {C_{ox}}^{2/3}$, and QM is a physical constant ($QM_N = 5.951993$ $Vm^{4/3}/C^{2/3}$ for electrons and $QM_P = 7.448711$ $Vm^{4/3}/C^{2/3}$ for holes). The quantum-mechanical carrier displacement can be taken into account in the above-derived charge model by replacing the total oxide capacitance C_{OX} by an effective oxide capacitance $C_{OX_{eff}}$ given by:

$$C_{OX_{eff}} = \frac{C_{OX}}{1 + QM_{t_{ox}} \cdot \left(\frac{C_{ox}}{\epsilon_{Si} \cdot E_{eff}}\right)^{1/3}} \tag{6.17}$$

In order to obtain the MM11 expression, we need to replace the effective field E_{eff} in (6.17) by an effective voltage $V_{eff} = \epsilon_{Si} \cdot \eta \cdot E_{eff}/C_{ox}$ and replace η by $1/\eta_{mob}$. Using no extra parameters, the implementation of $C_{OX_{eff}}$ results in an accurate charge description using the physical oxide thickness, see Fig. 6.3.

Non-Quasi-Static Effects: It should be noted here that the charge model as described above is quasi-static. The quasi-static approach assumes that a charge Q_X can be attributed to a terminal X and that this charge changes instantaneously with a changing voltage V_X. In other words, it assumes that the channel transit time is infinite, which is not physical. A finite transit time results, for example, in a phase shift (or delay) between channel current and gate voltage. This phase-shift is not taken into account in the quasi-static approach. This implies that for applications at high frequencies approaching the cut-off frequency, errors have to be expected due to non-quasi-static (NQS) effects. An accurate description of NQS-effects requires the solution of the continuity equation along the channel, which can only be obtained under certain approximations. Several methods to include NQS-effects have been proposed [140]-[143]. These methods are generally computation time intensive, do not include short-channel effects and in some cases consider the transient and ac small-signal behaviour separately, which can lead to inconsistent simulation results in the time domain and frequency domain. Another, less complex method is based on the small-signal approximation [144]-[147], the resulting models are not applicable to large-signal, transient and harmonic-balance simulations. An alternative simple and large-signal method makes use of the relaxation time approach [148], which

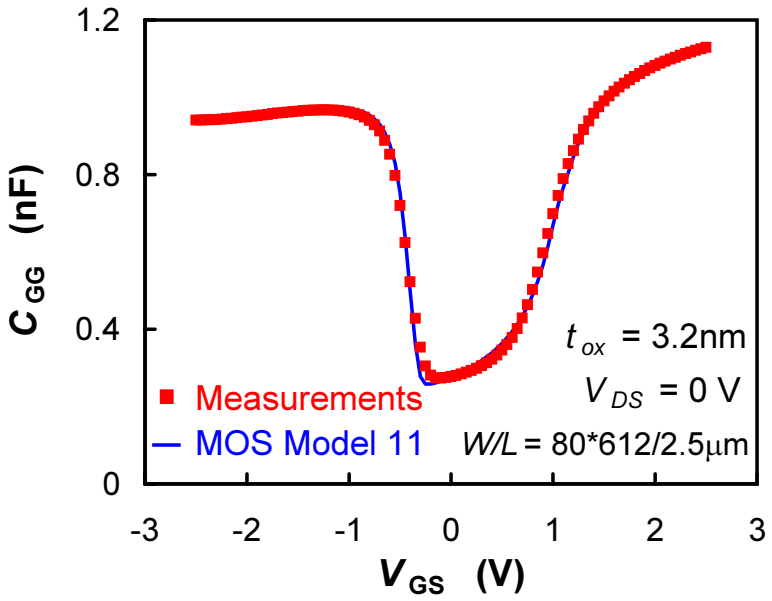


Figure 6.3: Same measurement results as used in Fig. 6.2. Quantum-mechanical effects have been included in the model, resulting in an electrical oxide thickness of 3.2nm, which corresponds to the physical oxide thickness.

is, however, inaccurate for certain bias conditions.
A simple yet accurate NQS-method that avoids most of the above-mentioned drawbacks, makes use of a so-called *segmentation model* [138]. In this method, the continuity equation is solved by breaking down the MOSFET channel into N equal segments in series, each described by a conventional quasi-static model. The resulting segmentation model is very accurate and consistent for dc steady-state, ac small-signal, large-signal and transient simulations. Of course, the use of N segments for each MOSFET will result in an increase in computation time. Bearing in mind, however, that in RF-circuits only a limited number of transistors are crucial for the high-frequency performance, the trade-off between the high accuracy and the computational penalty is nonetheless very positive.

6.2 Extrinsic Charges

For short-channel transistors, a major part of the total input capacitance C_{GG} will be determined by the gate-to-source and gate-to-drain overlap capacitances. An accurate modelling of these bias-dependent overlap capacitances is thus important. As discussed in Section 2.2, the overlap regions can be treated as n^+-gate/oxide/n^+-bulk MOS capacitances, where the source or drain acts as bulk terminal. The total charge in the overlap region Q_{OV} can simply be given by:

$$Q_{OV} = -C_{OV} \cdot V_{ov} \tag{6.18}$$

where V_{ov} is given by (2.39), C_{OV} is the oxide capacitance of the overlap region given by $C_{ox} \cdot W \cdot \Delta L_{ov}$, and ΔL_{ov} is the length of the gate/source or gate/drain overlap region. Again, the influence of quantum-mechanical effects could be taken into account in the overlap charge by making use of a bias-dependent effective oxide thickness. However, in order to limit calculation time, the quantum-mechanical effects are neglected in this case. Equation (6.18) gives an accurate description of the bias-dependent overlap capacitance, as is shown in Fig. 6.4. Replacing V_{ov} and C_{OV} in (6.18) by V_{ov_0} and C_{GSO}, respectively, we obtain the MM11-expression for Q_{OV_0}. In the same way, replacing V_{ov} and C_{OV} by V_{ov_L} and C_{GDO}, respectively, we obtain the MM11-expression for Q_{OV_L}.

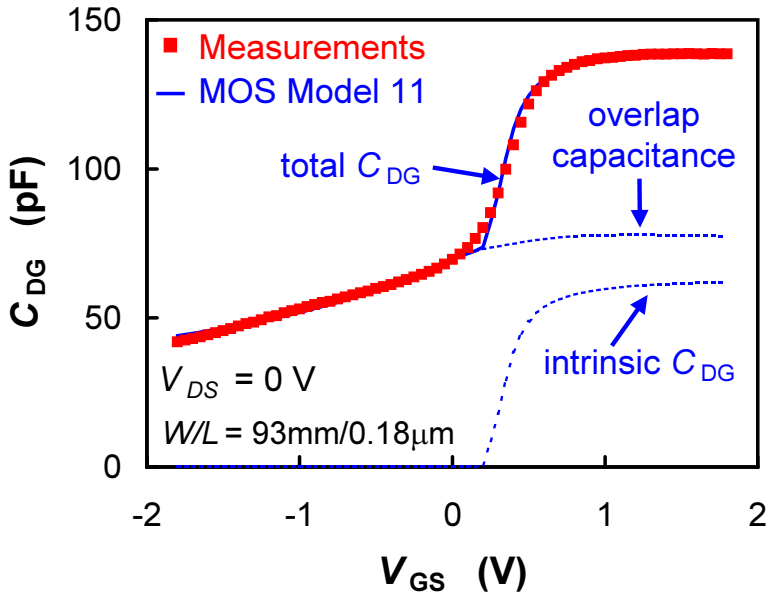


Figure 6.4: Drain-to-gate capacitance C_{DG} as a function of gate bias for a short-channel n-type transistor. The total C_{DG} consists of the intrinsic C_{DG} and the bias-dependent overlap capacitance.

7 Noise Modelling

So far, it has been assumed that the currents of a MOSFET vary with time only if one or more of the terminal voltages vary with time. However, irrespective of the presence of externally applied signals, a MOSFET shows spontaneous fluctuations in the terminal currents, referred to as *noise*. Such fluctuations can interfere with weak signals when the MOSFET is part of an analogue or RF circuit, and as a consequence, an accurate modelling of noise behaviour in circuit simulation is thus essential.

In a MOSFET, generally three different types of noise can be observed, see Fig. 7.1: $1/f$-noise, thermal noise and induced gate noise. These types of noise are all related to the channel current. In reality, the gate tunnel current and the bulk avalanche current will also exhibit noisy behaviour (due to shot noise), however, this has been neglected in MOS Model 11, Level 1101.

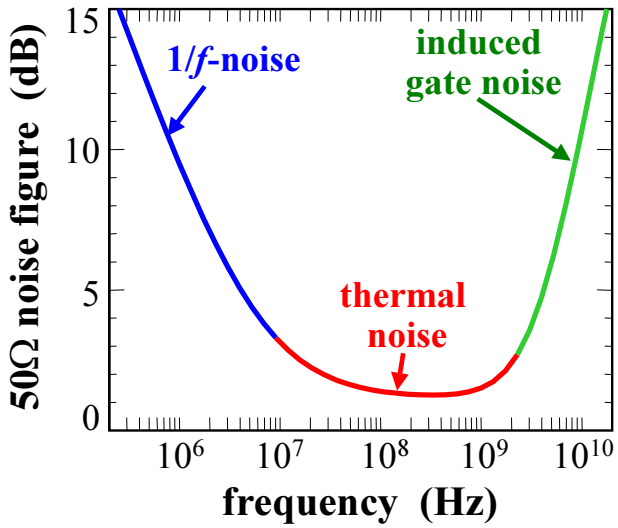
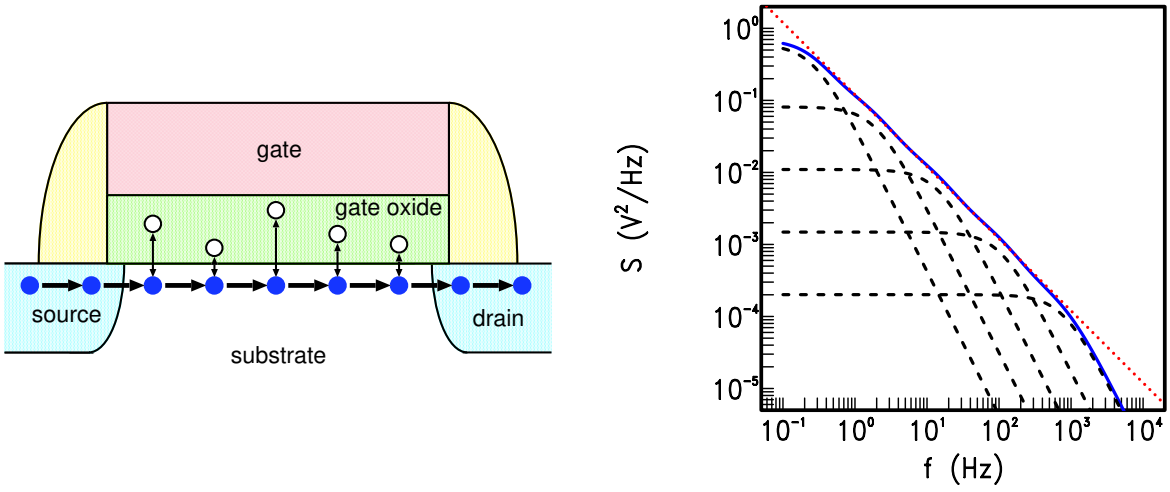


Figure 7.1: The 50Ω noise figure as a function of frequency for a typical MOSFET. Three different types of noise with different frequency dependence can be observed: $1/f$-noise, thermal noise and induced gate noise.

7.1 $1/f$-Noise

At low frequencies, flicker or $1/f$-noise becomes dominant in MOSFETs. In the past, this type of noise was interpreted either in terms of trapping and detrapping of charge carriers in the gate oxide or in terms of mobility fluctuations. Nowadays, a general $1/f$-noise model by Hung *et al* which combines both number and mobility fluctuations [149, 150], has found wide acceptance in the field of MOS modelling. The model assumes that the carrier number in the channel fluctuates due to trapping/detrapping in the gate oxide, see Fig. 7.2, and that these number fluctuations also affect the carrier mobility resulting in (correlated) mobility fluctuations. The model was originally formulated for V_T-based models, but it can easily be derived in terms of ψ_s resulting in an accurate expression for all operating regions. As a starting point we use eq. (22) in [150]:

$$\begin{aligned} S_{I_D}(f) &= \frac{k_B \cdot T \cdot {I_{DS}}^2}{\gamma_{ox} \cdot f \cdot W \cdot L^2} \cdot \int_0^L n_t(E_{fn}) \cdot \left[\frac{R}{N(x)} \pm \alpha_s \cdot \mu_{eff}\right]^2 \cdot dx \\ &= \frac{k_B \cdot T \cdot q \cdot I_{DS} \cdot \mu_{eff}}{\gamma_{ox} \cdot f \cdot L^2} \cdot \int_{V_{SB}}^{V_{DB}} n_t(E_{fn}) \cdot \frac{R^2}{N} \cdot \left[1 \pm \alpha_s \cdot \mu_{eff} \cdot \frac{N}{R}\right]^2 \cdot dV \end{aligned} \tag{7.1}$$



(a) (b)

Figure 7.2: (a) Flicker or $1/f$-noise in MOSFETs is attributed to the trapping/detrapping of carriers in the gate oxide, resulting in carrier number fluctuations and correlated mobility fluctuations. (b) The fluctuations caused by a single trap result in a Lorentzian-type noise spectral density. A distribution of traps in the oxide result in a distribution of Lorentzians. The addition of only 5 Lorentzians (dashed lines) already leads to a $1/f$-like spectrum (solid line) over a considerable frequency range. For reference, the dotted line represents true $1/f$ noise (corresponding to a spatially unifrom distribution of traps).

where $n_t(E_{fn})$ is the approximated distribution of oxide traps over energy, γ_{ox} is the attenuation coefficient of the electron wave function in the oxide, α_s is a scattering coefficient responsible for the mobility fluctuations, N is the number of channel carrier per unit area, $N = -Q_{inv}/q$, and R is the ratio of the fluctuations in carrier number (ΔN) to fluctuations in occupied trap number (ΔN_t):

$$R \equiv \frac{\partial \Delta N}{\partial \Delta N_t} \tag{7.2}$$

If ΔN_t fluctuates by an amount $\partial \Delta N_t$, then in strong inversion $\partial \Delta N = -\partial \Delta N_t$ (i.e., $R = -1$), and in weak inversion $|\partial \Delta N| \ll |\partial \Delta N_t|$ (i.e., $|R| \ll 1$). In order to find an expression for R in weak inversion, we note that when the trapped charge Q_t fluctuates, this produces fluctuations in the gate charge ∂Q_g, the inversion charge ∂Q_{inv} and the bulk charge ∂Q_b. Charge conservation requires:

$$\partial Q_g + \partial Q_{inv} + \partial Q_b + \partial Q_t = 0 \tag{7.3}$$

or:

$$R \equiv \frac{\partial \Delta N}{\partial \Delta N_t} = -\frac{\partial Q_{inv}}{\partial Q_g + \partial Q_b + \partial Q_{inv}} \tag{7.4}$$

$$\approx \frac{Q_{inv}/\phi_T}{C_g + C_b - Q_{inv}/\phi_T}$$

where $\partial Q_{inv}/\partial \psi_s$ in weak inversion has been approximated by $-Q_{inv}/\phi_T$. Eq. (7.4) can be rewritten as:

$$R = -\frac{N}{N + N^*} \tag{7.5}$$

with

$$N^* = \frac{\phi_T}{q} \cdot \left[C_g + C_b\right] \approx -\frac{\phi_T}{q} \cdot C_{inv} \tag{7.6}$$

wheNote that, although (7.5) is valid in weak inversion, R approaches the theoretical value of -1 in strong inversion. As a result, we can use (7.5) over the whole operation region.
To further proceed in solving (7.1), one needs to know the bias dependency of α_s, μ_{eff} and $n_t\,(E_{fn})$. In order to keep things mathematically feasible, the following parametrisation is made:

$$n_t^*\,(E_{fn}) = n_t\,(E_{fn}) \cdot \left[1 \pm \alpha_s \cdot \mu_{eff} \cdot \frac{N}{R}\right]^2 = A + B \cdot N + C \cdot N^2 \tag{7.7}$$

where A, B and C are treated as empirical parameters. Furthermore $\mathrm{d}V$ can be written in terms of $\mathrm{d}N$:

$$\mathrm{d}V = \frac{\partial V}{\partial \psi_s} \cdot \frac{\partial \psi_s}{\partial N} \cdot \mathrm{d}N = \frac{q}{C_{inv}} \cdot \frac{\partial V}{\partial \psi_s} \cdot \mathrm{d}N \tag{7.8}$$

where $\partial V/\partial \psi_s$ is given by (3.5), which can be approximated by, using $partial Q_{inv}/\partial \psi_s \approx -C_{inv}$:

$$\frac{\partial V}{\partial \psi_s} \approx \frac{N + N^*}{N} \tag{7.9}$$

Now using (7.7), (7.8) and (7.9), we can rewrite (7.1) as:

$$S_{I_D} = -\frac{{\phi_T}^2 \cdot q^2 \cdot I_{DS} \cdot \mu_{eff}}{\gamma_{ox} \cdot f \cdot L^2 \cdot N^*} \cdot \int_{N_0}^{N_L} \frac{N + N^*}{N} \cdot \frac{R^2}{N} \cdot \left[A + B \cdot N + C \cdot N^2\right] \cdot \mathrm{d}N \tag{7.10}$$

which results in the following equation:

$$S_{I_D} = \frac{{\phi_T}^2 \cdot q^2 \cdot I_{DS} \cdot \mu_{eff}}{\gamma_{ox} \cdot f \cdot L^2 \cdot N^*} \cdot \left\{A \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right) + B \cdot \left[N_0 - N_L - N^* \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right)\right]\right.$$

$$\left. + C \cdot \left[\frac{{N_0}^2 - {N_L}^2}{2} - N^* \cdot (N_0 - N_L) + {N^*}^2 \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right)\right]\right\} \tag{7.11}$$

In the saturation operation region, channel length modulation may affect the noise density. In order to take this effect into account, we rewrite (7.1):

$$S_{I_D} = \frac{k_B \cdot T \cdot {I_{DS}}^2}{\gamma_{ox} \cdot f \cdot W \cdot L^2} \cdot \left[\int_0^{L - \Delta L} n_t^*\,(E_{fn}) \cdot \frac{R^2}{N^2} \cdot \mathrm{d}x + \int_{L - \Delta L}^{L} n_t^*\,(E_{fn}) \cdot \frac{R^2}{N^2} \cdot \mathrm{d}x\right] \tag{7.12}$$

where the first term results in (7.11) and the second term is due to channel length modulation. In order to calculate the latter, we assume that both the electron quasi-Fermi level and carrier density are uniform in the saturation region ΔL and equal to those at the saturation point (in other words $N = N_L$ and $V = V_{SB} + V_{DS_{sat}}$). In this case, it is straightforward to show that the total noise power becomes:

$$S_{I_D} = \frac{{\phi_T}^2 \cdot q^2 \cdot I_{DS} \cdot \mu_{eff}}{\gamma_{ox} \cdot f \cdot L^2 \cdot N^*} \cdot \left\{A \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right) + B \cdot \left[N_0 - N_L - N^* \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right)\right]\right.$$

$$+C \cdot \left[\frac{{N_0}^2 - {N_L}^2}{2} - N^* \cdot (N_0 - N_L) + {N^*}^2 \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right)\right]\Bigg\} \quad (7.13)$$

$$+ \Delta L \cdot \frac{k_B \cdot T \cdot {I_{DS}}^2}{\gamma_{ox} \cdot f \cdot W \cdot L^2} \cdot \frac{A + B \cdot N_L + C \cdot {N_L}^2}{(N_L + N^*)^2}$$

In order to simplify (7.13), the following parameters are defined:

$$N_{FA} = \frac{q \cdot A}{\gamma_{ox} \cdot L \cdot W} \quad (7.14)$$

$$N_{FB} = \frac{q \cdot B}{\gamma_{ox} \cdot L \cdot W} \quad (7.15)$$

$$N_{FC} = \frac{q \cdot C}{\gamma_{ox} \cdot L \cdot W} \quad (7.16)$$

Replacing μ_{eff} by μ_0 / G_{vsat} in order to take into account both mobility reduction and velocity saturation, equation (7.13) can now be rewritten as:

$$S_{I_D} = \frac{q \cdot {\phi_T}^2 \cdot \beta \cdot I_{DS}}{f \cdot C_{ox} \cdot G_{mob} \cdot N^*} \cdot \left[\left(N_{FA} - N^* \cdot N_{FB} + {N^*}^2 \cdot N_{FC}\right) \cdot \ln\left(\frac{N_0 + N^*}{N_L + N^*}\right)\right.$$

$$\left. + \left(N_{FB} - N^* \cdot N_{FC}\right) \cdot (N_0 - N_L) + \frac{N_{FC}}{2} \cdot \left({N_0}^2 - {N_L}^2\right)\right] \quad (7.17)$$

$$+ \frac{\phi_T \cdot {I_{DS}}^2}{f} \cdot (1 - G_{\Delta L}) \cdot \left[\frac{N_{FA} + N_{FB} \cdot N_L + N_{FC} \cdot {N_L}^2}{(N_L + N^*)^2}\right]$$

Equation (7.17) is the expression for the $1/f$-noise spectral density S_{fl} as used in MM11 [17].

7.2 Thermal Noise

Thermal (or Nyquist) noise is caused by the random thermal (or Brownian) motion of carriers. In a MOSFET, the channel current I_{DS} exhibits thermal noise. As can be shown, see Appendix M.1, the thermal-noise spectral density of a MOSFET is given by [35, 152, 153]:

$$S_{I_D}(f) = \frac{4 \cdot k_B \cdot T}{I_{DS} \cdot L^2} \cdot \int_{V_{SB}}^{V_{DB}} g^2(V) \cdot \mathrm{d}V \quad (7.18)$$

where $g(V)$ is the channel conductance at a specific point along the channel, given by:

$$g(V) = -\mu(V) \cdot W \cdot Q_{inv}(V) \quad (7.19)$$

For n-type transistors, the mobility $\mu(V)$, including velocity saturation, is given by:

$$\mu(V) = \frac{\mu_{eff}}{\sqrt{1 + \left(\frac{\mu_{eff}}{v_{sat}} \cdot \frac{\partial \psi_s}{\partial x}\right)^2}} \quad (7.20)$$

where the effective mobility μ_{eff} is given by (3.32). Using eqs. (3.5) and (7.20), we can rewrite the expression (3.1) for channel current into:

$$I_{\text{DS}} = -\frac{\mu_{\text{eff}} \cdot W \cdot Q^*_{\text{inv}}}{\sqrt{1 + \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot \frac{\partial \psi_{\text{s}}}{\partial x}\right)^2}} \cdot \frac{\partial \psi_{\text{s}}}{\partial x} \tag{7.21}$$

Solving $\partial\psi_{\text{s}}/\partial x$ in (7.21) yields an explicit expression:

$$\frac{\partial \psi_{\text{s}}}{\partial x} = \frac{I_{\text{DS}}}{\sqrt{{\mu_{\text{eff}}}^2 \cdot W^2 \cdot {Q^*_{\text{inv}}}^2 - \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot I_{\text{DS}}\right)^2}} \tag{7.22}$$

Now, we can evaluate $g(V)$ by inserting the above equation into (7.20) and then next into (7.19), which results in:

$$g(V) = \sqrt{{\mu_{\text{eff}}}^2 \cdot W^2 \cdot {Q^*_{\text{inv}}}^2 - \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot I_{\text{DS}}\right)^2} \cdot \frac{Q_{\text{inv}}}{Q^*_{\text{inv}}} \tag{7.23}$$

For an evaluation of the thermal noise according to (7.18), we need to find an analytical expression for $g^2(V) \cdot \mathrm{d}V$. Using (7.23) for $g(V)$ and (3.5) for $\mathrm{d}V$, we can write:

$$\begin{aligned} g^2(V) \cdot \mathrm{d}V &= g^2(\psi_{\text{s}}) \cdot \frac{Q^*_{\text{inv}}}{Q_{\text{inv}}} \cdot \mathrm{d}\psi_{\text{s}} \\ &= \left[{\mu_{\text{eff}}}^2 \cdot W^2 \cdot Q_{\text{inv}} \cdot Q^*_{\text{inv}} - \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot I_{\text{DS}}\right)^2 \cdot \frac{Q_{\text{inv}}}{Q^*_{\text{inv}}}\right] \cdot \mathrm{d}\psi_{\text{s}} \end{aligned} \tag{7.24}$$

The last term of the above equation is determined by the effect of velocity saturation, and is thus only of importance in the strong inversion region, where drift current is dominant. Simplifying the influence of diffusion on the velocity saturation term, eq. (7.24) can be approximated by:

$$g^2(V) \cdot \mathrm{d}V \approx \left[{\mu_{\text{eff}}}^2 \cdot W^2 \cdot Q_{\text{inv}} \cdot Q^*_{\text{inv}} - \left(\frac{\mu_{\text{eff}}}{v_{\text{sat}}} \cdot I_{\text{DS}}\right)^2 \cdot \frac{\bar{Q}_{\text{inv}}}{\bar{Q}^*_{\text{inv}}}\right] \cdot \mathrm{d}\psi_{\text{s}} \tag{7.25}$$

The integration in (7.18) can now simply be performed:

$$\begin{aligned} S_{\text{I}_\text{D}} &= \frac{4 \cdot k_{\text{B}} \cdot T}{I_{\text{DS}} \cdot L^2} \cdot {\mu_{\text{eff}}}^2 \cdot W^2 \cdot \int_{\psi_{\text{s}_0}}^{\psi_{\text{s}_\text{L}}} \left[Q_{\text{inv}} \cdot Q^*_{\text{inv}} - \left(\frac{I_{\text{DS}}}{v_{\text{sat}} \cdot W}\right)^2 \cdot \frac{\bar{Q}_{\text{inv}}}{\bar{Q}^*_{\text{inv}}}\right] \cdot \mathrm{d}\psi_{\text{s}} \\ &= \frac{4 \cdot k_{\text{B}} \cdot T}{I_{\text{DS}} \cdot L^2} \cdot {\mu_{\text{eff}}}^2 \cdot W^2 \cdot \left[\bar{Q}_{\text{inv}} \cdot \bar{Q}^*_{\text{inv}} + \frac{{C_{\text{inv}}}^2}{12} \cdot \Delta\psi^2 - \left(\frac{I_{\text{DS}}}{v_{\text{sat}} \cdot W}\right)^2 \cdot \frac{\bar{Q}_{\text{inv}}}{\bar{Q}^*_{\text{inv}}}\right] \cdot \Delta\psi \end{aligned} \tag{7.26}$$

Using $\mu_{\text{eff}} = \mu_0 / G_{\text{mob}}$ and $\theta_{\text{sat}} = \mu_0/(v_{\text{sat}} \cdot L)$, we obtain:

$$S_{\text{I}_\text{D}} = \frac{4 \cdot k_{\text{B}} \cdot T}{{G_{\text{mob}}}^2} \cdot \left[\left(\frac{\beta}{C_{\text{ox}}}\right)^2 \cdot \frac{\bar{Q}_{\text{inv}} \cdot \bar{Q}^*_{\text{inv}} + {C_{\text{inv}}}^2/12 \cdot \Delta\psi^2}{I_{\text{DS}}} - {\theta_{\text{sat}}}^2 \cdot I_{\text{DS}} \cdot \frac{\bar{Q}_{\text{inv}}}{\bar{Q}^*_{\text{inv}}}\right] \cdot \Delta\psi \tag{7.27}$$

The above equation can further be developed by inserting the drain current expression, neglecting the influence of series resistance and self-heating for the moment:

$$I_{\text{DS}} = -\frac{\beta}{G_{\text{vsat}}} \cdot \frac{\bar{Q}^*_{\text{inv}}}{C_{\text{ox}}} \cdot \Delta\psi \tag{7.28}$$

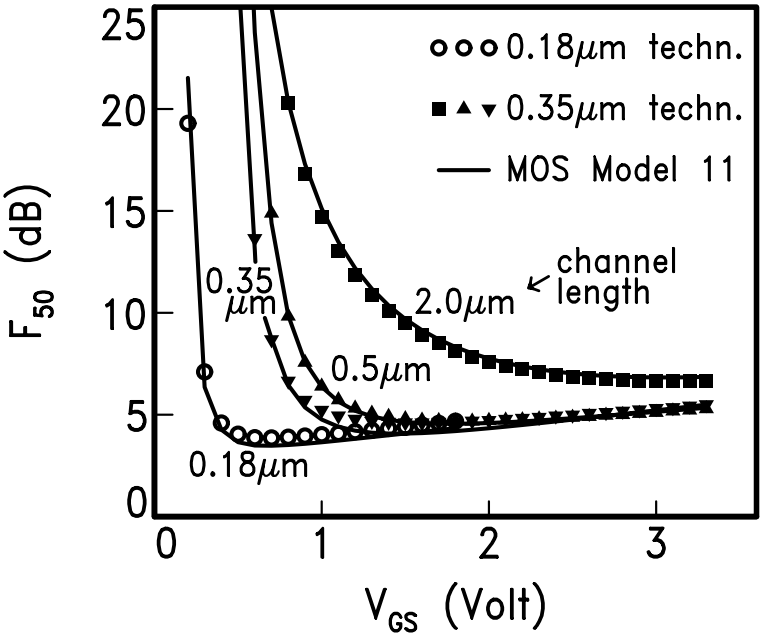


Figure 7.3: The 50Ω noise figure F_{50} as a function of gate bias for n-type MOSFETs in 0.35μm technology and for a short-channel n-type MOSFET in 0.18μm technology [8] ($V_{DS} = V_{DD}$, $W = 16 \times 10\mu$m).

Expression (7.27) can now be simplified to:

$$S_{I_D} = \frac{4 \cdot k_B \cdot T}{{G_{mob}}^2} \cdot \left[-\frac{\beta}{C_{ox}} \cdot G_{vsat} \cdot \left(\bar{Q}_{inv} + \frac{{C_{inv}}^2 \cdot \Delta\psi^2}{12 \cdot \bar{Q}^*_{inv}} \right) - {\theta_{sat}}^2 \cdot I_{DS} \cdot \Delta\psi \right] \tag{7.29}$$

For an accurate description of the total thermal noise spectral density, we need to include the influence of series resistance[14] as well. Using the derivation as outlined in Appendix M.2, the following expression for the spectral density S_{I_D} including the impact of series resistance can be derived:

$$S_{I_D} \approx \frac{4 \cdot k_B \cdot T}{{G_{mob}}^2} \cdot \left[-\frac{\beta}{C_{ox}} \cdot \frac{{G_{vsat}}^2}{G_{tot}} \cdot \left(\bar{Q}_{inv} + \frac{{C_{inv}}^2 \cdot \Delta\psi^2}{12 \cdot \bar{Q}^*_{inv}} \right) - {\theta_{sat}}^2 \cdot I_{DS} \cdot \Delta\psi \right] \tag{7.30}$$

This basically is the expression for the thermal noise spectral density S_{th} as used in MM11, where the term $4 \cdot k_B \cdot T$ has been replaced by parameter N_T. This thermal noise model, making use of the appropriate equation for velocity saturation, has been found to accurately describe experimental results for various CMOS technologies [8, 15], see Fig. 7.3. Some publications [154, 155] have reported an increased thermal noise in submicron MOSFETs with respect to long-channel thermal noise predictions, which has been ascribed to hot-carrier effects. The above noise model, nevertheless, gives an accurate description for all channel lengths without having to invoke carrier heating effects.

The above derivation for thermal noise holds for n-type MOSFETs. For p-type MOSFETs, a different expression (3.38) for velocity saturation has to be used. Using (3.38) complicates the derivation of thermal noise, but it can be simplified, along the same lines as was done in Section 3.3.2. In other words, the n-type MOS equations can simply be used for p-type MOS equations when θ_{sat} is replaced by $\theta_{sat}/(1 + {\theta_{sat}}^2 \cdot \Delta\psi^2)^{1/4}$, resulting in an accurate expression for thermal noise [8].

[14]The thermal noise contributed by the drain and the source series resistance is neglible, and both series resistances are consequently considered noiseless.

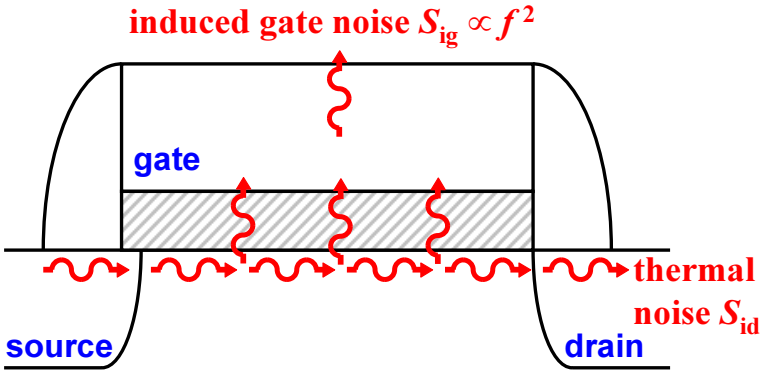


Figure 7.4: The fluctuating channel current induces noise in the gate current S_{I_G} owing to the capacitive coupling between gate and channel. Since the thermal noise in the channel is frequency independent, the resulting induced gate noise S_{I_G} is proportional to f^2 and consequently only becomes important at high frequencies.

7.3 Induced Gate Noise Modelling

Owing to capacitive coupling between the gate and the channel, the fluctuating channel current (as calculated in Section 7.2) induces a noise current in the gate terminal at high frequencies, see Fig. 7.4. This type of noise is generally referred to as induced gate noise. Hence, apart from the channel current thermal noise spectral density S_{I_D}, the high-frequency noise also consists of the induced gate noise spectral density S_{I_G}. Unfortunately the calculation of this component from first principles is too complicated to provide a result applicable to circuit simulation. It is more practical to derive the desired result from an equivalent circuit presentation given in Fig. 7.5. Owing to the mentioned capacitive coupling, a part of the channel is present as a resistance in series with the gate input capacitance. In saturation, this resistance is approximately equal to:

$$R_i = \frac{1}{3 \cdot g_m} \tag{7.31}$$

where g_m is the transconductance. It can be easily shown that the latter resistance produces an input noise current with a spectral density given by:

$$S_{I_G} = 4 \cdot k_B \cdot R_i \cdot \omega^2 \cdot {C_{OX}}^2 \tag{7.32}$$

where ω is the angular frequency ($= 2 \cdot \pi \cdot f$), and C_{OX} is the total oxide capacitance ($C_{ox} \cdot W \cdot L$. In addition, since Δi_d and Δi_g have the same physical source, both spectral densities are correlated. In saturation, it can be shown that the correlation $\rho_{I_G \cdot I_D}$ is approximately equal to [152]:

$$\rho_{I_G \cdot I_D} = 0.4 \cdot j \tag{7.33}$$

The cross-correlation spectral density $S_{I_G \cdot I_D}$ is simply expressed as:

$$S_{I_G \cdot I_D} = \rho_{I_G \cdot I_D} \cdot \sqrt{S_{I_G} \cdot S_{I_D}} \tag{7.34}$$

In order to obtain the MM11-expressions for induced gate noise, we need to replace the variables S_{I_G}, $\rho_{I_G \cdot I_D}$ and $S_{I_G \cdot I_D}$ by S_{ig}, ρ_{igth} and S_{igth}, respectively.

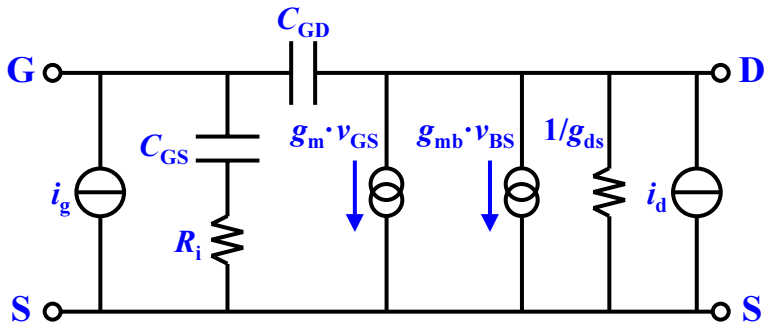


Figure 7.5: Noise current sources in the electrical scheme of the MOS transistor: i_d is the thermal noise current, i_g is the correlated induced gate noise current, and R_i is the intrinsic channel resistance.

The induced gate noise basically is a non-quasi static (NQS) effect. As a result, the use of an NQS segmentation model [138] (see Section 6.1) where each segment solely includes the channel-current thermal noise description, automatically results in a correct description of induced gate noise for all bias conditions [15]. For this purpose, the induced gate noise S_{I_G} can be made equal to zero by using parameter GATENOISE in MM11.

Appendices

A Smoothing Functions

In this Appendix, the various mathematical smoothing functions used in MOS Model 11 are defined. These smoothing functions are used to obtain a continuous transition from one mathematical function to the other at a certain transition point. Not only the smoothing function should be continuous at the transition point, but its higher-order derivatives as well. In other words, it should be C_∞-continuous.

A.1 Hyp Functions

In some cases, it is essential to use a function $f(x)$ which changes C_∞-continuously from zero for $x < 0$ to x for $x < 0$. This can be realized using a so-called hyp-function:

$$\mathrm{hyp}_1\{x;\epsilon\} = \frac{1}{2}\cdot\left(x+\sqrt{x^2+4\cdot\epsilon^2}\right) \tag{A.1}$$

where ϵ is a parameter that determines the smoothness of the transition, see Fig. A.1.

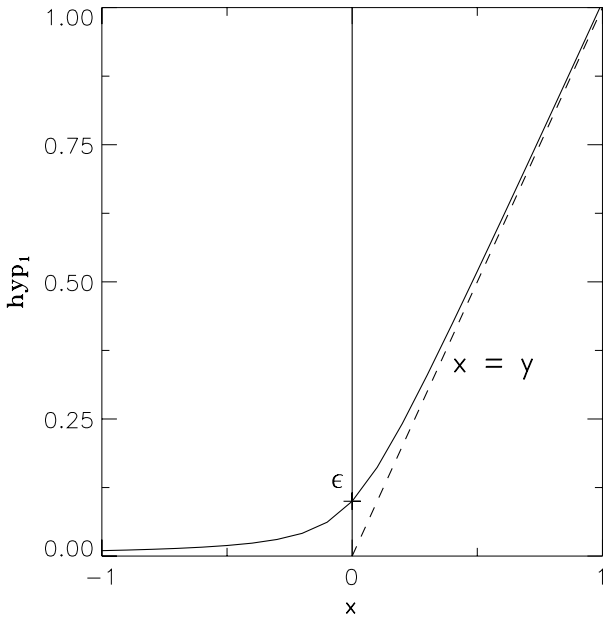


Figure A.1: The hyp-function $\mathrm{hyp}_1\{x;\epsilon\}$ realizes a C_∞-continuous transition from zero for $x < 0$ to x for $x < 0$. Parameter ϵ determines the smoothness of the transition.

A.2 Ohmic/Saturation Smoothing Function

For analog circuit simulation the transition from the ohmic region to the saturation region should not only be continuous for drain current, but it should also be C_∞-continuous. In most compact MOSFET models a smooth transition is obtained by using an empirical function, a so-called smoothing function. This smoothing function $V_{\mathrm{DS_x}}$ is ideally equal to V_{DS} in the ohmic region and to $V_{\mathrm{DS_{sat}}}$ in the saturation region. In order to preserve model symmetry at $V_{\mathrm{DS}} = 0$, the smoothing function $V_{\mathrm{DS_x}}$ should ensure that *i)* $\partial V_{\mathrm{DS_x}}/\partial V_{\mathrm{DS}} = 1$ at $V_{\mathrm{DS}} = 0$ and *ii)* $\partial^{\mathrm{n}} V_{\mathrm{DS_x}}/\partial {V_{\mathrm{DB}}}^{\mathrm{n}} = -\partial^{\mathrm{n}} V_{\mathrm{DS_x}}/\partial {V_{\mathrm{SB}}}^{\mathrm{n}}$ at $V_{\mathrm{DS}} = 0$ for $\mathrm{n} = 2$ and higher [7, 26]. A smoothing function that satisfies the above requirements has been introduced in [26]:

$$V_{\mathrm{DS_x}} = \frac{V_{\mathrm{DS}}\cdot V_{\mathrm{DS_{sat}}}}{\left[{V_{\mathrm{DS}}}^{2\cdot m} + {V_{\mathrm{DS_{sat}}}}^{2\cdot m}\right]^{\frac{1}{2\cdot m}}} \tag{A.2}$$

where m is an empirical parameter, which determines the smoothness of the transition, see Fig. A.2.

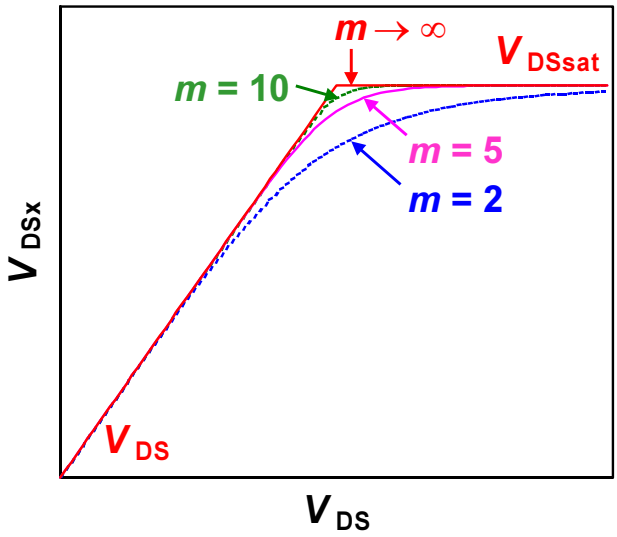


Figure A.2: Smoothing function V_{DS_x} realizes a C_∞-continuous transition from V_{DS} in the ohmic region (i.e., for $V_{DS} < V_{DS_{sat}}$) to $V_{DS_{sat}}$ in the saturation region (i.e., for $V_{DS} > V_{DS_{sat}}$). Parameter m determines the smoothness of the transition.

B Calculation of Surface Potential

B.1 Poly-Depletion Effect

The use of a polysilicon gate results in the formation of a depletion layer at the gate/SiO_2-interface (for $V_{GB} > V_{FB}$). A potential drop ψ_p falls across this polysilicon depletion layer (i.e., the electrostatic potential at the gate/gate-oxide interface with respect to the neutral gate, see Fig. 2.3 (a)) affecting the calculation of surface potential ψ_s. Including the poly-depletion potential drop ψ_p, equation (2.6) can be written as:

$$Q_s = -C_{ox} \cdot \left(V^*_{GB} - \psi_s - \psi_p\right) \tag{B.1}$$

In order to calculate ψ_p, the same derivations as above can be done for the n^+-type polysilicon gate. Since the gate is very highly n-type doped, we assume that the hole density p' in the gate can be neglected for practical bias conditions[15]. In this case, the gate space charge $\rho'(x, y)$ in the polysilicon gate is given by:

$$\rho'(x, y) = q \cdot \left[N_P - n'(x, y)\right] \tag{B.2}$$

where N_P is the net donor doping concentration in the polysilicon gate, and the polysilicon electron density n' is given by Maxwell-Boltzmann statistics[16]:

$$n'(x, y) \approx N_P \cdot \exp\left(-\frac{\psi'(x,y)}{\phi_T}\right) \tag{B.3}$$

Here ψ' is the electrostatic potential with respect to the neutral gate. Under the gradual channel approximation, the 1-D Poisson equation for the gate is written as:

$$\frac{\partial^2\psi'}{\partial y^2} \approx \frac{q \cdot N_P}{\epsilon_{Si}} \cdot \left[1 - \exp\left(-\psi'\phi_T\right)\right] \tag{B.4}$$

Again as boundary conditions both ψ' and $\partial\psi'/\partial y$ are taken to be equal to zero deep into the neutral gate. From Gauss' law the gate charge density Q_g is given by:

$$\begin{aligned} Q_g &= \epsilon_{Si} \cdot \left.\frac{\partial\psi'}{\partial y}\right|_{y=-t_{ox}} \\ &= \pm k_P \cdot C_{ox} \cdot \sqrt{\psi_p + \phi_T \cdot \left[\exp\left(-\frac{\psi_p}{\phi_T}\right) - 1\right]} \end{aligned} \tag{B.5}$$

where k_p is the gate body effect coefficient given by $\sqrt{2 \cdot q \cdot \epsilon_{Si} \cdot N_P}/C_{ox}$.
Due to charge neutrality $Q_g = -Q_s$, and as a result (2.5), (2.8) and (B.5) can be equated, resulting in two implicit equations from which both ψ_s and ψ_p can be calculated. In order to simplify matters a bit, a distinction is made between the accumulation (i.e., $V^*_{GB} < 0$) and the inversion (i.e., $V^*_{GB} > 0$) operation region. In the accumulation region, an accumulation layer of holes is formed in the substrate

[15] In other words it is assumed that only depletion and accumulation can occur in the polysilicon gate. The possibility of strong inversion is neglected. This implies that no inversion layer is formed in the polysilicon gate, which holds true for practical operation conditions.

[16] In practice the polysilicon gate is degenerately doped and as a result Fermi-Dirac statistics should be used.

and an accumulation layer of electrons is formed in the gate. In this case both ψ_s and ψ_p are negative, and consequently (2.5) can be approximated by:

$$Q_s \approx -k_0 \cdot C_{ox} \cdot \sqrt{\psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right]} \tag{B.6}$$

Equating (B.5) and (B.6), and bearing in mind that the polysilicon gate is much more highly doped than the silicon subtrate, in other words $k_0 \ll k_P$, it is easily seen that $|\psi_s| \gg |\psi_p|$. In the accumulation region ψ_p is thus approximately equal to zero.
In the inversion region, an inversion layer of electrons and a depletion layer of ionized acceptor atoms is formed in the substrate, and a depletion layer of ionized donor atoms is formed in the gate. In this case both ψ_s and ψ_p are positive, and consequently (B.5) can be approximated by:

$$Q_g \approx k_P \cdot C_{ox} \cdot \sqrt{\psi_p} \tag{B.7}$$

The potential ψ_p can now be solved by equating (2.6) and (B.7), which results in the following simple expression for ψ_p valid over all operation regions:

$$\psi_p = \begin{cases} 0 & \text{for: } V^*_{GB} \le 0 \\ \left(\sqrt{V^*_{GB} - \psi_s + {k_P}^2/4} - k_P/2\right)^2 & \text{for: } V^*_{GB} > 0 \end{cases} \tag{B.8}$$

Note that in the ideal case (e.g., metal gate) $k_P \to \infty$, and $\psi_p = 0$; no poly-depletion occurs. Rewriting (2.7), the implicit relation for ψ_s including poly-depletion becomes:

$$\left(\frac{V^*_{GB} - \psi_s - \psi_p}{k_0}\right)^2 = \psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right] + \phi_T \cdot \exp\left(-\frac{V + \phi_B}{m_S \cdot \phi_T}\right) \cdot \left[\exp\left(\frac{\psi_s}{m_S \cdot \phi_T}\right) - 1\right] \tag{B.9}$$

where theoretically $m_S = 1$.

B.2 Gate Overlap Region

In order to be able to calculate the surface potential in the gate/source and gate/drain overlap regions, the overlap regions are treated as n^+-gate/oxide/n^+-bulk MOS capacitances where the source and drain, respectively, act as bulk, see Fig. 2.6. In the following derivation, we denote the source or drain terminal by X.
Although the impurity doping concentration in the n^+-source/drain extension region is non-uniform in both lateral and transversal direction, it is assumed that an effective donor doping concentration N_{OV} can be defined for this structure. In this case, the gate space charge $\rho(x, y)$ in the extension is given by:

$$\rho(y) = q \cdot [N_{OV} - n(y)] \tag{B.10}$$

where it has been assumed that only accumulation and depletion occur in the n^+ overlapped region[17], and as a result the influence of holes has been neglected. The electron density n is given by Maxwell-Boltzmann statistics:

$$n(y) \approx N_{\mathrm{OV}} \cdot \exp\left(\frac{\psi(y)}{\phi_{\mathrm{T}}}\right) \tag{B.11}$$

Here, ψ is the electrostatic potential with respect to the neutral overlap region. Under the gradual channel approximation, the 1-D Poisson equation for this region is written as:

$$\frac{\partial^2\psi}{\partial y^2} \approx -\frac{q \cdot N_{\mathrm{OV}}}{\epsilon_{\mathrm{Si}}} \cdot \left[1 - \exp\left(\frac{\psi'}{\phi_{\mathrm{T}}}\right)\right] \tag{B.12}$$

Again, as boundary conditions both ψ and $\partial\psi/\partial y$ are taken to be equal to zero deep in the neutral source region. From Gauss' law, the charge density Q_{ov} in the overlapped source extension is given by:

$$Q_{\mathrm{ov}} = \epsilon_{\mathrm{Si}} \cdot \left.\frac{\partial\psi}{\partial y}\right|_{y=0} = \pm k_{\mathrm{ov}} \cdot C_{\mathrm{ox}} \cdot \sqrt{\phi_{\mathrm{T}} \cdot \left[\exp\left(\frac{\psi_{\mathrm{s_{ov}}}}{\phi_{\mathrm{T}}}\right) - 1\right] - \psi_{\mathrm{s_{ov}}}} \tag{B.13}$$

where k_{ov} is the body effect coefficient of the overlap region given by $\sqrt{2 \cdot q \cdot \epsilon_{\mathrm{Si}} \cdot N_{\mathrm{OV}}}/C_{\mathrm{ox}}$. Applying Gauss' theorem, Q_{ov} can also be related to the applied gate bias:

$$Q_{\mathrm{ov}} = -C_{\mathrm{ox}} \cdot \left(V_{\mathrm{GX}} - V_{\mathrm{FBov}} - \psi_{\mathrm{p_{ov}}} - \psi_{\mathrm{s_{ov}}}\right) \tag{B.14}$$

where V_{FBov} is the effective flat-band voltage of the source extension and $\psi_{\mathrm{p_{ov}}}$ is the potential drop in the polysilicon gate due to the poly-depletion effect. The latter can be derived along the same lines as equations (B.1) through (B.8), resulting in:

$$\psi_{\mathrm{p_{ov}}} = \begin{cases} 0 & \text{for: } V_{\mathrm{GX}} \le V_{\mathrm{FBov}} \\ \left(\sqrt{V_{\mathrm{GX}} - V_{\mathrm{FBov}} - \psi_{\mathrm{s_{ov}}} + {k_{\mathrm{P}}}^2/4} - k_{\mathrm{P}}/2\right)^2 & \text{for: } V_{\mathrm{GX}} > V_{\mathrm{FBov}} \end{cases}$$

Finally, equating (B.13) and (B.14), an implicit expression for the surface potential $\psi_{\mathrm{s_{ov}}}$ can be determined:

$$\left(\frac{V_{\mathrm{GX}} - V_{\mathrm{FBov}} - \psi_{\mathrm{p_{ov}}} - \psi_{\mathrm{s_{ov}}}}{k_{\mathrm{ov}}}\right)^2 = -\psi_{\mathrm{s_{ov}}} + \phi_{\mathrm{T}} \cdot \left[\exp\left(\frac{\psi_{\mathrm{s_{ov}}}}{\phi_{\mathrm{T}}}\right) - 1\right]$$

[17] Since the source/drain extension has a very high doping concentration, an inversion layer in the overlapped region will only be formed at very negative gate bias V_{GX}. For practical bias conditions, this will not occur and therefore this effect has been neglected.

C Explicit Approximation of Surface Potential

The implicit relation (2.10) for surface potential ψ_s cannot be solved analytically. It can only be solved iteratively, which is generally considered to be computationally expensive. In order to reduce computation time, an accurate approximation of ψ_s is required. In this Appendix, the explicit approximation for ψ_s as used in MOS Model 11 is discussed. It has already partly been treated in [9].
Three distinct regions of operation can be observed: the accumulation region, the weak-inversion or depletion region and the strong-inversion region. Different approximations of surface potential can be made in the various operation regions.

Accumulation Region: Accumulation occurs when the influence of holes is dominant. In other words, when $\psi_s < 0$, which corresponds to $V^*_{GB} < 0$. In this case, neglecting the influence of electrons, the implicit relation (2.10) can be approximated by:

$$V^*_{GB} - \psi_s \approx -k_0 \cdot \sqrt{\psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right]} \tag{C.1}$$

Bearing in mind the exponential term is dominant for $\psi_s < 0$, it is convenient to rewrite this equation into:

$$\psi_s = -\phi_T \cdot \ln\left[\frac{\left(V^*_{GB} - \psi_s\right)^2 \Big/ {k_0}^2 - \psi_s + \phi_T}{\phi_T}\right] \tag{C.2}$$

Since ψ_s is almost equal to zero in accumulation, the surface potential can be approximated by:

$$\psi_s \approx -\phi_T \cdot \ln\left[\frac{\left(V^*_{GB} \big/ k_0\right)^2 + \phi_T}{\phi_T}\right] \tag{C.3}$$

From the above, it is clear that ψ_s is independent of quasi-Fermi potential V, and only weakly dependent on gate bias V^*_{GB}. For an explicit calculation of ψ_s, the use of a constant value of zero in (C.3) does not give accurate results. In Fig. 2.3 (b) at negative values of V^*_{GB}, the surface potential seems to saturate at a value which is several ϕ_T lower than 0, and which is about $-4 \cdot \phi_T$ for conventional MOSFETs. For $\psi_s < 0$, the surface potential changes from 0 at $V^*_{GB} = 0$ to $-4 \cdot \phi_T$ at $V^*_{GB} \ll 0$. A simple empirical function ψ^*_{acc} can be defined which realizes this change:

$$\psi^*_{acc} = \frac{\mathrm{Acc} \cdot V^*_{GB}}{\sqrt{1 + \left(\frac{\mathrm{Acc} \cdot V^*_{GB}}{4 \cdot \phi_T}\right)^2}} \tag{C.4}$$

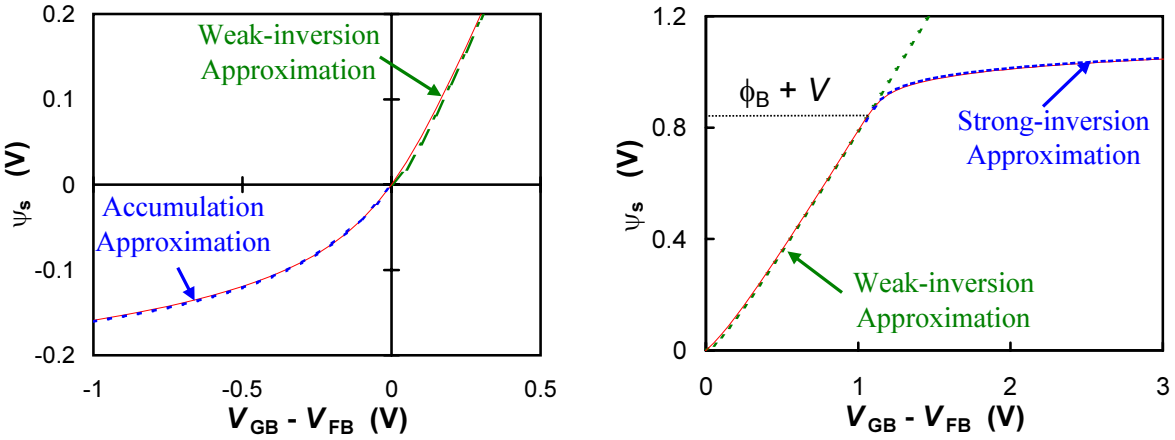
where Acc is given by:

$$\mathrm{Acc} = \left.\frac{\partial \psi_s}{\partial V_{GB}}\right|_{V^*_{GB}=0} = \frac{1}{1 + k_0 \Big/ \sqrt{2 \cdot \phi_T}} \tag{C.5}$$

ensuring that $\partial \psi^*_{acc} / \partial V^*_{GB} = \partial \psi_s / \partial V^*_{GB}$ at $V^*_{GB} = 0$. Equation (C.2) can now be rewritten as:

$$\psi_s \approx -\phi_T \cdot \ln\left[\frac{\left(V^*_{GB} - \psi^*_{acc}\right)^2 \Big/ {k_0}^2 - \psi^*_{acc} + \phi_T}{\phi_T}\right] \tag{C.6}$$

As can be seen in Fig. C.1 (a), the above expression gives an accurate description of ψ_s in the whole accumulation region.



(a) (b)

Figure C.1: The surface potential ψ_s calculated from the implicit relation (2.10) (solid line) as a function of gate bias V^*_{GB} in (a) the accumulation region and (b) the inversion region. The approximations used (dashed lines) are (C.6) in accumulation, (C.10) in weak inversion and (C.16) in strong inversion. (n-MOS, $V = 0$, $N_A = 2 \times 10^{23}\text{m}^{-3}$, $N_P \to \infty$, $t_{ox} = 3.2\text{nm}$ and $m_S = 1$)

Weak-Inversion or Depletion Region: Weak inversion (or depletion) occurs when the influence of ionized acceptor atoms is dominant, in other words when $0 < \psi_s < \phi_B + V$. In this case, neglecting the influence of electrons, the implicit relation (2.10) can be approximated by:

$$V^*_{GB} - \psi_s - \psi_p \approx k_0 \cdot \sqrt{\psi_s + \Delta_{acc}} \tag{C.7}$$

where Δ_{acc} is a function which takes into account the influence of holes:

$$\Delta_{acc} = \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T} \right) - 1 \right] \tag{C.8}$$

The function Δ_{acc} is equal to 0 at $V^*_{GB} = 0$ and it approaches a value of $-\phi_T$ for $V^*_{GB} \gg 0$. An approximate expression which ensures these conditions and which is no longer dependent on ψ_s, is given by:

$$\Delta_{acc} \approx \phi_T \cdot \left[\exp\left(-\text{Acc} \cdot \frac{V^*_{GB}}{\phi_T} \right) - 1 \right] \tag{C.9}$$

Note that the above approximate expression not only gives the exact value but also the exact derivative $\partial \Delta_{acc} / \partial V_{GB}$ at $V^*_{GB} = 0$ as given by (C.8). Solving ψ_s from (C.7) results in a simple expression:

$$\psi_s \approx \left(\frac{\sqrt{P_D \cdot \left(V^*_{GB} + \Delta_{acc} \right) + {k_0}^2/4} - k_0/2}{P_D} \right)^2 - \Delta_{acc} = \psi_{sat} \tag{C.10}$$

where $P_D = 1 + (k_0/k_P)^2$. From the above expression, it is clear that ψ_s is independent of quasi-Fermi potential V, and approximately proportional to gate bias V^*_{GB}. From (C.10), we can furthermore determine the boundaries of the weak inversion region in terms of V^*_{GB}: $0 < V^*_{GB} < P_D \cdot (\phi_B + V - \phi_T) + k_0 \cdot \sqrt{\phi_B + V - \phi_T} + \phi_T = V^*_{GB_T}$. As can be seen in Fig. C.1, the above expression gives an accurate description of ψ_s in the whole weak-inversion region.

Strong-Inversion Region: Strong inversion occurs when the influence of electrons is dominant, i.e., when $\psi_s > \phi_B + V$, which corresponds to $V^*_{GB} > P_D \cdot (\phi_B + V - \phi_T) + k_0 \cdot \sqrt{\phi_B + V - \phi_T} + \phi_T = V^*_{GB_T}$. In this case, the implicit relation (2.10) can be approximated by (assuming that $\phi_B + V \gg \phi_T$):

$$V^*_{GB} - \psi_s - \psi_p \approx k_0 \cdot \sqrt{\psi_s + \phi_T \cdot \exp\left(\frac{\psi_s - \phi_B - V}{m_S \cdot \phi_T}\right) + \Delta_{acc}} \tag{C.11}$$

Bearing in mind the exponential term is dominant for $\psi_s > \phi_B + V$, it is convenient to rewrite this equation into:

$$\psi_s = \phi_B + V + m_S \cdot \phi_T \cdot \ln\left(\frac{\left[\frac{2/k_0 \cdot (V^*_{GB} - \psi_s)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot (V^*_{GB} - \psi_s)}}\right]^2 - \psi_s - \Delta_{acc}}{\phi_T}\right) \tag{C.12}$$

It is clear that ψ_s is strongly dependent on quasi-Fermi potential V and only weakly dependent on gate bias V^*_{GB}. In a first-order approximation we can assume that ψ_s remains constant and equal to $\phi_B + V$. Replacing ψ_s in the right-hand side of (C.12) by $\phi_B + V$, we obtain:

$$\psi_s \approx \phi_B + V + m_S \cdot \phi_T \cdot \ln\left(\frac{\left[\frac{2/k_0 \cdot (V^*_{GB} - \phi_B - V)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot (V^*_{GB} - \phi_B - V)}}\right]^2 - \phi_B - V - \Delta_{acc} + \phi_T}{\phi_T}\right) \tag{C.13}$$

where the extra ϕ_T in the logarithm has been added so that (C.13) equals (C.10) for the threshold condition $V^*_{GB} = V^*_{GB_T}$, nonetheless, its influence overall is negligible. For an explicit calculation of ψ_s, however, the use of a constant value of $\phi_B + V$ in (C.13) does not give accurate results [9]. In Fig. C.1 (b) at high values of V^*_{GB}, the surface potential seems to saturate at a value which is several ϕ_T higher than $\phi_B + V$, and which is about $4 \cdot \phi_T$ for conventional MOSFETs. For $\psi_s > \phi_B + V$, the surface potential changes from $\phi_B + V$ at threshold to $\phi_B + V + 4 \cdot \phi_T$ at high gate bias values. A simple empirical function ψ^* can be defined which realizes this change:

$$\psi^* = \phi_B + V + \frac{\psi_{sat} - \phi_B - V}{\sqrt{1 + \left(\frac{\psi_{sat} - \phi_B - V}{4 \cdot \phi_T}\right)^2}} \tag{C.14}$$

Here ψ_{sat} is given by (C.10). In the logarithmic term in (C.13), $\phi_B + V$ can now be replaced by the above function ψ^*:

$$\psi_s \approx \phi_B + V + m_S \cdot \phi_T \cdot \ln\left(\frac{\left[\frac{2/k_0 \cdot (V^*_{GB} - \psi^*)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot (V^*_{GB} - \psi^*)}}\right]^2 - \psi^* - \Delta_{acc} + \phi_T}{\phi_T}\right) \tag{C.15}$$

where the quadratic term in the logarithm is dominant, and as a result, the above expression can be further approximated by:

$$\psi_s \approx \phi_B + V + m_S \cdot \phi_T \cdot \ln\left(\frac{\left[\frac{2/k_0 \cdot (V^*_{GB} - \psi^*)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot (V^*_{GB} - \psi^*)}}\right]^2 - \phi_B - V - \Delta_{acc} + \phi_T}{\phi_T}\right) \tag{C.16}$$

As can be seen in Fig. C.1 (b), the above expression gives an accurate description of ψ_s in the whole strong inversion region.

The above-derived approximate relations are only valid in their respective operation regions. An accurate and C_∞-continuous transition from accumulation to weak inversion and from weak inversion to strong inversion is needed.
The accumulation approximation (C.6) is only valid for $V^*_{GB} \leq 0$, and in order to ensure it smoothly goes to zero for $V^*_{GB} > 0$, we replace V^*_{GB} by $V^*_{GB} - V_{GB_{eff}}$ in the derivation of the accumulation approximation. Here, the function $V_{GB_{eff}}$ smoothly changes from V^*_{GB} for $V^*_{GB} < 0$ to 0 for $V^*_{GB} > 0$ by using a hyp-smoothing function, see Appendix A:

$$V_{GB_{eff}} = \mathrm{hyp}_1\left\{V^*_{GB}; \epsilon_1\right\} \tag{C.17}$$

where ϵ_1 is a smoothing factor, fixed at a value of 2×10^{-2}. Using the above method, we obtain:

$$\psi^*_{acc} = \frac{\mathrm{Acc} \cdot \left(V^*_{GB} - V_{GB_{eff}}\right)}{\sqrt{1 + \left(\frac{\mathrm{Acc} \cdot \left(V^*_{GB} - V_{GB_{eff}}\right)}{4 \cdot \phi_T}\right)^2}} \tag{C.18}$$

$$\psi_{s_{acc}} = -\phi_T \cdot \ln\left[\frac{\left(V^*_{GB} - V_{GB_{eff}} - \psi^*_{acc}\right)^2 \Big/ {k_0}^2 - \psi^*_{acc} + \phi_T}{\phi_T}\right] \tag{C.19}$$

In addition, to ensure that approximation (C.10) smoothly goes to zero for $V^*_{GB} < 0$, we replace V^*_{GB} by $V_{GB_{eff}}$ in the derivation of the weak-inversion approximation, resulting in:

$$\psi_{sat} = \left(\frac{\sqrt{P_D \cdot \left(V_{GB_{eff}} + \Delta_{acc}\right) + {k_0}^2/4} - k_0/2}{P_D}\right)^2 - \Delta_{acc} \tag{C.20}$$

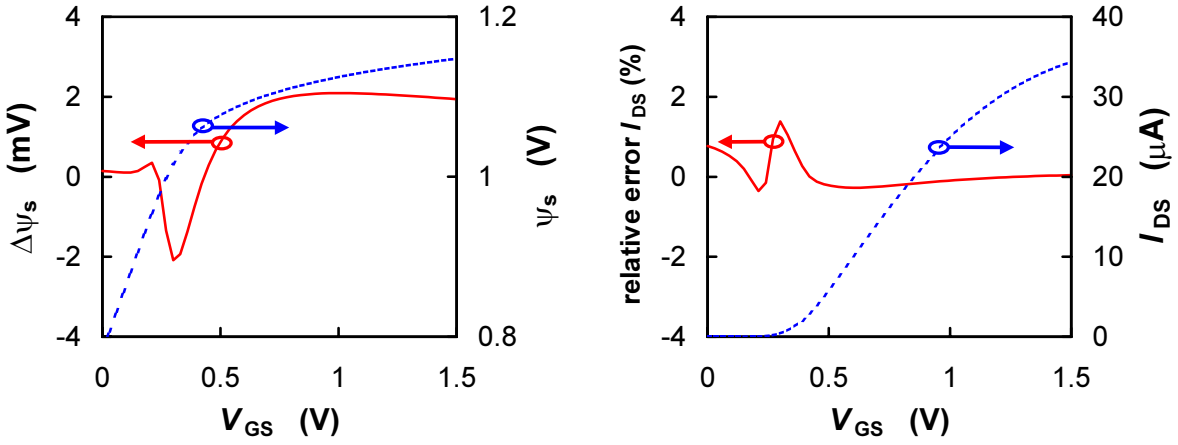
A continuous transition from weak to strong inversion can be acquired by replacing $\phi_B + V$ in (C.14) and (C.16) by a smoothing function f_1 that changes from ψ_{sat} in weak inversion to $\phi_B + V$ in strong inversion:

$$f_1 = \psi_{sat} - \mathrm{hyp}_1\left\{\psi_{sat} - \phi_B - V; \epsilon_1\right\} \tag{C.21}$$

Using the above method, we obtain:

$$\psi^* = f_1 + \frac{\psi_{sat} - f_1}{\sqrt{1 + \left(\frac{\psi_{sat} - f_1}{4 \cdot \phi_T}\right)^2}} \tag{C.22}$$

$$\psi_{s_{inv}} = f_1 + m_S \cdot \phi_T \cdot \ln\left(\frac{\left[\frac{2/k_0 \cdot \left(V^*_{GB} - \psi^*\right)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GB} - \psi^*\right)}}\right]^2 - f_1 - \Delta_{acc} + \phi_T}{\phi_T}\right) \tag{C.23}$$



(a) (b)

Figure C.2: (a) Surface potential ψ_s (dashed line) and corresponding absolute deviation between the explicit solution (C.24) and the implicit solution (2.10) of surface potential $\Delta\psi_s$ (solid line) as a function of gate bias V_{GS} (n-MOS, $V = 0$V, $N_A = 5 \times 10^{23}\text{m}^{-3}$, $N_P \to \infty$, $t_{ox} = 2$nm and $m_S = 1$). (b) Channel current I_{DS} (dashed line) and corresponding relative deviation between the explicit (approximate) and the implicit calculation (solid line) as a function of gate bias V_{GS} ($W/L = 10\mu\text{m}/10\mu\text{m}$, $V_{SB} = 0$V and $V_{DS} = 0.1$V).

The resulting expression is valid in both the weak and the strong inversion region, and goes to zero in the accumulation region. The explicit approximation of surface potential can now be written as:

$$\psi_s = \psi_{s_{acc}} + \psi_{s_{inv}} \tag{C.24}$$

As can be seen in Fig. C.2 (a), the above explicit formulation of ψ_s typically results in a maximum absolute error of 2 mV with respect to the implicit solution (2.10). This, in turn, leads to a maximum relative error in channel current I_{DS} of about 1.5% around threshold. Considering that the statistical variation in electrical behaviour of transistors on different batches, wafers or even dies is typically higher than this maximum value of 1.5%, we can conclude that the accuracy of the explicit approximation is sufficient.

D Quantum-Mechanical Effects

In modern CMOS technologies, the combination of a thinner gate oxide t_{ox} and a higher channel doping N_A results in a very high normal field at the Si-SiO_2-interface, which in turns leads to a significant bending of the energy bands at the interface. The resulting potential well can become sufficiently narrow so that the motion of carriers in the direction perpendicular to the interface is quantized. This gives rise to *i*) splitting of the conduction energy band into discrete subbands and *ii*) a displacement of the inversion-layer carrier distribution from the interface [39]-[42], see Fig. D.1.

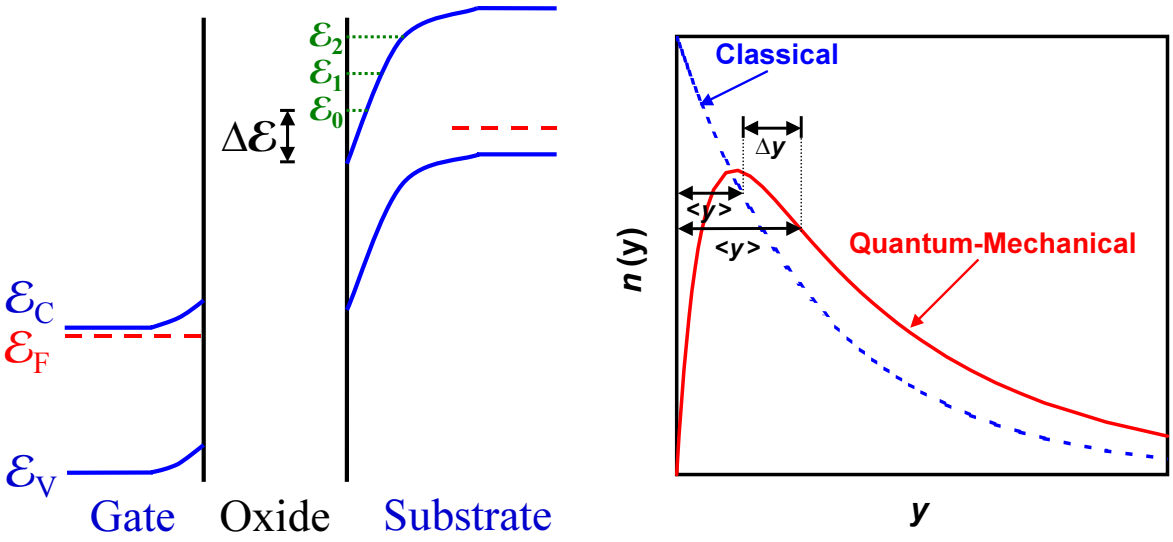


Figure D.1: (a) Energy-band diagram (in transversal direction) of an n-type MOS transistor for $V_{GB} > V_{FB}$. In the quantum-mechanical picture, the energy specrum consists of a discrete set of energy levels in the potential well (formed by the band bending at the Si-SiO_2-interface). The first energy level $\mathcal{E}_0$ does not coincide with the bottom of the conduction band resulting in an energy difference $\Delta\mathcal{E}$. (b) Electron density $n(y)$ as a function of transversal position y for the classical (dashed line) and the quantum-mechanical (solid line) case. In the quantum-mechnical case, the average distance to the interface $\langle y \rangle$ is larger by an amount of Δy compared to the classical case.

Concerning the first effect, quantum-mechanical self-consistent calculations show that energy subbands of electrons and holes are formed in the different energy valleys. As indicated in Fig. D.1, the energy spectrum consists of a set of discrete energy levels, where the first allowed energy level $\mathcal{E}_0$ does not coincide with the bottom of the conduction band $\mathcal{E}_C$, resulting in an energy difference $\Delta\mathcal{E}$. Under the approximation of a triangular potential well, the energy difference $\Delta\mathcal{E}$ can be solved from the Schrödinger equation resulting in [39]:

$$\Delta\mathcal{E} = \left(\frac{\hbar^2}{2 \cdot m_{Si}}\right)^{1/3} \cdot \left(\frac{9}{8} \cdot \pi \cdot q \cdot E_{Si}\right)^{2/3} = Z_0 \cdot \left(\frac{\hbar \cdot q \cdot E_{Si}}{\sqrt{2 \cdot m_{Si}}}\right)^{2/3} \tag{D.1}$$

where m_{Si} is the effective electron mass in silicon normal to the interface, Z_0 is the zero of the zero-order Airy function (i.e., $Z_0 \approx 2.32$) and E_{Si} is the normal electric field at the Si/SiO_2-interface. From the above expression, it is clear that the spacing between the energy levels increases with increasing normal electric field. As a result, in the weak inversion region, the splitting of the energy

levels is small compared to the thermal voltage ϕ_T, many subbands are occupied and quantum effects are washed out. In the strong inversion region, on the other hand, the spacing of energy levels is large, only a few subbands are occupied and consequently quantum-mechanical effects become important. The energy difference $\Delta\mathcal{E}$ effectively widens the bandgap for all temperatures, a larger surface potential is needed for a given channel charge and hence the threshold voltage increases with respect to the classical case.

The second quantum-mechanical effect to be taken into account is the different shape of the wave function. The electron density $n(y)$ has to vanish at the Si-SiO_2-interface and the average distance $\langle y \rangle$ to the interface increases by an amount of Δy compared to the classical solution, see Fig. D.1. Assuming that only the first subband is occupied, we can write for Δy [39]:

$$\Delta y = \frac{2}{3} \cdot \frac{\Delta\mathcal{E}}{q \cdot E_{Si}} - \frac{\phi_T}{E_{Si}} \approx \frac{2}{3} \cdot \frac{\Delta\mathcal{E}}{q \cdot E_{Si}} \tag{D.2}$$

It is clear that Δy decreases with increasing normal field E_{Si}. The displacement of the electron distribution effectively increases the oxide thickness:

$$t_{ox_{eff}} = t_{ox} + \frac{\epsilon_{ox}}{\epsilon_{Si}} \cdot \Delta y \tag{D.3}$$

which results in a bias-dependent reduction of gate capacitance, especially for thin gate oxides.

Both of the above effects should be taken into account self-consistently in Poisson's equation and Gauss' law. This is, however, quite complicated, and in order to simplify matters, both effects are generally incorporated in the modelling of an effective bandgap widening $\Delta\mathcal{E}_g$ [37]:

$$\Delta\mathcal{E}_g = \Delta\mathcal{E} + q \cdot E_{Si} \cdot \Delta y \tag{D.4}$$

The QM effects are accounted for by using a correction for the intrinsic carrier density:

$$n_i^{QM} = n_i \cdot \exp\left(-\frac{\Delta E_g/q}{2 \cdot \phi_T}\right) \tag{D.5}$$

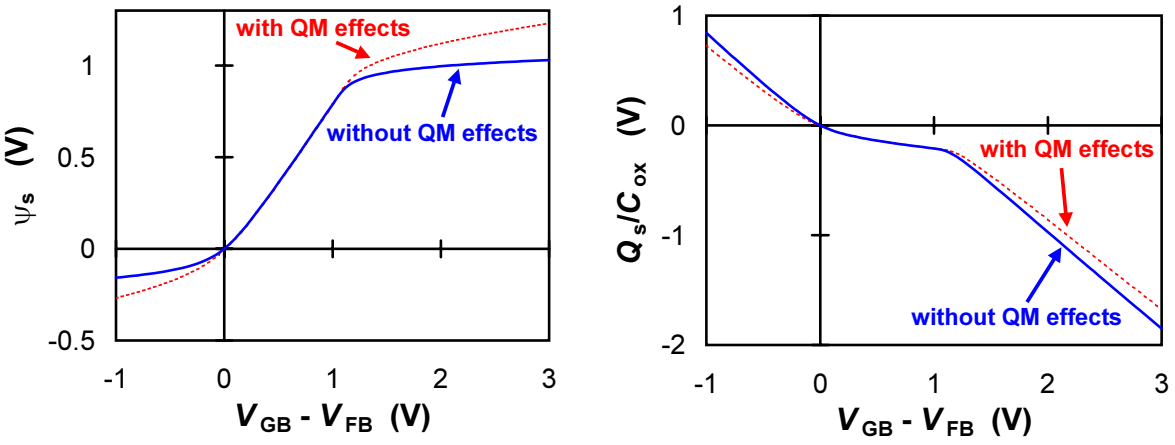
Using the above, the implicit surface potential expression (B.9) can be rewritten to (assuming that $m_S = 1$ for simplicity's sake):

$$\left(\frac{V_{GB}^* - \psi_s - \psi_p}{k_0}\right)^2 = \psi_s + \phi_T \cdot \left[\exp\left(-\frac{\Delta E_g/q}{2 \cdot \phi_T}\right) \cdot \exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right] \tag{D.6}$$

$$+ \phi_T \cdot \exp\left(-\frac{V + \phi_B}{\phi_T}\right) \cdot \left[\exp\left(-\frac{\Delta E_g/q}{2 \cdot \phi_T}\right) \cdot \exp\left(\frac{\psi_s}{\phi_T}\right) - 1\right]$$

The impact of the quantum-mechanical effects on surface potential ψ_s and on total charge Q_s can be seen in Fig. D.2. QM-effects result in an increase in $|\psi_s|$ and a decrease in $|Q_s|$, particularly in the accumulation and the strong-inversion region. It is clear that QM-effects cannot be neglected. Although the above method results in an accurate implementation of QM-effects[18], its implementation is not straightforward. It has been found that the increase in $|\psi_s|$ can simply be compensated by adjusting certain model parameters (such as k_0 and ϕ_B). In this case, the implementation of QM-effects can be

[18] Although the inclusion of both $\Delta\mathcal{E}$ and Δy in an effective $\mathcal{E}_g$ results in an accurate description of surface potential ψ_s, it nevertheless results in a slight overestimation of the depletion charge in the strong inversion ($Q_b \approx -k_0 \cdot \sqrt{\psi_s}$). This does not lead to a significant inaccuracy.



(a) (b)

Figure D.2: (a) The surface potential ψ_s and (b) the total charge density Q_s as a function of gate bias V^*_{GB} with (solid line) and without (dashed line) quantum-mechanical effects (calculated by (D.6)). (n-MOS, $V = 0$, $N_A = 2 \times 10^{23}\text{m}^{-3}$, $N_P = 1 \times 10^{26}\text{m}^{-3}$, $t_{ox} = 3.2\text{nm}$ and $m_S = 1$)

limited to the inclusion of the effective oxide thickness $t_{ox_{eff}}$ in the model [41], which will affect both I-V and CV characteristics. The I-V characteristics may still be compensated by adjusting certain parameters (such as the mobility parameters), errors in gate capacitances, however, will still remain. For an accurate charge modelling, we need to include the effective oxide thickness $t_{ox_{eff}}$.

In the above derivation, it was assumed that the potential well has a triangular shape, which is a somewhat crude approximation. A more accurate approximation can be found under the assumption that only the first allowed subband has carriers in it. In this case, the energy difference $\Delta\mathcal{E}$ can be calculated to be [39]:

$$\Delta\mathcal{E} = \frac{3}{2}\cdot\left(\frac{3\cdot q\cdot\hbar}{2\cdot\epsilon_{Si}\cdot\sqrt{m_{Si}}}\right)^{2/3}\cdot\frac{Q_b+\frac{55}{96}\cdot Q_{inv}}{\left(Q_b+\frac{11}{32}\cdot Q_{inv}\right)^{1/3}}$$

$$\approx \frac{3}{2}\cdot\left(\frac{3\cdot q\cdot\hbar\cdot E_{eff}}{2\cdot\sqrt{m_{Si}}}\right)^{2/3} = q\cdot\frac{3}{5}\cdot QM\cdot(\epsilon_{Si}\cdot E_{eff})^{2/3} \tag{D.7}$$

where QM is a physical constant ($QM_N = 5.951993\ \text{V}\cdot\text{m}^{4/3}/\text{C}^{2/3}$ for electrons and $QM_P = 7.448711\ \text{V}\cdot\text{m}^{4/3}/\text{C}^{2/3}$ for holes) and E_{eff} is the effective normal field in the potential well, given by:

$$E_{eff} = -\frac{Q_b+\frac{1}{3}\cdot Q_{inv}}{\epsilon_{Si}} \tag{D.8}$$

Note that E_{eff} is less dependent on inversion charge Q_{inv} than E_{Si}, resulting in a slightly different dependence of the QM-effects on back and gate bias. The effective oxide thickness increase Δt_{ox} ($= t_{ox_{eff}} - t_{ox}$) can be obtained from (D.2) and (D.3) replacing E_{Si} by E_{eff}:

$$\frac{\Delta t_{ox}}{t_{ox}} = \frac{2}{3}\cdot\frac{\Delta\mathcal{E}}{q}\cdot\frac{C_{ox}}{\epsilon_{Si}\cdot E_{eff}} = QM_{t_{ox}}\cdot\left(\frac{C_{ox}}{\epsilon_{Si}\cdot E_{eff}}\right)^{1/3} \tag{D.9}$$

where $QM_{t_{ox}}$ is equal to $2/5 \cdot QM \cdot {C_{ox}}^{2/3}$. The above equation is used in the charge description as used in Section 6.1.
For this case, the effective bandgap widening $\Delta\mathcal{E}_g$ can be recalculated according to (D.4):

$$\Delta\mathcal{E}_g = \Delta\mathcal{E} + q \cdot E_{eff} \cdot \Delta t_{ox} \cdot \frac{\epsilon_{Si}}{\epsilon_{ox}} = q \cdot QM_{\psi} \cdot \left(\frac{\epsilon_{Si} \cdot E_{eff}}{C_{ox}}\right)^{2/3} \tag{D.10}$$

where QM_{ψ} is equal to $QM \cdot {C_{ox}}^{2/3}$. In case of gate tunnelling, see Chapter 5, the bandgap widening can also be interpreted as an effective decrease in the oxide potential barrier $\Delta\chi_B$ simply given by $\Delta\mathcal{E}_g/q$.

E Charge Sheet Approximation and Its Impact

The total charge density Q_s in the semiconductor, as given by (2.5) and (2.8), is the sum of the inversion-layer charge density Q_{inv} consisting of electrons and the bulk charge density Q_b consisting of holes and (positively charged) ionized impurity atoms. In order to find an accurate expression for Q_{inv}, we need to integrate the electron density n as given by (2.2) over transversal coordinate y from deep into the neutral bulk to the Si/SiO_2-interface:

$$Q_{inv} = -q \cdot \int n \cdot dy \tag{E.1}$$

Bearing in mind that the transversal field E_y is given by $-\partial\psi/\partial y$ and using (2.2) for n, we can write:

$$Q_{inv} = -q \cdot N_A \cdot \int_0^{\psi_s} \frac{\exp\left(\left[\psi - \phi_B - V\right]/\phi_T\right)}{E_y(\psi, V)} \cdot d\psi \tag{E.2}$$

where, according to (2.5), E_y is given by (for $V_{GB}^* > 0$):

$$E_y = \frac{k_0 \cdot C_{ox}}{\epsilon_{Si}} \cdot \sqrt{\psi + \phi_T \cdot \left[\exp\left(-\frac{\psi}{\phi_T}\right) - 1\right] + \phi_T \cdot \exp\left(-\frac{V + \phi_B}{m_S \cdot \phi_T}\right) \cdot \left[\exp\left(\frac{\psi}{m_S \cdot \phi_T}\right) - 1\right]} \tag{E.3}$$

Unfortunately expression (E.2) cannot be solved explicitly. In order to obtain an explicit yet approximate expression for Q_{inv}, often use is made of the so-called charge sheet approximation. In the charge sheet approximation, it is assumed that the inversion layer is infinitesimally thin, in other words, it is confined to a (charge) sheet. Under this assumption, the bulk charge density can be calculated using (2.4) where the influence of electrons is neglected:

$$\frac{\partial^2\psi}{\partial y^2} \approx \frac{q \cdot N_A}{\epsilon_{Si}} \cdot \left[1 - \exp\left(-\frac{\psi}{\phi_T}\right)\right] \tag{E.4}$$

Integrating the above, Q_b can be obtained from Gauss' law:

$$Q_b = \epsilon_{Si} \cdot \left.\frac{\partial\psi}{\partial y}\right|_{y=0^+} = \pm k_0 \cdot C_{ox} \cdot \sqrt{\psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right]} \tag{E.5}$$

where Q_b is negative for $V_{GB}^* > 0$ (i.e., inversion) and positive for $V_{GB}^* < 0$ (i.e., accumulation). The inversion-layer charge density Q_{inv} is now simply given by (for $V_{GB}^* > 0$):

$$Q_{inv} = Q_s - Q_b = -C_{ox} \cdot \left(V_{GB}^* - \psi_s - \psi_p - k_0 \cdot \sqrt{\psi_s + \phi_T \cdot \left[\exp\left(-\frac{\psi_s}{\phi_T}\right) - 1\right]}\right) \tag{E.6}$$

The above expressions for Q_b and Q_{inv} can be subsequently used to derive expressions for the channel current I_{DS}, see Section 3, and the intrinsic charges, see Section 6.1. The impact of the use of the charge sheet approximation on the calculation of channel current I_{DS} and the intrinsic charges will be discussed in Sections E.1 and E.2, respectively.

E.1 Impact on Channel Current

Under the assumptions that the hole current as well as recombination/generation can be neglected and that the current flow is limited to the x-direction only, the current equation and the continuity equation in the MOSFET can be written as:

$$J_{\mathrm{n}} = -q \cdot n \cdot \mu \cdot \frac{\partial V}{\partial x} \tag{E.7}$$

$$\frac{\partial J_{\mathrm{n}}}{\partial x} = 0 \tag{E.8}$$

where J_{n} is the current density in the x-direction, and μ is the carrier mobility. From the above, integrating over y and z, it follows that the drain-source channel current I_{DS} can be written as:

$$I_{\mathrm{DS}} = -\int\int J_{\mathrm{n}} \cdot \mathrm{d}y \cdot \mathrm{d}z = q \cdot W \cdot \int n \cdot \mu \cdot \frac{\partial V}{\partial x} \cdot \mathrm{d}y \tag{E.9}$$

For the moment, μ is supposed to be bias and position independent, so that we can rewrite:

$$I_{\mathrm{DS}} = -W \cdot \mu \cdot Q_{\mathrm{inv}} \cdot \frac{\partial V}{\partial x} \tag{E.10}$$

Using the accurate expression (E.2) for Q_{inv} and integrating along the channel from source to drain, we obtain:

$$I_{\mathrm{DS}} = -q \cdot \frac{W}{L} \cdot \mu \cdot N_{\mathrm{A}} \cdot \int_{V_{\mathrm{SB}}}^{V_{\mathrm{DB}}} \int_{0}^{\psi_{\mathrm{s}}} \frac{\exp\left(\left[\psi - \phi_{\mathrm{B}} - V\right]/\phi_{\mathrm{T}}\right)}{E_{\mathrm{y}}(\psi, V)} \cdot \mathrm{d}\psi \cdot \mathrm{d}V \tag{E.11}$$

This double integral equation for I_{DS} is commonly referred to as the Pao-Sah model [43], it can only be solved numerically. The double integral can be rewritten into a completely equivalent single integral [46]. This single integral, nevertheless, still has to be solved numerically, making the Pao-Sah model unsuitable for use in circuit simulators.
The use of numerical integration is circumvented by reverting to the charge sheet approximation [44, 45], where Q_{b} and Q_{inv} are given by (E.5) and (E.6), respectively. In order to be able to integrate (E.10) along the channel, we need to find a relation between $\mathrm{d}V$ and $\mathrm{d}\psi_{\mathrm{s}}$. This can be done using the implicit relation (2.10):

$$\left(\frac{V_{\mathrm{ox}}}{k_0}\right)^2 = \psi_{\mathrm{s}} + \phi_{\mathrm{T}} \cdot \left[\exp\left(-\frac{\psi_{\mathrm{s}}}{\phi_{\mathrm{T}}}\right) - 1\right] + \phi_{\mathrm{T}} \cdot \exp\left(-\frac{V + \phi_{\mathrm{B}}}{m_{\mathrm{S}} \cdot \phi_{\mathrm{T}}}\right) \cdot \left[\exp\left(\frac{\psi_{\mathrm{s}}}{m_{\mathrm{S}} \cdot \phi_{\mathrm{T}}}\right) - 1\right] \tag{E.12}$$

where $V_{\mathrm{ox}} = V_{\mathrm{GB}}^{*} - \psi_{\mathrm{s}} - \psi_{\mathrm{p}}$ and ψ_{p} is given by (2.9). The above expression can be differentiated with respect to V resulting in:

$$\frac{\partial \psi_{\mathrm{s}}}{\partial V} = \frac{\frac{{k_0}^2}{m_{\mathrm{S}}} \cdot \exp\left(\frac{\psi_{\mathrm{s}} - V - \phi_{\mathrm{B}}}{m_{\mathrm{S}} \cdot \phi_{\mathrm{T}}}\right)}{{k_0}^2 \cdot \left[1 + \frac{1}{m_{\mathrm{S}}} \cdot \exp\left(\frac{\psi_{\mathrm{s}} - V - \phi_{\mathrm{B}}}{m_{\mathrm{S}} \cdot \phi_{\mathrm{T}}}\right) - \exp\left(-\frac{\psi_{\mathrm{s}}}{\phi_{\mathrm{T}}}\right)\right] - 2 \cdot V_{\mathrm{ox}} \cdot \frac{\partial V_{\mathrm{ox}}}{\partial \psi_{\mathrm{s}}}} \tag{E.13}$$

Using the implcit relation (E.12), we can rewrite the above into:

$$\frac{\partial V}{\partial \psi_{\mathrm{s}}} \approx 1 + m_{\mathrm{S}} \cdot \phi_{\mathrm{T}} \cdot \frac{{k_0}^2 \cdot \left[1 - \exp\left(-\frac{\psi_{\mathrm{s}}}{\phi_{\mathrm{T}}}\right)\right] - 2 \cdot V_{\mathrm{ox}} \cdot \frac{\partial V_{\mathrm{ox}}}{\partial \psi_{\mathrm{s}}}}{{V_{\mathrm{ox}}}^2 - {k_0}^2 \cdot \left(\psi_{\mathrm{s}} + \phi_{\mathrm{T}} \cdot \left[\exp\left(-\frac{\psi_{\mathrm{s}}}{\phi_{\mathrm{T}}}\right) - 1\right]\right)} \tag{E.14}$$

which in terms of $Q_{\rm inv}$ and $Q_{\rm b}$ becomes:

$$\frac{\partial V}{\partial \psi_{\rm s}} \approx 1 + m_{\rm S} \cdot \phi_{\rm T} \cdot \frac{C_{\rm ox}}{Q_{\rm inv}} \cdot \left(\frac{\partial V_{\rm ox}}{\partial \psi_{\rm s}} + \frac{Q_{\rm inv} \cdot \frac{\partial V_{\rm ox}}{\partial \psi_{\rm s}} + {k_0}^2 \cdot C_{\rm ox} \cdot \left[1 - \exp\left(-\frac{\psi_{\rm s}}{\phi_{\rm T}} \right) \right]}{Q_{\rm inv} + 2 \cdot Q_{\rm b}} \right) \tag{E.15}$$

This expression is still rather complex, but it can be simplified by noting that the second term between the round brackets only becomes important when $Q_{\rm inv}$ becomes very small. As a result, the above expression can be approximated by setting $Q_{\rm inv} = 0$ in the round bracket term [44, 45]:

$$\frac{\partial V}{\partial \psi_{\rm s}} \approx 1 + m_{\rm S} \cdot \phi_{\rm T} \cdot \frac{C_{\rm ox}}{Q_{\rm inv}} \cdot \left(\frac{\partial V_{\rm ox}}{\partial \psi_{\rm s}} + \frac{{k_0}^2 \cdot C_{\rm ox} \cdot \left[1 - \exp\left(-\frac{\psi_{\rm s}}{\phi_{\rm T}} \right) \right]}{2 \cdot Q_{\rm b}} \right)$$

$$\approx 1 + m_{\rm S} \cdot \phi_{\rm T} \cdot \frac{C_{\rm ox}}{Q_{\rm inv}} \cdot \left(\frac{\partial V_{\rm ox}}{\partial \psi_{\rm s}} - \frac{1}{C_{\rm ox}} \cdot \frac{\partial Q_{\rm b}}{\partial \psi_{\rm s}} \right) = 1 - \frac{m_{\rm S} \cdot \phi_{\rm T}}{Q_{\rm inv}} \cdot \frac{\partial Q_{\rm inv}}{\partial \psi_{\rm s}} \tag{E.16}$$

Finally, replacing $\mathrm{d}V$ in (E.10) by $\mathrm{d}\psi_{\rm s}$ and using the ideal value $m_{\rm S} = 1$, we obtain:

$$I_{\rm DS} = -\mu \cdot W \cdot Q_{\rm inv} \cdot \frac{\partial \psi_{\rm s}}{\partial x} + \mu \cdot W \cdot \phi_{\rm T} \cdot \frac{\partial Q_{\rm inv}}{\partial x} = I_{\rm drift} + I_{\rm diff} \tag{E.17}$$

where we can distinguish a drift component $I_{\rm drift}$ and a diffusion component $I_{\rm diff}$. The above expression (E.17) for $I_{\rm DS}$ is often reffered to as the charge-sheet model, and it is used throughout this report. It has been found that the charge-sheet model predicts $I_{\rm DS}$ within 1% of that calculated using the Pao-Sah model under most operating conditions [44].

E.2 Impact on Charge Model

In general, the dynamic capacitances $C_{\rm ij}$ of a MOSFET, defined by (6.5), are non-reciprocal, i.e., $C_{\rm ij} \neq C_{\rm ji}$. However, at $V_{\rm DS} = 0$, the MOSFET is a passive device, and as a consequence its dynamic capacitances $C_{\rm ij}$ are reciprocal (i.e., $C_{\rm ij} = C_{\rm ji}$) for this specific condition. Since at $V_{\rm DS} = 0$, $V = V_{\rm SB} = V_{\rm DB}$ and $\psi_{\rm s_L} = \psi_{\rm s_0}$, we can simply write for the total inversion charge $Q_{\rm INV}$:

$$Q_{\rm INV} = Q_{\rm S} + Q_{\rm D} = W \cdot L \cdot Q_{\rm inv} \tag{E.18}$$

and for the total gate charge $Q_{\rm G}$:

$$Q_{\rm G} = W \cdot L \cdot Q_{\rm g} = W \cdot L \cdot C_{\rm ox} \cdot V_{\rm ox} \tag{E.19}$$

Making use of the accurate expression (E.2) for $Q_{\rm inv}$ and the implicit relation (E.12), it can be shown that indeed:

$$C_{\rm CG} = -\frac{\partial Q_{\rm INV}}{\partial V^*_{\rm GB}} = -\frac{\partial Q_{\rm G}}{\partial V} = C_{\rm GC} \tag{E.20}$$

where $C_{\rm CG}$ ($= C_{\rm SG} + C_{\rm DG}$) is the channel-to-gate capacitance and $C_{\rm GC}$ ($= C_{\rm GS} + C_{\rm GD}$) is the gate-to-channel capacitance. The same can be shown for $C_{\rm GB} = C_{\rm BG}$ and $C_{\rm CB} = C_{\rm BC}$. The capacitances are indeed reciprocal at $V_{\rm DS} = 0$.

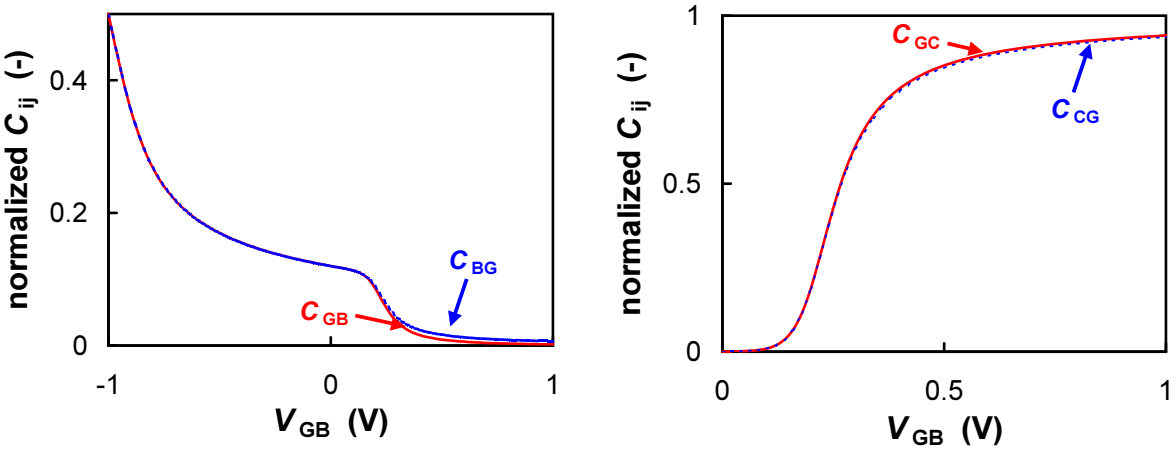


Figure E.1: (a) The gate-to-bulk capacitance C_{GB} and bulk-to-gate capacitance C_{BG}, and (b) the gate-to-channel capacitance C_{GC} and channel-to-gate capacitance C_{CG} as a function of gate bias V_{GB} for $V_{SB} = V_{DB} = 0$ as calculated with the charge sheet approximation, cf. (E.5) and (E.6). The capacitances are not exactly reciprocal, particularly in the stong inversion region, due to the use of the charge sheet approximation. (n-MOS, $N_A = 2 \times 10^{23}\text{m}^{-3}$, $N_P \to \infty$, $t_{ox} = 3.2\text{nm}$ and $m_S = 1$)

Making use of the charge sheet approximation (E.6) instead of the accurate expression (E.2), however, a slight difference occurs between C_{ij} and C_{ji}, see Fig. E.1. This difference particularly occurs in the strong inversion region where Q_{inv} becomes non-negligible. The use of the charge sheet approximation thus results in non-reciprocal capacitances at $V_{DS} = 0$, nevertheless, the difference between C_{ij} and C_{ji} is almost negligible and as a result it is generally felt that the use of the charge sheet approximation is allowed.

F Impact of Velocity Saturation

In this Appendix, the impact of velocity saturation on channel current and on the charge model are investigated in Section F.1 and F.2, respectively.

F.1 Impact on Channel Current

In an n-type MOSFET, the influence of velocity saturation is given by (3.40), repeated here for the sake of completeness:

$$I_{\mathrm{DS}} = -\frac{\mu_{\mathrm{eff}} \cdot W \cdot Q^*_{\mathrm{inv}}}{\sqrt{1 + \left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}}} \cdot \frac{\partial \psi_{\mathrm{s}}}{\partial x}\right)^2}} \cdot \frac{\partial \psi_{\mathrm{s}}}{\partial x} \tag{F.1}$$

This equation can be rewritten to:

$$I_{\mathrm{DS}} \cdot \mathrm{d}x = \sqrt{\left(\mu_{\mathrm{eff}} \cdot W \cdot Q^*_{\mathrm{inv}}\right)^2 - \left(\mu_{\mathrm{eff}} \,/ v_{\mathrm{sat}} \cdot I_{\mathrm{DS}}\right)^2} \cdot \mathrm{d}\psi_{\mathrm{s}} \tag{F.2}$$

Bearing in mind that $\mathrm{d}Q^*_{\mathrm{inv}} = -C_{\mathrm{inv}} \cdot \mathrm{d}\psi_{\mathrm{s}}$, the above becomes:

$$I_{\mathrm{DS}} \cdot \mathrm{d}x = -\sqrt{\left(\mu_{\mathrm{eff}} \cdot W \cdot Q^*_{\mathrm{inv}}\right)^2 - \left(\mu_{\mathrm{eff}} \,/ v_{\mathrm{sat}} \cdot I_{\mathrm{DS}}\right)^2} \cdot \frac{\mathrm{d}Q^*_{\mathrm{inv}}}{C_{\mathrm{inv}}} \tag{F.3}$$

which can be integrated from source to drain, resulting in:

$$I_{\mathrm{DS}} = -\frac{1}{2 \cdot C_{\mathrm{inv}} \cdot L} \cdot \left[Q^*_{\mathrm{inv}} \cdot \sqrt{\left(\mu_{\mathrm{eff}} \cdot W \cdot Q^*_{\mathrm{inv}}\right)^2 - \left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}}} \cdot I_{\mathrm{DS}}\right)^2} \right. \tag{F.4}$$

$$\left. + \frac{\left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}}} \cdot I_{\mathrm{DS}}\right)^2}{\mu_{\mathrm{eff}} \cdot W} \cdot \ln\left(-\mu_{\mathrm{eff}} \cdot W \cdot Q^*_{\mathrm{inv}} + \sqrt{\left(\mu_{\mathrm{eff}} \cdot W \cdot Q^*_{\mathrm{inv}}\right)^2 - \left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}}} \cdot I_{\mathrm{DS}}\right)^2} \right) \right] \Bigg|_{Q^*_{\mathrm{inv_0}}}^{Q^*_{\mathrm{inv_L}}}$$

The above relation gives an exact description of the impact of velocity saturation on the channel current I_{DS}. Unfortunately I_{DS} is only given implicitly, and it cannot be calculated analytically. In order to simplify (F.4), we define $\Delta Q^*_{\mathrm{inv}} = Q^*_{\mathrm{inv_L}} - Q^*_{\mathrm{inv_0}} = \Delta Q_{\mathrm{inv}}$. Next, using $Q^*_{\mathrm{inv_0}} = \bar{Q}^*_{\mathrm{inv}} - \Delta Q^*_{\mathrm{inv}}/2$ and $Q^*_{\mathrm{inv_L}} = \bar{Q}^*_{\mathrm{inv}} + \Delta Q^*_{\mathrm{inv}}/2$, we can develop a third-order Taylor series of the right-hand side of (F.4) around $I_{\mathrm{DS}} = 0$:

$$I_{\mathrm{DS}} \approx \mu_{\mathrm{eff}} \cdot \frac{W}{L} \cdot \bar{Q}^*_{\mathrm{inv}} \cdot \frac{\Delta Q^*_{\mathrm{inv}}}{C_{\mathrm{inv}}} + \frac{\left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}}} \cdot I_{\mathrm{DS}}\right)^2}{2 \cdot C_{\mathrm{inv}} \cdot \mu_{\mathrm{eff}} \cdot W \cdot L} \cdot \ln\left(\frac{2 - \Delta Q^*_{\mathrm{inv}}/\bar{Q}^*_{\mathrm{inv}}}{2 + \Delta Q^*_{\mathrm{inv}}/\bar{Q}^*_{\mathrm{inv}}}\right) \tag{F.5}$$

from which I_{DS} can be solved:

$$I_{\mathrm{DS}} \approx -\frac{2 \cdot \mu_{\mathrm{eff}} \cdot \frac{W}{L} \cdot \bar{Q}^*_{\mathrm{inv}} \cdot \Delta\psi}{1 + \sqrt{1 - 2 \cdot \left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}} \cdot L}\right)^2 \cdot \frac{\bar{Q}^*_{\mathrm{inv}}}{C_{\mathrm{inv}}} \cdot \frac{\Delta Q^*_{\mathrm{inv}}}{C_{\mathrm{inv}}} \cdot \ln\left(\frac{2 - \Delta Q^*_{\mathrm{inv}}/\bar{Q}^*_{\mathrm{inv}}}{2 + \Delta Q^*_{\mathrm{inv}}/\bar{Q}^*_{\mathrm{inv}}}\right)}} \tag{F.6}$$

The above equation can be further simplified by approximating the logarithmic term:

$$\ln\left(\frac{2-\Delta Q^*_{\rm inv}/\bar{Q}^*_{\rm inv}}{2+\Delta Q^*_{\rm inv}/\bar{Q}^*_{\rm inv}}\right) \approx -\frac{\Delta Q^*_{\rm inv}}{\bar{Q}^*_{\rm inv}} - \frac{1}{12}\cdot\left(\frac{\Delta Q^*_{\rm inv}}{\bar{Q}^*_{\rm inv}}\right)^3 \tag{F.7}$$

resulting in:

$$\begin{aligned} I_{\rm DS} &\approx -\frac{2\cdot\mu_{\rm eff}\cdot\frac{W}{L}\cdot\bar{Q}^*_{\rm inv}\cdot\Delta\psi}{1+\sqrt{1+2\cdot\left(\frac{\mu_{\rm eff}}{v_{\rm sat}\cdot L}\cdot\Delta\psi\right)^2\cdot\left[1+\frac{1}{12}\cdot\left(\frac{\Delta Q^*_{\rm inv}}{\bar{Q}^*_{\rm inv}}\right)^2\right]}} \\ &\approx -\frac{2\cdot\beta\cdot\bar{Q}^*_{\rm inv}/C_{\rm ox}\cdot\Delta\psi}{G_{\rm mob}+\sqrt{{G_{\rm mob}}^2+2\cdot(\theta_{\rm sat}\cdot\Delta\psi)^2\cdot\left[1+\frac{1}{12}\cdot\left(\frac{\Delta Q^*_{\rm inv}}{\bar{Q}^*_{\rm inv}}\right)^2\right]}} \end{aligned} \tag{F.8}$$

where $\theta_{\rm sat}$ is defined as $\mu_0/(v_{\rm sat}\cdot L)$.

F.2 Impact on Charge Model

In order to study the influence of velocity saturation on the overall charge model, we study the impact on the total inversion-layer charge $Q_{\rm INV}$ ($= Q_{\rm S} + Q_{\rm D}$):

$$Q_{\rm INV} = W\cdot\int_0^L Q_{\rm inv}\cdot{\rm d}x \tag{F.9}$$

Using (F.3), the above integral can be rewritten to:

$$Q_{\rm INV} = -W\cdot\int_0^L\left(Q^*_{\rm inv}-\phi_{\rm T}\cdot C_{\rm inv}\right)\cdot\sqrt{\left(\mu_{\rm eff}\cdot W\cdot Q^*_{\rm inv}\right)^2-\left(\frac{\mu_{\rm eff}}{v_{\rm sat}}\cdot I_{\rm DS}\right)^2}\cdot\frac{{\rm d}Q^*_{\rm inv}}{C_{\rm inv}\cdot I_{\rm DS}} \tag{F.10}$$

Although the above integral can be analytically solved, it does not result in a transparent equation. Therefore, we simplify the integral by developing the square root in a second-order Taylor series around $I_{\rm DS}=0$:

$$Q_{\rm INV} \approx \frac{W}{I_{\rm DS}}\cdot\int_0^L\left(Q^*_{\rm inv}-\phi_{\rm T}\cdot C_{\rm inv}\right)\cdot\left[\mu_{\rm eff}\cdot W\cdot Q^*_{\rm inv}-\frac{\left(\frac{\mu_{\rm eff}}{v_{\rm sat}}\cdot I_{\rm DS}\right)^2}{2\cdot\mu_{\rm eff}\cdot W\cdot Q^*_{\rm inv}}\right]\cdot\frac{{\rm d}Q^*_{\rm inv}}{C_{\rm inv}} \tag{F.11}$$

This integral can be easily solved, resulting in:

$$\begin{aligned} Q_{\rm INV} \approx \frac{W}{2\cdot C_{\rm inv}}\cdot\Bigg(&\frac{\mu_{\rm eff}\cdot W\cdot Q^*_{\rm inv}\cdot\left[2\cdot Q^*_{\rm inv}-3\cdot C_{\rm inv}\cdot\phi_{\rm T}\right]}{3\cdot I_{\rm DS}} \\ &-\frac{\left[\frac{\mu_{\rm eff}}{v_{\rm sat}}\right]^2\cdot I_{\rm DS}\cdot\left[Q^*_{\rm inv}-\phi_{\rm T}\cdot C_{\rm inv}\cdot\ln\left(Q^*_{\rm inv}\right)\right]}{\mu_{\rm eff}\cdot W}\Bigg)\Bigg|_{Q^*_{\rm inv_0}}^{Q^*_{\rm inv_L}} \end{aligned} \tag{F.12}$$

Using $Q^*_{\mathrm{inv}_0} = \bar{Q}^*_{\mathrm{inv}} - \Delta Q^*_{\mathrm{inv}}/2$ and $Q^*_{\mathrm{inv}_L} = \bar{Q}^*_{\mathrm{inv}} + \Delta Q^*_{\mathrm{inv}}/2$, and replacing I_{DS} by (F.6), we can develop a fifth-order Taylor series of (F.12) around $\Delta Q^*_{\mathrm{inv}} = 0$:

$$Q_{\mathrm{INV}} \approx W \cdot L \cdot \left[\bar{Q}^*_{\mathrm{inv}} - \phi_{\mathrm{T}} \cdot C_{\mathrm{inv}} + \frac{1}{12} \cdot \frac{{\Delta Q^*_{\mathrm{inv}}}^2}{\bar{Q}^*_{\mathrm{inv}}} + \frac{1}{12} \cdot \left(\frac{\mu_{\mathrm{eff}}}{v_{\mathrm{sat}} \cdot L} \right)^2 \cdot \frac{{\Delta Q^*_{\mathrm{inv}}}^4}{{C_{\mathrm{inv}}}^2 \cdot \bar{Q}^*_{\mathrm{inv}}} \right] \tag{F.13}$$

The impact of velocity saturation only appears in the fourth-order term, and can thus be neglected. Using (6.9), the above equation reduces to:

$$Q_{\mathrm{INV}} \approx W \cdot L \cdot \left[\bar{Q}^*_{\mathrm{inv}} - \phi_{\mathrm{T}} \cdot C_{\mathrm{inv}} - \frac{\Delta Q^*_{\mathrm{inv}}}{6} \cdot F_{\mathrm{j}} \right] = W \cdot L \cdot \left[\bar{Q}_{\mathrm{inv}} - \frac{\Delta Q^*_{\mathrm{inv}}}{6} \cdot F_{\mathrm{j}} \right] \tag{F.14}$$

which is exactly equal to the summation of eqs. (6.10) and (6.11), where velocity saturation has been neglected.

G Implementation of Series Resistance

In this Appendix, the impact of series resistance on the channel current and on the charge model are investigated in Section G.1 and G.2, respectively.

G.1 Impact on Channel Current

In the presence of source (R_S) and drain (R_D) resistance, the internal bias V_{GB} is unaffected, but the internal bias values V_{SB} and V_{DB} are changed by a voltage drop across these respective resistors and have to be replaced by:

$$\begin{aligned} V_{SB} &\rightarrow V_{SB} + I_{DS} \cdot R_S \\ V_{DB} &\rightarrow V_{DB} - I_{DS} \cdot R_D \end{aligned} \tag{G.1}$$

Since a MOSFET is a symmetrical device with respect to source and drain, it is valid to assume that source and drain resistance are equal (i.e., $R_S = R_D$). In order to determine the impact of series resistance on channel current I_{DS} given by (3.42), the surface potentials ψ_{s_0} and ψ_{s_L} have to be recalculated using (G.1). This results in an implicit relation for I_{DS}. To simplify the solution of I_{DS}, we will focus on the superthreshold region where I_{DS} and consequently the voltage drop across the series resistance can be significant. In the subthreshold region, on the other hand, I_{DS} will be small and the above-mentioned voltage drop can be neglected.
In strong inversion, the surface potentials ψ_{s_0} and ψ_{s_L} are in a zero-order approximation (see Appendix C) equal to $\phi_B + V_{SB} + I_{DS} \cdot R_S$ and $\phi_B + V_{DB} - I_{DS} \cdot R_S$, respectively. Neglecting higher-order effects, we can thus replace ψ_{s_0} and ψ_{s_L} by:

$$\begin{aligned} \psi_{s_0} &\rightarrow \psi_{s_0} + I_{DS} \cdot R_S \\ \psi_{s_L} &\rightarrow \psi_{s_L} - I_{DS} \cdot R_S \end{aligned} \tag{G.2}$$

In terms of variables $\bar{\psi}$ and $\Delta\psi$, we can write:

$$\begin{aligned} \bar{\psi} &\rightarrow \bar{\psi} \\ \Delta\psi &\rightarrow \Delta\psi - 2 \cdot I_{DS} \cdot R_S \end{aligned} \tag{G.3}$$

This implies that variables such as, e.g., $\bar{Q}_{inv}$, C_{inv} and E_{eff}, which are only dependent on $\bar{\psi}$, are not affected by series resistance. Equation (G.3) can now be introduced in the expression (3.42) for I_{DS}:

$$I_{DS} = -\frac{\beta \cdot \bar{Q}^*_{inv}/C_{ox} \cdot (\Delta\psi - 2 \cdot I_{DS} \cdot R_S)}{G^*_{vsat}} \tag{G.4}$$

where G^*_{vsat} is the velocity saturation expression where the impact of series resistance has been taken into account. In order to keep the derivation manageable, we use the simplified expression (3.44) for G_{vsat}[19] and take a first-order Taylor expansion around $I_{DS} \cdot R_S = 0$:

$$\begin{aligned} G^*_{vsat} &= \sqrt{{G_{mob}}^2 + {\theta_{sat}}^2 \cdot (\Delta\psi - 2 \cdot I_{DS} \cdot R_S)^2} \\ &\approx \sqrt{{G_{mob}}^2 + (\theta_{sat} \cdot \Delta\psi)^2} - \frac{2 \cdot {\theta_{sat}}^2 \cdot \Delta\psi \cdot I_{DS} \cdot R_S}{\sqrt{{G_{mob}}^2 + (\theta_{sat} \cdot \Delta\psi)^2}} \end{aligned} \tag{G.5}$$

[19] The small difference between (3.44) and (3.45) is neglected here, see Fig. 3.8.

$$\approx G_{\mathrm{vsat}} - 2 \cdot {\theta_{\mathrm{sat}}}^2 \cdot \Delta\psi \cdot I_{\mathrm{DS}} \cdot \frac{R_{\mathrm{S}}}{G_{\mathrm{vsat}}}$$

Equation (G.4) can now be rewritten to a second-order polynomial:

$$-\frac{2 \cdot {\theta_{\mathrm{sat}}}^2 \cdot R_{\mathrm{S}} \cdot \Delta\psi}{G_{\mathrm{vsat}}} \cdot {I_{\mathrm{DS}}}^2 + [G_{\mathrm{R}} + G_{\mathrm{vsat}}] \cdot I_{\mathrm{DS}} + \beta \cdot \frac{\bar{Q}^*_{\mathrm{inv}}}{C_{\mathrm{ox}}} \cdot \Delta\psi = 0 \tag{G.6}$$

where:

$$G_{\mathrm{R}} = -2 \cdot \beta \cdot \frac{\bar{Q}^*_{\mathrm{inv}}}{C_{\mathrm{ox}}} \cdot R_{\mathrm{S}} \tag{G.7}$$

The solution to (G.6) of I_{DS} is simply given by:

$$I_{\mathrm{DS}} = -\frac{\beta}{G_{\mathrm{tot}}} \cdot \frac{\bar{Q}^*_{\mathrm{inv}}}{C_{\mathrm{ox}}} \cdot \Delta\psi \tag{G.8}$$

where:

$$G_{\mathrm{tot}} = \frac{G_{\mathrm{vsat}} + G_{\mathrm{R}}}{2} \cdot \left[1 + \sqrt{1 - 4 \cdot \frac{G_{\mathrm{R}}}{G_{\mathrm{vsat}}} \cdot \left(\frac{\theta_{\mathrm{sat}} \cdot \Delta\psi}{G_{\mathrm{vsat}} + G_{\mathrm{R}}} \right)^2} \right] \tag{G.9}$$

Using (3.44), we can rewrite ${\theta_{\mathrm{sat}}}^2 \cdot \Delta\psi^2$ as ${G_{\mathrm{vsat}}}^2 - {G_{\mathrm{mob}}}^2$, and as a consequence G_{tot} becomes:

$$G_{\mathrm{tot}} = \frac{G_{\mathrm{vsat}} + G_{\mathrm{R}}}{2} \cdot \left[1 + \sqrt{1 - \frac{4 \cdot G_{\mathrm{R}} / G_{\mathrm{vsat}}}{(G_{\mathrm{vsat}} + G_{\mathrm{R}})^2} \cdot \left({G_{\mathrm{vsat}}}^2 - {G_{\mathrm{mob}}}^2 \right)} \right] \tag{G.10}$$

For reasons of simplicity, G_{vsat} can simply be replaced by (3.45) instead of (3.44).
Note that, when the impact of series resistance on velocity saturation is neglected, equation (G.10) reduces to:

$$G_{\mathrm{tot}} = G_{\mathrm{vsat}} + G_{\mathrm{R}} \tag{G.11}$$

This approximate expression will be used in the calculation of the saturation voltage, see Appendix H.

G.2 Impact on Charge Model

The presence of source and drain resistance not only affects the (steady-state) channel current but also has an impact on the charge model. We need to determine the impact of series resistance on the total (intrinsic) charges Q_{D}, Q_{S}, Q_{G} and Q_{B} given by (6.10), (6.11), (6.13) and (6.14), respectively. The total charges can be written as a function of average charge densities, such as $\bar{Q}^*_{\mathrm{inv}}$ and $\bar{Q}_{\mathrm{g}}$, and the inversion-layer charge difference $\Delta Q^*_{\mathrm{inv}}$. The average charge densities are only dependent on $\bar{\psi}$, and consequently, as pointed out in Section G.1, they are not affected by series resistance. Only $\Delta Q^*_{\mathrm{inv}}$ is affected by series resistance. Using (G.3), we can write:

$$\Delta Q^*_{\mathrm{inv}} = -C_{\mathrm{inv}} \cdot (\Delta\psi - 2 \cdot I_{\mathrm{DS}} \cdot R_{\mathrm{S}}) \tag{G.12}$$

which can be simplified using (G.7) and (G.8):

$$\Delta Q^*_{\rm inv} = -C_{\rm inv} \cdot \Delta\psi \cdot \left(1 + \frac{2 \cdot \beta \cdot \bar{Q}^*_{\rm inv} \cdot R_{\rm S}}{C_{\rm ox} \cdot G_{\rm tot}}\right) = \Delta Q_{\rm inv} \cdot \left(1 - \frac{G_{\rm R}}{G_{\rm tot}}\right) \tag{G.13}$$

In MM11, we have approximated the function $1 - x$ by $1/(1 + x)$, resulting in:

$$\Delta Q^*_{\rm inv} \approx \frac{\Delta Q_{\rm inv}}{1 + \frac{G_{\rm R}}{G_{\rm tot}}} \tag{G.14}$$

Finally, to incorporate the impact of series resistance on the charge model, we replace $\Delta Q^*_{\rm inv}$ in eqs. (6.10), (6.11), (6.13) and (6.14) by the above expression for $\Delta Q^*_{\rm inv}$.

H Calculation of Saturation Voltage

In literature [35], the saturation voltage for short-channel transistors is usually calculated by defining the saturated drain-source channel current to be equal to the product of the free carrier concentration at the drain end and the saturated drift velocity:

$$I_{\mathrm{DS_{sat}}} = -W \cdot Q_{\mathrm{inv_L}} \cdot v_{\mathrm{sat}}$$

Here, $Q_{\mathrm{inv_L}}$ is given by (2.13) where $\psi_{\mathrm{s}} = \psi_{\mathrm{s_L}}$ and Δ_{acc} is neglected:

$$Q_{\mathrm{inv_L}} = -C_{\mathrm{ox}} \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}} - k_0 \cdot \sqrt{\psi_{\mathrm{s_L}}} \right] \tag{H.1}$$

In order to take into account the impact of series resistance on $Q_{\mathrm{inv_L}}$, $\psi_{\mathrm{s_L}}$ has to be replaced by $\psi_{\mathrm{s_L}} - I_{\mathrm{DS_{sat}}} \cdot R_{\mathrm{S}}$ in the above relation along the same lines as done in Appendix G. Taking a first-order Taylor series around $I_{\mathrm{DS_{sat}}} \cdot R_{\mathrm{S}} = 0$, we can write:

$$Q_{\mathrm{inv_L}} = -C_{\mathrm{ox}} \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}} - k_0 \cdot \sqrt{\psi_{\mathrm{s_L}}} \right] + C_{\mathrm{inv_L}} \cdot I_{\mathrm{DS_{sat}}} \cdot R_{\mathrm{S}} \tag{H.2}$$

where:

$$C_{\mathrm{inv_L}} = -\left. \frac{\partial Q_{\mathrm{inv}}}{\partial \psi_{\mathrm{s}}} \right|_{\psi_{\mathrm{s}} = \psi_{\mathrm{s_L}}} = -C_{\mathrm{ox}} \cdot \left[\frac{1}{\sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}} + \frac{k_0}{2 \cdot \sqrt{\psi_{\mathrm{s_L}}}} \right] \tag{H.3}$$

The saturated channel current becomes:

$$I_{\mathrm{DS_{sat}}} = \frac{W \cdot C_{\mathrm{ox}} \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}} - k_0 \cdot \sqrt{\psi_{\mathrm{s_L}}} \right] \cdot v_{\mathrm{sat}}}{1 + W \cdot C_{\mathrm{inv_L}} \cdot v_{\mathrm{sat}} \cdot R_{\mathrm{S}}} \tag{H.4}$$

$$= \frac{\beta \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}} - k_0 \cdot \sqrt{\psi_{\mathrm{s_L}}} \right]}{\theta_{\mathrm{sat}} + C_{\mathrm{inv_L}}/C_{\mathrm{ox}} \cdot \theta_{\mathrm{R}}/2}$$

On the other hand, the channel current I_{DS} is also given by expression (3.50). Equating (H.4) and (3.50), an expression is obtained from which the saturation $\psi_{\mathrm{s_L}}$ value can be calculated:

$$\frac{\beta \cdot \bar{Q}^*_{\mathrm{inv}} \cdot \left(\psi_{\mathrm{s_L}} - \psi_{\mathrm{s_0}}\right)}{C_{\mathrm{ox}} \cdot G_{\mathrm{tot}}} = \frac{\beta \cdot \left[\frac{2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}{1 + \sqrt{1 + 4/{k_{\mathrm{P}}}^2 \cdot \left(V^*_{\mathrm{GB}} - \psi_{\mathrm{s_L}}\right)}} - k_0 \cdot \sqrt{\psi_{\mathrm{s_L}}} \right]}{\theta_{\mathrm{sat}} + C_{\mathrm{inv_L}}/C_{\mathrm{ox}} \cdot \theta_{\mathrm{R}}/2} \tag{H.5}$$

where G_{tot} is given by (G.9). Unfortunately $\psi_{\mathrm{s_L}}$ cannot be solved analytically from the above relation. In order to find an analytical and explicit expression for the saturation voltage, some approximations have to be made.

If we focus our attention merely on the strong inversion region, we can approximate $\psi_{\mathrm{s_0}}$ by $\phi_{\mathrm{B}} + V_{\mathrm{SB}}$

and ψ_{s_L} by $\phi_B + V_{SB} + V_{DS}$. In this case, at the saturation point, by definition the drain-source bias V_{DS} is equal to the saturation voltage $V_{DS_{sat}}$. Furthermore, in strong inversion we will neglect the diffusion component, and only take the drift component into account. In addition, the impact of series resistance on velocity saturation is neglected in order to keep the derivation manageable. In this case, according to (G.11), G_{tot} becomes equal to $G_{vsat} + G_R$. Taking into account all of the above approximations, we can rewrite (H.5) to:

$$\frac{\beta \cdot \bar{V}_{GT} \cdot V_{DS_{sat}}}{G_{vsat} + \theta_R \cdot \bar{V}_{GT}} = \frac{\beta \cdot V_{GT_L}}{\theta_{sat} + \theta_R/2 \cdot C_{inv_L}/C_{ox}} \tag{H.6}$$

where ($V^*_{GS} = V^*_{GB} - V_{SB}$):

$$\bar{V}_{GT} = \frac{2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat}}/2\right)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat}}/2\right)}} - k_0 \cdot \sqrt{\phi_B + V_{SB} + \frac{V_{DS_{sat}}}{2}} \tag{H.7}$$

$$V_{GT_L} = \frac{2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat}}\right)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat}}\right)}} - k_0 \cdot \sqrt{\phi_B + V_{SB} + V_{DS_{sat}}} \tag{H.8}$$

Expression (H.6), however, is still not analytically solvable for $V_{DS_{sat}}$. For an infinitely long channel the impact of velocity saturation and series resistance is negligible, and in this case the saturation voltage $V_{DS_{sat\infty}}$ corresponds to the pinch-off voltage, given by (3.12). For short-channel devices, the impact of velocity saturation and series resistance cannot be neglected, and the resulting $V_{DS_{sat}}$ will differ from the ideal case $V_{DS_{sat\infty}}$. If we assume that the difference $\Delta = V_{DS_{sat}} - V_{DS_{sat\infty}}$ is only small, we can simplify (H.6) by linearizing all variables around $V_{DS_{sat}} = V_{DS_{sat\infty}}$:

$$\bar{V}_{GT} \approx \frac{2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat\infty}}/2\right)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat\infty}}/2\right)}} - k_0 \cdot \sqrt{\phi_B + V_{SB} + \frac{V_{DS_{sat\infty}}}{2}} - \frac{\xi^*_\infty}{2} \cdot \Delta \tag{H.9}$$

$$\approx \frac{\xi^*_\infty}{2} \cdot V_{DS_{sat\infty}} - \frac{\xi^*_\infty}{2} \cdot \Delta$$

$$V_{GT_L} \approx \frac{2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat\infty}}\right)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat\infty}}\right)}} - k_0 \cdot \sqrt{\phi_B + V_{SB} + V_{DS_{sat\infty}}} - \xi^*_{L\infty} \cdot \Delta \tag{H.10}$$

$$\approx -\xi^*_{L\infty} \cdot \Delta$$

where:

$$\xi^*_\infty = -2 \cdot \left.\frac{\partial \bar{V}_{GT}}{\partial V_{DS_{sat}}}\right|_{V_{DS_{sat}} = V_{DS_{sat\infty}}} \tag{H.11}$$

$$= \frac{1}{\sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GS} - \phi_B - V_{DS_{sat\infty}}/2\right)}} + \frac{k_0}{2 \cdot \sqrt{\phi_B + V_{SB} + V_{DS_{sat\infty}}/2}}$$

$$\xi^*_{\mathrm{L}\infty} = -\left.\frac{\partial V_{\mathrm{GT_L}}}{\partial V_{\mathrm{DS_{sat}}}}\right|_{V_{\mathrm{DS_{sat}}}=V_{\mathrm{DS_{sat}\infty}}} = -\frac{C_{\mathrm{inv_L}}}{C_{\mathrm{ox}}} \tag{H.12}$$

$$= \frac{1}{\sqrt{1+4/{k_\mathrm{P}}^2\cdot\left(V^*_{\mathrm{GS}}-\phi_\mathrm{B}-V_{\mathrm{DS_{sat}\infty}}\right)}} + \frac{k_0}{2\cdot\sqrt{\phi_\mathrm{B}+V_{\mathrm{SB}}+V_{\mathrm{DS_{sat}\infty}}}}$$

In order to keep the derivation manageable, we use the simplified expression (3.44) for G_{vsat}[20], resulting in:

$$G_{\mathrm{vsat}} \approx \sqrt{{G_\mathrm{mob}}^2+\left(\theta_\mathrm{sat}\cdot V_{\mathrm{DS_{sat}\infty}}\right)^2} + \left.\frac{\partial G_{\mathrm{vsat}}}{\partial V_{\mathrm{DS_{sat}}}}\right|_{V_{\mathrm{DS_{sat}}}=V_{\mathrm{DS_{sat}\infty}}}\cdot\Delta \tag{H.13}$$

$$= G_{\mathrm{vsat}\infty} + \frac{{\theta_\mathrm{sat}}^2\cdot V_{\mathrm{DS_{sat}\infty}}}{\sqrt{{G_\mathrm{mob}}^2+\left(\theta_\mathrm{sat}\cdot V_{\mathrm{DS_{sat}\infty}}\right)^2}}\cdot\Delta \approx G_{\mathrm{vsat}\infty}+\theta_\mathrm{sat}\cdot\Delta$$

Using (H.9), (H.10) and (H.13), we can rewrite (H.6) to

$$\frac{\beta\cdot\xi^*_\infty/2\cdot\left(V_{\mathrm{DS_{sat}\infty}}-\Delta\right)\cdot\left(V_{\mathrm{DS_{sat}\infty}}+\Delta\right)}{G_{\mathrm{vsat}\infty}+\theta_\mathrm{sat}\cdot\Delta+\theta_\mathrm{R}\cdot\xi^*_\infty/2\cdot\left(V_{\mathrm{DS_{sat}\infty}}-\Delta\right)} = -\frac{\beta\cdot\xi^*_{\mathrm{L}\infty}\cdot\Delta}{\theta^*_\mathrm{sat}} \tag{H.14}$$

where $\theta^*_\mathrm{sat}=\theta_\mathrm{sat}-\theta_\mathrm{R}/2\cdot\xi^*_{\mathrm{L}\infty}$. We further approximate (H.14) by assuming $\xi^*_\infty\approx\xi^*_{\mathrm{L}\infty}\approx 1$, which eventually results in a second-order polynomial:

$$\frac{\theta^*_\mathrm{sat}}{2}\cdot\Delta^2+\left(G_{\mathrm{vsat}\infty}+\frac{\theta_\mathrm{R}}{2}\cdot V_{\mathrm{DS_{sat}\infty}}\right)\cdot\Delta+\frac{\theta^*_\mathrm{sat}}{2}\cdot{V_{\mathrm{DS_{sat}\infty}}}^2=0 \tag{H.15}$$

From the above equation, we can calculate an expression for Δ:

$$\Delta = -\frac{\theta^*_\mathrm{sat}\cdot{V_{\mathrm{DS_{sat}\infty}}}^2}{\left(G_{\mathrm{vsat}\infty}+\theta_\mathrm{R}/2\cdot V_{\mathrm{DS_{sat}\infty}}\right)\cdot\left(1+\sqrt{1-\left(\frac{\theta^*_\mathrm{sat}\cdot V_{\mathrm{DS_{sat}\infty}}}{G_{\mathrm{vsat}\infty}+\theta_\mathrm{R}/2\cdot V_{\mathrm{DS_{sat}\infty}}}\right)^2}\right)} \tag{H.16}$$

Bearing in mind that $V_{\mathrm{DS_{sat}}}=V_{\mathrm{DS_{sat}\infty}}+\Delta$, the saturation voltage can finally be written as:

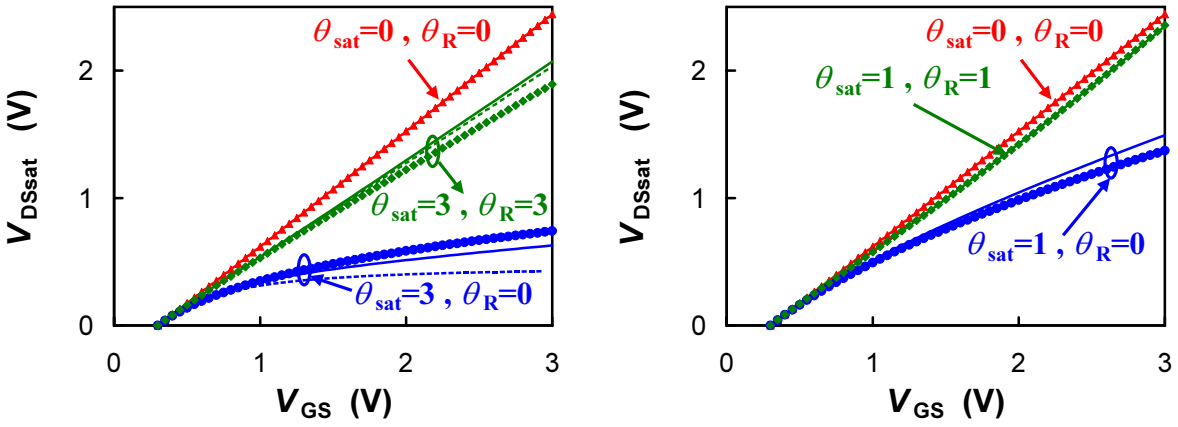
$$V_{\mathrm{DS_{sat}}} = V_{\mathrm{DS_{sat}\infty}}\cdot\left[1-\frac{\Delta_\mathrm{SAT}}{1+\sqrt{1-{\Delta_\mathrm{SAT}}^2}}\right] \tag{H.17}$$

where:

$$\Delta_\mathrm{SAT} = \frac{\theta^*_\mathrm{sat}\cdot V_{\mathrm{DS_{sat}\infty}}}{\sqrt{{G_\mathrm{mob}}^2+\left(\theta_\mathrm{sat}\cdot V_{\mathrm{DS_{sat}\infty}}\right)^2}+\theta_\mathrm{R}/2\cdot V_{\mathrm{DS_{sat}\infty}}} \tag{H.18}$$

It can easily be seen that expression (H.17) for $V_{\mathrm{DS_{sat}}}$ reduces to $V_{\mathrm{DS_{sat}\infty}}$ for long-channel devices. In Fig. H.1 (a), results from (H.17) are compared to results from the implicit relation (H.5) for n-type transistors. The influence of velocity saturation results in a decrease in $V_{\mathrm{DS_{sat}}}$, whereas the influence of

[20] The small difference between (3.44) and (3.45) is neglected here, see Fig. 3.8.



(a) (b)

Figure H.1: Saturation voltage $V_{\mathrm{DS_{sat}}}$ as a function of gate bias V_{GS} for (a) n-MOS and (b) p-MOS. Symbols are calculated using the implicit relation (H.5) (where $V_{\mathrm{DS_{sat}}} = \psi_{\mathrm{s_L}} - \psi_{\mathrm{s_0}}$), dashed lines are calculated using the approximation (H.17) and solid lines are calculated using the more accurate approximation (H.19). Results are given for the long-channel case ($\theta_{\mathrm{sat}} = 0$, $\theta_{\mathrm{R}} = 0$), a hypothetical short-channel case where the influence of series resisance is negligible ($\theta_{\mathrm{R}} = 0$) and a typical short-channel case. ($V_{\mathrm{FB}} = -1\mathrm{V}$, $k_0 = 0.3\mathrm{V}^{1/2}$, $\phi_{\mathrm{B}} = 1.0\mathrm{V}$, $k_{\mathrm{P}} \to \infty$)

series resistance results in an increase. It has been found that the accuracy of (H.17) can be improved by using the following empirical modification:

$$V_{\mathrm{DS_{sat}}} = V_{\mathrm{DS_{sat\infty}}} \cdot \left[1 - \frac{9}{10} \cdot \frac{\Delta_{\mathrm{SAT}}}{1 + \sqrt{1 - {\Delta_{\mathrm{SAT}}}^2}} \right] \tag{H.19}$$

As can be seen in Fig. H.1 (a), this results in a somewhat better description, in particular for the case where velocity saturation is dominant. Although the obtained accuracy seems not optimal, we have to bear in mind that the use of saturation voltage is an artefact of the gradual channel approximation and as a consequence, its definition is somewhat arbitrary. The obtained accuracy has been found to be adequate.

For p-type MOSFETs, θ_{sat} is replaced by $\theta_{\mathrm{sat}}/(1 + {\theta_{\mathrm{sat}}}^2 \cdot {V_{\mathrm{DS_{sat\infty}}}}^2)^{1/4}$ along the same lines as indicated in Section 3.3.2. In Fig. H.1 (b), results from (H.17) are compared to results from the implicit relation (H.5) for p-type transistors. Since the low-field mobility μ_0 of holes is typically about a factor 3 smaller than that of electrons (i.e., both θ_{sat} and θ_{R} are three times smaller), both velocity saturation and series resistance have less impact on $V_{\mathrm{DS_{sat}}}$ in p-MOS. Equation (H.19) gives an accurate description of saturation voltage for p-MOS transistors.

I Channel Length Modulation

The analysis of channel length modulation, as discussed in this appendix, is based on a pseudo two-dimensional analysis as was first proposed in [83] and subsequently modified by others who also took into account the shape of the source/drain structures [157]-[161].
When V_{DS} is increased beyond $V_{DS_{sat}}$, the velocity saturation point moves towards the source, causing effectively that the channel length L is shortened by a length ΔL. This movement is referred to as channel length modulation, and it effectively causes the conductance to be non-zero in the saturation region. An accurate calculation of ΔL is needed, but this is not easily accomplished, owing to the fact that it requires a two-dimensional solution of Poisson's equation for the channel saturation region. The pseudo two-dimensional analysis is based on the application of Gauss' law to a specific area at the drain end of the channel. In order to get an explicit solution of ΔL, the shape is assumed to be rectangular, see Fig. I.1.

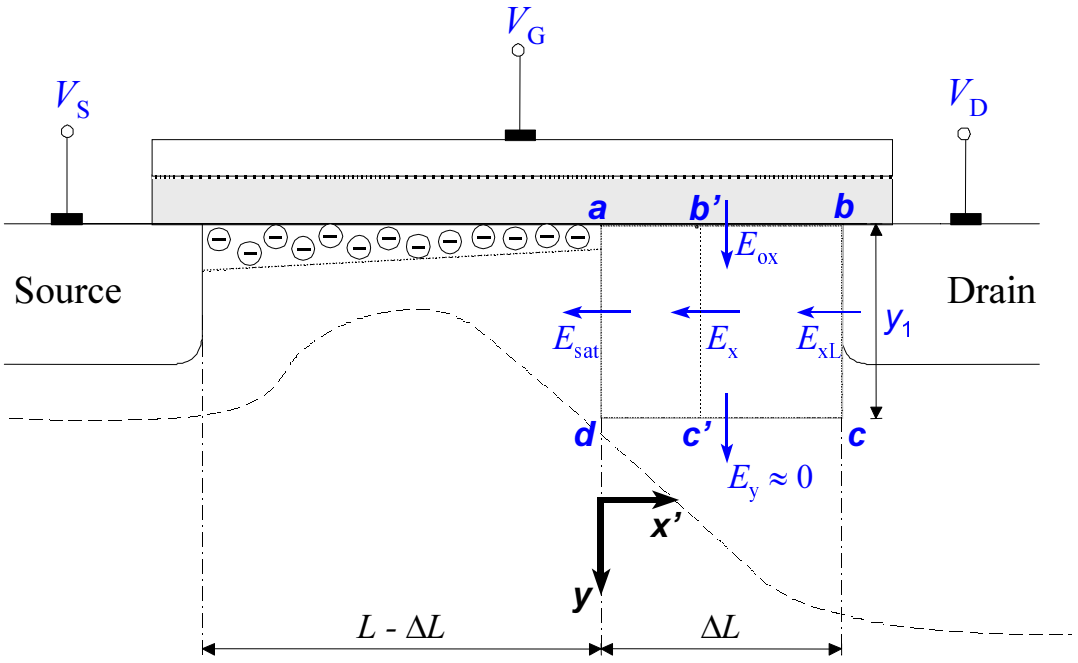


Figure I.1: Cross-section of an n-type MOS structure used for pseudo two-dimensional analysis of channel length modulation ΔL.

Here, y_1 is the depth of the box, which is taken to be so large that the normal electric field E_y at the boundary **c-d** can be assumed to be zero. The lateral electric field E_{sat} is the electric field for which the carrier velocity saturates, and it is assumed to be bias independent. The channel length modulation ΔL is simply given by:

$$\Delta L = \int_0^{\Delta L} \mathrm{d}x' = \int_{\psi_{s_{sat}}}^{\psi_{s_L}} \frac{\mathrm{d}\psi_s}{\partial\psi_s / \partial x'} = \int_{\psi_{s_{sat}}}^{\psi_{s_L}} \frac{\mathrm{d}\psi_s}{-E_x} \tag{I.1}$$

where $\psi_{s_{sat}}$ is the surface potential at the saturation point (i.e., $x' = 0$) and ψ_{s_L} is the surface potential at the drain junction. In order to solve (I.1), an expression for lateral electric field E_x in terms of ψ_s has to be found. Applying Gauss' law to the volume with sidewall **a-b'-c'-d** and unit width W, we can write:

$$\int_0^{y_1} \left(\frac{\partial\psi}{\partial x'} + E_{sat} \right) \cdot \mathrm{d}y + \int_0^{x'} \frac{\epsilon_{ox}}{\epsilon_{Si}} \cdot E_{ox} \cdot \mathrm{d}x' = \frac{1}{\epsilon_{Si}} \cdot \int_0^{x'} (q \cdot N_A \cdot y_1 - Q_{inv}) \cdot \mathrm{d}x' \tag{I.2}$$

where the electric field E_{ox} at the Si-SiO_2-interface is given by Q_g/ϵ_{ox}. If it is further assumed that the distribution of longitudinal field $\partial\psi/\partial x'$ over y in the space-charge region does not vary with x', we can write:

$$\int_0^{y_1} \frac{\partial \psi}{\partial x'} \cdot \mathrm{d}y = \kappa \cdot y_1 \cdot \frac{\partial \psi_s}{\partial x'} \tag{I.3}$$

where κ is a parameter depending on the distribution of the space charge, e.g., $\kappa = 1/3$ for a parabolic distribution. Using this relation in (I.2) and differentiating both sides with respect to x', we get:

$$\frac{\partial^2 \psi_s}{\partial {x'}^2} = -\frac{1}{\kappa \cdot \epsilon_{Si} \cdot y_1} \cdot \left(Q_g - q \cdot N_A \cdot y_1 + Q_{inv}\right) \tag{I.4}$$

Next, realizing that:

$$\frac{\partial^2 \psi_s}{\partial {x'}^2} = \frac{1}{2} \cdot \frac{\partial}{\partial \psi_s} \left(\frac{\partial \psi_s}{\partial x'}\right)^2 \tag{I.5}$$

we can integrate (I.4) resulting in:

$$\int_{{E_{sat}}^2}^{{E_x}^2} \mathrm{d}{E_x}^2 = -\frac{2}{\kappa \cdot \epsilon_{Si} \cdot y_1} \cdot \int_{\psi_{s_{sat}}}^{\psi_s} \left(Q_g - q \cdot N_A \cdot y_1 + Q_{inv}\right) \cdot \mathrm{d}\psi_s \tag{I.6}$$

Here, the gate charge density Q_g given by (2.17) can be approximated by:

$$Q_g \approx Q_{g_{sat}} - C_{g_{sat}} \cdot \left(\psi_s - \psi_{s_{sat}}\right) \tag{I.7}$$

where:

$$Q_{g_{sat}} = C_{ox} \cdot \left[\frac{2 \cdot \left(V^*_{GB} - \psi_{s_{sat}}\right)}{1 + \sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GB} - \psi_{s_{sat}}\right)}}\right] \tag{I.8}$$

$$C_{g_{sat}} = -\left.\frac{\partial Q_g}{\partial \psi_s}\right|_{\psi_s = \psi_{s_{sat}}} = C_{ox} \cdot \left(\frac{1}{\sqrt{1 + 4/{k_P}^2 \cdot \left(V^*_{GB} - \psi_{s_{sat}}\right)}}\right) \tag{I.9}$$

Expression (I.6) can now be solved, resulting in:

$${l_c}^2 \cdot {E_x}^2 = \frac{C_{g_{sat}}}{C_{ox}} \cdot \left(\psi_s - \psi_{s_{sat}}\right)^2 + B \cdot \left(\psi_s - \psi_{s_{sat}}\right) + C \tag{I.10}$$

where:

$$\begin{aligned} l_c &= \sqrt{\kappa \cdot \epsilon_{Si} \cdot y_1 / C_{ox}} \\ B &= -2 \cdot \left(Q_{g_{sat}} - q \cdot N_A \cdot y_1\right) / C_{ox} \\ C &= {l_c}^2 \cdot {E_{sat}}^2 - 2 \cdot \int_{\psi_{s_{sat}}}^{\psi_s} Q_{inv}/C_{ox} \cdot \mathrm{d}\psi_s \end{aligned} \tag{I.11}$$

The expression (I.10) for E_x can be introduced in the expression (I.1) for ΔL:

$$\Delta L = l_c \cdot \int_{\psi_{s_{sat}}}^{\psi_{s_L}} \frac{d\psi_s}{\sqrt{C_{g_{sat}}/C_{ox} \cdot \left(\psi_s - \psi_{s_{sat}}\right)^2 + B \cdot \left(\psi_s - \psi_{s_{sat}}\right) + C}} \tag{I.12}$$

The above integration can only be evaluated in closed form when the coefficient C is independent of I_{DS}. Therefore, in most evaluations it is assumed that $Q_{inv} \approx 0$ in the saturation region, so that C reduces to ${E_{sat}}^2 \cdot {l_c}^2$. Furthermore, y_1 is often assumed to be equal to the depletion width at the saturation point $\sqrt{2 \cdot \epsilon_{Si} \cdot \psi_{s_{sat}}/q \cdot N_A}$, in which case B reduces to zero. Under the above assumptions, equation (I.12) reduces to the following expression:

$$\Delta L = \frac{l_c}{\sqrt{C_{g_{sat}}/C_{ox}}} \cdot \ln\left[\frac{\psi_{s_L} - \psi_{s_{sat}} + \sqrt{\left(\psi_{s_L} - \psi_{s_{sat}}\right)^2 + C_{ox}/C_{g_{sat}} \cdot {E_{sat}}^2 \cdot {l_c}^2}}{\sqrt{C_{ox}/C_{g_{sat}}} \cdot E_{sat} \cdot l_c}\right] \tag{I.13}$$

In order to further simplify the above equation, we approximate $\psi_{s_{sat}}$ by $V_{DS_{sat}} + V_{SB} + \phi_B$ and ψ_{s_L} by $V_{DS} + V_{SB} + \phi_B$. In addition we neglect the impact of the poly-depletion effect (i.e., $C_{g_{sat}}/C_{ox} = 1$), and as a result (I.13) reduces to the commonly used expression:

$$\Delta L = l_c \cdot \ln\left[\frac{V_{DS} - V_{DS_{sat}} + \sqrt{\left(V_{DS} - V_{DS_{sat}}\right)^2 + {E_{sat}}^2 \cdot {l_c}^2}}{E_{sat} \cdot l_c}\right] \tag{I.14}$$

In view of the numerous simplifications made to obtain (I.14), E_{sat} and l_c are often considered as empirical parameters and we finally obtain:

$$\frac{\Delta L}{L} = \alpha \cdot \ln\left[\frac{V_{DS} - V_{DS_{sat}} + \sqrt{\left(V_{DS} - V_{DS_{sat}}\right)^2 + {V_P}^2}}{V_P}\right] \tag{I.15}$$

where α and V_P are model parameters, and α is inversely proportional with channel length L.

J Derivation of Gate-Induced Drain Leakage Equation

Gate-induced drain leakage occurs for negative gate-drain bias V_{GD} when a depletion region is formed underneath the gate-to-drain overlap region, a high transversal field is created in the depletion region and electron-hole pairs are generated by the band-to-band tunnelling of valence band electrons into the conduction band, see Fig. J.1. For an accurate description of GIDL, a precise expression of band-to-band tunnelling in the depletion layer is needed.

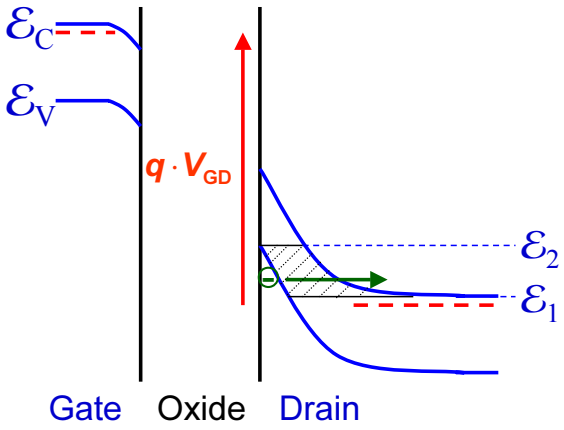


Figure J.1: Energy band diagram of gate/oxide/drain-extension structure for $V_{GD} \ll 0$. Electron-hole pairs are generated by tunnelling of valence band electrons into the conduction band resulting in a leakage current between drain and bulk, generally referred to as gate-induced drain leakage. Only valence band electrons between energies $\mathcal{E}_1$ and $\mathcal{E}_2$ can tunnel to the conduction band, and as a result, no tunnelling occurs for the case where $\mathcal{E}_1 > \mathcal{E}_2$ (i.e., $-q \cdot \psi_{s_{ov}} < \mathcal{E}_g$).

A general theory of band-to-band tunnelling has been developed in [163]-[167]. Under the assumption of a constant electric field E in the tunnelling direction, an expression for the tunnelling current density per unit energy $\mathrm{d}J_{BBT}/\mathrm{d}\mathcal{E}$ can be obtained:

$$\frac{\mathrm{d}J_{BBT}}{\mathrm{d}\mathcal{E}} \propto E^{\sigma} \cdot D\left(E, \mathcal{E}\right) \cdot \exp\left(-\frac{B}{E}\right) \tag{J.1}$$

where B is determined by physical constants and proportional to ${\mathcal{E}_g}^{3/2}$, and $\mathcal{E}$ is the energy of the incoming electron. In [163], it can be found that $\sigma = 1$ for direct transitions and $\sigma = 5/2$ for indirect transitions. Since silicon is an indirect semiconductor whose direct bandgap is much larger than its indirect gap, indirect transitions including electron-phonon interaction are predominant and we use $\sigma = 5/2$. In the above equation, the function $D(E, \mathcal{E})$ accounts for the fact that only valence band electrons between energies $\mathcal{E}_1$ and $\mathcal{E}_2$ can tunnel directly to the conduction band, see Fig. J.1. For $\mathcal{E}_1 < \mathcal{E} < \mathcal{E}_2$, the function D approximately equals unity. For $\mathcal{E} < \mathcal{E}_1$ or $\mathcal{E}_1 > \mathcal{E}_2$, on the other hand, the function D equals zero. The tunnelling current density J_{BBT} can now be calculated by integrating (J.1) over energy from $\mathcal{E}_1$ to $\mathcal{E}_2$ (i.e., the shaded area in Fig. J.1):

$$J_{BBT} \propto \int_{\mathcal{E}_1}^{\mathcal{E}_2} E^{\sigma} \cdot D\left(E, \mathcal{E}\right) \cdot \exp\left(-\frac{B}{E}\right) \cdot \mathrm{d}\mathcal{E} \tag{J.2}$$

The electric field E in the depletion layer is not constant. In order to perform the above integration, we need to replace $\mathrm{d}\mathcal{E}$ by $-q \cdot \mathrm{d}\psi$ and subsequently transform it to an integration over E $(= -\partial\psi/\partial y)$. The resulting integral can, however, not be solved analytically, although an approximate solution can

be obtained [164]. The latter results in a quite complex equation.
In most papers on GIDL, nevertheless, the current density J_{BBT} is based on [162] and simply assumed to be given by:

$$J_{\mathrm{BBT}} \propto {E_{\mathrm{tov}}}^2 \cdot \exp\left(-\frac{B}{E_{\mathrm{tov}}}\right) \tag{J.3}$$

where E_{tov} is the maximum electric field in the overlap region, and B is again a parameter theoretically proportional to ${\mathcal{E}_{\mathrm{g}}}^{3/2}$. Since the field dependence of J_{BBT} is dominated by the exponential term $\exp(-B/E_{\mathrm{tov}})$, the quadratic field dependence of the prefactor ${E_{\mathrm{tov}}}^2$ is not that important and consequently the use of approximation (J.3) is allowed.
The maximum electric field E_{tov} occurs at the Si/SiO_2-interface and consists not only of a (dominant) transversal component E_{y}, simply given by $-Q_{\mathrm{ov}}/\epsilon_{\mathrm{Si}}$, but also of a lateral component E_{x}. An expression for the lateral component in the overlapped drain extension is hard to come by. Empirically, we assume this component is constant along the x-direction and proportional to the voltage drop V_{DB}. The maximum electric field E_{tov} can now be written as:

$$E_{\mathrm{tov}} = \frac{C_{\mathrm{ox}}}{\epsilon_{\mathrm{Si}}} \cdot \sqrt{{V_{\mathrm{ov}}}^2 + (C \cdot V_{\mathrm{DB}})^2} \tag{J.4}$$

where C is an empirical parameter.

K Derivation of Gate Current Equations

K.1 General Formula for Gate Current Density

In a typical MOS structure, three major tunnelling mechanisms can be distinguished, namely, electron conduction-band tunnelling (ECB), electron valence-band tunnelling (EVB) and hole valence-band tunnelling (HVB), as illustrated in Fig. K.1. ECB tunnelling is the most important mechanism for an n-type MOSFET, whereas HVB tunnelling is most important for p-type MOSFETs. EVB tunnelling only occurs when the band bending is so strong that part of the valence band in the substrate silicon overlaps the conduction band in the gate polysilicon, see Fig. K.1, or vice-versa. In other words it occurs for $V_{ox} > \mathcal{E}_g/q \approx 1.15$ V, and since the supply voltage V_{DD} is 1.2 V or less for technologies where gate tunnelling becomes important, EVB tunnelling has been neglected in MOS Model 11.

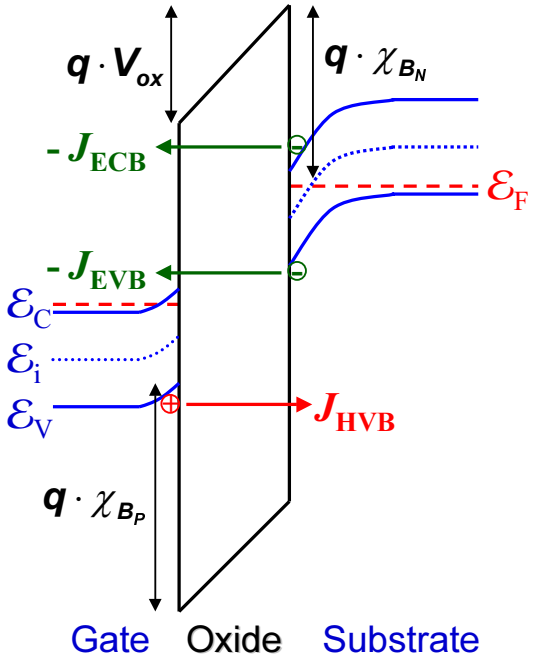


Figure K.1: The energy-band diagram of an n-MOS in inversion ($V^*_{GB} > 0$) where χ_{B_N} and χ_{B_P} are the oxide potential barriers for electrons and holes, respectively. The three major mechanisms of gate dielectric tunnelling are indicated: electron conduction-band tunnelling (J_{ECB}), electron valence-band tunnelling (J_{EVB}) and hole valence-band tunnelling (J_{HVB}).

In the first instance, focusing on the ECB tunnelling in an n-MOSFET in inversion (i.e., $V^*_{GB} > 0$), the gate current is caused by electrons tunnelling from the inversion layer to the gate. Assuming that the electrons in the lowest energy subband determine the tunnelling current, the gate current density J_G can be given by [118]:

$$J_G = \theta \cdot Q_{inv} \cdot P_{tun} \tag{K.1}$$

where θ is the fraction of electrons in the inversion layer residing in the lowest energy subband and P_{tun} is the transmission probability per electron per second. In the case of direct tunnelling, using the

WKB-approximation, the transmission probability P_{tun} can be given by:

$$P_{\text{tun}} = \frac{2}{{Z_0}^{3/2}} \cdot \frac{\sqrt{\frac{m_{\text{ox}}}{m_{\text{Si}}} \cdot \frac{\Delta\mathcal{E}}{q \cdot \chi_{\text{B}}^*}}}{1 + \frac{m_{\text{ox}}}{m_{\text{Si}}} \cdot \frac{\Delta\mathcal{E}}{q \cdot \chi_{\text{B}}^*}} \cdot \frac{\Delta\mathcal{E}}{\hbar} \cdot \exp\left[-2 \cdot \int_0^{t_{\text{ox}}} k(y) \cdot \mathrm{d}y\right] \tag{K.2}$$

where Z_0 is the zero of the zero-order Airy function defined in (D.1), m_{ox} and m_{Si} are the effective electron masses normal to the interface in silicon and silicon-oxide, respectively, $k(y)$ is the wave vector and χ_{B}^* is the effective oxide energy barrier:

$$\chi_{\text{B}}^* = \chi_{\text{B}} - \frac{\Delta\mathcal{E}}{q} \tag{K.3}$$

Here, χ_{B} is the oxide potential barrier between the conduction band of the silicon and the silicon-oxide (i.e., χ_{B_N}), and $\Delta\mathcal{E}$ is the energy of the lowest subband energy level with respect to the conduction band given by (D.1). Assuming that $\Delta\mathcal{E} \ll q \cdot \chi_{\text{B}_\text{N}}$ and bearing in mind that $E_{\text{Si}} = \epsilon_{\text{ox}}/\epsilon_{\text{Si}} \cdot E_{\text{ox}}$, we can approximate (K.2) by:

$$P_{\text{tun}} = \frac{q}{m_{\text{Si}}} \cdot \frac{\epsilon_{\text{ox}}}{\epsilon_{\text{Si}}} \cdot \sqrt{\frac{2 \cdot m_{\text{ox}}}{q \cdot \chi_{\text{B}}}} \cdot E_{\text{ox}} \cdot \exp\left[-2 \cdot \int_0^{t_{\text{ox}}} k(y) \cdot \mathrm{d}y\right] \tag{K.4}$$

The exponential term is the WKB approximation for the tunnelling probability. For parabolic dispersion, we can write $k(y) = \sqrt{2 \cdot m_{\text{ox}} \cdot \mathcal{E}(y)}/\hbar$ where $\mathcal{E}$ is the positive energy measured from the tunnelling electron's energy level to the bottom of the SiO_2 conduction band. The integral in the exponential term can now be written as:

$$-2 \cdot \int_0^{t_{\text{ox}}} k(y) \cdot \mathrm{d}y = \begin{cases} -\frac{B}{V_{\text{ox}}} \cdot \left[1 - \left(1 - \frac{V_{\text{ox}}}{\chi_{\text{B}}}\right)^{3/2}\right] & \text{for: } V_{\text{ox}} < \chi_{\text{B}} \\ -\frac{B}{V_{\text{ox}}} & \text{for: } V_{\text{ox}} \geq \chi_{\text{B}} \end{cases} \tag{K.5}$$

where B is the probability factor given by:

$$B = \frac{4}{3} \cdot \frac{\sqrt{2 \cdot q \cdot m_{\text{ox}}}}{\hbar} \cdot t_{\text{ox}} \cdot {\chi_{\text{B}}}^{3/2} \tag{K.6}$$

In (K.5) a distinction can be made between the direct-tunnelling regime (i.e., $V_{\text{ox}} < \chi_{\text{B}_\text{N}}$) and the Fowler-Nordheim regime (i.e., $V_{\text{ox}} \geq \chi_{\text{B}_\text{N}}$). For practical conditions in a typical MOSFET, only the direct-tunnelling regime is of importance.
Since holes have a different effective mass m_{ox} and a different oxide potential barrier χ_{B_P}, the expression for tunnelling probability P_{tun} generalized for both elecrons and holes is a function of oxide voltage V_{ox}, oxide potential barrier χ_{B} and probability factor B:

$$P_{\text{tun}}(V_{\text{ox}}, \chi_{\text{B}}, B) = A \cdot V_{\text{ox}} \cdot \exp\left(-\frac{B}{V_{\text{ox}}} \cdot \left[1 - \left(1 - \frac{V_{\text{ox}}}{\chi_{\text{B}}}\right)^{3/2}\right]\right) \tag{K.7}$$

where A is theoretically given by:

$$A = \frac{q}{m_{\text{Si}}} \cdot \frac{C_{\text{ox}}}{\epsilon_{\text{Si}}} \cdot \sqrt{\frac{2 \cdot m_{\text{ox}}}{q \cdot \chi_{\text{B}}}} \cdot \theta \tag{K.8}$$

The general expression for the gate current density is given by:

$$J_G = q \cdot N \cdot P_{tun}\,(V_{ox},\, \chi_B,\, B) \tag{K.9}$$

where N is the number of mobile carriers per unit area.

K.2 General Formula for the Source/Drain Partitioning of Gate Current

In order to derive the partitioning of the gate-to-channel current I_{GC} into a source component I_{GS} and a drain component I_{GD}, the following coupled differential equations have to be solved:

$$I_{DS}(x) = g(V) \cdot \frac{\partial V}{\partial x} \tag{K.10}$$

$$\frac{\partial I_{DS}(x)}{\partial x} = -W \cdot J_G(x) \tag{K.11}$$

where I_{DS} is the channel current, V is the quasi-Fermi potential ranging from V_{SB} at the source side to V_{DB} at the drain side, J_G is the gate leakage current density (in A/m^2) at a certain position x along the channel and $g(V)$ is the channel conductance given by:

$$g(V) = -\mu(V) \cdot W \cdot Q_{inv}(V) \tag{K.12}$$

Since the gate leakage current density J_G is a complicated function of the quasi-Fermi potential V, see K.9, the coupled equations (K.10) and (K.11) cannot be solved analytically. Nevertheless, an expression for the source/drain partitioning of the gate-to-channel current can easily be derived. Let us consider the case where $J_G = 0$. In this case, eqs. (K.10) and (K.11) can be solved, resulting in:

$$V(x) = V_0(x) \tag{K.13}$$

$$I_{DS}(x) = I_0 = \frac{1}{L} \cdot \int_{V_{SB}}^{V_{DB}} g(V) \cdot \mathrm{d}V \tag{K.14}$$

where I_0 is independent of position x, and is given by the conventional MOS channel current description, see Chapter 3. Next, let us consider the case where $J_G \neq 0$, we can write:

$$V(x) = V_0(x) + v(x) \tag{K.15}$$

$$I_{DS}(x) = I_0 + i(x) \tag{K.16}$$

where v and i are the corrections on the initial solution. The boundary conditions are given by:

$$v(0) = v(L) = 0 \tag{K.17}$$

As a consequence, from (K.10) and (K.14), we can find:

$$\int_0^L (I_0 + i) \cdot \mathrm{d}x = \int_{V_{SB}}^{V_{DB}} g\,(V_0 + v) \cdot \mathrm{d}\,(V_0 + v) = I_0 \cdot L \tag{K.18}$$

which results in:

$$\int_0^L i \cdot \mathrm{d}x = 0 \tag{K.19}$$

For the source/drain partitioning, the required solutions are:

$$I_{\mathrm{GS}} = i(0) \tag{K.20}$$

$$I_{\mathrm{GD}} = -i(L) \tag{K.21}$$

Using (K.14), equation (K.11) can be simplfied to:

$$\frac{\partial i}{\partial x} = -W \cdot J_{\mathrm{G}}(x) \tag{K.22}$$

which, integrating from 0 to x, can be rewritten as:

$$i(x) - i(0) = -W \cdot \int_0^x J_{\mathrm{G}}(x) \cdot \mathrm{d}x \tag{K.23}$$

This equation can be integrated from $x = 0$ to $x = L$:

$$\int_0^L i(x) \cdot \mathrm{d}x - \int_0^L i(0) \cdot \mathrm{d}x = -W \cdot \int_0^L \int_0^x J_{\mathrm{G}}(\hat{x}) \cdot \mathrm{d}\hat{x} \cdot \mathrm{d}x \tag{K.24}$$

which can be simplified using (K.20) and (K.19):

$$I_{\mathrm{GS}} = \frac{W}{L} \cdot \int_0^L \int_0^x J_{\mathrm{G}}(\hat{x}) \cdot \mathrm{d}\hat{x} \cdot \mathrm{d}x \tag{K.25}$$

The double integral is of the form $\int_0^L G(x) \cdot \mathrm{d}x$, where $G(x) = \int_0^x J_{\mathrm{G}}(\hat{x}) \cdot \mathrm{d}\hat{x}$. Applying integration by parts to $\int_0^L G(x) \cdot \mathrm{d}x$, with G and x the two variables involved, we can rewrite (K.25) as:

$$I_{\mathrm{GS}} = W \cdot \int_0^L \left[1 - \frac{x}{L}\right] \cdot J_{\mathrm{G}}(x) \cdot \mathrm{d}x \tag{K.26}$$

Since the gate-to-channel current I_{GC} is given by:

$$I_{\mathrm{GC}} = W \cdot \int_0^L J_{\mathrm{G}}(x) \cdot \mathrm{d}x \tag{K.27}$$

the gate-to-drain current I_{GD} ($= I_{\mathrm{GC}} - I_{\mathrm{GS}}$) can be written as:

$$I_{\mathrm{GD}} = W \cdot \int_0^L \frac{x}{L} \cdot J_{\mathrm{G}}(x) \cdot \mathrm{d}x \tag{K.28}$$

Note that (K.26) and (K.28) are similar to the Ward-Dutton scheme for the charge partitioning of Q_{S} and Q_{D}.
The above-derived partitioning scheme for the gate-to-channel current has been published in [14]. Shih *et al.* have shown that the above partitioning scheme can also be derived using Green's function technique [168].

L Auxiliary Derivations for Integration along the Channel

In the derivation of the gate current model, the charge model and the noise model, certain electrical variables have to be integrated along the channel. In order to so, we need to rewrite $\mathrm{d}x$ and x in terms of surface potential ψ_s or inversion-layer charge density Q_inv.
From (2.20), we find $\partial Q_\mathrm{inv} = -C_\mathrm{inv} \cdot \partial \psi_\mathrm{s}$, and as a result the equation for channel current (3.1) can be rewritten as:

$$I_\mathrm{DS} = \mu \cdot W \cdot \frac{Q^*_\mathrm{inv}}{C_\mathrm{inv}} \cdot \frac{\partial Q^*_\mathrm{inv}}{\partial x} = \mu \cdot \frac{W}{L} \cdot \frac{\bar{Q}^*_\mathrm{inv}}{C_\mathrm{inv}} \cdot \Delta Q^*_\mathrm{inv} = -\mu \cdot \frac{W}{L} \cdot \bar{Q}^*_\mathrm{inv} \cdot \Delta\psi \tag{L.1}$$

where $\Delta Q^*_\mathrm{inv} = Q^*_{\mathrm{inv}_\mathrm{L}} - Q^*_{\mathrm{inv}_0}$. From (L.1), neglecting the influence of velocity saturation, one can easily derive that:

$$\frac{\mathrm{d}x}{L} = \frac{Q^*_\mathrm{inv}}{\bar{Q}^*_\mathrm{inv}} \cdot \frac{\mathrm{d}Q^*_\mathrm{inv}}{\Delta Q^*_\mathrm{inv}} = \frac{Q^*_\mathrm{inv}}{\bar{Q}^*_\mathrm{inv}} \cdot \frac{\mathrm{d}\psi_\mathrm{s}}{\Delta\psi} \tag{L.2}$$

and one can furthermore deduce that:

$$\int_0^x I_\mathrm{DS} \cdot \mathrm{d}x = \mu \cdot W \cdot \int_{Q^*_{\mathrm{inv}_0}}^{Q^*_\mathrm{inv}} \frac{Q^*_\mathrm{inv}}{C_\mathrm{inv}} \cdot \mathrm{d}Q^*_\mathrm{inv} \tag{L.3}$$

resulting in:

$$\frac{x}{L} = \frac{\int_0^x I_\mathrm{DS} \cdot \mathrm{d}x}{\int_0^L I_\mathrm{DS} \cdot \mathrm{d}x} = \frac{1}{2} \cdot \frac{Q^*_\mathrm{inv} + Q^*_{\mathrm{inv}_0}}{\bar{Q}^*_\mathrm{inv}} \cdot \frac{Q^*_\mathrm{inv} - Q^*_{\mathrm{inv}_0}}{\Delta Q^*_\mathrm{inv}} \tag{L.4}$$

M Derivation of Thermal Noise Expressions

M.1 Derivation of General Thermal Noise

In this section, we will present a detailed derivation of formula (7.97) in [35]. As a starting point for this calculation, we use the following equation, a derivation of which is found in [152, 153]:

$$S_{\mathrm{I}}(f) = \frac{1}{L^2} \cdot \int_0^L S_{I,\Delta x}(x, f) \cdot \Delta x \cdot \mathrm{d}x \tag{M.1}$$

where $S_{I,\Delta x}(x, f)$ is the current noise associated with an infinitesimal part of the inversion channel at position x, having a width Δx. This formula is valid when the current noise sources, associated with different positions x in the inversion channel, are uncorrelated. This is usually assumed to be true for thermal noise.
It is well known that a resistor R exhibits thermal noise, given by:

$$S_{\mathrm{I}}(f) = \frac{4 \cdot k_{\mathrm{B}} \cdot T}{R} \tag{M.2}$$

Consider now a piece Δx of a MOSFET. The current through this piece of MOSFET is given by:

$$I_{\mathrm{DS}} = g(x) \cdot \frac{\Delta V}{\Delta x} \tag{M.3}$$

where ΔV is the potential drop over the section Δx, and the channel conductance $g(x)$ is given by:

$$g(x) = -\mu(x) \cdot W \cdot Q_{\mathrm{inv}}(x) \tag{M.4}$$

From (M.3), we find that this piece of MOSFET has a differential resistance $\Delta x / g(x)$. Applying (M.2), this differential resistance exhibits a thermal noise spectral density given by:

$$S_{I,\Delta x}(x, f) = \frac{4 \cdot k_{\mathrm{B}} \cdot T \cdot g(x)}{\Delta x} \tag{M.5}$$

Thus for thermal noise, the basic formula (M.1) may be rewritten as:

$$S_{\mathrm{I_D}}(f) = \frac{4 \cdot k_{\mathrm{B}} \cdot T}{L^2} \cdot \int_0^L g(x) \cdot \mathrm{d}x \tag{M.6}$$

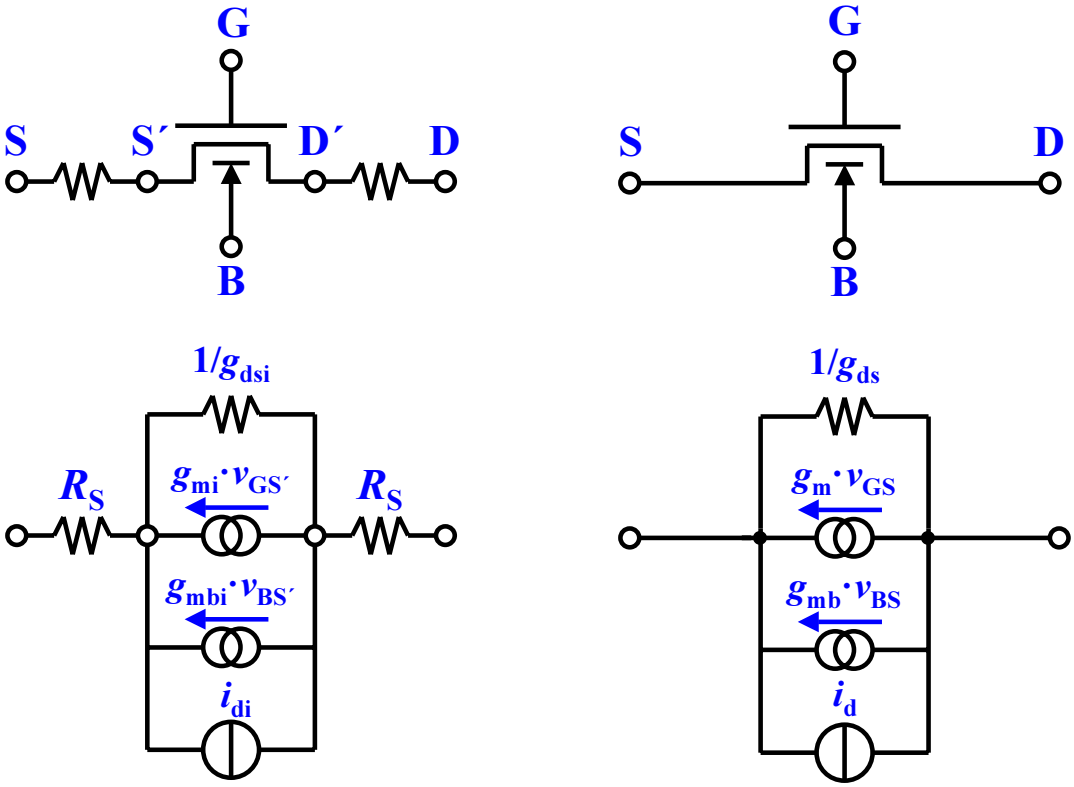
This is equation (7.95) in [35], which can be rewritten in a more convenient form. The integration over x is replaced by integration over the quasi Fermi-potential V using (M.3). Now, we arrive at:

$$S_{\mathrm{I_D}}(f) = \frac{4 \cdot k_{\mathrm{B}} \cdot T}{I_{\mathrm{DS}} \cdot L^2} \cdot \int_0^{V_{\mathrm{DS}}} g^2(V) \cdot \mathrm{d}V \tag{M.7}$$

This is equation (7.96) in [35], it is valid both in all regions of operation.

M.2 Impact of Series Resistance on Thermal Noise

The thermal noise as calculated by (7.29) has been calculated including the impact of velocity saturation and neglecting the impact of series resistance. In order to include the latter, we need to consider



(a) (b)

Figure M.1: (a) The MOSFET with external series resistances and the equivalent small-signal model, where conductance $g_{\mathrm{dsi}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{D'S'}}$, transconductance $g_{\mathrm{mi}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{GS'}}$ and substrate transconductance $g_{\mathrm{mbi}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{BS'}}$ are used. The thermal noise source i_{di} only incorporates the impact of velocity saturation. (b) The MOSFET with internal series resistances and the equivalent small-signal model, where $g_{\mathrm{ds}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{DS}}$, $g_{\mathrm{m}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{GS}}$ and $g_{\mathrm{mb}} = \partial I_{\mathrm{DS}}/\partial V_{\mathrm{BS}}$ are used. The thermal noise source i_{d} incorporates both the impact of velocity saturation and series resistance.

the MOSFET with external and internal series resistances, see Fig. M.1 (a) and (b) respectively. Of course, in reality, the drain and source series resistances are also subject to thermal noise. This thermal noise, however, is neglible, and both series resistances are consequently considered noiseless.
As mentioned in Fig. M.1, the thermal noise source i_{di} only includes velocity saturation and its spectral density $S_{\mathrm{I_{Di}}}$ is thus given by (7.29):

$$S_{\mathrm{I_{Di}}} = \frac{4 \cdot k_{\mathrm{B}} \cdot T}{{G_{\mathrm{mob}}}^2} \cdot \left[-\frac{\beta}{C_{\mathrm{ox}}} \cdot G_{\mathrm{vsat}} \cdot \left(\bar{Q}_{\mathrm{inv}} + \frac{{C_{\mathrm{inv}}}^2 \cdot {\Delta\psi_{\mathrm{eff}}}^2}{12 \cdot \bar{Q}^*_{\mathrm{inv}}} \right) - {\theta_{\mathrm{sat}}}^2 \cdot I_{\mathrm{DS}} \cdot \Delta\psi_{\mathrm{eff}} \right] \tag{M.8}$$

where we have simply replaced the variable $\Delta\psi$ in (7.29) by $\Delta\psi_{\mathrm{eff}}$ given by, see Appendix G:

$$\Delta\psi_{\mathrm{eff}} = \Delta\psi - 2 \cdot I_{\mathrm{DS}} \cdot R_{\mathrm{S}} = \Delta\psi \cdot \left(1 - \frac{\theta_{\mathrm{R}} \cdot \bar{Q}^*_{\mathrm{inv}}}{C_{\mathrm{ox}} \cdot G_{\mathrm{tot}}} \right) \tag{M.9}$$

In order to calculate the spectral density S_{I_D} of noise current i_d, we need to calculate the small-signal channel current i_{DS}. Using the MOSFET with external series resistances, we can write:

$$i_{DS} = \frac{v_{SS'}}{R_S} = g_{dsi} \cdot v_{D'S'} + g_{mi} \cdot v_{GS'} + g_{mbi} \cdot v_{BS'} + i_{di} = \frac{v_{DD'}}{R_S} \tag{M.10}$$

Eliminating $v_{S'}$ and $v_{D'}$, we obtain:

$$i_{DS} = \frac{g_{dsi} \cdot v_{DS} + g_{mi} \cdot v_{GS} + g_{mbi} \cdot v_{BS} + i_{di}}{1 + (2 \cdot g_{dsi} + g_{mi} + g_{mbi}) \cdot R_S} \tag{M.11}$$

Using the MOSFET with internal series resistances, the small-signal current i_{DS} can be calculated as well:

$$i_{DS} = g_{ds} \cdot v_{DS} + g_m \cdot v_{GS} + g_{mb} \cdot v_{BS} + i_d \tag{M.12}$$

Comparing (M.11) and (M.11), we can write the external quantities g_{ds}, g_m, g_{mb} and i_d in terms of the internal quantities g_{dsi}, g_{mi}, g_{mbi} and i_{di}:

$$g_{ds} = \frac{g_{dsi}}{1 + (2 \cdot g_{dsi} + g_{mi} + g_{mbi}) \cdot R_S} \tag{M.13}$$

$$g_m = \frac{g_{mi}}{1 + (2 \cdot g_{dsi} + g_{mi} + g_{mbi}) \cdot R_S} \tag{M.14}$$

$$g_{mb} = \frac{g_{mbi}}{1 + (2 \cdot g_{dsi} + g_{mi} + g_{mbi}) \cdot R_S} \tag{M.15}$$

$$i_d = \frac{i_{di}}{1 + (2 \cdot g_{dsi} + g_{mi} + g_{mbi}) \cdot R_S} \tag{M.16}$$

The latter equation can also be written in terms of the external quantities g_{ds}, g_m and g_{mb}:

$$i_d = [1 - (2 \cdot g_{ds} + g_m + g_{mb}) \cdot R_S] \cdot i_{di} \tag{M.17}$$

In the above, we can approximate the term $2 \cdot g_{ds} + g_m + g_{mb}$ by $2 \cdot \beta / C_{ox} \cdot \bar{Q}^*_{inv} / G_{tot}$ resulting in:

$$i_d \approx \left[1 - \left(2 \cdot \frac{\beta}{C_{ox}} \cdot \frac{\bar{Q}^*_{inv}}{G_{tot}}\right) \cdot R_S\right] \cdot i_{di} = \left[1 - \frac{G_R}{G_{tot}}\right] \cdot i_{di} \tag{M.18}$$

Furthermore, approximating G_{tot} by $G_{vsat} + G_R$, we can rewrite the above into:

$$i_d \approx \frac{G_{vsat}}{G_{tot}} \cdot i_{di} \tag{M.19}$$

Using the above expression and (M.8), the spectral density S_{I_D} can now be calculated to be:

$$S_{I_D} = \frac{4 \cdot k_B \cdot T \cdot {G_{vsat}}^2}{{G_{mob}}^2 \cdot {G_{tot}}^2} \cdot \left[-\frac{\beta}{C_{ox}} \cdot G_{vsat} \cdot \left(\bar{Q}_{inv} + \frac{{C_{inv}}^2 \cdot {\Delta\psi_{eff}}^2}{12 \cdot \bar{Q}^*_{inv}}\right) - {\theta_{sat}}^2 \cdot I_{DS} \cdot \Delta\psi_{eff}\right] \tag{M.20}$$

It has been found that this equation can be approximated by:

$$S_{I_D} \approx \frac{4 \cdot k_B \cdot T}{{G_{mob}}^2} \cdot \left[-\frac{\beta}{C_{ox}} \cdot \frac{{G_{vsat}}^2}{G_{tot}} \cdot \left(\bar{Q}_{inv} + \frac{{C_{inv}}^2 \cdot \Delta\psi^2}{12 \cdot \bar{Q}^*_{inv}}\right) - {\theta_{sat}}^2 \cdot I_{DS} \cdot \Delta\psi\right] \tag{M.21}$$

References

[1] R.M.D.A. Velghe, D.B.M. Klaassen, F.M. Klaassen, "MOS Model 9," *NL-UR 003/94*, Philips Electronics N.V., 1994.
internet: www.semiconductors.philips.com/Philips_Models

[2] internet: www-device.eecs.berkeley.edu

[3] C. Enz, F. Krummenacher and E. Vittoz, "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications," *J. Analog Integrated Circ. and Sign. Processing*, Vol. 8, pp. 83-114, 1995.
internet: legwww.epfl.ch/ekv

[4] R. van Langevelde and F.M. Klaassen, "Influence of Mobility Degradation on Distortion Analysis in MOSFETs," *in Proc. ESSDERC*, pp. 667-670, 1996.

[5] R. van Langevelde and F.M. Klaassen, "Effect of Gate-Field Dependent Mobility Degradation on Distortion Analysis in MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-44, No. 11, pp. 2044-2052, 1997.

[6] R. van Langevelde and F.M. Klaassen, "Accurate Drain Conductance Modeling for Distortion Analysis in MOSFETs," *in IEDM Tech. Dig.*, pp. 313-316, 1997.

[7] R. van Langevelde, "A Compact MOSFET Model for Distortion Analysis in Analog Circuit Design," *Ph.D. Thesis*, Eindhoven University of Technology, Eindhoven, 1998.
Downloadable from: http://alexandria.tue.nl/extra2/980343.pdf

[8] A.J. Scholten *et al.*, "Accurate Thermal Noise Model for Deep-Submicron CMOS," *in IEDM Tech. Dig.*, pp. 155-158, 1999.

[9] R. van Langevelde and F.M. Klaassen, "An Explicit Surface-Potential Based MOSFET Model for Circuit Simulation," *Solid-State Electron.*, Vol. 44, pp. 409-418, 2000.

[10] R. van Langevelde *et al.*, "RF-Distortion in Deep-Submicron CMOS Technologies," *in IEDM Tech. Dig.*, pp. 807-810, 2000.

[11] A.J. Scholten, R. van Langevelde, L.F. Tiemeijer, R.J. Havens and D.B.M. Klaassen, "Compact MOS Modelling for RF CMOS Circuit Simulation," *in Proc. SISPAD*, pp. 194-201, 2001.

[12] R. van Langevelde, A.J. Scholten, R.J. Havens, L.F. Tiemeijer and D.B.M. Klaassen, "Advanced Compact MOS Modelling," *in Proc. ESSDERC*, pp. 81-88, 2001.

[13] A.J. Scholten, R. Duffy, R. van Langevelde and D.B.M. Klaassen, "Compact Modelling of Pocket-Implanted MOSFETs," *in Proc. ESSDERC*, pp. 311-314, 2001.

[14] R. van Langevelde, A.J. Scholten, R. Duffy, F.N. Cubaynes, M.J. Knitel and D.B.M. Klaassen, "Gate Current: Modeling, ΔL Extraction and Impact on RF Performance," *in IEDM Tech. Dig.*, pp. 289-292, 2001.

[15] A.J. Scholten *et al.*, "Compact Modeling of Drain and Gate Current Noise for RF CMOS," *in IEDM Tech. Dig.*, pp. 129-132, 2002.

[16] R. van Langevelde, "MOS Model 11, Level 1100," *NL-UR 2001/813*, Philips Electronics N.V., 2001.
internet: www.semiconductors.philips.com/Philips_Models

[17] R. van Langevelde, A.J. Scholten and D.B.M. Klaassen, "MOS Model 11, Level 1101," *NL-UR 2002/802*, Philips Electronics N.V., 2002.
internet: www.semiconductors.philips.com/Philips_Models

[18] N.D. Arora, R. Rios, C.-L. Huang and K. Raol, "PCIM: A Physically Based Continuous Short-Channel IGFET Model for Circuit Simulation," *IEEE Trans. Electron Devices*, Vol. ED-41, No. 6, pp. 988-997, 1994.

[19] H.K. Gummel and K. Singhal, "Inversion Charge Modeling," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 8, pp. 1585-1593, 2001.

[20] H.K. Gummel and K. Singhal, "Intrinsic MOSFET Capacitance Coefficients," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 10, pp. 2384-2393, 2001.

[21] A.R. Boothroyd, S.W. Tarasewicz and C. Slaby, "MISNAN - A Physically Based Continuous MOSFET Model for CAD Applications," *IEEE Trans. Computer-Aided Des.*, Vol. CAD-10, pp. 1512-1529, 1991.

[22] M. Miura-Mattausch, "Analytical MOSFET Model for Quarter Micron Technologies," *IEEE Trans. Computer-Aided Des.*, Vol. CAD-13, No. 5, pp. 610-615, 1994.

[23] M. Miura-Mattausch *et al.*, "Unified Complete MOSFET Model for Analysis of Digital and Analog Circuits," *IEEE Trans. Computer-Aided Des.*, Vol. CAD-15, pp. 1-7, 1996.

[24] H.J. Mattausch *et al.*, "HiSIM: The First Complete Drift-Diffusion MOSFET Model for Circuit Simulation," *in Proc. ICSICT*, pp. 861-866, 2001.

[25] M. Miura-Mattausch *et al.*, "HiSIM: A MOSFET Model for Circuit Simulation Connecting Circuit Performance with Technology," *in IEDM Tech. Dig.*, pp. 109-112, 2002.

[26] K. Joardar, K.K. Gullapulli, C.C. McAndrew, M.E. Burnham and A. Wild, "An Improved MOSFET Model for Circuit Simulation," *IEEE Trans. Electron Devices*, Vol. ED-45, No. 1, pp. 134-148, 1998.

[27] G. Gildenblat, T.-L. Chen and P. Bendix, "Analytical Approximation for Perturbation of MOSFET Surface Potential by Polysilicon Depletion Layer," *Electron. Lett.*, Vol. 35, No. 22, pp. 1974-1976, 1999.

[28] G. Gildenblat, T.-L. Chen and P. Bendix, "Closed-form Approximation for the Perturbation of MOSFET Surface Potential by Quantum-Mechanical Effects," *Electron. Lett.*, Vol. 36, No. 12, pp. 1072-1073, 2000.

[29] G. Gildenblat and T.-L. Chen, "Analytical Approximation for the MOSFET Surface Potential," *Solid-State Electron.*, Vol. 45, pp. 335-341, 2001.

[30] T.-L. Chen and G. Gildenblat, "Symmetric Bulk Charge Linearisation in Charge-Sheet MOSFET Model," *Electron. Lett.*, Vol. 37, No. 12, pp. 791-793, 2001.

[31] B. Gu *et al.*, "A Surface-Potential-Based Extrinsic Compact MOSFET Model," *in Proc. Nanotech*, Vol. 2, pp. 364-367, 2003.

[32] G. Groenewold and W.J. Lubbers, "Systematic Distortion Analysis for MOSFET Integrators with Use of New MOSFET Model," *IEEE Trans. Circuits Syst.*, Vol. 41, pp. 569-580, 1994.

[33] W. Liu, "Model Quality Needs to Be Job One," *IEEE Circuits & Devices Magazine*, pp. 29-35, November 2002.

[34] Y.P. Tsividis, *Operation and Modeling of the MOS Transistor,* New York: McGraw-Hill, 1987.

[35] H.C. de Graaff and F.M. Klaassen, *Compact Transistor Modelling for Circuit Design,* Vienna/New York: Springer-Verlag, 1990.

[36] N.D. Arora, *MOSFET Models for VLSI Circuit Simulation,* Vienna/New York: Springer-Verlag, 1993.

[37] M.J. van Dort, P.H. Woerlee and A.J. Walker, "A Simple Model for Quantization Effects in Heavily-Doped Silicon MOSFETs at Inversion Conditions," *Solid-State Electron.*, Vol. 37, pp. 411-414, 1994.

[38] N.D. Arora, R. Rios and C.-L. Huang, "Modeling the Polysilicon Depletion Effect and Its Impact on Submicrometer CMOS Circuit Performance," *IEEE Trans. Electron Devices*, Vol. ED-42, No. 5, pp. 935-943, 1995.

[39] F. Stern, "Quantum Properties of Surface Space-Charge Layers," *CRC Crit. Rev. Solid State Sci.*, pp. 499-514, 1974.

[40] C. Moglestue, "Self-Consistent Calculation of Electron and Hole Inversion Charges at Silicon-Silicon Dioxide Interfaces," *J. Appl. Phys.*, Vol. 59, No. 9, pp. 3175-3183, 1986.

[41] R. Rios and N.D. Arora, "Determination of Ultra-Thin Gate Oxide Thicknesses for CMOS Structures Using Quantum Effects," *in IEDM Tech. Dig.*, pp. 613-316, 1994.

[42] R. Rios *et al.*, "A Physical Compact MOSFET Model, Including Quantum Mechanical Effects for Statistical Circuit Design Applications," *in IEDM Tech. Dig.*, pp. 937-940, 1995.

[43] H.C. Pao and C.T. Sah, "Effects of Diffusion Current on Characteristics of Metal-Oxide (Insulator)-Semiconductor Transistors," *Solid-State Electron.*, Vol. 9, pp. 927-937, 1966.

[44] J.R. Brews, "A Charge-Sheet Model of the MOSFET," *Solid-State Electron.*, Vol. 21, pp. 345-355, 1978.

[45] F. van de Wiele, "A Long-Channel MOSFET Model," *Solid-State Electron.*, Vol. 22, pp. 991-997, 1979.

[46] R.F. Pierret and J.A. Shields, "Simplified Long-Channel MOSFET Theory," *Solid-State Electron.*, Vol. 26, No. 2, pp. 143-147, 1983.

[47] R. Troutman, "VLSI Limitations from Drain Induced Barrier Lowering," *IEEE J. Solid-State Circuits*, Vol. SC-14, No. 2, pp. 383-391, 1979.

[48] Z.-H. Liu *et al.*, "Threshold Voltage Model for Deep-Submicrometer MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-40, No. 1, pp. 86-95, 1993.

[49] T. Skotnicki and W. Marciniak, "A New Approach to Threshold Voltage Modeling of Short-Channel MOSFETs," *Solid-State Electron.*, Vol. 29, pp. 1115-1127, 1985.

[50] T. Skotnicki, G. Merckel and T. Pedron, "The Voltage-Doping Transformation: A New Approach to Modeling of MOSFET Short-Channel Effects," *IEEE Electron Device Lett.*, Vol. 9, No. 3, pp. 109-112, 1988.

[51] A.G. Sabnis and J.T. Clemens, "Characterization of the Electron Mobility in the Inverted <100> Si Surface," *in IEDM Tech. Dig.*, pp. 18-21, 1979.

[52] M.S. Lin, "The Classical versus the Quantum Mechanical Model of Mobility Degradation due to the Gate Field in MOSFET Inversion Layers," *IEEE Trans. Electron Devices*, Vol. ED-32, No. 3, pp. 700-710, 1985.

[53] A.J. Walker and P.H. Woerlee, "A Mobility Model for MOSFET Device Simulation," *J. Phys. Coll.*, Vol. C4, pp. 265-268, 1988.

[54] T.J. Krutsick and M.H. White, "Consideration of Doping Profiles in MOSFET Mobility Modeling," *IEEE Trans. Electron Devices*, Vol. ED-35, No. 7, pp. 1153-1155, 1988.

[55] M.S. Lin, "A Better Understanding of the Channel Mobility of Si MOSFET's Based on the Physics of Quantized Subbands," *IEEE Trans. Electron Devices*, Vol. ED-35, No. 12, pp. 2406-2411, 1988.

[56] S. Takagi, M. Iwase and A. Toriumi, "On the Universality of Inversion-Layer Mobility in N- and P-channel MOSFET's," *in IEDM Tech. Dig.*, pp. 398-401, 1988.

[57] D.S. Jeon and D.E. Burk, "MOSFET Electron Inversion Layer Mobilities-A Physically Based Semi-Empirical Model for a Wide Temperature Range," *IEEE Trans. Electron Devices*, Vol. ED-36, No. 8, pp. 1456-1463, 1989.

[58] C.-L. Huang and G.S. Gildenblat, "Measurements and Modeling of the n-Channel MOSFET Inversion Layer Mobility and Device Characteristics in the Temperature Range 60-300 K," *IEEE Trans. Electron Devices*, Vol. ED-37, No. 5, pp. 1289-1300, 1990.

[59] V.M. Agostinelli, H. Shin and A.F. Tasch, "A Comprehensive Model for Inversion Layer Hole Mobility for Simulation of Submicrometer MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-38, No. 1, pp. 151-159, 1991.

[60] K. Lee, J.-S. Choi, S.-P. Sim and C.-K. Kim, "Physical Understanding of Low-Field Carrier Mobility in Silicon MOSFET Inversion Layer," *IEEE Trans. Electron Devices*, Vol. ED-38, No. 8, pp. 1905-1911, 1991.

[61] S. Takagi, A. Toriumi, M. Iwase and H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFET's: Part I - Effects of Substrate Impurity Concentration," *IEEE Trans. Electron Devices*, Vol. ED-41, No. 12, pp. 2357-2362, 1994.

[62] S. Takagi, A. Toriumi, M. Iwase and H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFET's: Part II - Effects of Surface Orientation," *IEEE Trans. Electron Devices*, Vol. ED-41, No. 12, pp. 2363-2368, 1994.

[63] C.-L. Huang and N.D. Arora, "Characterization and Modeling of the n- and p-Channel MOSFETs Inversion-Layer Mobility in the Range 25-150 C," *Solid-State Electron.*, Vol. 37, No. 1, pp. 97-103, 1994.

[64] K.-S. Min and K. Lee, "Temperature-dependent hole and electron mobility models for CMOS circuit simulation," *IEEE Trans. Electron Devices*, Vol. ED-42, No. 11, pp. 1956-1961, 1995.

[65] D. Vasileska and D.K. Ferry, "Scaled Silicon MOSFET's: Universal Mobility Behavior," *IEEE Trans. Electron Devices*, Vol. ED-44, No. 4, pp. 577-583, 1997.

[66] S. Villa, A.L. Lacaita, L.M. Perron and R. Bez, "A Physically-Based Model of the Effective Mobility in Heavily-Doped *n*-MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-45, No. 1, pp. 110-115, 1998.

[67] J. Banqueri, J.A. López-Villanueva, P. Cartujo-Cassinello, S. Rodríguez and J.E. Carceller, "Experimental Determination of the Effective Mobility in NMOSFETs: a Comparative Study," *Solid-State Electron.*, Vol. 43, pp. 701-707, 1999.

[68] Y. Ma, L. Liu and Z. Li, "A Discussion on the Universality of Inversion Layer Mobility in MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-46, No. 9, pp. 1920-1922, 1999.

[69] A. Pirovano, A.L. Lacaita, G. Zandler and R. Oberhuber, "Explaining the Dependences of the Hole and Electron Mobilities in Si Inversion Layers," *IEEE Trans. Electron Devices*, Vol. ED-47, No. 4, pp. 718-724, 2000.

[70] M. Kondo and H. Tanimoto, "An Accurate Coulomb Mobility Model for MOS Inversion Layer and Its Application to NO-Oxynitride Devices," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 2, pp. 265-270, 1999.

[71] C. Jacoboni, C. Canalo, G. Ottaviani and A. Quaranta, "A Review of Some Charge Transport Properties of Silicon," *Solid-State Electron.*, Vol. 20, No. 1, pp. 77-89, 1977.

[72] D.L. Scharfetter and H.K. Gummel, "Large-Signal Analysis of a Silicon Read Diode Oscillator," *IEEE Trans. Electron Devices*, Vol. ED-16, No. 1, pp. 64-77, 1969.

[73] K. Yamaguchi, "A Mobility Model for Carriers in the MOS Inversion Layer," *IEEE Trans. Electron Devices*, Vol. ED-30, No. 6, pp. 658-663, 1983.

[74] J. Sim, "An Analytical Deep Submicron MOS Device Model Considering Velocity Overshoot Behavior Using Energy Balance Equation," *IEEE Trans. Electron Devices*, Vol. ED-42, No. 5, pp. 864-869, 1995.

[75] L. Ge, J.G. Fossum and B. Liu, "Physical Compact Modeling and Analysis of Velocity Overshoot in Extremely Scaled CMOS Devices and Circuits," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 9, pp. 2074-2080, 2001.

[76] H. Wang and G. Gildenblat, "Scattering Matrix Based Compact MOSFET Model," *in IEDM Tech. Dig.*, pp. 125-128, 2002.

[77] A.J. Scholten and D.B.M. Klaassen, "Anomalous Geometry Dependence of Source/Drain Resistance in Narrow-Width MOSFETs," *in Proc. IEEE ICMTS*, Vol. 11, pp. 77-82, 1998.

[78] G.J. Hu, C. Chang and Y.T. Chia, "Gate Voltage Dependent Effective Channel Length and Series Resistance of LDD MOSFETs," *IEEE Trans. Electron Devices*, Vol. ED-34, No. 12, pp. 2469-2475, 1987.

[79] K.L. Peng, S.Y. Oh, M.A. Afromowitz and J.L. Moll, "Basic Parameter Measurement and Channel Broadening Effect in the Submicron MOSFET," *IEEE Electron Device Lett.*, Vol. EDL-5, No. 8, pp. 473-475, 1984.

[80] B.J. Sheu, C. Hu, P.K. Ko and F.-C. Hsu, "Source-and-Drain Series Resistance of LDD MOSFETs," *IEEE Electron Device Lett.*, Vol. EDL-5, No. 6, pp. 365-367, 1984.

[81] K.K. Ng and W.T. Lynch, "Analysis of the Gate-Voltage Dependent Series Resistance of MOSFETs," *IEEE Trans. Electron Devices*, Vol. ED-33, No. 6, pp. 965-972, 1986.

[82] J.A.M. Otten and F.M. Klaassen, "A Novel Technique to Determine the Gate and Drain Bias Dependent Series Resistance in Drain Engineered MOSFET's Using One Single Device," *IEEE Trans. Electron Devices*, Vol. ED-43, No. 9, pp. 1478-1488, 1996.

[83] Y.A. El-Mansy and A.R. Boothroyd, "A Simple Two-Dimensional Model for IGFET Operation in the Saturation Region," *IEEE Trans. Electron Devices*, Vol. ED-24, No. 3, pp. 254-262, 1977.

[84] S.R. Hofstein and G. Warfield, "Carrier Mobility and Current Saturation in the MOS Transistor," *IEEE Trans. Electron Devices*, Vol. ED-12, No. 3, pp. 129-138, 1965.

[85] T. Poorter, J.H. Satter and V.A. Satyadharma, "A Simple and Accurate Model for an MOS-Transistor in the Triode and Saturation Region," *Journées d'Electronique*, pp. 285-301, 1977.

[86] F.M. Klaassen and W.C.J. De Groot, "Modelling of Scaled-Down MOS Transistors," *Solid-State Electron.*, Vol. 23, No. 3, pp. 237-242, 1980.

[87] F.M. Klaassen and R.M.D.A. Velghe, "Compact Modelling of the MOSFET Drain Conductance," *in Proc. ESSDERC*, pp. 418-422, 1989.

[88] S.-L. Jang, M.-C. Hu and Y.-S. Chen, "Current-Voltage Model of Short-Channel MOSFETs Operated in the Linear Region," *Solid-State Electron.*, Vol. 38, No. 6, pp. 1239-1245, 1995.

[89] D. Sharma, J. Gautier and G. Merckel, "Negative Dynamic Resistance in MOS Devices," *IEEE J. Solid-State Circuits*, Vol. SC-13, No. 3, pp. 378-380, 1978.

[90] Y.-G. Chen, S.-Y. Ma, J.B. Kuo, Z. Yu and R.W. Dutton, "An Analytical Drain Current Model Considering Both Electron and Lattice Temperatures Simultaneously for Deep Submicron Ultra-thin SOI NMOS Devices with Self-Heating," *IEEE Trans. Electron Devices*, Vol. ED-42, No. 5, pp. 899-906, 1995.

[91] Y. Cheng and T.A. Fjeldly, "Unified Physical I-V Model Including Self-Heating Effect for Fully Depleted SOI/MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-43, No. 8, pp. 1291-1296, 1996.

[92] T. Hori, "A 0.1-μm CMOS Technology with Tilt-Implanted Punchthrough Stopper (TIPS)," *in IEDM Tech. Dig.*, pp. 75-78, 1994.

[93] B. Yu *et al.*, "Short-Channel Effect Improved by Lateral Channel-Engineering in Deep-Submicronmeter MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-44, No. 4, pp. 627-634, 1997.

[94] H. Ueno *et al.*, "Impurity-Profile-Based Threshold-Voltage Model of Pocket-Implanted MOSFETs for Circuit Simulation," *IEEE Trans. Electron Devices*, Vol. ED-49, No. 10, pp. 1783-1789, 2002.

[95] R. Rios *et al.*, "A Three-Transistor Threshold Voltage Model for Halo Processes," *in IEDM Tech. Dig.*, pp. 113-116, 2002.

[96] A. Chatterjee *et al.*, "Transistor Design Issues in Integrating Analog Functions with High Performance Digital CMOS," *in Proc. VLSI Symp.*, pp. 147-148, 1999.

[97] K.M. Cao *et al.*, "Modeling of Pocket Implanted MOSFETs for Anomalous Analog Behavior," *in IEDM Tech. Dig.*, pp. 171-174, 1999.

[98] P. Klein and S. Chladek, "A New Mobility Model for Pocket Implanted Quarter Micron *n*-MOSFETs and Below," *in Proc. ICECS2001*, Vol. 3, pp. 1587-1590, 2001.

[99] H. van Meer *et al.*, "Limitations of Shift-and-Ration Based L_{eff} Extraction Techniques for MOS Transistors with Halo or Pocket Implants," *IEEE Electron Device Lett.*, Vol. EDL-21, No. 3, pp. 133-136, 2000.

[100] M. Minondo, G. Gouget and A. Juge, "New Length Scaling of Current Gain Factor and Characterization Method for Pocket Implanted MOSFET's," *in Proc. ICMTS*, pp. 263-267, 2001.

[101] T.S. Hsieh, Y.W. Chang, W.J. Tsai and T.C. Lu, "A New Leff Extraction Approach for Devices with Pocket Implants," *in Proc. ICMTS*, pp. 15-18, 2001.

[102] S. Hong *et al.*, "Novel Direct-Tunneling-Current (DTC) Method for Channel Length Extraction beyond Sub-50nm Gate CMOS," *in IEDM Tech. Dig.*, pp. 297-300, 2001.

[103] T. Toyabe, K. Yamaguchi, S. Asai and M.S. Mock, "A Two-Dimensional Avalanche Breakdown Model of Submicron MOSFETs," *in IEDM Tech. Dig.*, pp. 432-435, 1977.

[104] S. Selberherr, *Analysis and Simulation of Semiconductor Devices,* Vienna/New York: Springer-Verlag, 1984.

[105] H.J.M. Veendrick, *Deep-Submicron CMOS ICs,* Deventer: Kluwer, 1998.

[106] J. Chen, T.Y. Chan, P.K. Ko and C. Hu, "Subbreakdown Drain Leakage Current in MOSFET," *IEEE Electron Device Lett.*, Vol. EDL-8, No. 11, pp. 515-517, 1987.

[107] T. Endoh, R. Shirota, M. Momodomi and F. Masuoka, "An Accurate Model of Subbreakdown Due to Band-to-Band Tunneling and Some Applications," *IEEE Trans. Electron Devices*, Vol. ED-37, No. 1, pp. 290-296, 1990.

[108] S.A. Parke, J.E. Moon, H.C. Wann, P.K. Ko and C. Hu, "Design for Suppression of Gate-Induced Drain Leakage in LDD MOSFET's Using a Quasi-Two-Dimensional Analytical Model," *IEEE Trans. Electron Devices*, Vol. ED-39, No. 7, pp. 1694-1703, 1992.

[109] S. Tanaka "Theory of the Drain Leakage Current in Silicon MOSFETs," *Solid-State Electron.*, Vol. 38, No. 3, pp. 683-691, 1995.

[110] A. Bouhada, S. Bakkali and A. Touhami "Modelling of Gate-Induced Drain leakage in Relation to Technological Parameters and Temperature," *Microelectron. Reliab.*, Vol. 37, No. 4, pp. 649-652, 1997.

[111] J.-C. Guo, Y.-C. Liu, M.H. Chou, M.T. Wang and F. Shone, "A Three-Terminal Band-Trap-Band Tunneling Model for Drain Engineering and Substrate Bias Effect on GIDL in MOSFET," *IEEE Trans. Electron Devices*, Vol. ED-45, No. 7, pp. 1518-1523, 1998.

[112] L. Huang, P.T. Lai, J.P. Xu and Y.C. Cheng "Mechanism Analysis of Gate-Induced Drain Leakage in Off-State n-MOSFET," *Microelectron. Reliab.*, Vol. 38, pp. 1425-1431, 1998.

[113] K.-F. You and C.-Y. Wu, "A New Quasi-2-D Model for Hot-Carrier Band-to-Band Tunneling Current," *IEEE Trans. Electron Devices*, Vol. ED-46, No. 6, pp. 1174-1179, 1999.

[114] M. Rosar, B. Leroy and G. Schweeger, "A New Model for the Description of Gate Voltage and Temperature Dependence of Gate Induced Drain Leakage (GIDL) in the Low Electric Field Region," *IEEE Trans. Electron Devices*, Vol. ED-47, No. 1, pp. 154-159, 2000.

[115] J.-H. Chen, S.-C. Wong and Y.-H. Wang, "An Analytic Three-Terminal Band-to-Band Tunneling Model on GIDL in MOSFET," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 7, pp. 1400-1405, 2001.

[116] P.J. Wright and K.C. Saraswat, "Thickness Limitations of SiO_2 Gate Dielectrics for MOS ULSI," *IEEE Trans. Electron Devices*, Vol. ED-37, No. 8, pp. 1884-1892, 1990.

[117] C.-H. Choi, K.-Y. Nam, Z. Yu and R.W. Dutton, "Impact of Gate Direct Tunneling Current on Circuit Performance: A Simulation Study," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 12, pp. 2823-2829, 2001.

[118] Z.A. Weinberg, "On Tunneling in Metal-Oxide-Silicon Structures," *J. Appl. Phys.*, Vol. 53, No. 7, pp. 5052-56, 1982.

[119] N. Garry Tarr, D.L. Pulfrey and D.S. Camporese, "An Analytic Model for the MIS Tunnel Junction," *IEEE Trans. Electron Devices*, Vol. ED-30, No. 12, pp. 1760-1770, 1983.

[120] M. Depas, R.L. Van Meirhaeghe, W.H. Laflère and F. Cardon, "Electrical Characteristics of Al/SiO_2/*n*-Si Tunnel Diodes with an Oxide Layer Grown by Rapid Thermal Oxidation," *Solid-State Electron.*, Vol. 37, No. 3, pp. 433-441, 1994.

[121] K.F. Schuegraf and C. Hu, "Reliability of Thin SiO_2," *Semicond. Sci. Technol.*, Vol. 9, pp. 989-1004, 1994.

[122] M. Depas, B. Vermeire, P.W. Mertens, R.L. Van Meirhaeghe and M.M. Heyns, "Determination of Tunnelling Parameters in Ultra-Thin Oxide Layer Poly-Si/SiO_2/Si Structures," *Solid-State Electron.*, Vol. 38, No. 8, pp. 1465-1471, 1995.

[123] S.-H. Lo, D.A. Buchanan, Y. Taur and W. Wang, "Quantum-Mechanical Modeling of Electron Tunneling Current from the Inversion Layer of Ultra-Thin-Oxide *n*MOSFET's," *IEEE Electron Device Lett.*, Vol. EDL-18, No. 5, pp. 209-211, 1997.

[124] A. Schenk and G. Heiser, "Modeling and Simulation of Tunneling through Ultra-Thin Gate Dielectrics," *J. Appl. Phys.*, Vol. 81, No. 12, pp. 7900-7908, 1997.

[125] W.-K. Shih, E.X. Wang, S. Jallepalli, F.Leon, C.M. Maziar and A.F. Tasch Jr., "Modeling Gate Leakage Current in *n*MOS Structures due to Tunneling through an Ultra-Thin Oxide," *Solid-State Electron.*, Vol. 42, No. 6, pp. 997-1006, 1998.

[126] L.F. Register, E. Rosenbaum and K. Yang, "Analytic Model for Direct Tunneling Current in Polycrystalline Silicon-Gate Metal-Oxide-Semiconductor Devices," *Appl. Phys. Lett.*, Vol. 74, No. 3, pp. 457-459, 1999.

[127] N. Yang, W.K. Henson, J.R. Hauser and J.J. Wortman, "Modeling Study of Ultrathin Gate Oxides Using Direct Tunneling Current and Capacitance-Voltage Measurements in MOS Devices," *IEEE Trans. Electron Devices*, Vol. ED-46, No. 7, pp. 1464-1471, 1999.

[128] Khairurrijal, W. Mizubayashi, S. Miyazaki and M. Hirose, "Analytic Model of Direct Tunnel Current through Ultrathin Gate Oxides," *J. Appl. Phys.*, Vol. 87, No. 6, pp. 3000-3005, 2000.

[129] N. Yang, W.K. Henson and J.J. Wortman, "A Comparative Study of Gate Direct Tunneling and Drain Leakage Currents in *n*-MOSFET's with Sub-2-nm Gate Oxides," *IEEE Trans. Electron Devices*, Vol. ED-47, No. 8, pp. 1636-1644, 2000.

[130] K.-N. Yang *et al.*, "A Physical Model for Hole Direct Tunneling Current in p^+ Poly-Gate PMOSFETs with Ultrathin Gate Oxides," *IEEE Trans. Electron Devices*, Vol. ED-47, No. 11, pp. 2161-2666, 2000.

[131] W.-C. Lee and C. Hu, "Modeling CMOS Tunneling Currents through Ultrathin Gate Oxide due to Conduction- and Valence-Band Electron and Hole Tunneling," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 7, pp. 1366-1373, 2001.

[132] P. O'Sullivan *et al.*, "Towards a Compact Model for MOSFETs with Direct Tunneling Gate Dielectrics," *in Proc. ESSDERC*, pp. 488-491, 1999.

[133] C.-H. Choi *et al.*, "Direct Tunneling Current Model for Circuit Simulation," *in IEDM Tech. Dig.*, pp.735-738, 1999.

[134] K.M. Cao *et al.*, "BSIM4 Gate Leakage Model Including Source-Drain Partition," *in IEDM Tech. Dig.*, pp.815-818, 2000.

[135] R. Clerc, P. O'Sullivan, K.G. McCarthy, G. Ghibaudo, G. Pananakakis and A. Mathewson, "A Physical Compact Model for Direct Tunneling from NMOS Inversion Layers," *Solid-State Electron.*, Vol. 45, pp. 1705-1716, 2001.

[136] B. Gu *et al.*, "A Surface-Potential-Based Compact Model of NMOSFET Gate Tunneling Current," *in Proc. Nanotech*, Vol. 2, pp. 318-321, 2003.

[137] Q. Ngo *et al.*, "Gate Current Partitioning in MOSFET Models for Circuit Simulation," *in Proc. Nanotech*, Vol. 2, pp. 322-325, 2003.

[138] A.J. Scholten, L.F. Tiemeijer, P.W.H. de Vreede and D.B.M. Klaassen, "A Large Signal Non-Quasi-Static MOS Model for RF Circuit Simulation," *in IEDM Tech. Dig.*, pp. 163-166, 1999.

[139] S.-Y. Oh, D.E. Ward and R.W. Dutton, "Transient Analysis of MOS Transistors," *IEEE J. Solid-State Circuits*, Vol. SC-15, pp. 636-643, 1980.

[140] C. Turchetti, P. Mancini and G. Masetti, "A CAD-Oriented Non-Quasi-Static Approach for the Transient Analysis of MOS IC's," *IEEE J. Solid-State Circuits*, Vol. SC-21, No. 5, pp. 827-836, 1986.

[141] H.-J. Park, P.K. Ko and C. Hu, "A Charge Conserving Non-Quasi-Static (NQS) MOSFET Model for SPICE Transient Analysis," *IEEE Trans. Computer-Aided Des.*, Vol. CAD-10, No. 5, pp. 629-642, 1991.

[142] E. Dubaois and E. Robilliart, "Efficient Non-Quasi-Static MOSFET's Model for Circuit Simulation," *in IEDM Tech. Dig.*, pp. 945-948, 1995.

[143] W. Liu, C. Bowen and M.-C. Chang, "A CAD-Compatible Non-Quasi-Static MOSFET Model," *in IEDM Tech. Dig.*, pp. 151-154, 1996.

[144] M. Bagheri and Y. Tsividis, "A Small-Signal DC-to-High-Frequency Non-Quasi-Static Model for the Four-Terminal MOSFET Valid in all Regions of Operation," *IEEE Trans. Electron Devices*, Vol. ED-32, No. 11, pp. 2383-2391, 1985.

[145] T. Smedes and F.M. Klaassen, "An Analytical Model for the Non-Quasi-Static Small-Signal Behaviour of Submicron MOSFETs," *Solid-State Electron.*, Vol. 38, No. 1, pp. 121-130, 1995.

[146] J.M. Sallese and A.-S. Porret, "Novel Approach to Charge-Based Non-Quasi-Static Model of the MOS Transistors Valid in all Modes of Operation," *Solid-State Electron.*, Vol. 44, No. 6, pp. 887-894, 2000.

[147] A.-S. Porret, J.M. Sallese and C.C. Enz, "A Compact Non-Quasi-Static Extension of a Charge-Based MOS Model," *IEEE Trans. Electron Devices*, Vol. ED-48, No. 8, pp. 1647-1654, 2001.

[148] M. Chan, K.Y. Hui, C. Hu and P.K. Ko, "A Robust and Physical BSIM3 Non-Quasi-Static Transient and AC Small-Signal Model for Circuit Simulation," *IEEE Trans. Electron Devices*, Vol. ED-45, No. 4, pp. 834-841, 1998.

[149] K.K. Hung, P.K. Ko, C. Hu and Y.C. Cheng, "A Unified Model for the Flicker Noise in Metal-Oxide-Semiconductor Field-Effect Transistors," *IEEE Trans. Electron Devices*, Vol. ED-37, No. 3, pp. 654-665, 1990.

[150] K.K. Hung, P.K. Ko, C. Hu and Y.C. Cheng, "A Physics-Based MOSFET Noise Model for Circuit Simulators," *IEEE Trans. Electron Devices*, Vol. ED-37, No. 5, pp. 1323-1333, 1990.

[151] A.J. Scholten and D.B.M. Klaassen, "New $1/f$ Noise Model in MOS Model 9, Level 903," *NL-UR 816/98*, Philips Electronics N.V., 1998. internet: www.semiconductors.philips.com/Philips_Models

[152] A. van der Ziel, *Noise: Sources, Characterization, Measurement,* Prentice-Hall, 1970.

[153] A. van der Ziel, *Noise in Measurements,* John Wiley & Sons, 1974.

[154] A.A. Abidi, "High-Frequency Noise Measurements on FET's with Small Dimensions," *IEEE Trans. Electron Devices*, Vol. ED-33, pp. 1801-1805, 1986.

[155] P. Klein, "An Analytical Thermal Noise Model of Deep Submicron MOSFET's," *IEEE Electron Device Lett.*, Vol. EDL-20, pp. 399-401, 1999.

[156] A. van der Ziel, *Noise in Solid State Devices and Circuits,* John Wiley & Sons, 1986.

[157] P.K. Ko, R.S. Muller and C. Hu, "A Unified Model for Hot-Electron Currents in MOSFETs," *in IEDM Tech. Dig.*, pp. 600-603, 1981.

[158] P.P. Guebels and F. van de Wiele, "A Small Geometry MOSFET Model for CAD Applications," *Solid-State Electron.*, Vol. 26, No. 4, pp. 267-273, 1983.

[159] K. Mayaram, J.C. Lee and C. Hu, "A Model for the Electric Field in Lightly Doped Drain Structures," *IEEE Trans. Electron Devices*, Vol. ED-34, No. 7, pp. 1509-1518, 1987.

[160] M. El-Banna and M.A. El-Nokali, "A Simple Analytical Model for Hot-Carrier MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-36, No. 5, pp. 979-986, 1989.

[161] A.T. Dejenfelt, "An Analytical Model for the Internal Electric Field in Submicrometer MOSFET's," *IEEE Trans. Electron Devices*, Vol. ED-37, No. 5, pp. 1352-1363, 1990.

[162] E.O. Kane, "Zener Tunneling in Semiconductors," *J. Phys. Chem. Solids*, Vol. 12, pp. 181-188, 1959.

[163] E.O. Kane, "Theory of Tunneling," *J. Appl. Phys.*, Vol. 32, No. 1, pp. 83-91, 1961.

[164] G.A.M. Hurkx "On the Modelling of Tunnelling Currents in Reverse-Biased p-n Junctions," *Solid-State Electron.*, Vol. 32, No. 8, pp. 665-668, 1989.

[165] G.A.M. Hurkx, D.B.M. Klaassen and M.P.G. Knuvers, "A New Recombination Model for Device Simulation Including Tunneling," *IEEE Trans. Electron Devices*, Vol. ED-39, No. 2, pp. 331-338, 1992.

[166] G.A.M. Hurkx, H.C. de Graaff, W.J. Kloosterman and M.P.G. Knuvers, "A New Analytical Diode Model Including Tunneling and Avalanche Breakdown," *IEEE Trans. Electron Devices*, Vol. ED-39, No. 9, pp. 2090-2098, 1992.

[167] F. Hurkx, H.L. Peek, J.W. Slotboom and R.A. Windgassen, "Anomalous Behavior of Surface Leakage Currents in Heavily Doped Gated-Diodes," *IEEE Trans. Electron Devices*, Vol. ED-40, No. 12, pp. 2273-2281, 1993.

[168] W.-K. Shih, R. Rios, P. Packan, K. Mistry and T. Abbott, "A General Partition Scheme for Gate Leakage Current Suitable for MOSFET Compact Models," *in IEDM Tech. Dig.*, pp. 293-296, 2001.

