CHARGE-SHEET AND NON-QUASISTATIC MOSFET MODELS FOR SPICE

by

Hong June Park

Memorandum No. UCB/ERL M89/20

24 February 1989

CHARGE-SHEET AND NON-QUASISTATIC MOSFET MODELS FOR SPICE

by

Hong June Park

Memorandum No. UCB/ERL M89/20

24 February 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

CHARGE-SHEET AND NON-QUASISTATIC MOSFET MODELS FOR SPICE

by

Hong June Park

Memorandum No. UCB/ERL M89/20

24 February 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Charge-Sheet and Non-Quasistatic MOSFET Models for SPICE

Ph.D Hong June Park Department of EECS

Ping K. Ko
Committee Chairman

ABSTRACT

Analytic equations have been derived for the charge-sheet capacitance model and the non-quasistatic MOSFET model, and both models have been implemented into SPICE3. The charge sheet capacitance model for short channel MOSFETs, is based on the surface potential formulation which uses an approximation to find surface potentials without iterations. The currents, charges and their first and second derivatives are continuous under all operating regions. An automatic gate capacitance measurement system with r.m.s. resolution of 14aF has been set up to extract model parameters. This model shows good agreements with the measured gate capacitances for long and short channel MOSFETs respectively.

The non-quasistatic MOSFET model has been derived based on the approximate solution of current continuity equation. Different sets of model equations have been derived for the transient analysis and the AC analysis, respectively. Both the channel transit time problem and the uncertainty in channel charge partitioning scheme, which are two major drawbacks of conventional quasistatic MOSFET models, have been solved using this model. Several example circuits including the channel charge injection of switched analog circuits, have been simulated using this model, and the simulation results are compared with those from other models. Good agreements have been observed among this model, PISCES, and 1-D numerical solution of current continuity equation.

ACKNOWLEDGEMENTS

First of all, I would like to appreciate the guidance and the support of my research advisor Prof. Ping K. Ko throughout this work. Without his support, much of this work would not have been possible.

I would like to thank my co-research advisor Prof. Chenming Hu for his guidance, so many kind corrections of my papers, and personal advices. I had always been impressed by his enthusiasm and thoughtfulness in guiding students.

I would like to thank Prof. Donald O. Pederson for his being the committee member of my Ph.D qualifying exam and also for the encouragement of this work. And the help from Prof. R. Sherman Lehman of Math. Department is greatly acknowledged for his being the thesis committee member and for his kind help in the Ph.D. qualifying exam.

I met many wonderful Professors in U.C.Berkeley who gave me the support and the encouragement in this work and from whom I had learned technical skills which had been the foundation of this work. I would like to express my appreciations to Prof. Paul R. Gray for technical discussions, his enthusiasm in the non-quasistaic part of this work for analog applications and proof readings of papers, to Prof. David A. Hodges for his support of this work and invaluable advices, to Prof. Robert G. Meyer for his help in my Ph.D preliminary exam, to Prof. A. Richard Newton for giving me computer accounts in CAD group machines and introducing me to the project of implementing BSIM in SPICE3, to Prof. Leon O. Chua for his teaching me in the course EECS 219 on circuit simulations, to Prof.'s Andrew R. Neureuther, Richard Muller, Richard M. White and William Oldham for their kind encouragements.

I appreciate the constant financial support from S.R.C., INTEL and California State MICRO, which made this work possible.

And I would like to express my thanks to my colleague students in U.C.Berkeley, to Thomas L. Quarles for discussions on BSIM implementation in SPICE3, to Wayne A. Christopher for discussions on NUTMEG in SPICE3, to Beomsup Kim, Yumin Lin, Joey Doernberg, Hyunjong Shin and Cormac Conroy for discussions on I.C. circuits, to Seungho Hwang for his friendship, to my office mates Minchie Jeng, Peter Lee, and Visiting Scholar Kwok K. Hung for the friendly office environments, and to all of my friends in U.C.Berkeley who made my stay in Berkeley enjoyable.

I always appreciate the unfailing support of my parents who have sacrificed their lives for my long education.

Last, but not the least, I thank my wife Hyangmee for her patience, sacrifice and love throughout this work. Also I thank my twin sons, Younghoon Andrew and Jeehoon Michael, for their patience in not playing with them often.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

As MOSFET is widely used in integrated circuits, it becomes essential to predict the circuit performance before fabrication using circuit simulation programs. Hence the device modeling work which predicts the characteristics of a single MOSFET as accurately as possible, is important. Many works have been done on MOSFET DC models so far, but considerably less works have been done on MOSFET capacitance models. One reason for this is the fact that the measurement of MOSFET capacitances is difficult especially for small geometry MOSFETs. Another reason for this is the fact that capacitance models can be fully evaluated only if they are implemented in circuit simulation programs. Compared to capacitance models, DC models are easier to implement in existing circuit simulation programs. Implementation of capacitance models in a circuit simulation program requires in-depth understanding of the program. Hence only a few capacitance models have been successfully implemented.

1.1. Charge sheet capacitance model

All the MOSFET models available in SPICE are piecewise-sectional models in which different sets of model equations are used for different operating regions. Usually, current and charges are continuous at boundaries between operating regions, but derivatives of current and charges (conductances and capacitances) are discontinuous at these boundaries. These discontinuities cause convergence problems in circuit simulations. The opposite approaches to the above mentioned piecewise-sectional models are the double integration model by Pao and Sah, and the charge sheet model by Brews. In these models, current, charges, and their derivatives with respect to bias are continuous under

all operating regions, since only one equation is used for all operating regions.

The double integration model is an accurate model at least for long channel MOSFETs, but too much computation time is required.

The charge sheet model is a simplified form of the double integration model. It assumes that the channel carriers form a infinitely thin layer of conducting plane at the surface of silicon between gate oxide and substrate. Hence, the computation time of the charge sheet model is less than that of the double integration model.

Some works have been reported on the charge sheet MOSFET models. Compared to the conventional piecewise multi-section models available in circuit simulation programs, the charge sheet model still takes much longer computation time, since many time consuming iterations are required to find surface potentials at source and drain. Hence, in spite of all the advantages of the charge sheet model, it is not widely used in real circuit simulation programs.

In this thesis, an approximation is made to find surface potentials without iterations or losing accuracy. Using this approximation, a charge sheet capacitance model for short channel MOSFETs has been derived and implemented in SPICE.

Chapter 2 describes the computation of the surface potential, the derivation of the charge sheet capacitance model, its implementation in SPICE, and the automatic gate capacitance measurement system.

1.2. Non-quasistatic MOSFET model

All the models available in circuit simulation programs such as SPICE are QS(quasistatic) models. Since these QS models do not account for the channel transit time effect, they fail for fast signals.

Most of the recent MOSFET models are charge-based models which guarantee charge conservation. In these QS charge-based models, the channel charge must be partitioned into drain and source charges to compute the drain and source currents, in transient and

AC analyses. Usually a constant channel charge partitioning ratio between drain and source charges is assumed in saturation region. Many channel charge partitioning schemes have been reported but none is valid under all operating conditions.

The NQS (non-quasistatic) MOSFET model accounts for the channel transit time effect. Also the NQS model doesn't require any constant channel charge partitioning scheme in saturation region. The model partitions the channel charge automatically following physical rules so that it is valid under all operating conditions.

This channel charge partitioning scheme used in a circuit simulation, deeply affects the simulation results of the charge injection problem of switched analog circuits. Channel charge injection is one of the major distortion sources for low-distortion switched capacitor filters, and it is one of the major bottlenecks for high-speed high-resolution MOS A/D converters.

Chapter 3 describes a NQS MOSFET model for transient analysis, where the current continuity equation is approximated into a diffusion equation and analytic equations have been derived for node currents from the solution of this diffusion equation.

Chapter 4 describes a NQS MOSFET model for AC analysis, where analytic equations for small signal node currents have been derived from the solution of the above mentioned diffusion equation.

Chapter 5 describes a charge conserving NQS MOSFET model for transient analysis, where analytic equations have been derived for node charges to guarantee the charge conservation.

Chapter 6 concludes this thesis.

Chapter 2

A CHARGE-SHEET CAPACITANCE MODEL OF SHORT CHANNEL MOSFETS FOR SPICE

2.1. Introduction

Conventional analytical charge based capacitance models are based on classical long-channel approximations that are not accurate enough for today's small geometry VLSI devices [2.1] [2.2] [2.3]. In addition, the capacitances are not continuous at the threshold voltage because of the division into the inversion and the depletion regions. This might cause numerical instabilities in transient circuit simulations. In contrast, both charge and capacitance are inherently continuous in the charge sheet model [2.4] [2.5]. This continuity property is important for the convergence property of circuit simulation programs [2.6]. The charge sheet model is a simplified form of the double integration model [2.7] by assuming the inversion layer to be a conducting plane with zero thickness. Long-channel charge-sheet capacitance models have been published in the literature[2.8], but time consuming iterations are needed to compute the surface potential thus prohibits their use in a circuit simulation environment. In this work, analytic expressions for surface potential and its derivatives with respect to bias are derived. No iterations are needed to find the surface potential. Adopting a linear approximation of the square root term for the bulk charge [2.9] [2.10], analytic expressions for current and charges including the drift velocity saturation effect and the diffusion current component, have been derived. This model has been implemented in SPICE3 and simulation results are shown.

Section 2.2 describes the schemes to find surface potentials without iterations, for long channel and short channel MOSFETs respectively.

Section 2.3 describes the derivation steps of model equations for DC drain current and node charges. The drift velocity saturation effect and the diffusion current component have been included analytically in the model equations. The channel length modulation effect and the channel side fringing field capacitance effect have been included as semi-empirical terms.

Section 2.4 describes the automatic gate capacitance measurement system which is used to extract model parameters and whose r.m.s. resolution is about 14aF.

Section 2.5 shows the comparison of this model with measured data. Comparison has been done for surface potentials, their derivatives with respect to bias, and gate capacitances of short and long channel MOSFETs.

Section 2.6 shows the comparison of the simulation results between this work implemented in SPICE3 and other capacitance models available in SPICE.

Appendix 1 shows the program to compute the surface potentials of long channel MOSFETs without iterations.

Appendix 2 shows the implementation of this model in SPICE3, for DC, transient, and AC analyses.

Appendix 3 shows the capacitance measurement program for the automatic gate capacitance measurement setup.

Appendix 4 describes the algorithm to find surface potentials of long channel MOSFETs with non-uniform substrate doping concentration, and the results are compared with measured data.

Appendix 5 shows the limiting routines used in the implementation of this model in SPICE3.

Appendix 6 shows the SPICE3 input listings for the simulation examples used in this Chapter.

Appendix 7 shows programs to compute MOSFET capacitances using the Meyer model or the Ward-Dutton model which are implemented in SPICE2.

2.2. Formulation of surface potential

2.2.1. Long channel MOSFET

Fig.2.1 shows an energy band diagram of MOS system. Ψ_S is the surface potential referenced to the bulk. We want to find the surface potential as a function of $(V_{GB}-V_{FB})$ and V_{CB}. V_{GB} is the voltage applied between gate and bulk. V_{FB} is the flat band voltage. V_{CB} is the voltage applied between channel and bulk. Expressions for charge densities can be derived as functions of Ψ_S and $(V_{GB}-V_{FB})$. Fig.2.2 shows a cross section of NMOSFET. Inside the silicon, we have the 2-D Poisson equation.

$$\frac{\partial^2\Psi(x,y)}{\partial x^2}+\frac{\partial^2\Psi(x,y)}{\partial y^2}=-\frac{q}{\varepsilon_S}(-N_A(x,y)+p(x,y)-n(x,y)) \tag{2.1}$$

where $\Psi(x,y)$ is the electro-static potential at a point (x,y), $N_A(x,y)$ is the effective ionized acceptor concentration, $p(x,y)$ and $n(x,y)$ are hole and electron concentrations respectively. $N_A(x,y)$, $p(x,y)$ and $n(x,y)$ can be represented in terms of n_i, $\Psi(x,y)$ and $V_{CB}(y)$, where n_i is the intrinsic carrier concentration.

$$N_A(x,y)=n_i\cdot e^{\Phi_F/V_t} \tag{2.2}$$

$$n(x,y)=n_i\cdot e^{(\Psi(x,y)-\Phi_F-V_{CB}(y))/V_t} \tag{2.3}$$

$$p(x,y)=n_i\cdot e^{(-\Psi(x,y)+\Phi_F)/V_t} \tag{2.4}$$

$$\textit{where}\qquad \Phi_F=V_t\cdot\ln(N_A/n_i) \tag{2.5}$$

where uniform bulk (substrate) doping concentration is assumed and V_t is the thermal voltage kT/q. $V_{CB}(y)$ is the applied bias between the channel at y and bulk.

If we assume the gradual channel approximation, that is , assuming

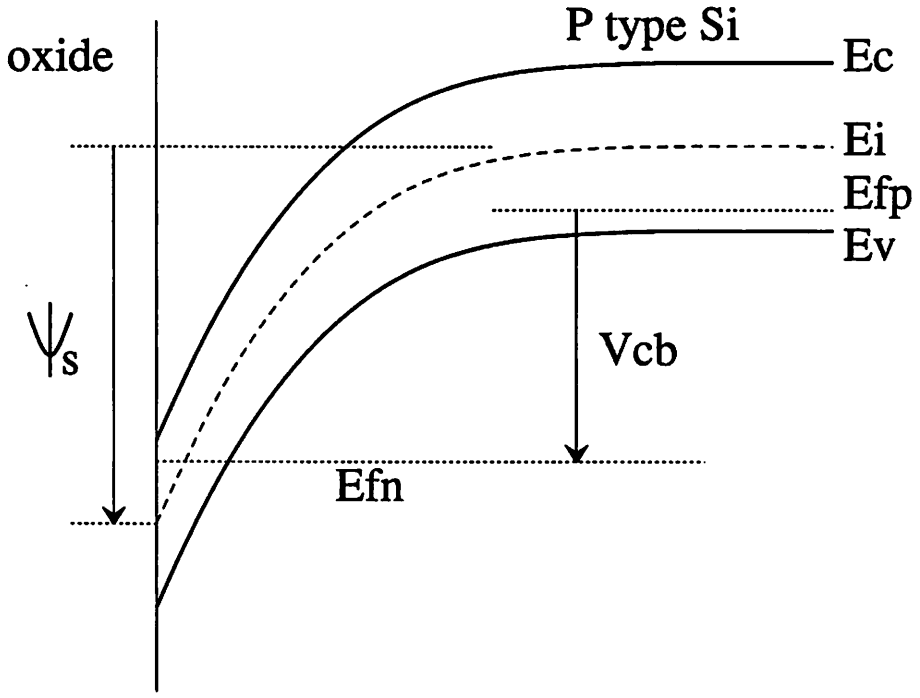


Fig.2.1. Cross section of a NMOS system. Gate is not shown for clarity. **Efn** and **Efp** are quasi Fermi levels of electrons and holes respectively. **Ec** and **Ev** represent the edges of the conduction band and the valence band respectively.

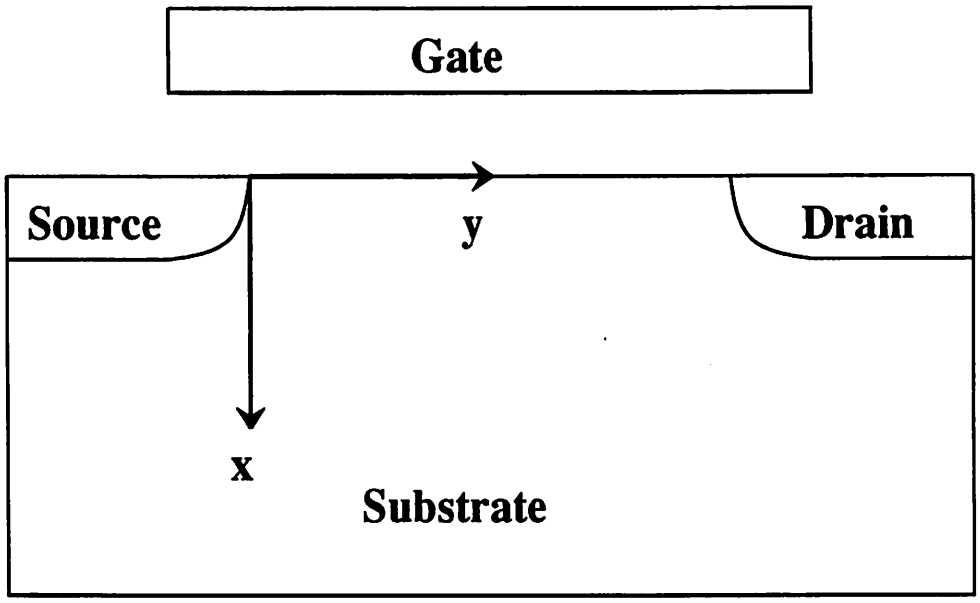


Fig.2.2. Cross section of a NMOSFET in the channel length direction. y is the lateral dimension from source (y=0) toward drain (y=L). x is the vertical dimension from the surface (x=0) toward the bulk substrate.

$$\left| \frac{\partial^2 \Psi(x,y)}{\partial x^2} \right| >> \left| \frac{\partial^2 \Psi(x,y)}{\partial y^2} \right| \tag{2.6}$$

then we can neglect the y dependence of $\Psi(x,y)$.

We now integrate both sides of (2.1) with respect to Ψ from the neutral bulk to the surface of silicon using (2.2), (2.3), (2.4), (2.5) and also the following relation (2.7). [2.11]

$$\int_0^{\Psi_S(y)} \frac{\partial^2 \Psi(x,y)}{\partial x^2} d\Psi = \int_0^{E_S(y)} \frac{\partial E(x,y)}{\partial x} \cdot E(x,y) dx = \frac{1}{2} \cdot (E_S(y))^2 \tag{2.7}$$

where $E(x,y)$ is x component of the electric field at (x,y), $E_S(y)$ and $\Psi_S(y)$ are the x component of the electric field and the surface potential at the surface of silicon respectively. The vertical (x component) electric field $E_S(y)$ can be represented in terms of V_{OX}, the voltage across gate oxide, as

$$E_S(y) = \frac{C_{OX}}{\varepsilon_S} \cdot V_{OX}(y) = \frac{C_{OX}}{\varepsilon_S} \cdot (V_{GB} - V_{FB} - \Psi_S(y)) \tag{2.8}$$

where C_{OX} is the gate oxide capacitance per unit area. Observing ($2\Phi_F >> V_t$) and ($V_{CB}(y) \geq 0$) and using (2.7) and (2.8), the integration of (2.1) yields

$$V_{GB} - V_{FB} - \Psi_S(y) =$$

$$\pm \gamma \sqrt{\Psi_S(y) - V_t + V_t \cdot e^{(\Psi_S(y) - 2\Phi_F - V_{CB}(y))/V_t} + V_t \cdot e^{-\Psi_S(y)/V_t}} \tag{2.9}$$

$$\textit{where} \qquad \gamma = \frac{\sqrt{2\varepsilon_S q N_{SUB}}}{C_{OX}} \tag{2.10}$$

The (+) sign before the square root term in (2.9) is used for $(V_{GB} - V_{FB}) \geq 0$ and the (-) sign is used for $(V_{GB} - V_{FB}) < 0$ (accumulation region). The term $(\Psi_S - V_t)$ inside the square root term in (2.9) is originated from the bulk charge and the first exponential term in (2.9) is due to the channel inversion charge and the second exponential term is due to the accumulation charge.

Some researchers tried to find the surface potential from this exact equation (eq.(2.9)) but many time consuming iterations are needed. This time consuming iteration is almost prohibitive in circuit simulation programs because, in some circuits, the surface potential

needs to be evaluated millions of times in one simulation.

To bypass this difficulty, we made an approximation to find the surface potential without iterations. Fig.2.3 shows the surface potential versus $(V_{GB}-V_{FB})$ for $V_{CB}=0$ and $V_{CB}=V1$, where $V1$ is non-negative. To derive the curve for $V_{CB}=0$, $(V_{GB}-V_{FB})$ is computed from (2.9) for given Ψ_S by simple substitution. Then the whole curve for $V_{CB}=0$ is stored as a form of cubic spline function [2.12].

The cubic spline function is a piecewise cubic polynomial and this guarantees the continuity of surface potential and its first and second derivatives with respect to applied biases. In this work the curve for $V_{CB}=0$ is divided into 150 pieces with $\Delta\Psi_S = 0.01V$ for the range of Ψ_S from -0.3V to 1.2V and the linear (natural) end condition is used [2.12] [2.13]. The memory requirement for storing the cubic spline coefficients are 600 double precision numbers per model, which is not a major concern in the modern computers.

We now consider how to find the surface potential for $V_{CB} = V1$. The curve for $V_{CB} = V1$ is divided into 3 regions which are shown in Fig.2.3.

(a) Region I (accumulation or depletion region)

($\Psi_S < \Phi_F$, that is, $(V_{GB}-V_{FB}) < (\Phi_F+\gamma\sqrt{\Phi_F-V_t})$)

This region corresponds to either accumulation or depletion region where the surface of silicon is accumulated with holes or is depleted of mobile carriers. In this region Ψ_S is independent of V_{CB} and it can be computed directly from the stored cubic spline function.

(b) Region II (depletion region)

($\Phi_F \leq \Psi_S < (\Phi_F+V_{CB})$, that is, $(\Phi_F+\gamma\sqrt{\Phi_F-V_t}) \leq (V_{GB}-V_{FB}) < (\Phi_F+V_{CB}+\gamma\sqrt{\Phi_F+V_{CB}-V_t})$)

This region corresponds to the depletion region where the surface of silicon is depleted of mobile carriers. Neglecting the two exponential terms in (2.9) which are related to inversion or accumulation carriers, we can find that (2.9) is reduced to a quadratic

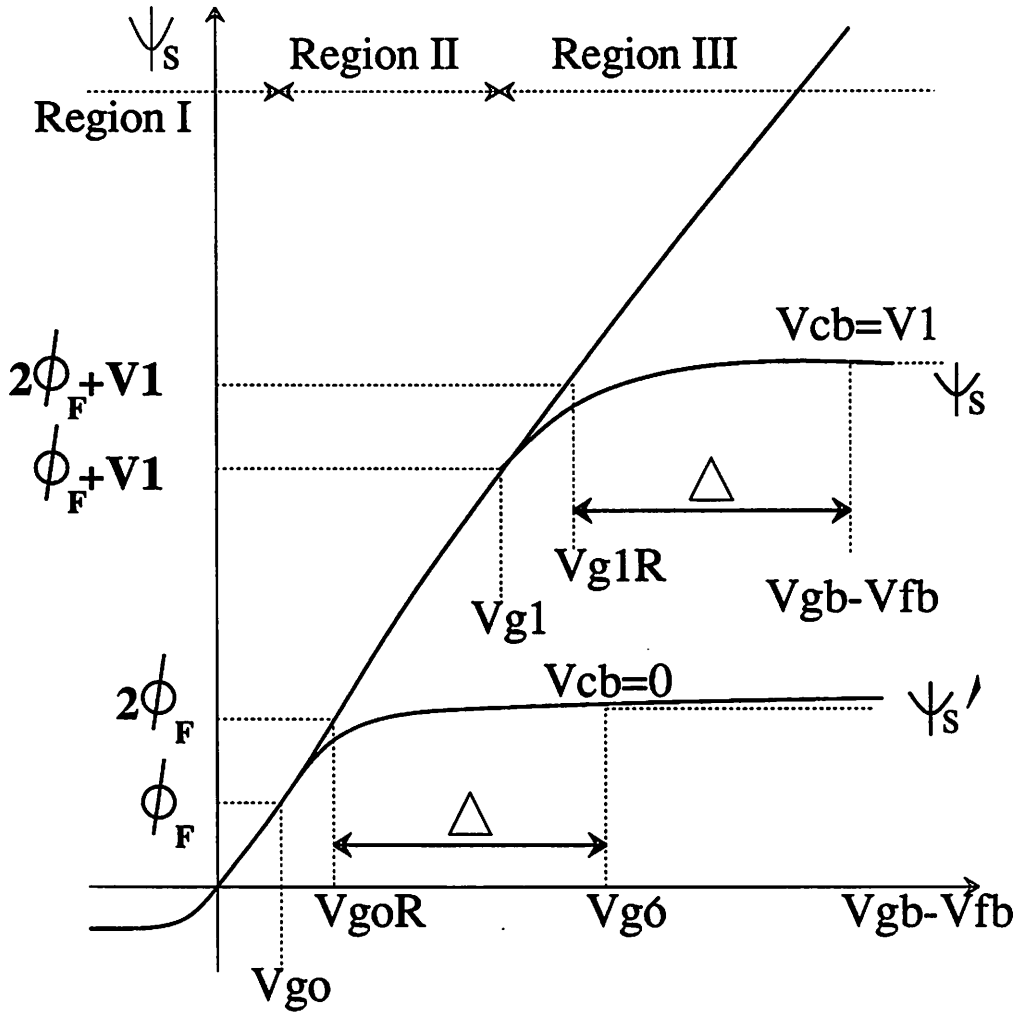


Fig.2.3. Surface potential versus $(V_{GB}-V_{FB})$ with V_{CB} as parameters. For $V_{CB}=V1$, Region I corresponds to the region of $(V_{GB}-V_{FB}) < Vgo$, Region II corresponds to the region of $Vgo \leq (V_{GB}-V_{FB}) < Vg1$, Region III corresponds to the region of $(V_{GB}-V_{FB}) \geq Vg1$. Vgo is $(V_{GB}-V_{FB})$ where the surface potential is Φ_F. $Vg1$ is $(V_{GB}-V_{FB})$ where the surface potential is $\Phi_F+V_{CB}(V1)$. $VgoR$ and $Vg1R$ are reference gate voltages $(V_{GB}-V_{FB})$ where the surface potential becomes $2\Phi_F$ and $2\Phi_F+V_{CB}$ respectively. Hence, $Vgo=\Phi_F+\gamma\sqrt{\Phi_F-V_t}$, $VgoR=2\Phi_F+\gamma\sqrt{2\Phi_F-V_t}$, $Vg1=\Phi_F+V1+\gamma\sqrt{\Phi_F+V1-V_t}$, and $Vg1R=2\Phi_F+V1+\gamma\sqrt{2\Phi_F+V1-V_t}$.

equation in Ψ_S. Hence an analytic equation for Ψ_S can be derived as

$$\Psi_S = V_{GB} - V_{FB} + 0.5\gamma^2 - \gamma\sqrt{V_{GB}-V_{FB}+0.25\gamma^2-V_t} \tag{2.11}$$

(c) Region III (inversion region)

($\Psi_S \geq (\Phi_F+V_{CB})$, that is, $(V_{GB}-V_{FB}) \geq (\Phi_F+V_{CB}+\gamma\sqrt{\Phi_F+V_{CB}-V_t})$)

This region corresponds to the inversion region where the surface of silicon is inverted with electrons. In this region, Ψ_S can be computed from (2.12) if the normalized inversion charge density Q_n' is known.

$$Q_n' = V_{GB} - V_{FB} - \Psi_S - \gamma\sqrt{\Psi_S-V_t} \tag{2.12}$$

For $V_{CB} = 0$, Q_n' can be computed from the stored cubic spline function. For non-zero V_{CB}, Q_n' can be found if the same inversion charge density Q_n' is assumed for the same Δ as shown in Fig.2.3. Δ is the displacement in gate voltage from the reference point where $(V_{GB}-V_{FB}) = 2\Phi_F + V_{CB} + \gamma\sqrt{2\Phi_F-V_t+V_{CB}}$ for each V_{CB}. Hence,

$$\Delta = V_{GB} - V_{FB} - 2\Phi_F - V_{CB} - \gamma\sqrt{2\Phi_F-V_t+V_{CB}} \tag{2.13.a}$$

Once Δ is found from (2.13.a), from the curve for some non-zero $V_{CB} = V1$, Q_n' can be computed for the same Δ in the (V_{CB}=0) curve using the cubic spline function. Then Ψ_S can be computed from (2.12) as

$$\Psi_S = V_{GB}-V_{FB}-Q_n'+0.5\gamma^2-\gamma\sqrt{V_{GB}-V_{FB}-Q_n'+0.25\gamma^2-V_t} \tag{2.13.b}$$

The algorithm for finding Ψ_S is summarized in Fig.2.4. The program for computing Ψ_S is shown in Appendix 1.

---

if $(V_{GB}-V_{FB}) < (\Phi_F+\gamma\sqrt{\Phi_F-V_t})$

  $\Psi_S = CSF(V_{GB}-V_{FB})$

else if $(V_{GB}-V_{FB}) < (\Phi_F+V_{CB}+\gamma\sqrt{\Phi_F+V_{CB}-V_t})$

  $\Psi_S = V_{GB}-V_{FB}+0.5\gamma^2-\gamma\sqrt{V_{GB}-V_{FB}+0.25\gamma^2-V_t}$

else

$$(V_{GB}-V_{FB})' = V_{GB}-V_{FB}-V_{CB}-\gamma\sqrt{2\Phi_F+V_{CB}-V_t}+\gamma\sqrt{2\Phi_F-V_t}$$

$$\Psi_S' = CSF((V_{GB}-V_{FB})')$$

$$Q_n' = (V_{GB}-V_{FB})'-\Psi_S'-\gamma\sqrt{\Psi_S'-V_t}$$

$$\Psi_S = V_{GB}-V_{FB}-Q_n'+0.5\gamma^2-\gamma\sqrt{V_{GB}-V_{FB}-Q_n'+0.25\gamma'-V_t}$$

Fig.2.4. The algorithm for finding the surface potential Ψ_S as a function of $(V_{GB}-V_{FB})$ and V_{CB} for long channel MOSFETs.

Fig.2.5.(a) shows the approximated and exact values of the surface potential versus $(V_{GB}-V_{FB})$ with V_{CB} as parameters. The approximated values are computed using the method described in Fig.2.4 and the exact values are directly computed from (2.9). In Fig.2.5.(a), T_{OX} is 400 *Aungstrom* and N_{SUB} is $2.5*10^{16}\ cm^{-3}$. Fig.2.5.(b) shows the normalized inversion charge density Q_n' versus $(V_{GB}-V_{FB})$. The same T_{OX} and N_{SUB} as in Fig.2.5.(a) are used. Good agreements can be observed except a slight mismatch in subthreshold slopes for non-zero V_{CB}.

For long channel MOSFETs, Ψ_S for drain and source ends can be computed by taking $V_{CB} = V_{DB}$ and $V_{CB} = V_{SB}$ respectively, where V_{DB} is the applied drain-to-bulk voltage and V_{SB} is the applied source-to-bulk voltage.

2.2.2. Short channel MOSFET

For short channel MOSFETs, the surface potential at the source end, Ψ_{SO}, is computed using the schemes discussed for the long channel case. The surface potential at the drain end, Ψ_{SL}, is computed from V_{DS} and V_{DSSAT}, where V_{DS} is the applied drain-to-source voltage and V_{DSSAT} is the drain saturation voltage. Hence Ψ_{SL} can be written as

$$\Psi_{SL} = \Psi_{SO} + VDS \tag{2.14}$$

where VDS is the effective drain to source voltage. Conceptually $VDS = V_{DS}$ in linear region and $VDS = V_{DSSAT}$ in saturation region. To make the smooth transition for VDS between linear and saturation regions, a smooth function which we call 'Saturation

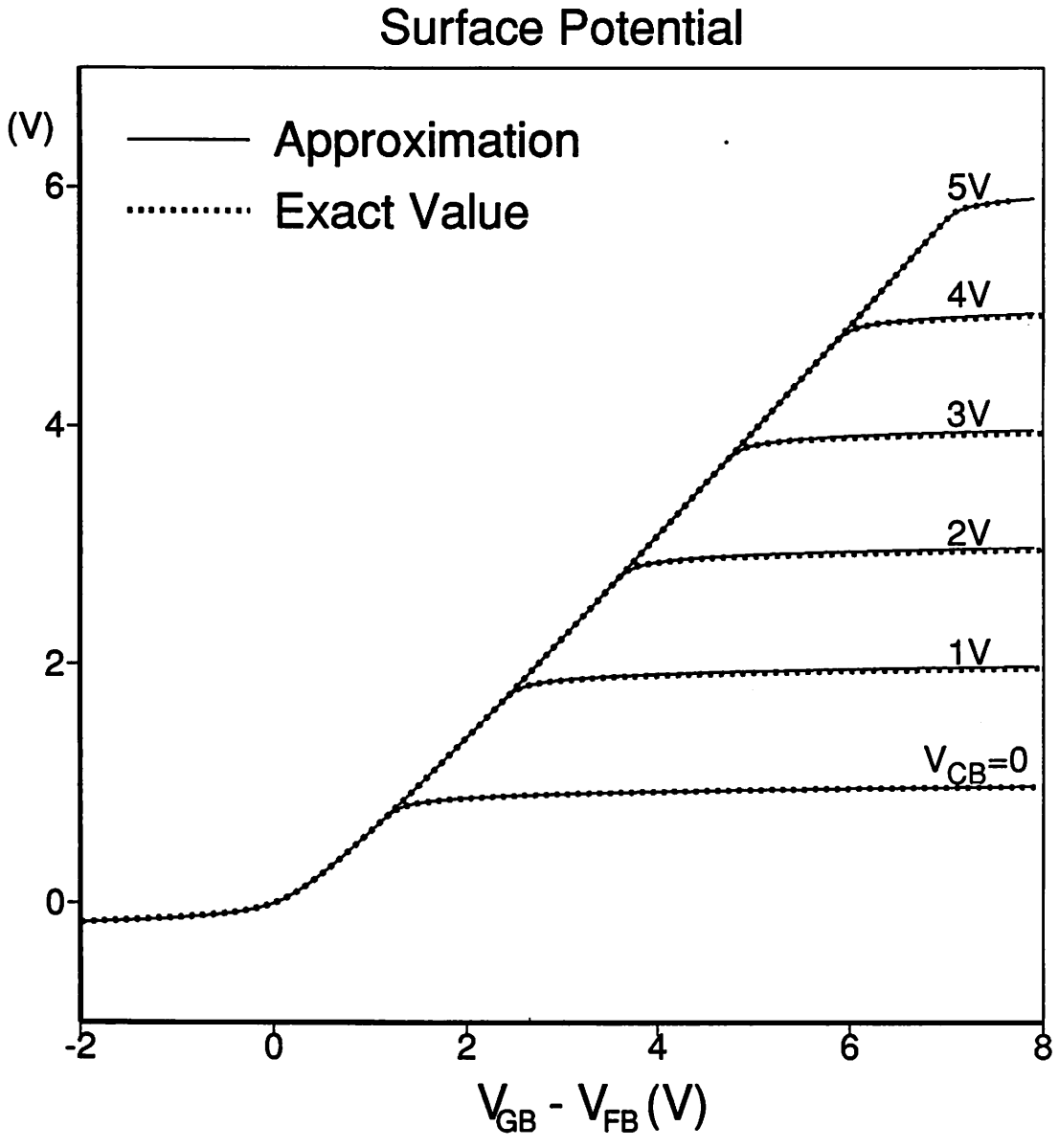


Fig.2.5.(a). Comparison of the surface potential, between the values computed using the approximation schems in Fig.2.4, and the exact values computed directly from (2.9).

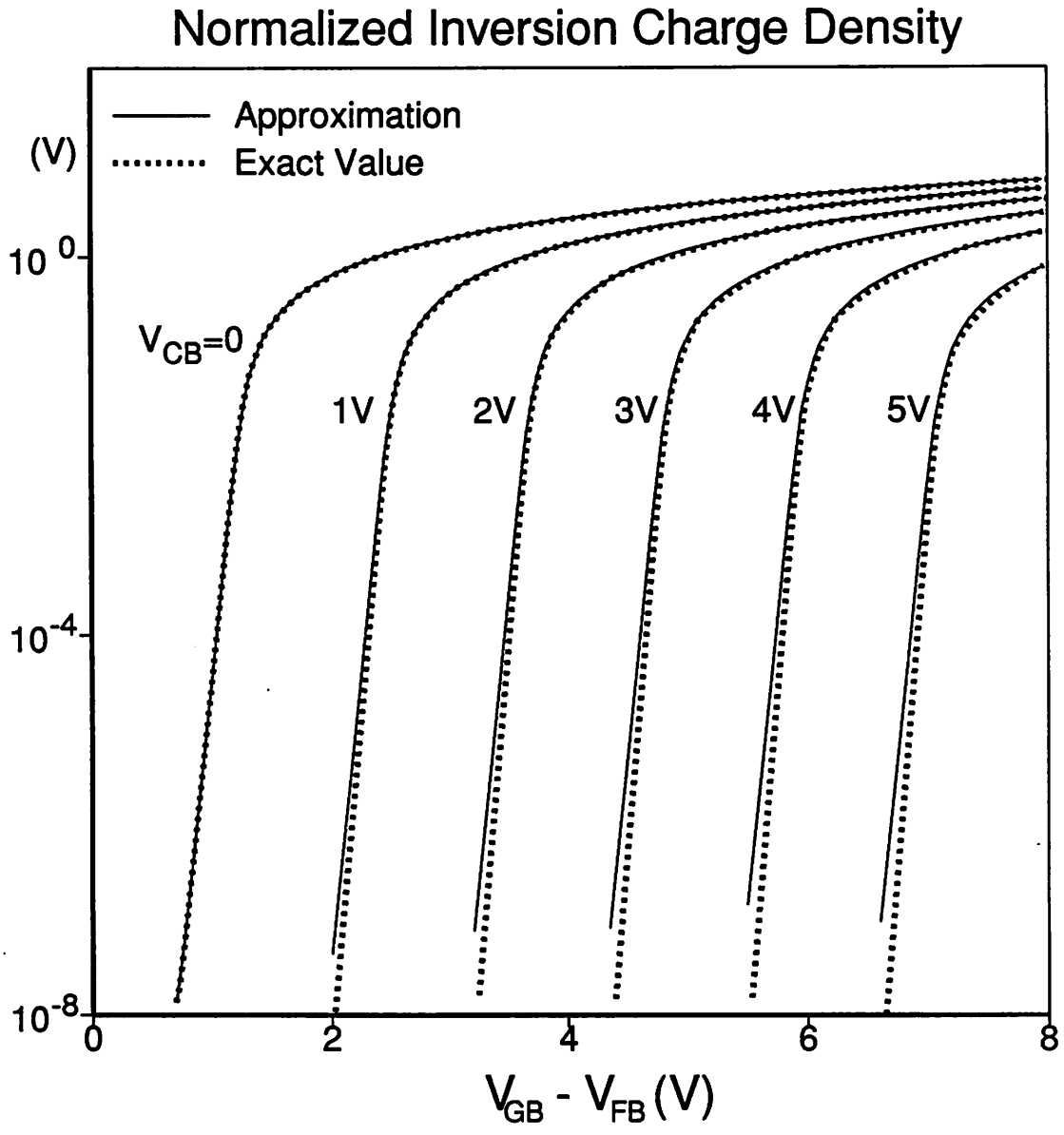


Fig.2.5.(b). Comparison of the inversion charge density, between the values computed using the approximation schemes in Fig.2.4, and the exact values computed directly from (2.9).

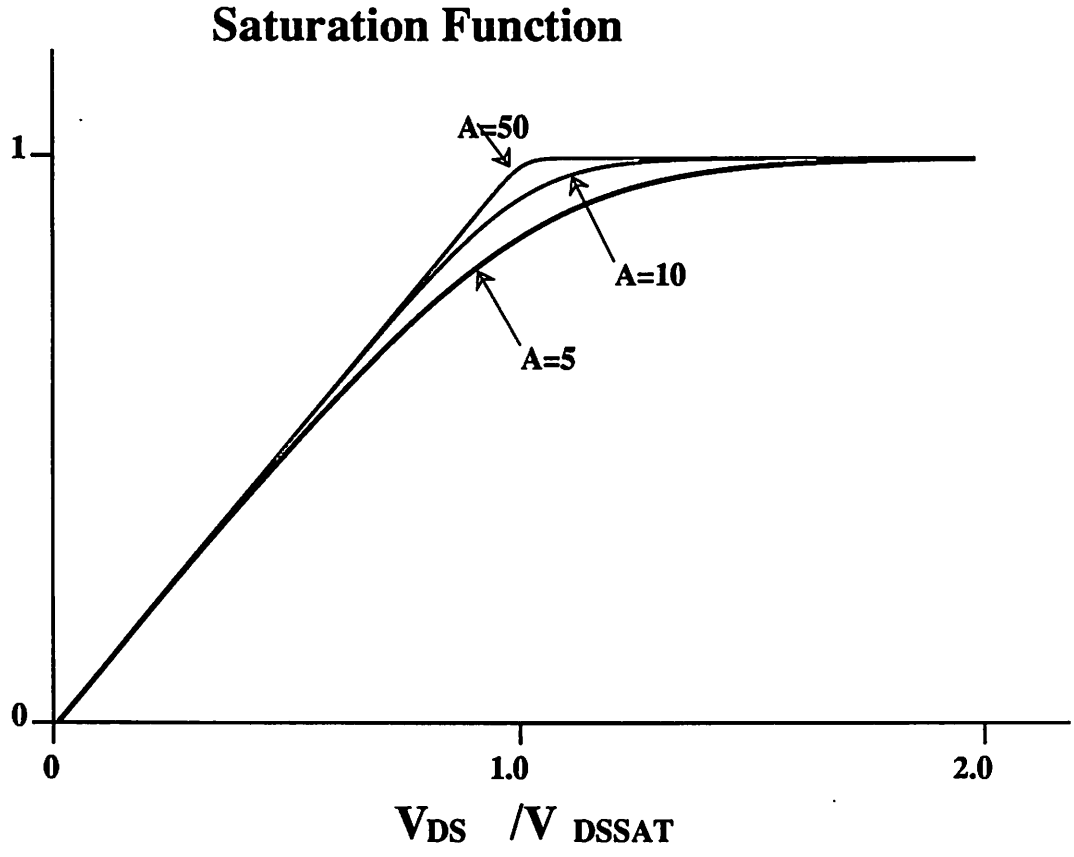


Fig.2.6. The saturation function in eq. (2.15) for different values of A.

Function' is devised as.

$$Saturation\_Function(z) = 1 - \frac{\log_e (1 + e^{A(1-z)})}{A} \tag{2.15}$$

Fig.2.6 shows the saturation function with different values of A. Large A gives a steep transition between linear and saturation regions and small A gives a smooth transition. This parameter A is set to be a model parameter in the implementation and its default value is 40. But the computation of (2.15) is costly in computer time because it includes time consuming exponential and logarithmic functions. For the computational efficiency, another cubic spline function $SF(\ )$ is derived from (2.15). The saturation function (2.15) is computed only once in the SPICE setup stage to compute the cubic spline coefficients of $SF(\ )$, and $SF(\ )$ is repeatedly used for the model computation.

Hence the effective drain to source voltage VDS is computed from

$$VDS = SF\left[\frac{V_{DS}}{V_{DSSAT}}\right] \cdot V_{DSSAT} \cdot \frac{\partial \Psi_{SO}}{\partial V_{CB}} \tag{2.16}$$

$\partial \Psi_{SO} / \partial V_{CB}$ is added in (2.16) to guarantee the smooth transition between strong inversion and cutoff region. In the strong inversion region $\partial \Psi_{SO} / \partial V_{CB}$ is 1.0 and it is 0.0 in cutoff region and it changes smoothly between the two regions.

2.3. Derivation of current and charge equations

The derivation steps for the model equations are shown here and the implementation details of this model in SPICE3 is shown in Appendix 2.

2.3.1. Bulk charge approximation

In the inversion and depletion regions where $\Psi_S > \Phi_F$, the normalized bulk charge density $q_B'(y)$ can be represented in terms of surface potential Ψ_S as

$$Q_B'(y) = \gamma \sqrt{\Psi_S(y) - V_t} \tag{2.17}$$

To find simple analytic charge equations which include the velocity saturation effect, we

need a linear relationship between $Q_B'(y)$ and $\Psi_S(y)$ [2.9] [2.10]. For this purpose, the square root term in (2.17) is approximated as

$$\sqrt{\Psi_S(y)-V_t} = \sqrt{(\Psi_{SO}-V_t) + (\Psi_S(y)-\Psi_{SO})}$$

$$\approx \sqrt{\Psi_{SO}-V_t} + \frac{0.5g}{\sqrt{\Psi_{SO}-V_t}}\cdot(\Psi_S(y)-V_t) \tag{2.18}$$

$$\textit{where} \quad g = 1 - \frac{1}{1.744+0.8364\cdot(\Psi_{SO}-V_t)} \tag{2.19}$$

This approximation was introduced by H.C.Poon et. al [2.10] and was used also in BSIM [2.14]. Using the approximation (2.18), the normalized inversion charge density $Q_n'(y)$ for $(\Psi_{SO} > \Phi_F)$ can be represented as

$$Q_n'(y) = V_{GB}-V_{FB}-\Psi_S(y) - \gamma\sqrt{\Psi_S(y)-V_t}$$

$$= V_{GST} - F_B\cdot(\Psi_S(y) - \Psi_{SO}) \tag{2.20}$$

$$\textit{where} \quad V_{GST} = V_{GB} - V_{FB} - \Psi_{SO} - \gamma\sqrt{\Psi_{SO}-V_t} \tag{2.21}$$

$$F_B = 1 + \frac{0.5\ g\ \gamma}{\sqrt{\Psi_{SO}-V_t}} \tag{2.22}$$

where F_B has a value which is slightly larger than 1.0 and is usually between 1.1 and 1.3. In (2.18), the two terms $(\Psi_{SO}-V_t)$ and $(\Psi_S(y)-\Psi_{SO})$ are symmetric before the approximation but they are no longer symmetric after the approximation. This causes an asymmetry problem between source and drain capacitances when V_{DS} becomes 0, but the overall effect of this asymmetry on circuit simulation results is believed to be negligible.

2.3.2. Current equations

The DC drain current I_D can be written as the sum of drift and diffusion components, following the charge sheet formulation [2.4].

$$I_D = WC_{OX}\cdot\mu_n(y)\cdot\left[Q_n'(y)\cdot\frac{d\Psi_S(y)}{dy} - V_t\cdot\frac{dQ_n'(y)}{dy}\right] \tag{2.23}$$

The Einstein relation $(D = \mu\cdot V_t)$ is used in this derivation. The velocity saturation

effect on the mobility is modeled as

$$\mu_n(y) = \frac{\mu_{no}}{1 + \mu_{no} \cdot \dfrac{\dfrac{d\Psi_S(y)}{dy}}{v_{SAT}}} \tag{2.24}$$

where μ_{no} is the mobility at a given gate voltage and v_{SAT} is the carrier saturation velocity. Substituting (2.24) into (2.23) and integrating both sides of the resulting equation from source to drain, we can derive an analytic equation for DC drain current I_D. Including the empirical channel length modulation factor λ, we can get the final equation for I_D as

$$I_D = \mu_{no} C_{OX} \frac{W}{L} \cdot \frac{1 + \lambda V_{DS}}{1 + \dfrac{\mu_{no}(\Psi_{SL} - \Psi_{SO})}{v_{SAT} L}} \cdot$$

$$(V_{GST} + F_B V_t - 0.5 \cdot F_B \cdot (\Psi_{SL} - \Psi_{SO})) \cdot (\Psi_{SL} - \Psi_{SO}) \tag{2.25}$$

where $$\mu_{no} = \frac{\mu_o}{1 + \dfrac{E_{EFF}}{E_{CRIT}}} \tag{2.26}$$

$$E_{EFF} = \frac{C_{OX}}{\varepsilon_S} \cdot (0.5 \cdot Q_n'(0) + Q_B'(0))$$

$$= \frac{C_{OX}}{\varepsilon_S} \cdot (0.5 \cdot V_{GST} + \gamma \sqrt{\Psi_{SO} - V_t}) \tag{2.27}$$

$$\lambda = \lambda_o + \lambda_D \cdot V_{DS} + \lambda_B \cdot V_{BS} \tag{2.28}$$

μ_o is the zero field mobility. E_{EFF} is the effective vertical electric field at the surface of silicon and it is computed following the schemes in [2.15].

E_{CRIT}, λ_o, λ_D and λ_B can be specified in SPICE input files as model parameters.

2.3.3. Drain saturation voltage V_{DSSAT}

To derive the equation for V_{DSSAT}, (2.23) is converted to

$$I_D = WC_{OX} \cdot v_{DRIFT}(y) \cdot (Q_n'(y) - V_t \cdot \frac{dQ_n'(y)}{d\Psi_S}) \tag{2.29}$$

$$\textit{where} \quad v_{DRIFT}(y) = \mu_n(y) \cdot \frac{d\Psi_S(y)}{dy} \tag{2.30}$$

Substituting ($y=L$) into (2.29) and using (2.20), we can find

$$I_D = WC_{OX} \cdot v_{DRIFT}(L) \cdot (V_{GST} + F_B V_t - F_B \cdot (\Psi_{SL} - \Psi_{SO})) \tag{2.31}$$

where L is the effective channel length.

The drain saturation voltage V_{DSSAT} is defined as

$$V_{DSSAT} = (\Psi_{SL} - \Psi_{SO})|_{v_{DRIFT}(L) = v_{SAT}} \tag{2.32}$$

Conceptually (2.32) implies that V_{DSSAT} is the drain-to-source voltage V_{DS} at which the drift velocity at the drain end of the channel reaches the saturation velocity v_{SAT}.

From (2.25), (2.31) and (2.32) we can find V_{DSSAT} as

$$V_{DSSAT} = -\frac{v_{SAT}L}{\mu_{no}} + \sqrt{\left[\frac{v_{SAT}L}{\mu_{no}}\right]^2 + 2 \cdot \frac{v_{SAT}L}{\mu_{no}} \cdot \frac{(V_{GST} + F_B V_t)}{F_B}} \tag{2.33}$$

In this derivation, the empirical channel length modulation effect is not included. For the very long channel MOSFETs, V_{DSSAT} in (2.33) is reduced to $(V_{GST}/F_B + V_t)$. This agrees with the SPICE level-3 model [2.16] except the thermal voltage term which is due to the diffusion component. Since V_{GST} is non-negative, V_{DSSAT} is always larger than or equal to V_t.

2.3.4. Charge equations

In circuit simulations for MOSFET, we need charge equations associated to each node, in order to get the transient or small-signal node currents. For this purpose, the operating regions are divided into two regions.

Region I: where $\Psi_{SO} \geq \Phi_F$

Region II: where $\Psi_{SO} < \Phi_F$

Region I corresponds to either inversion or depletion region and Region II corresponds to accumulation or depletion region.

(a) charge equations for $\Psi_{SO} \geq \Phi_F$ (Region I)

In this region, expressions for node charges can be derived from the following equations.

$$Q_G = WC_{OX} \cdot \int_0^L (V_{GB} - V_{FB} - \Psi_S(y))\, dy \tag{2.34}$$

$$Q_B = -WC_{OX} \cdot \int_0^L \gamma \sqrt{\Psi_S(y) - V_t}\, dy \tag{2.35}$$

$$Q_D = -WC_{OX} \cdot \int_0^L Q_n'(y) \cdot \frac{y}{L}\, dy \tag{2.36}$$

$$Q_S = -(Q_G + Q_B + Q_D) \tag{2.37}$$

where Q_G, Q_B, Q_D and Q_S are gate, bulk, drain and source charges respectively. To get Q_D in (2.36), we followed the Oh, Ward and Dutton's channel charge partitioning scheme [2.17] which can be derived analytically from the current continuity equation as shown in Appendix 10 and gives the 40/60 channel charge partitioning between drain and source charges in saturation region.

To derive analytic equations for node charges, we represent Ψ_S in terms of y. For this purpose, (2.23) is converted to

$$I_D\, dy = \mu_n WC_{OX} \cdot \left(Q_n'(y) - V_t \cdot \frac{dQ_n'(y)}{d\Psi_S} \right) \cdot d\Psi_S \tag{2.38}$$

Substituting (2.20) into (2.38) and integrating the resulting equation from source to some channel point, at which the lateral displacement from source is y and the surface potential is $\Psi_S(y)$, we can find

$$\Psi_S(y) - \Psi_{SO} = \frac{1}{F_B} \cdot \left[V_O - \sqrt{V_O^2 - 2F_B I_D' \cdot \frac{y}{L}} \right] \tag{2.39}$$

$$\textit{where} \quad V_O = V_{GST} + F_B V_t - \frac{\mu_{no} I_D'}{v_{SAT} L} \tag{2.40}$$

$$I_D' = \frac{I_D}{\mu_{no} \cdot \frac{W}{L} C_{OX} \cdot (1+\lambda V_{DS})} \tag{2.41}$$

I_D' is the normalized drain current which does not include the empirical channel length modulation effect. I_D' has a unit of $[V^2]$.

Substituting (2.39) into (2.34), we can find the gate charge Q_G as

$$Q_G = WLC_{OX} \cdot \left\{ V_{GB} - V_{FB} - \Psi_{SO} - \frac{V_O}{F_B} + \frac{V_O^3 - V_L^3}{3F_B^2 I_D'} \right\} \tag{2.42}$$

where $$V_L = \sqrt{V_O^2 - 2F_B I_D'} \tag{2.43}$$

Using the linear approximation of the square root term as shown in (2.18), the bulk charge Q_B in (2.35) can be re-written as

$$Q_B = -WC_{OX} \cdot \int_0^L \gamma\sqrt{\Psi_{SO} - V_t} + (F_B - 1) \cdot (\Psi_S(y) - \Psi_{SO})\ dy \tag{2.44}$$

Substituting (2.39) into (2.44), we can find

$$Q_B = -WLC_{OX} \cdot \left\{ \gamma\sqrt{\Psi_{SO} - V_t} + (1 - \frac{1}{F_B}) \cdot (V_O - \frac{V_O^3 - V_L^3}{3F_B I_D'}) \right\} \tag{2.45}$$

The drain charge Q_D can be derived by substituting (2.20) and (2.39) into (2.36), as

$$Q_D = -WLC_{OX} \cdot \left\{ \frac{\mu_{no} I_D'}{2v_{SAT} L} - \frac{F_B V_t}{2} + \frac{\frac{2}{3} V_O^2 \cdot (V_O^3 - V_L^3) - \frac{2}{5}(V_O^5 - V_L^5)}{(2F_B I_D')^2} \right\} \tag{2.46}$$

The source charge Q_S can be computed from the other three components as shown in (2.37). In the charge equations (2.42), (2.45) and (2.46), when the normalized drain current I_D' goes to 0, both nominator and denominator go to zero in some terms. Hence for very small I_D' where $(\Psi_{SL} - \Psi_{SO}) \approx 0$, the charge equations are approximated into their asymptotic forms as

$$Q_G = WLC_{OX} \cdot \left\{ V_{GB} - V_{FB} - 0.5 \cdot (\Psi_{SL} + \Psi_{SO}) \right\} \tag{2.47}$$

$$Q_B = -WLC_{OX}\cdot\left\{\gamma\sqrt{\Psi_{SO}-V_t} + 0.5\cdot(F_B-1)\cdot(\Psi_{SL}-\Psi_{SO})\right\} \tag{2.48}$$

$$Q_D = -WLC_{OX}\cdot\left\{0.5\cdot V_{GST} - \frac{F_B\cdot(\Psi_{SL}-\Psi_{SO})}{3}\right\} \tag{2.49}$$

$$Q_S = -(Q_G + Q_B + Q_D) \tag{2.50}$$

(b) Charge equations for $\Psi_{SO} < \Phi_F$ (Region II)

In this operating region, the equation (2.20) for the normalized inversion charge density Q_n' is not valid and so the previously derived equations (2.42), (2.45) and (2.46) can not be used. In this region, $\Psi_S(y)=\Psi_{SO}$ and $Q_n'(y)=0$ for all y $(0 \le y \le L)$. Hence,

$$Q_G = WLC_{OX}\cdot(V_{GB} - V_{FB} - \Psi_{SO}) \tag{2.51}$$

$$Q_B = -Q_G \tag{2.52}$$

$$Q_D = 0 \tag{2.53}$$

$$Q_S = 0 \tag{2.54}$$

2.3.5. Channel length modulation for charge components

For short channel MOSFETs, the total inversion charge changes appreciably as the effective channel length varies, as shown in Fig.2.7. By substituting (2.39) into (2.20), one can find the channel charge density profile $Q_n'(y)$ as

$$Q_n'(y) = \frac{I_D' L}{\mu_{no} v_{SAT}} - F_B V_t + \sqrt{V_O^2 - 2F_B I_D'\cdot\frac{y}{L}} \tag{2.55}$$

When the effective channel length L is changed to $(L-\Delta L)$, the profile of the difference can be approximated as

$$\Delta Q_n'(y) = Q_n'(y)|_{(L-\Delta L)} - Q_n'(y)|_L = -\frac{\partial Q_n'(y)}{\partial L}\cdot\Delta L = \frac{\partial Q_n'(y)}{\partial y}\cdot y\cdot\frac{\Delta L}{L} \tag{2.56}$$

In this derivation, the velocity saturation effect is not included for simplicity.

Using (2.56), one can derive the change of each node charge as

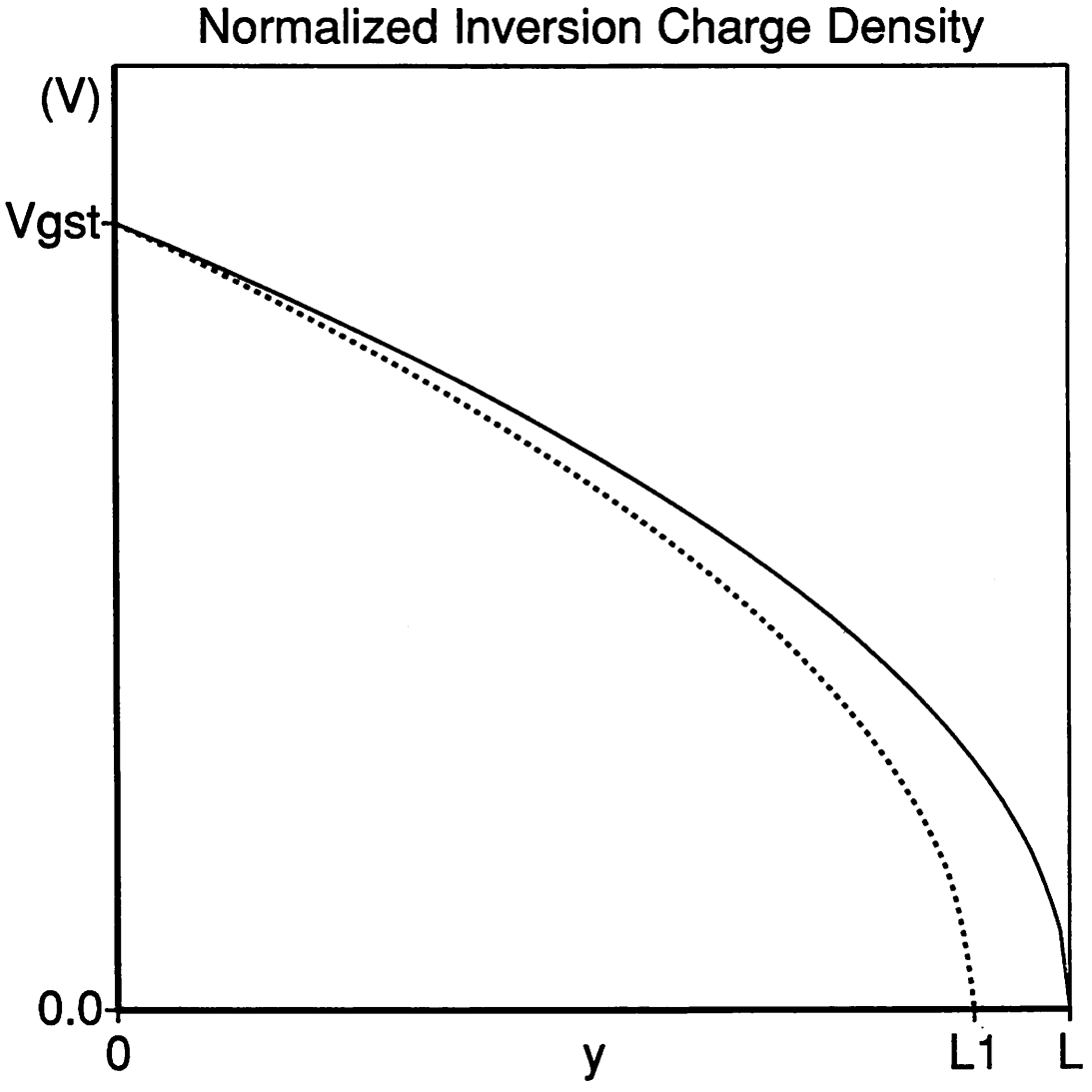


Fig.2.7. Change of the inversion charge density profile in the saturation region due to the channel length modulation. The y dependence of the profile in the saturation region can be represented as $\sqrt{1-y/L_{EFF}}$.

$$\Delta Q_D = - WC_{OX} \cdot \int_0^L \Delta Q_n'(y) \cdot \frac{y}{L}\, dy = - 2 \cdot Q_D \cdot \frac{\Delta L}{L} \tag{2.57}$$

$$\Delta Q_S = - WC_{OX} \cdot \int_0^L \Delta Q_n'(y) \cdot (1-\frac{y}{L})\, dy = - (Q_S - Q_D) \cdot \frac{\Delta L}{L} \tag{2.58}$$

Assuming $\Delta Q_B = 0$ for simplicity, then we have

$$\Delta Q_G = - (\Delta Q_D + \Delta Q_S) = (Q_D + Q_S) \cdot \frac{\Delta L}{L} \tag{2.59}$$

The change of the effective channel length ΔL for channel length modulation of charges, is modeled as

$$\Delta L = \lambda_Q \cdot (V_{DS} - VDS) \tag{2.60}$$

where λ_Q is a model parameter for the channel length modulation of charges. This λ_Q can be specified in SPICE input, independently of λ for the channel length modulation of DC drain current shown in (2.25). Conceptually the term $(V_{DS} - VDS)$ in (2.60) becomes zero in linear region and becomes $(V_{DS} - V_{DSSAT})$ in saturation region.

2.3.6. Channel side fringing field capacitance

For short channel MOSFETs, the channel side stray capacitances have significant contribution to the total gate capacitance [2.18]. Channel side stray capacitances have three components $C1$, $C2$ and $C3$ as shown in Fig.2.8. $C1$ is the outer fringing capacitance between gate and source or drain electrode. $C2$ is the direct overlap capacitance between gate and source or drain junction. $C3$ is the channel side fringing field capacitance between gate and the side wall of the source or drain junction. $C3$ becomes non-zero when the channel is depleted of mobile carriers and so some electric fields originating from gate is terminated at the side wall of source or drain junction.

For $C1$, $C2$ and $C3$, we followed the equations in [2.18].

$$C1 = W \cdot \frac{\varepsilon_{OX}}{\alpha} \cdot \ln(1 + \frac{T_{GATE}}{T_{OX}}) \tag{2.61}$$

$$C2 = W \cdot \frac{\varepsilon_{OX}}{T_{OX}} \cdot (LD + \Delta LD) \tag{2.62}$$

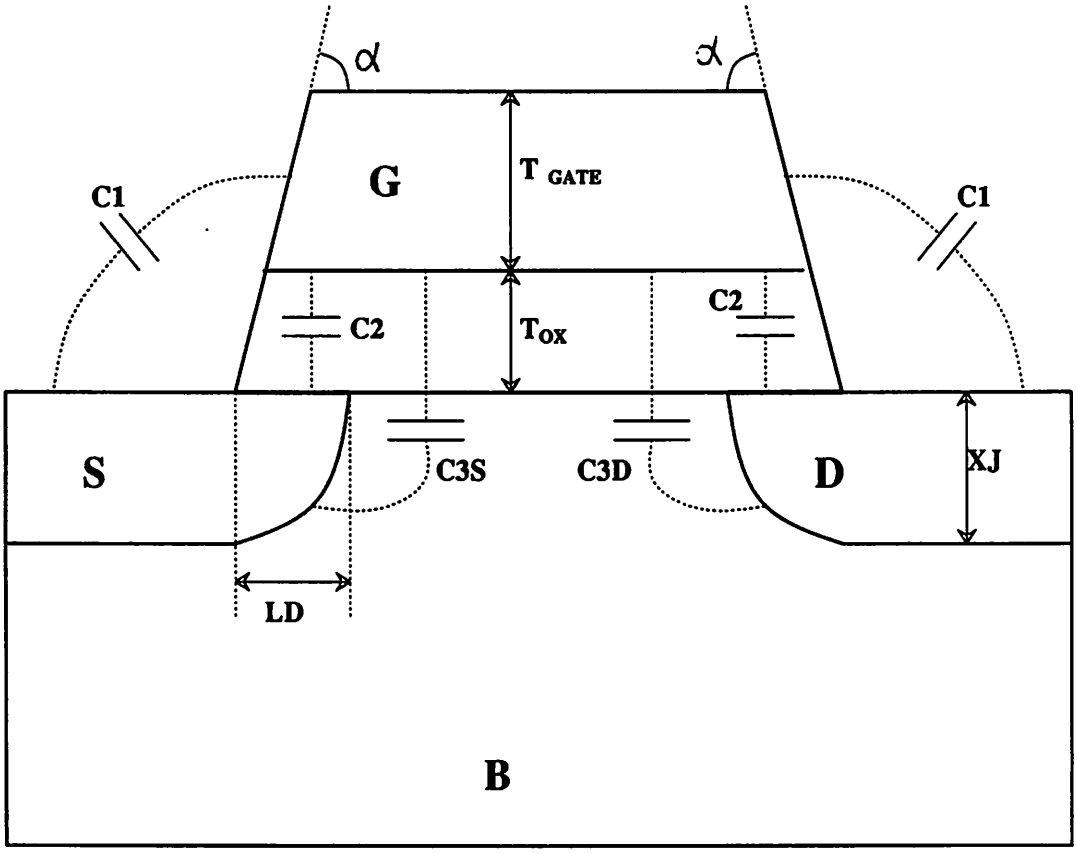


Fig.2.8. Cross section of a NMOSFET to show the three components of the channel side stray capacitances, $C1$, $C2$ and $C3$ [2.18].

$$\textit{where} \quad \Delta LD = 0.5 \cdot T_{OX} \cdot \left(\frac{1 - \cos\alpha}{\sin\alpha} + \frac{1 - \cos\delta}{\sin\delta} \right) \tag{2.63}$$

$$\delta = 0.5 \cdot \pi \cdot \frac{\varepsilon_{OX}}{\varepsilon_S} \tag{2.64}$$

where α is the slanting angle of gate electrode, T_{GATE} is the thickness of the gate electrode, W is the channel width, and LD is the metallurgical lateral diffusion of source, drain junction as shown in Fig.2.8. ε_{OX} and ε_S are the dielectric constants of oxide and silicon respectively. C_F is the maximum value of the channel side fringing capacitance component $C3$.

$$C_F = W \cdot \frac{\varepsilon_{OX}}{\delta} \cdot \log_e \left(1 + \frac{XJ \cdot \sin\alpha}{T_{OX}} \right) \tag{2.65}$$

The capacitance components $C1$ and $C2$ are bias-independent and they are added to gate-drain or gate-source overlap capacitances, C_{GDO} and C_{GSO}. The capacitance component $C3$ is bias-dependent and it is modeled as a charge based form. Hence,

$$Q_{D.F} = - C_F \cdot \frac{V_{DS} - VDS}{1 + \exp\left(- \frac{V_{GB} - V_{FB}}{30 \cdot V_t} \right)} \tag{2.66}$$

$$Q_{S.F} = - C_F \cdot \frac{V_{GST} - V_{GS} + V_{FB} + 2\Phi_F + \gamma \sqrt{2\Phi_F - V_{BS}}}{1 + \exp\left(- \frac{V_{GB} - V_{FB}}{30 \cdot V_t} \right)} \tag{2.67}$$

$Q_{D.F}$ is added to the drain charge Q_D, $Q_{S.F}$ is added to the source charge Q_S and $-(Q_{D.F}+Q_{S.F})$ is added to Q_G. The exponential term in (2.66) and (2.67) is added to guarantee that the channel side fringing field capacitance is 0 in the accumulation region. Conceptually, the nominator in (2.66) is 0 in linear region and it is linearly dependent on $(V_{DS}-V_{DSSAT})$ in saturation region. Similarly, the nominator in (2.67) is 0 when $V_{GS} > V_{TH}$ and it is proportional to $(-V_{GS}+V_{TH})$ when $V_{GS} < V_{TH}$. The constant 30 in (2.66) and (2.67) is a heuristic factor.

2.4. Capacitance measurement system

2.4.1. Measurement system

To extract model parameters, an automatic direct-on-wafer off-chip gate capacitance measurement system has been developed. The principle of the measurement system is shown in Fig.2.9. D.U.T. is the gate capacitance of a MOSFET to be measured. C_{P1} and C_{P2} are stray capacitances associated with wires of the measurement setup. V_S is the small-signal sinusoidal voltage source. Due to the configuration of the circuit, the effects of stray capacitances C_{P1} and C_{P2} on the output voltage V_{OUT} are minimized. Since C_{P1} is connected directly between the voltage source V_S and ground, the current through C_{P1} does not go through the feedback resistor R_F and so it does not affect V_{OUT}. If the gain of the OP Amp. is sufficiently high at the frequency of V_S, the voltage across C_{P2} is very small and so the current through C_{P2} can be made very small. To reduce C_{P2}, the length of the wire between D.U.T. and OP Amp should be minimized.

The block diagram of the measurement system is shown in Fig.2.10. The system in Fig.2.10 is similar to the system reported earlier [2.2]. However it has been fully automated now and the resolution has been enhanced by an order of magnitude to about 14aF. In Fig.2.10, HP9836 computer serves as a system controller, HP4145A is used as a programmable voltage source to set the biases of MOSFET to be measured. The switching matrix scanner connects the reference signal to gate, drain, source or bulk node of MOSFET for C_{GG}, C_{GD}, C_{GS} or C_{GB} measurement respectively.

Fig.2.11 shows the detailed signal path of the C_{GD} measurement. The reference AC signal is supplied from lock-in amplifier and goes through audio transformer and current booster and is applied to the drain node of MOSFET. The audio transformer is used for DC isolation and the current booster is used to sink or source high currents when the reference AC signal is connected to either drain or source node, that is, for the measurement of C_{GD} or C_{GS} respectively. The output voltage of the I-V converter goes to lock-in amplifier and the lock-in amplifier captures only the signal with the same

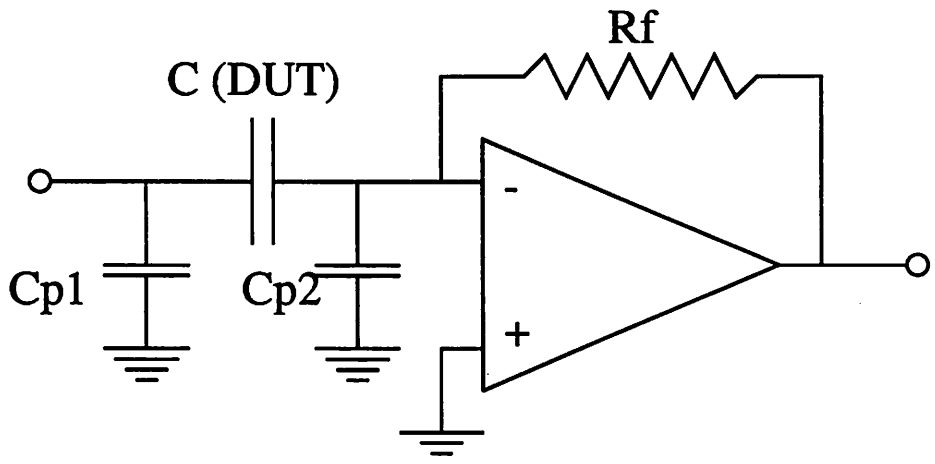


Fig.2.9. The principle of the gate capacitance measurement system.

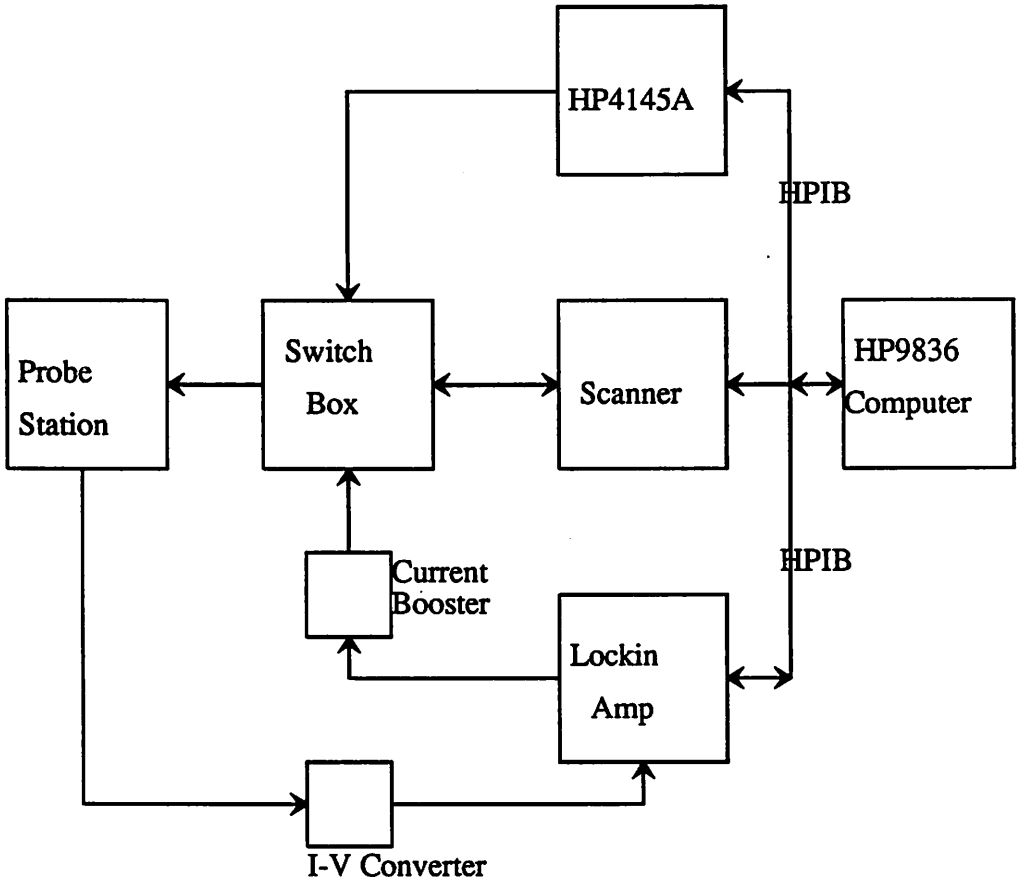


Fig.2.10. The block diagram of the gate capacitance measurement system.

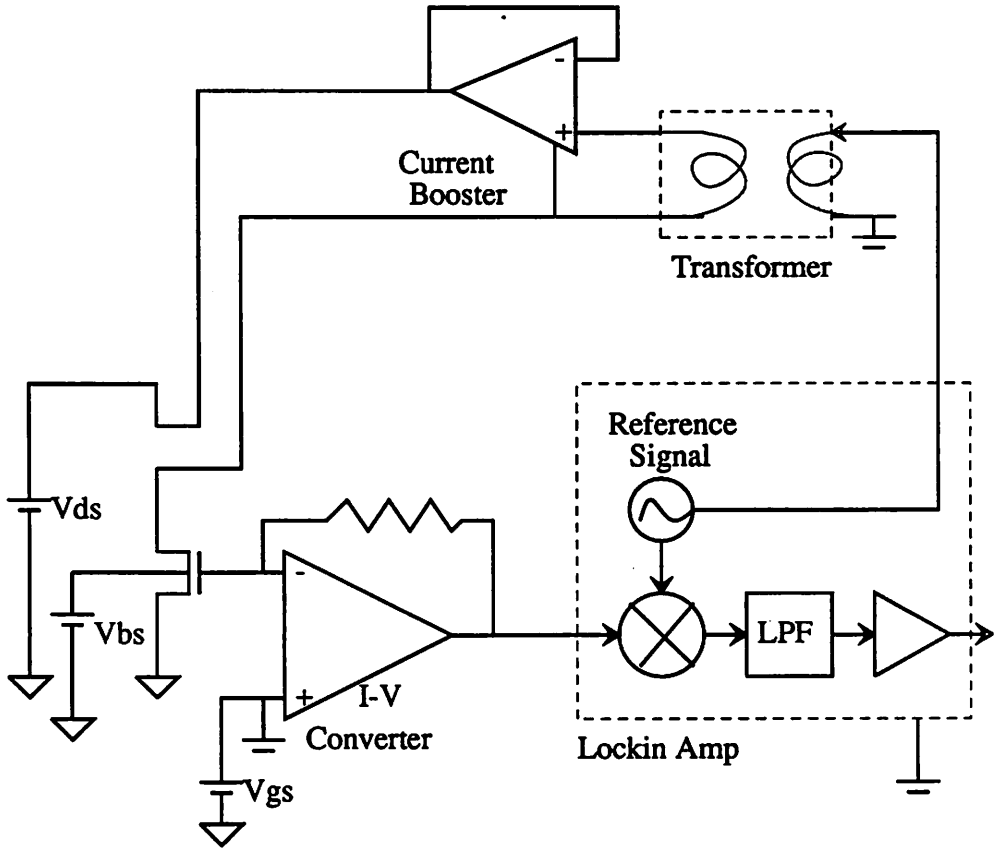


Fig.2.11. The detailed signal path of the C_{GD} measurement. The reference AC signal from the lock-in amplifier is superimposed to the drain node of MOSFET.

frequency as the reference signal and sends out the digital output to HP9836 computer. Ideally the output voltage amplitude of the I-V converter increases proportionately to the signal frequency ω. However ω cannot be increased without limits due to the frequency limitation of the I-V converter OP Amp, the current booster OP Amp and the stray wire capacitances as shown by C_{P2} in Fig.2.9. With some compromise the signal frequency is set to 6.5KHz in this work. The r.m.s. amplitude of the input signal is set to 60mV. EG&G model 5206 is used for the lock-in amplifier and Keithley model 706 is used for the switching matrix scanner.

Burr Brown 3551 J video OP Amp is used for the I-V converter. It has a single pole characteristics which is important for the stability of the measurement system. The DC gain of the video OP Amp is 100dB and the pole frequency is between 100Hz and 1KHz depending on the output load, according to the spec. sheet. Burr Brown 3571 power OP Amp is used for the current booster. It can drive up to 100mA, according to the spec. sheet.

The averaging technique has been used to reduce the random fluctuation of data and so to enhance the resolution. Fig.2.12 shows the r.m.s. resolution versus N the number of data points used in the averaging. Each point in the plot is the standard deviation of 200 capacitance values measured on a 9pF capacitor with each value being the average of N measurements. If the fluctuation of data is perfectly Gaussian, the r.m.s. resolution should follow $N^{-0.5}$. But in Fig.2.12, it followed $N^{-0.4}$. The r.m.s. resolution is about 14 aF at N=40 as shown in Fig.2.12. Fig.2.13 shows the measured raw data of C_{GS} for W/L=1.5μm/2μm (drawn dimension) NMOSFET with N=40. The peak to peak fluctuation is about 40 aF. For N=40, it takes about 3 seconds to get one measured value.

2.4.2. Measurement program

Measurement program is written in HP BASIC 3.0. Using the graphic facility, an interactive measurement is done. We can choose one of the measurement modes (C_{GG},C_{GD},C_{GS} or C_{GB}) within the program. According to each measurement mode, HP 4145 A is set up appropriately via HPIB(GPIB) communication cable. Then the

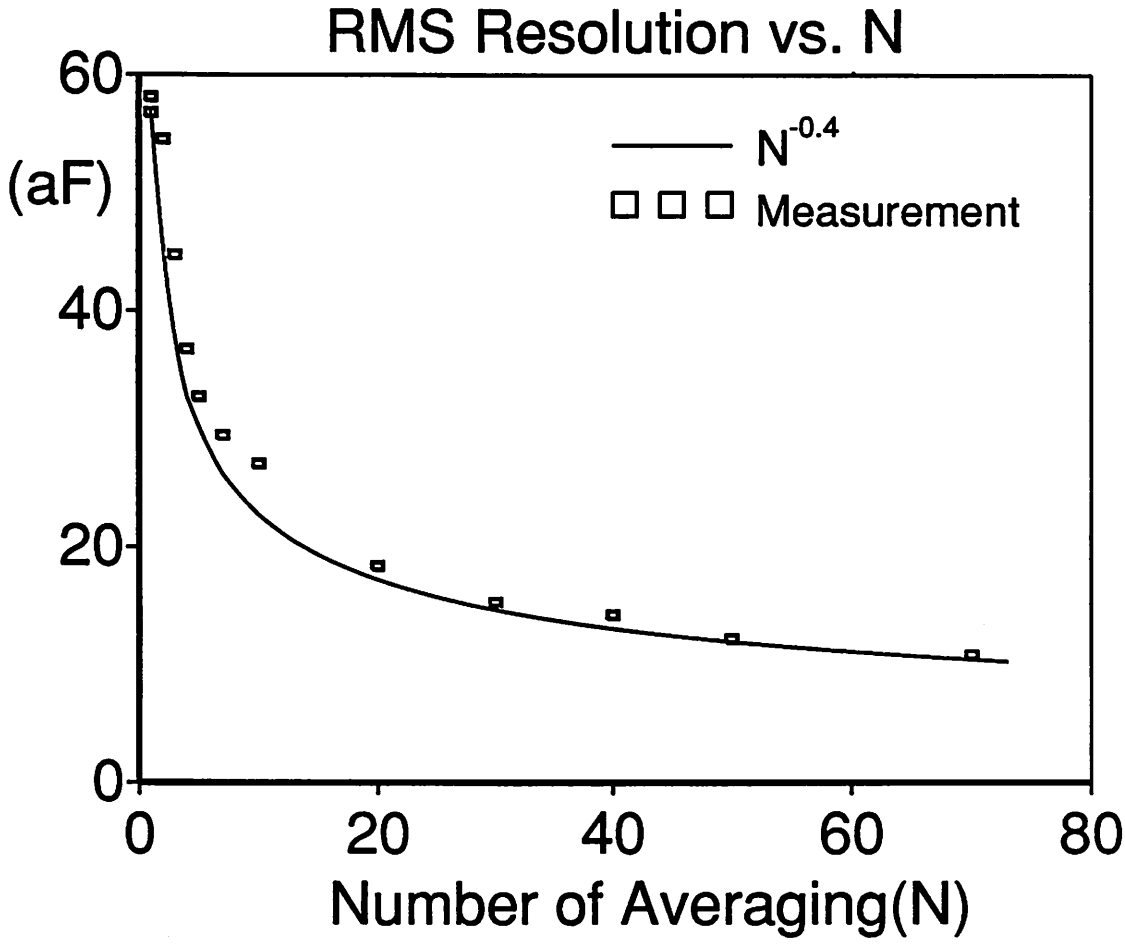


Fig.2.12. The rms resolution versus N (the number of averaging)

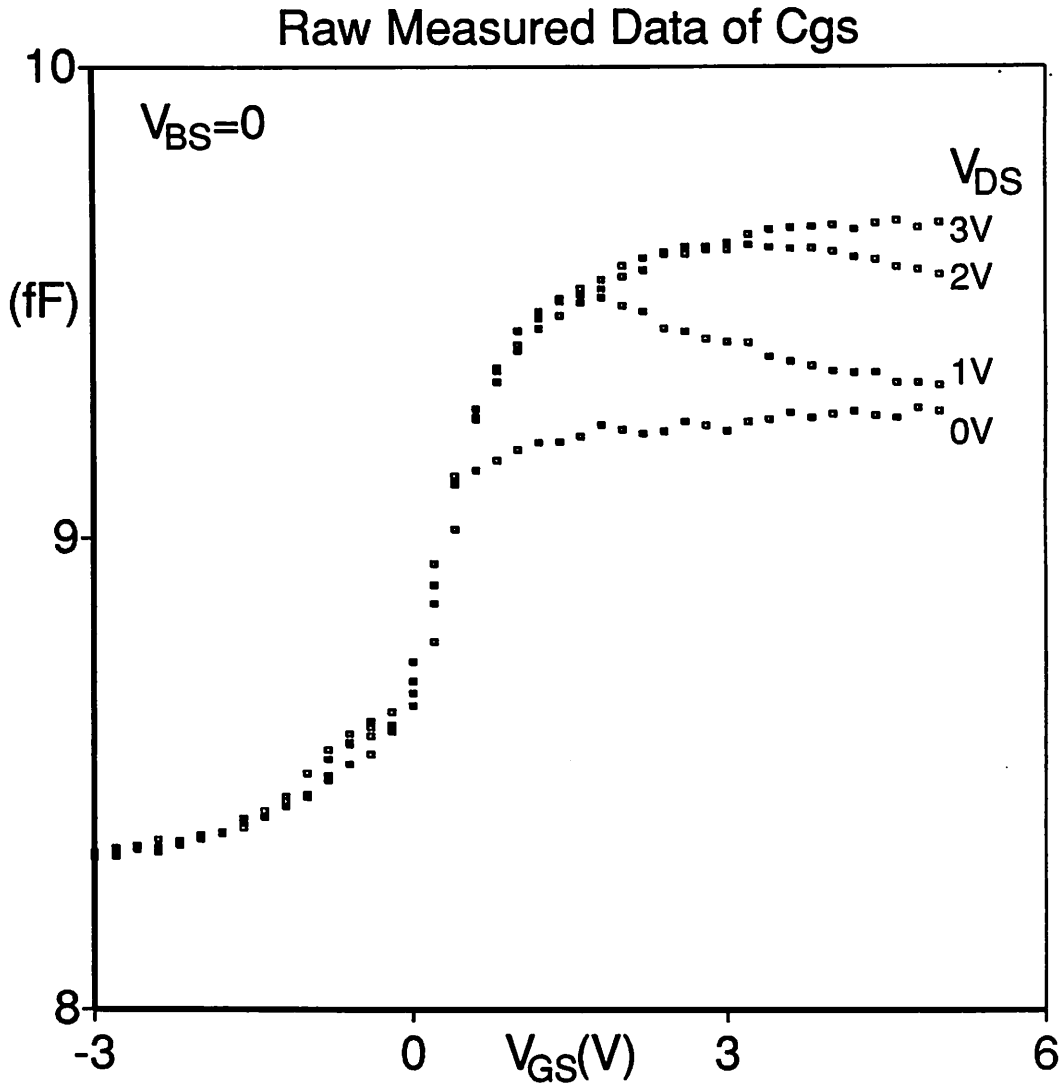


Fig.2.13. The raw data of measured C_{GS} for a NMOSFET with W/L=1.5μm/2μm (drawn dimension).

switching matrix scanner is set up corresponding to each measurement mode. We used only one scanner board in the switching matrix scanner.

In one scanner board, there are 4 Rows and 5 Columns. C1 (Column 1) is connected to the series connection of AC signal and SMU 1. C2 is connected to SMU 2, C3 to SMU 3, C4 to SMU 4 and C5 to Low(HP4145 GND). R 1 (Row 1) is connected to GND, R 2 to Bulk(Substrate), R3 to Source, and R4 to Drain. According to each measurement mode the Rows and Columns are connected appropriately. There are 3 lines(High, Low, Guard) for each Row and Column. Only the High line is used for the connection and the other two lines are used for shielding.

Lockin Amp is set up independently of the measurement mode. Lock in amp consists of an analog multiplier followed by a low pass filter(LPF) and an amplifier. The low pass filter is set up with the time constant of 30 ms and with the roll off factor of 12 dB/dec. Then the sensitivity of the lock in amp is set up. The sensitivity corresponds to the amplification factor of the final stage amplifier in the lock in amp.

Within one measurement mode, we decided not to change the sensitivity of the lockin Amp. So, in order to get the best resolution, the bias for the maximum capacitance for the measurement mode is applied. For the range of V_{GS} from -2 to 8 V and V_{DS} from 0 to 5 V and V_{BS} from 0 to -10 V , that bias point for C_{GS} is V_{GS}=V_{DS}=5V, V_{BS}=0, and that for CGD is V_{GS}=5V, V_{DS}=0, V_{BS} =0, and that for CGG(CGT) and CGB is V_{GS}=-5 V,V_{DS}=0, V_{BS}=0. The output of the Lockin amplifier is the digitized discrete value between 0 and 2000. Appling those biases for maximum capacitances to the MOSFET, the sensitivity is adjusted so that the lock in amp output lies between 800 and 2000. And then the measurement is performed. Period of 5 time constants(lockin amp LPF time constant) is inserted between each bias point to get the steady state value. At the abrupt change of bias values, at the change of a parameter value, a period of 20 time constants is inserted. The total measurement program consists of around 400 BASIC lines and the program listings are shown in Appendix 3.

2.5. Comparison of the model with the measured data

2.5.1. Surface potential and its derivatives

To check the validity of the surface potential formulation scheme shown in Fig.2.4, the calculated values using this scheme are compared with the measured data. The surface potential and its derivatives with respect to bias can be extracted from the MOSFET gate capacitance measurement. When drain and source nodes are tied together, the surface potential $\Psi_S(y)$ is constant all along the channel ($0 \le y \le L$). In this case the gate charge Q_G can be derived from (2.34) as

$$Q_G = WLC_{OX} \cdot (V_{GB} - V_{FB} - \Psi_S) \tag{2.68}$$

The gate capacitances C_{GG}, C_{GC} and C_{GB} are defined to be $\partial Q_G/\partial V_G$, $\partial Q_G/\partial V_C$ and $\partial Q_G/\partial V_B$ respectively, where C represents the channel node, that is, the tied drain, source node.

$$\frac{\partial \Psi_S}{\partial V_G} = 1 - \frac{C_{GG}}{WLC_{OX}} \tag{2.69}$$

$$\frac{\partial \Psi_S}{\partial V_C} = - \frac{C_{GC}}{WLC_{OX}} \tag{2.70}$$

$$\frac{\partial \Psi_S}{\partial V_B} = -1 - \frac{C_{GB}}{WLC_{OX}} \tag{2.71}$$

Hence, the derivatives of the surface potential can be computed from the measured gate capacitances. For the measurement of capacitances C_{GG}, C_{GC} and C_{GB}, the small signal voltage is applied at each node G, C and B respectively and the small signal gate current is measured. Ψ_S can be computed by integrating both sides of (2.69) with V_G.

$$\Psi_S = \int_{V_{FB}}^{V_{GB}} (1 - \frac{C_{GG}}{WLC_{OX}}) \, dV_G \tag{2.72}$$

where we used the fact that Ψ_S is 0 at $V_{GB} = V_{FB}$ (flat band voltage). V_{FB} can be extracted from the capacitance measurement [2.19] or the threshold voltage

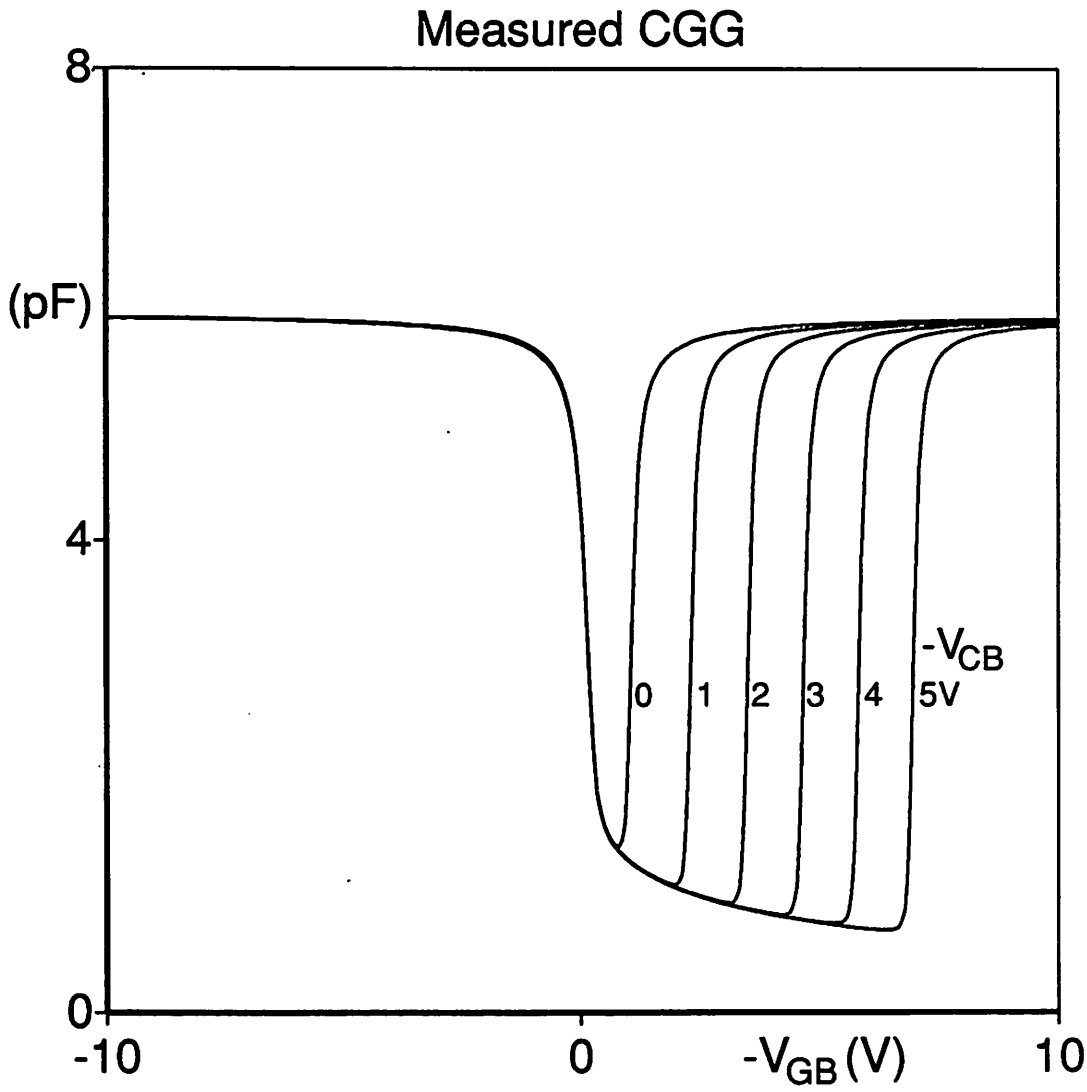


Fig.2.14.(a). The measured C_{GG} of a long channel PMOSFET with W/L=100μm/100μm. This C_{GG} data will be used to extract the measured surface potential curve shown in Fig.2.14.(b). Stray capacitances such as overlap capacitances are not included.

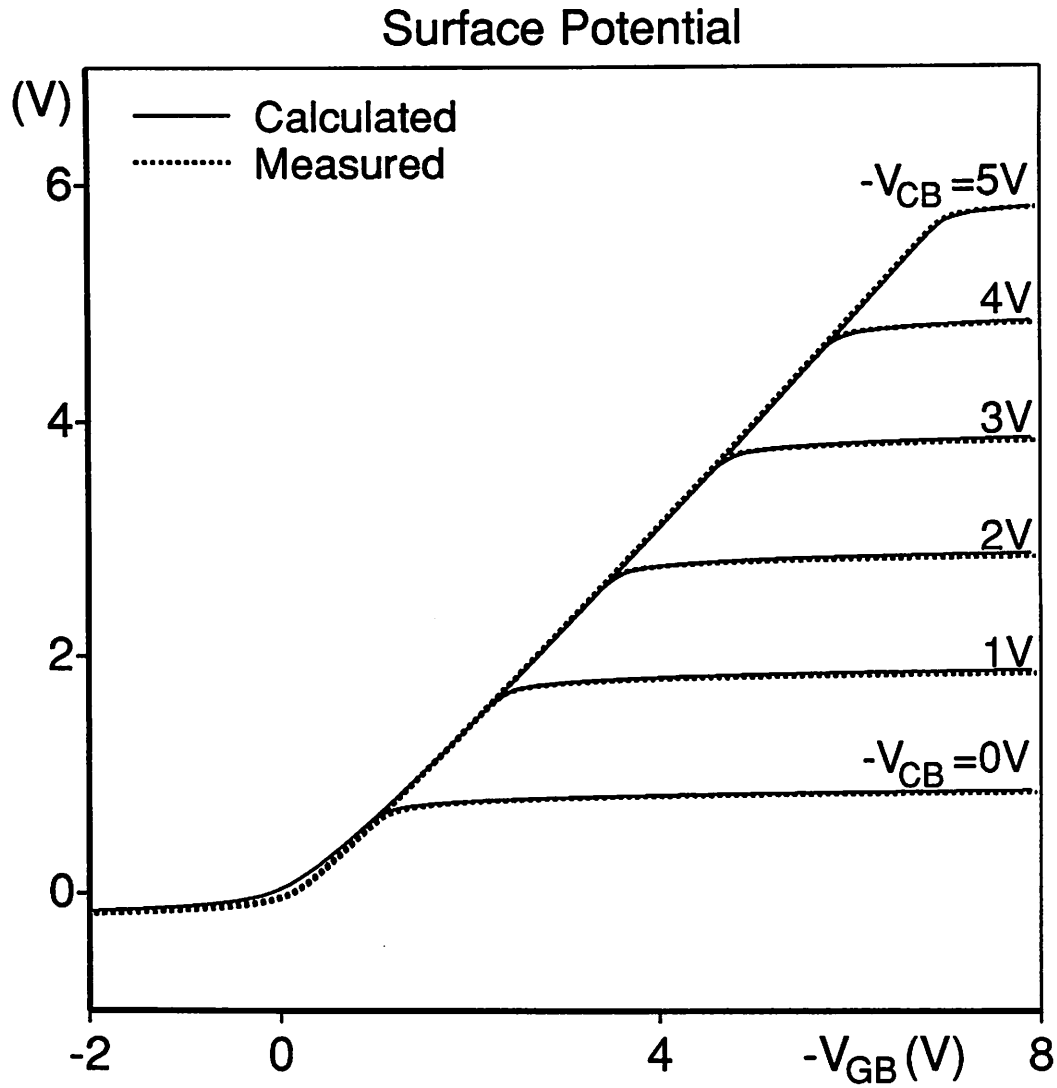


Fig.2.14.(b). Comparison of the measured and the calculated surface potential curves for the long channel PMOSFET. The measured data have been extracted from the C_{GG} data in Fig.2.14.(a) using (2.72). The model parameters used for the calculated curve are V_{FB}=0.11V, N_{SUB}=3.5·10$^{15}cm^{-3}$, T_{OX}=58.6nm, and γ=0.58$\sqrt{V}$.

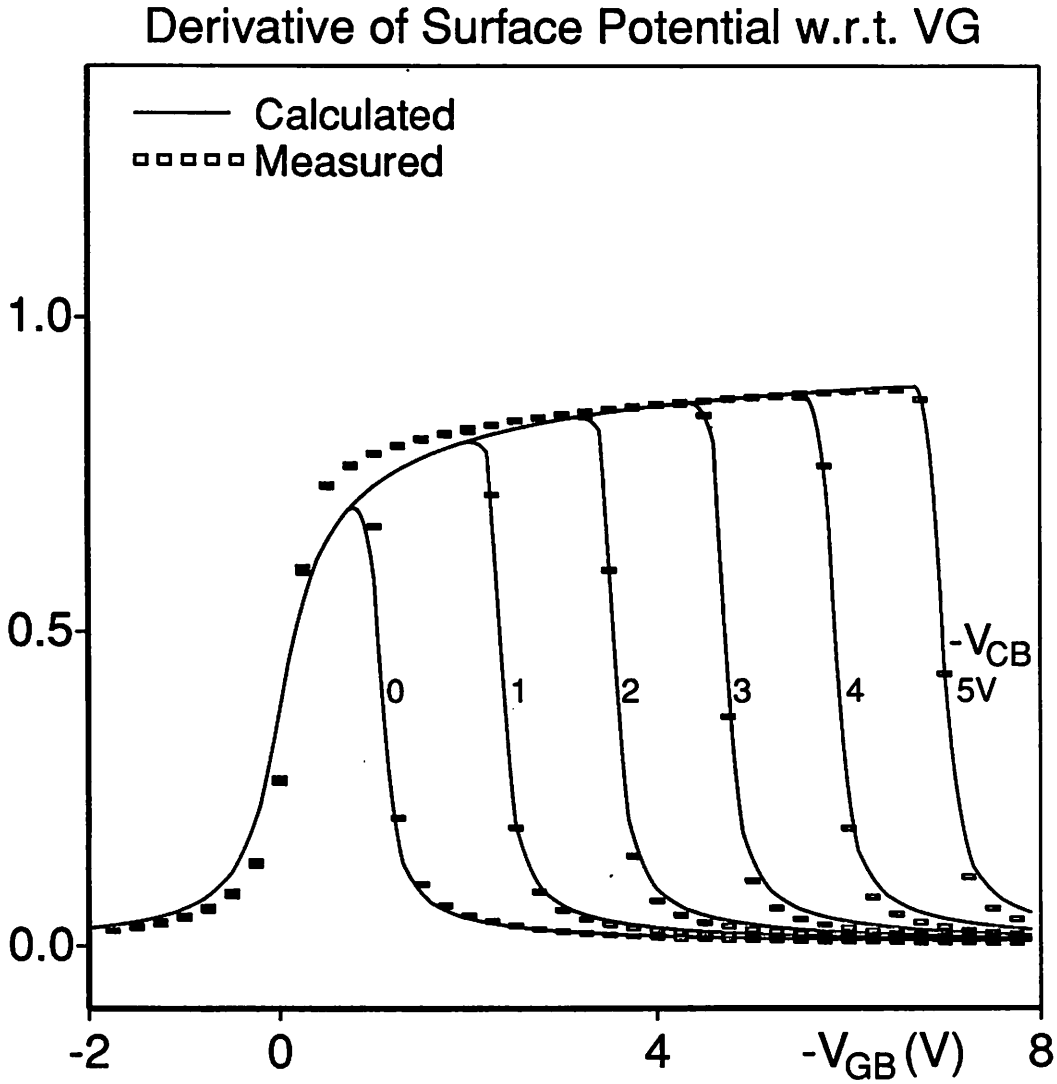


Fig.2.14.(c). Comparison of the measured and the calculated $\partial\Psi_S/\partial V_G$ for the long channel PMOSFET whose C_{GG} is shown in Fig.2.14.(a).

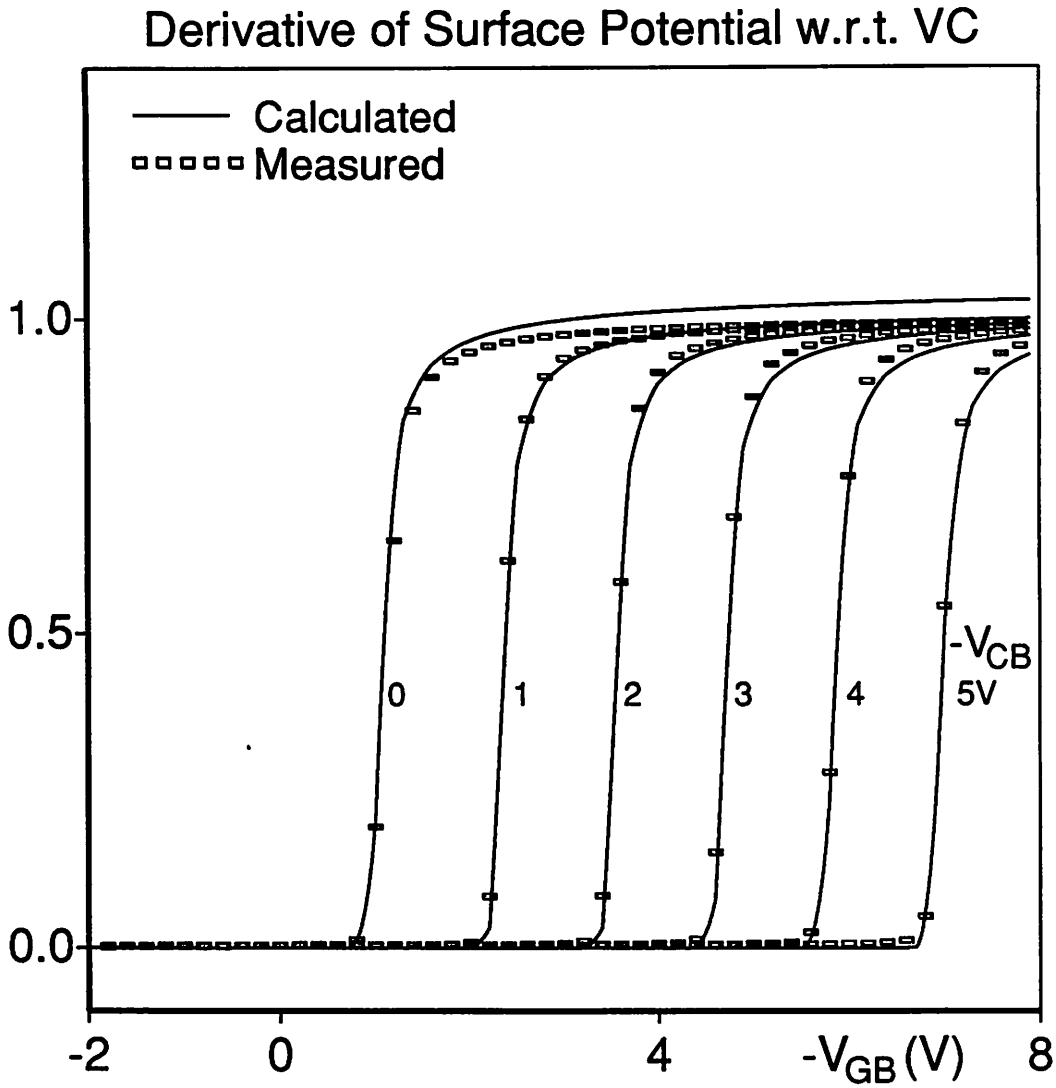


Fig.2.14.(d). Comparison of the measured and the calculated $\partial\Psi_S/\partial V_C$ for the long channel PMOSFET whose C_{GG} is shown in Fig.2.14.(a).

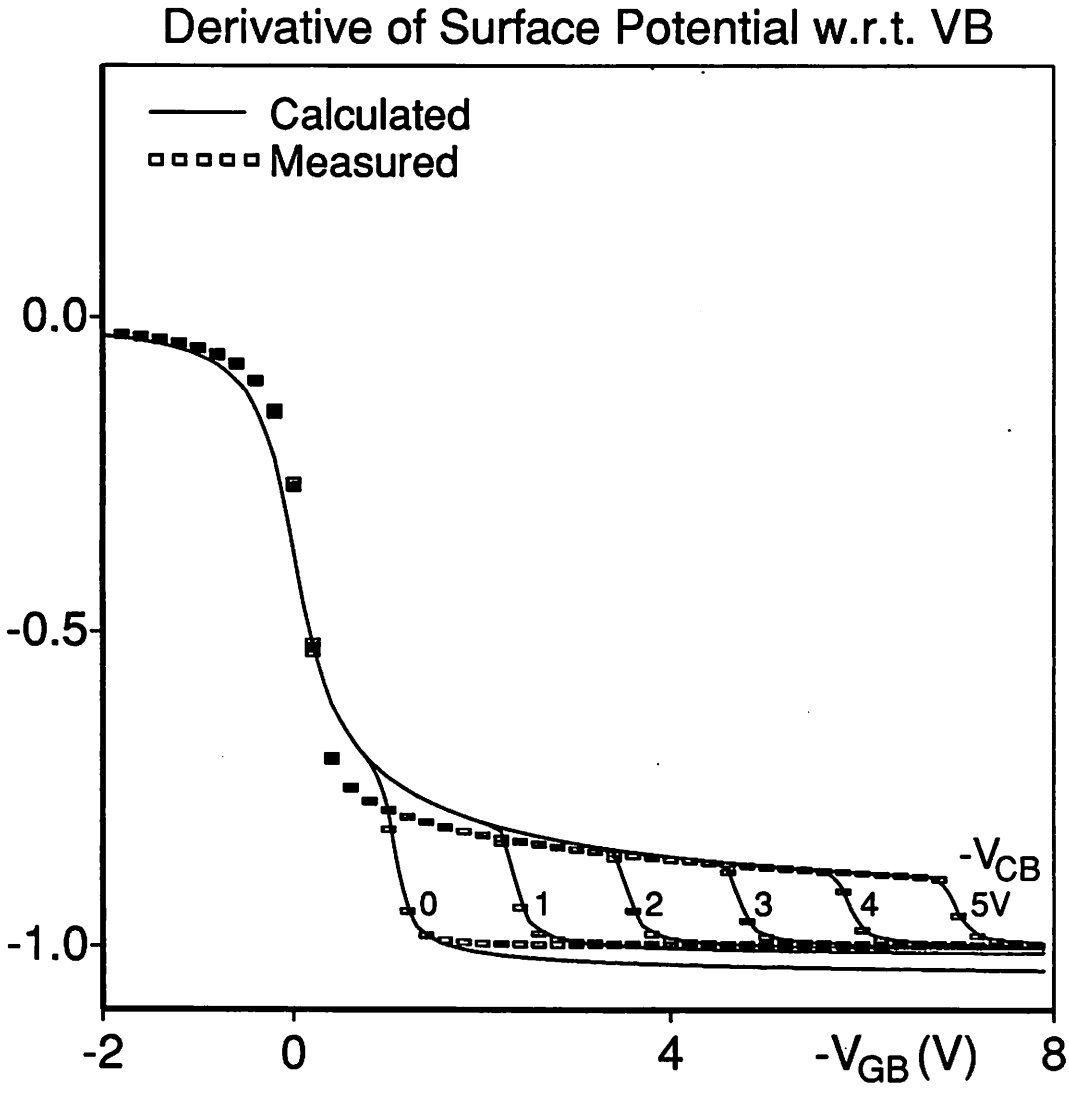


Fig.2.14.(e). Comparison of the measured and the calculated $\partial\Psi_S/\partial V_B$ for the long channel PMOSFET whose C_{GG} is shown in Fig.2.14.(a).

measurement. In this work V_{FB} has been adjusted to fit the calculated data.

Comparison has been made between calculated values and the measured data for PMOSFET with W/L=100μm/100μm and uniform substrate doping concentration. Fig.2.14.(a) shows the measured C_{GG} for computation of Ψ_S. Fig.2.14.(b) shows the measured and the calculated surface potentials for the PMOSFET. Fig.2.14.(c), Fig.2.14.(d) and Fig.2.14.(e) show $\partial\Psi_S/\partial V_G$, $\partial\Psi_S/\partial V_C$ and $\partial\Psi_S/\partial V_B$ respectively for the same PMOSFET.

In Fig.2.14.(b), slight discrepancies can be observed between measured and calculated surface potentials at around the flat band voltage ($-V_{GB} = -0.11V$). This tendency can also be observed in the derivative curves $\partial\Psi_S/\partial V_G$ and $\partial\Psi_S/\partial V_B$ in Fig.2.14.(c) and Fig.2.14.(e) respectively. This phenomenon is due to the slight decrease of the substrate doping concentration at the very surface of silicon due to the threshold-adjustment-implantation.

Derivatives $\partial\Psi_S/\partial V_C$ and $\partial\Psi_S/\partial V_B$ give discrepancies with maximum error of 4% between measured and calculated values, in strong inversion region for $V_{CB} = 0$. This discrepancy could be minimized by slightly modifying the surface potential formulation scheme shown in Fig.2.4, but this modification caused the discrepancy in the surface potential curve (maximum error less than 2%). Since the derivatives $\partial\Psi_S/\partial V_C$ and $\partial\Psi_S/\partial V_B$ correspond to conductances and capacitances, and the surface potential correspond to current and charges, the modification is not included in Fig.2.4 to emphasize the accuracy of the surface potential rather than that of its derivatives. In general, good agreements can be observed between measured and calculated values of the derivatives of the surface potential.

The algorithm for finding the surface potential Ψ_S for MOSFETs with non-uniform substrate doping concentrations and its comparison with measurement are shown in Appendix 4.

2.5.2. Gate capacitances of long channel MOSFETs

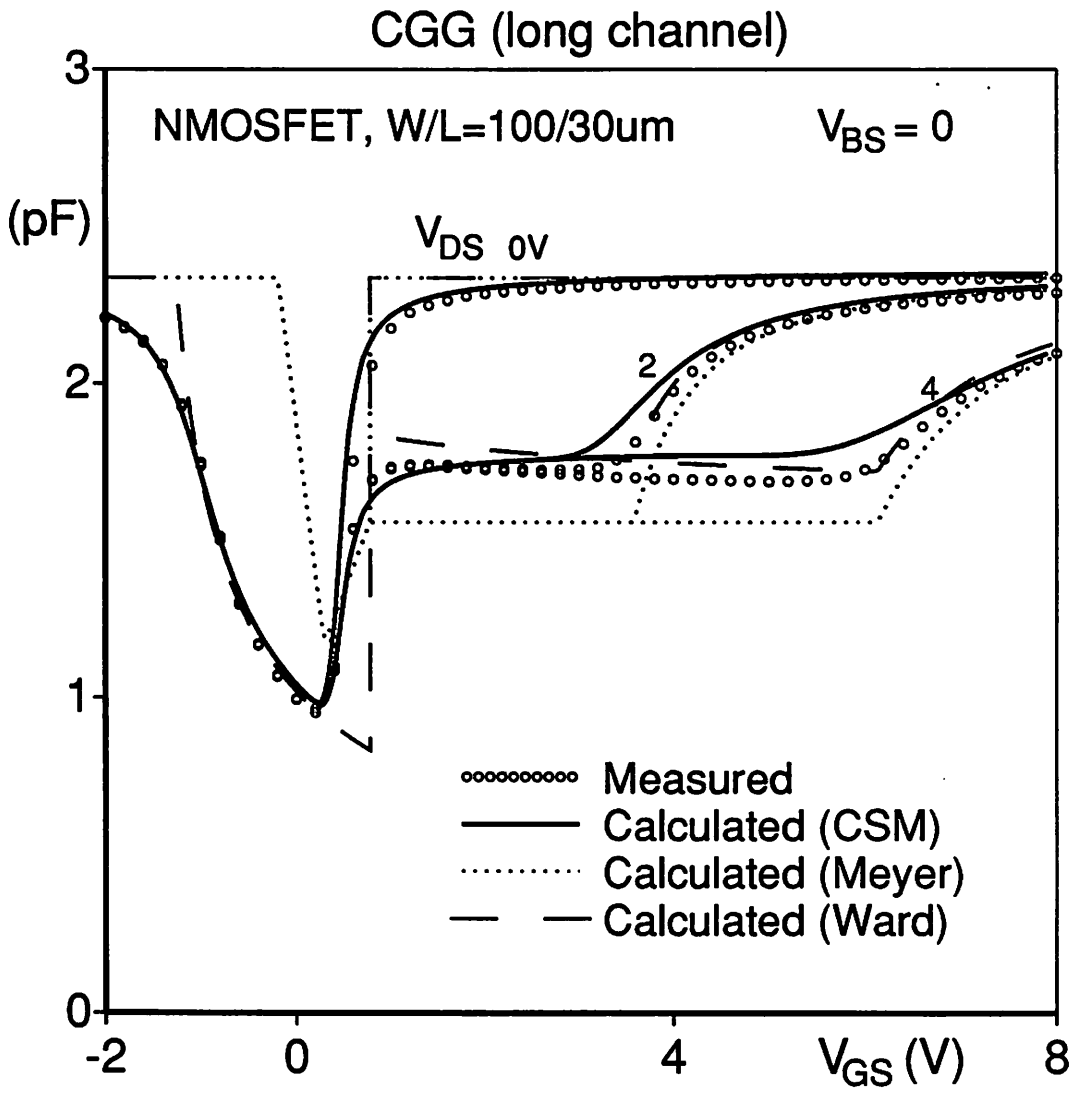


Fig.2.15.(a). Comparison of the measured and the calculated C_{GG} for a long channel NMOSFET. This work(CSM), the Meyer and the Ward-Dutton models in SPICE2 are used to get the calculated values. Model parameters used for the calculation are V_{FB}= −1.26V, γ=1.087$\sqrt{V}$, T_{OX}=43nm, PHI=0.967V, VTO=0.776V, *satfactor* (A in (2.15)) = 8.

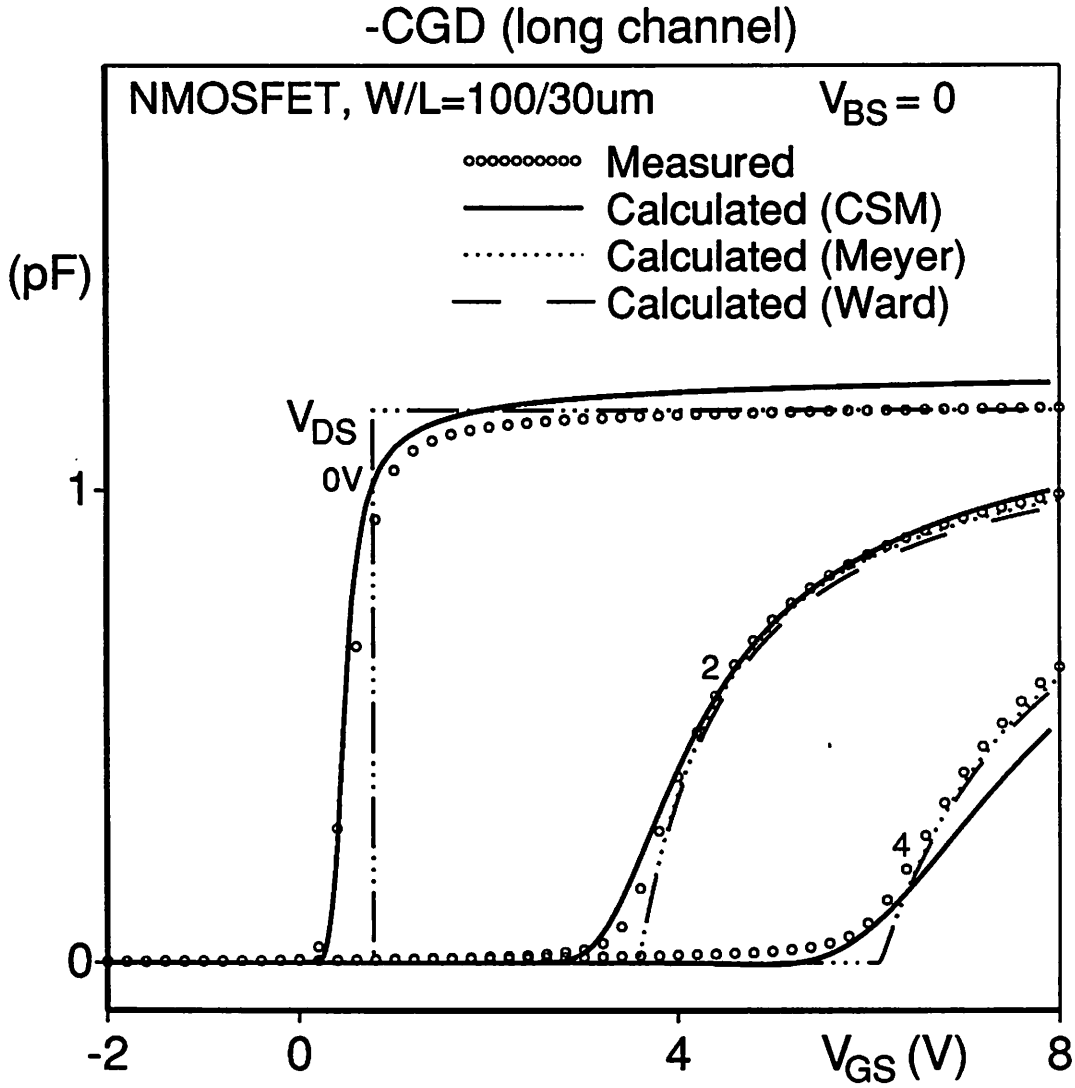


Fig.2.15.(b). Comparison of the measured and the calculated C_{GD} for the same long channel NMOSFET shown in Fig.2.15.(a).

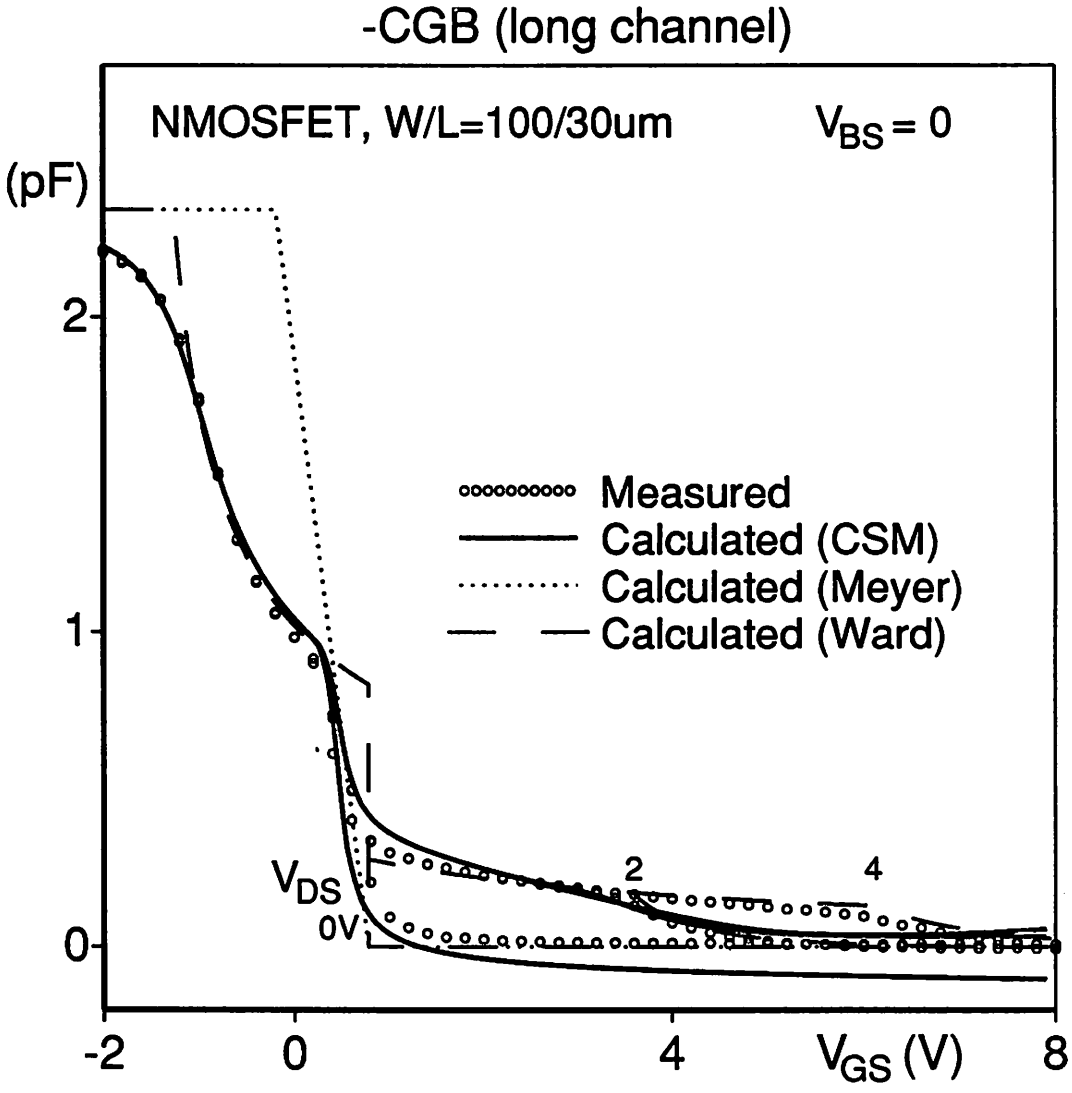


Fig.2.15.(c). Comparison of the measured and the calculated C_{GB} for the same long channel NMOSFET shown in Fig.2.15.(a).

Fig.2.15.(a), Fig.2.15.(b) and Fig.2.15.(c) show the comparison of the measured and the calculated gate capacitances for long channel MOSFETs with W/L=100μm/30μm (drawn dimension), T_{OX}=43nm, V_{FB}= -1.26V, N_{SUB}=2.3·$10^{16}cm^{-3}$. Calculated values using the Meyer model and the Ward-Dutton model in SPICE2, are also shown for comparison.

In the C_{GG} curve shown in Fig.2.15.(a), the Meyer model gives large errors from the measured data in the cutoff and saturation regions, since the bulk charge effect is neglected in the Meyer model. The Ward-Dutton model gives large errors in the subthreshold region and also it gives big discontinuities at the threshold voltage ($V_{GS} = V_{TH}$). This work shows good agreements with the measured data. Especially the continuity of the capacitance and its slope is excellent in this work.

In the C_{GD} curve shown in Fig.2.15.(b), all the models show fair agreements with the measured data except the big discontinuities of the Meyer and Ward-Dutton models at V_{GS}=V_{TH} for V_{DS} = 0. Also this work(CSM) and the measured data show smooth transitions between the linear and saturation regions, but the Meyer and the Ward-Dutton models show steep transitions between the two regions. This smooth transition is due to the diffusion current which is included in this work but is not included in the Meyer or the Ward-Dutton model.

In the C_{GB} curve shown in Fig.2.15.(c), the Meyer model gives large errors from the measured data in the cutoff region and it predicts that C_{GB}=0 for $V_{GS} \geq V_{TH}$ independent of V_{DS}. This is due to the neglect of the bulk charge effect in the Meyer model. The Ward-Dutton model gives good agreements with the measured data in the strong inversion region($V_{GS} > V_{TH}$) but it gives a discontinuity at $V_{GS} = V_{TH}$. This work gives good agreements with the measured data except slight mismatches in strong inversion region for V_{DS}=0 and 4V respectively. The slight mismatch for V_{DS}=0 is due to the surface potential formulation scheme described in Section 2.2.1. This same mismatch can be observed in Fig.2.14.(e) for V_{CB}=0. The slight mismatch for V_{DS}=4V is due to the linearization approximation of the bulk charge which is described in Section 2.3.1.

Although this linearization approximation gives a slight mismatch from the measured data for long channel MOSFETs with thick gate oxide, it gives good agreements for short channel MOSFETs as will be shown in Fig.2.16.(d).

Programs for computing MOSFET capacitances using the Meyer model or the Ward-Dutton model in SPICE2, are shown in Appendix 7.

2.5.3. Gate capacitances of short channel MOSFETs

Fig.2.16.(a), Fig.2.16.(b) and Fig.2.16.(c) show the comparison of the measured and calculated gate capacitances for a short channel NMOSFET with W/L=100μm/2.25μm. The measured data of a long channel NMOSFET are also shown for comparison. The capacitance components are normalized by the total gate oxide capacitance WLC_{OX}. The velocity saturation effect, the channel length modulation effect on charges, and the channel side fringing field capacitances which are described in Section 2.3, are also included in the calculation. Bias-independent stray capacitances such as $C1$ and $C2$ in Fig.2.8, are not included in the figures for clarity.

$< C_{GG} >$

In the C_{GG} curve shown in Fig.2.16.(a), this work shows good agreements with the measured data. In saturation region, appreciable differences can be observed between long and short channel C_{GG}. Also very smooth transitions between saturation and linear regions can be observed in this work and the measured data. Also the splitting of capacitance values can be observed in the cutoff region (for $V_{GS} < 0.5V$) for this work and the measured data. C_{GG} decreases as V_{DS} increases, in the cutoff region. This is due to the channel side fringing field capacitances which are shown as $C3S$ and $C3D$ in Fig.2.8. As V_{DS} increases, the depletion region of the drain junction becomes wider and the less bulk charge is under control of the gate electric field. Hence the gate capacitance C_{GG} becomes smaller as V_{DS} increases.

$< C_{GD} >$

In the C_{GD} curve shown in Fig.2.16.(b), appreciable short channel effects can be observed. The calculated and the measured C_{GD} for the short channel MOSFET show

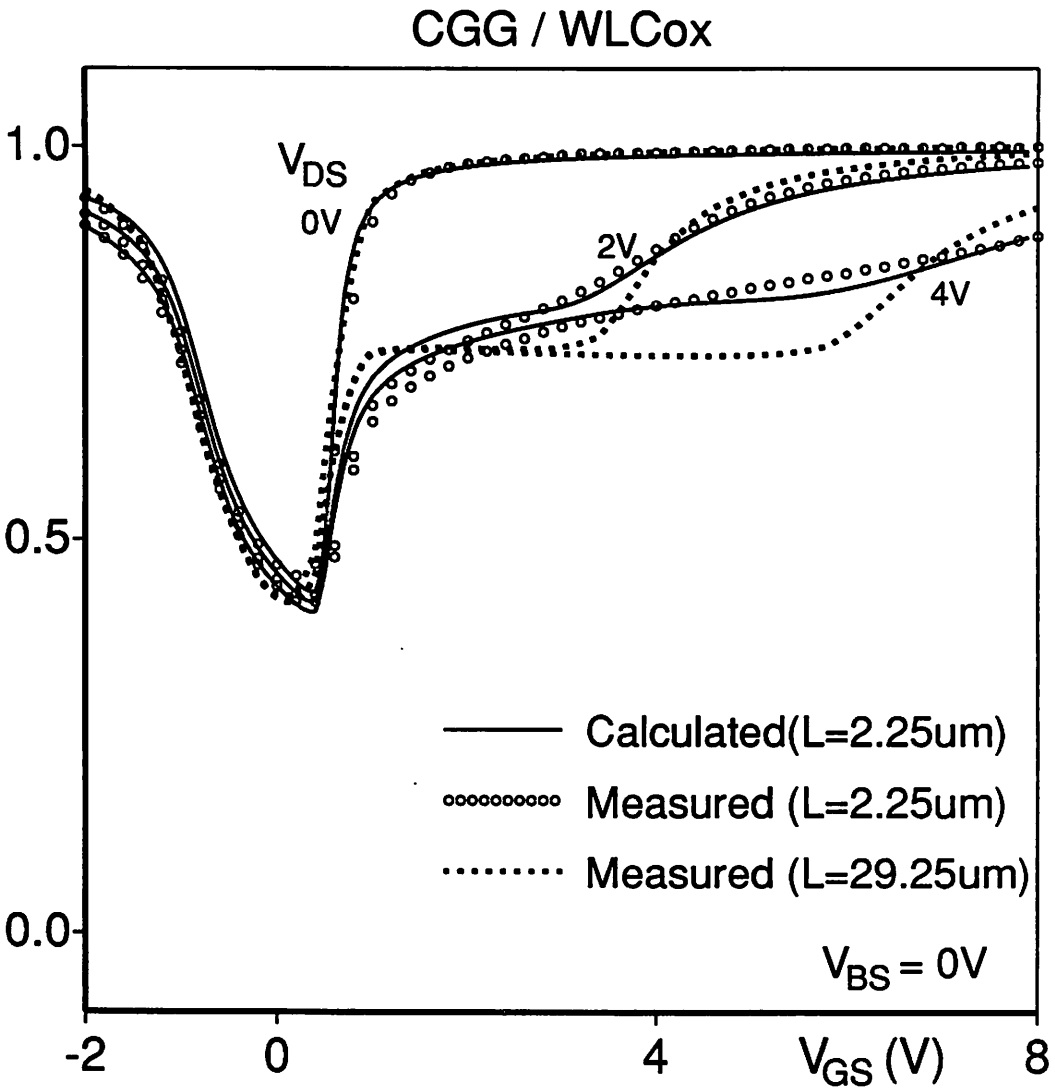


Fig.2.16.(a). Comparison of the measured and the calculated C_{GG} for a short channel NMOSFET with W/L=100μm/2.25μm (effective dimension). The measured C_{GG} of a long channel NMOSFET is also shown for comparison. Model parameters used for the calculation are $V_{FB}=-1.06V$, $T_{OX}=43nm$, $N_{SUB}=2\cdot10^{16}cm^{-3}$, $\mu_0=700cm^2/V-s$, $E_{CRIT}=6\cdot10^4V/cm$, $v_{SAT}=10^7cm/\text{sec}$, $\lambda_Q=0.01V^{-1}$, $XJ=0.17\mu m$, $T_{GATE}=0.7\mu m$, and *satfactor* (A in (2.15)) = 5.

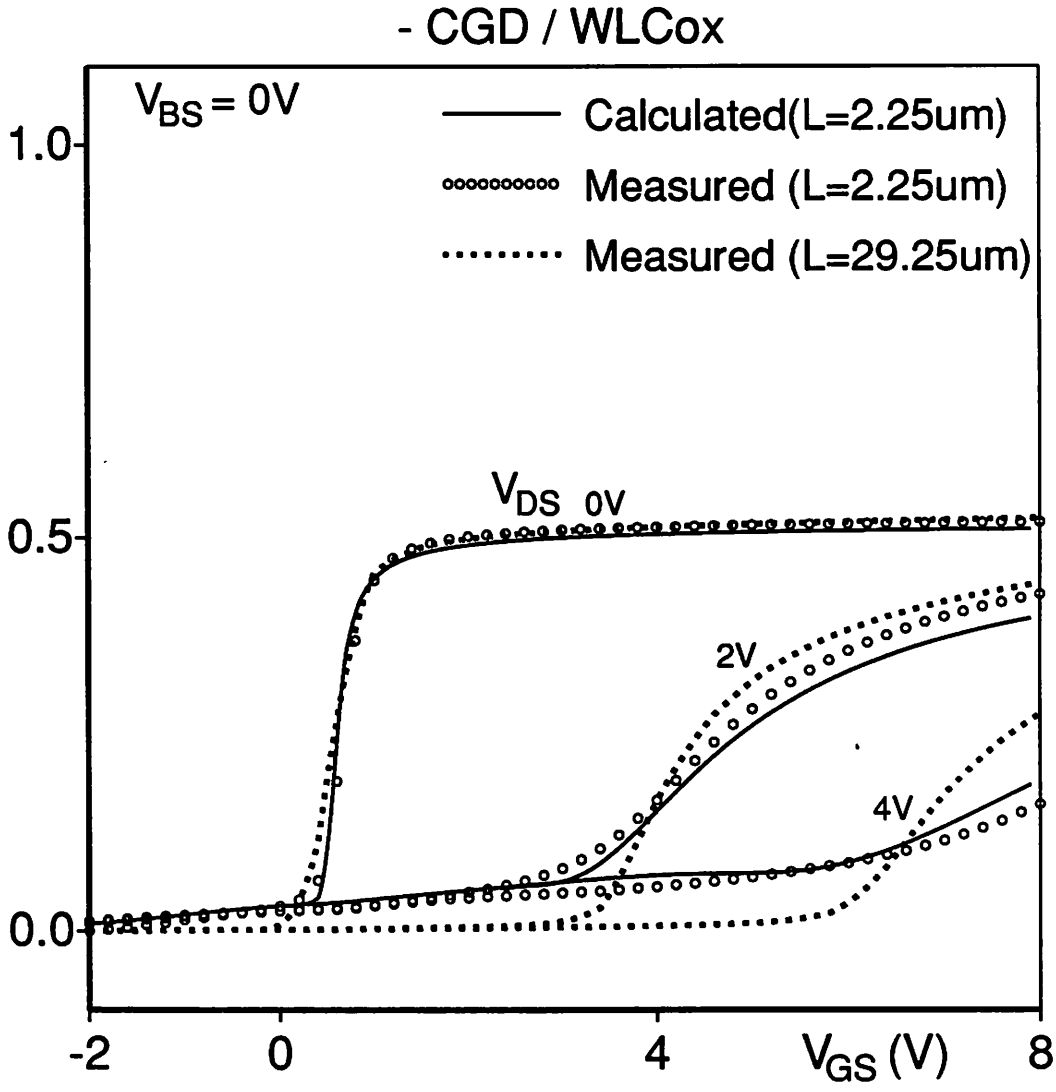


Fig.2.16.(b). Comparison of the measured and the calculated C_{GD} for a short channel NMOSFET with W/L=100μm/2.25μm (effective dimension), along with a measured C_{GD} for a long channel NMOSFET. Model parameters for the calculation are the same as those in Fig.2.16.(a).

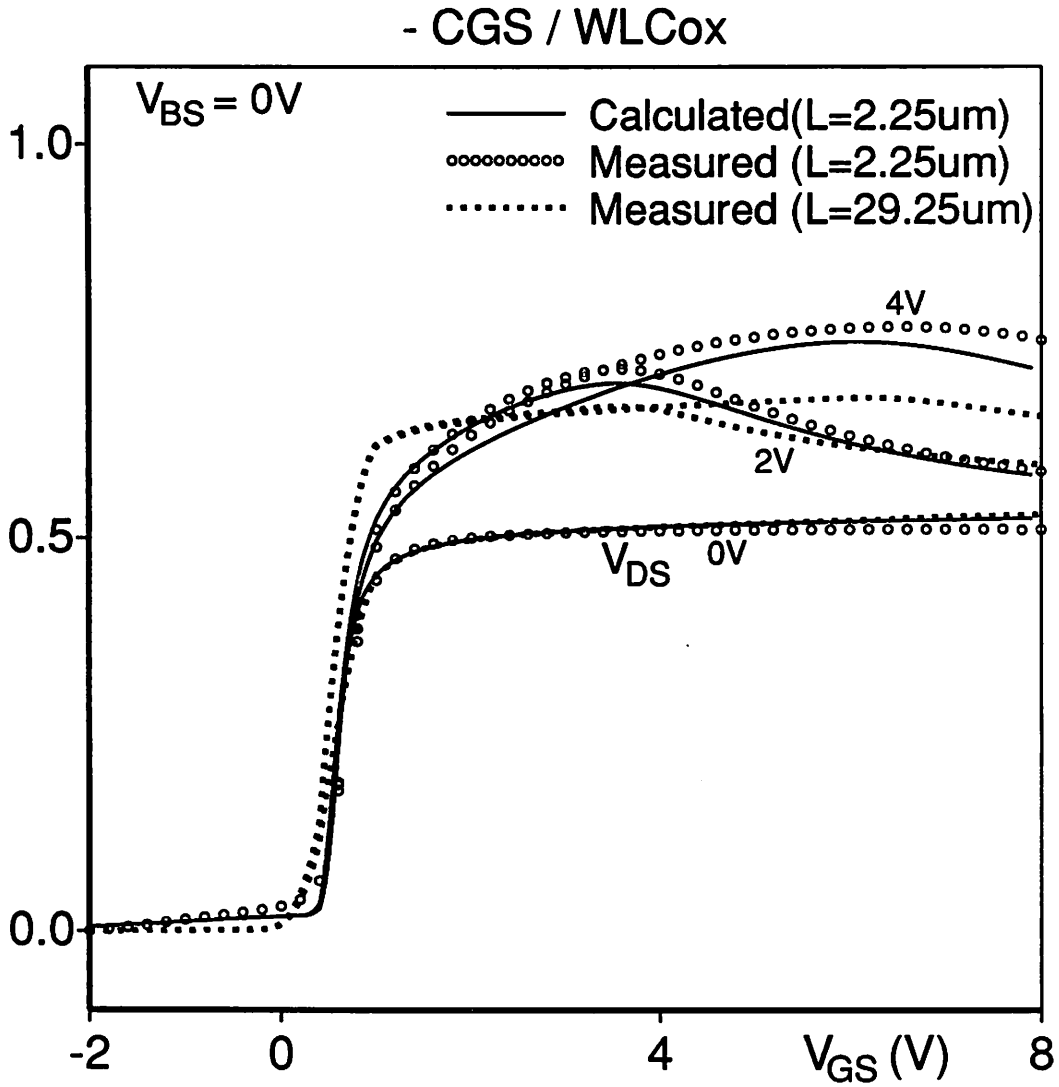


Fig.2.16.(c). Comparison of the measured and the calculated C_{GS} for a short channel NMOSFET with W/L=100μm/2.25μm (effective dimension), along with a measured C_{GS} for a long channel NMOSFET. Model parameters for the calculation are the same as those in Fig.2.16.(a).

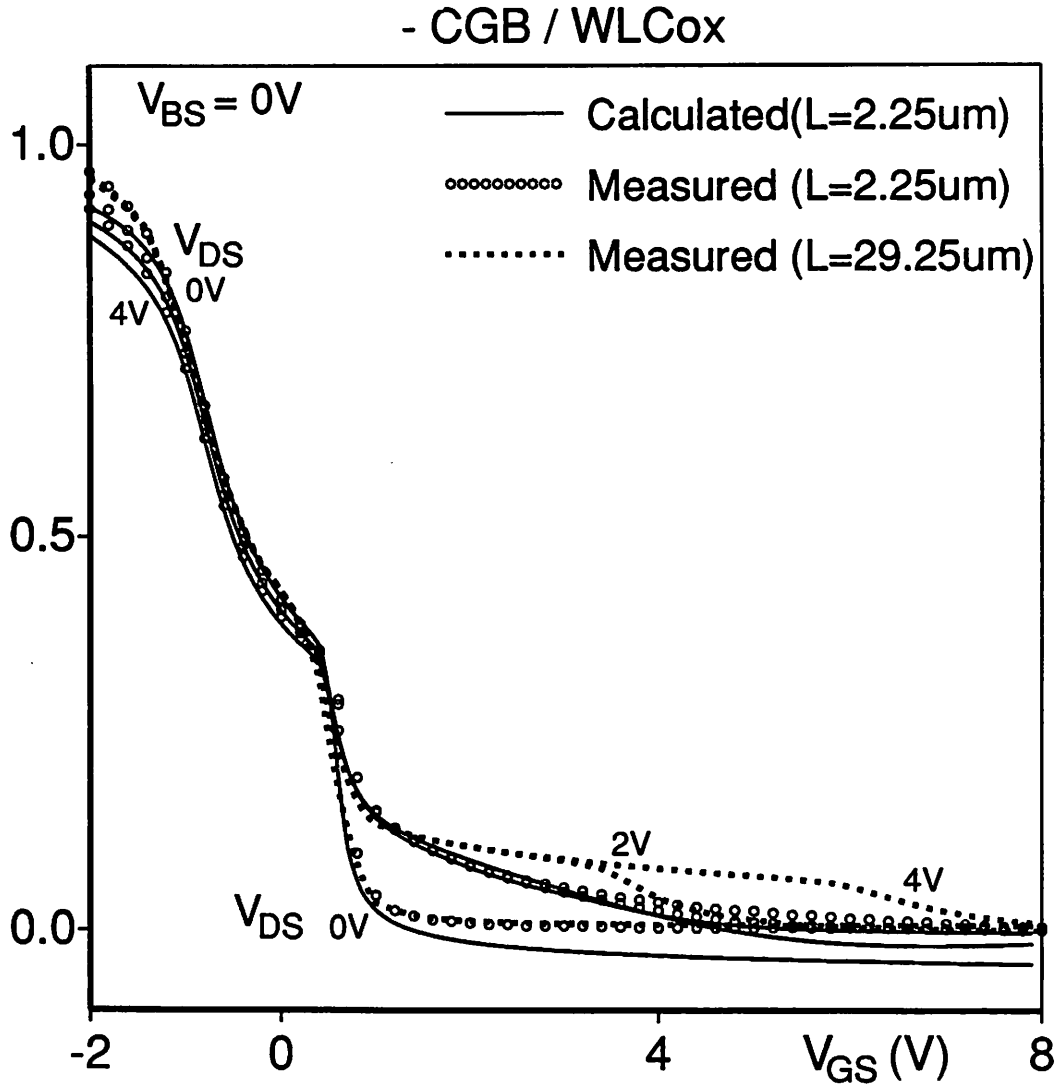


Fig.2.16.(d). Comparison of the measured and the calculated C_{GB} for a short channel NMOSFET with W/L=100μm/2.25μm (effective dimension), along with a measured C_{GB} for a long channel NMOSFET. Model parameters for the calculation are the same as those in Fig.2.16.(a).

good agreements between each other. But the measured C_{GD} for the long channel MOSFET shows appreciable differences from this work or the measured C_{GD} for the short channel MOSFET. The C_{GD} component is the most important among gate capacitances because its effect is multiplied by the voltage gain between drain and gate nodes due to the Miller effect.

In the cutoff region (where $V_{GS} < V_{TH}$), the calculation and the measurement for the short channel MOSFET show that C_{GD} increases with V_{GS}, while the long channel C_{GD} is constant at 0. This is due to the channel side fringing capacitance component $C3D$ in Fig.2.8. As V_{GS} increases in the cutoff region with V_{DS} fixed, the depletion depth increases and hence the total bulk charge increases with V_{GS}. Also the change of the bulk charge which is under the control of the gate electric field, with respect to the V_D change, increases. Hence C_{GD} increases with V_{GS} in the cutoff region. When V_{GS} becomes larger than V_{TH}, the depletion depth profile doesn't change with V_{GS} and so this component of C_{GD} becomes constant with V_{GS}.

In the saturation region, C_{GD} is non-zero and increases with V_{GS} for the short channel MOSFET, but C_{GD} is constant at 0 for the long channel MOSFET. This non-zero C_{GD} in the saturation region, is due to the combination of the two effects. One effect is due to the channel side fringing field capacitance component $C3D$ shown in Fig.2.8 which gives a constant capacitance C_F shown in (2.65) independent of V_{GS}. The other effect is due to the channel length modulation effect on charges, which is described in Section 2.3.5. From (2.59) and (2.60), C_{GD} due to the channel length modulation effect can be written as $(Q_D + Q_S)\cdot\lambda_Q/L$. Since the magnitudes of Q_D and Q_S increase with V_{GS} in saturation region, the magnitude of C_{GD} due to the channel length modulation, increases with V_{GS} in saturation region.

In the linear region, C_{GD} of short channel MOSFETs is smaller than that of long channel MOSFETs due to the velocity saturation effect. This effect is evident for the (V_{DS}=4V) curve in Fig.2.16.(b). The physical origin of this effect is due to the fact that the total channel charge of short channel MOSFETs is less sensitive to the drain voltage

than that of long channel MOSFETs in linear region, because of the velocity saturation effect.

$< C_{GS} >$

In the C_{GS} curve, appreciable short channel effects can be observed. In the cutoff region (where $V_{GS} < V_{TH}$), C_{GS} of the short channel MOSFET is non-zero and increases with V_{GS} due to the channel side fringing field capacitance $C3S$ shown in Fig.2.8. This $C3S$ component becomes 0 in the strong inversion region ($V_{GS} > V_{TH}$), as shown in (2.67). In saturation region, C_{GS} of the short channel MOSFET shows a slight dependence on V_{DS}, while C_{GS} of the long channel MOSFET is independent of V_{DS}. And the magnitude of C_{GS} for V_{DS} = 2V is slightly larger than that for V_{DS} = 4V. This is due to the channel length modulation effect. From (2.59) and (2.60), the change of C_{GS} due to the channel length modulation can be written as $\Delta C_{GS} \approx (C_{DS} + C_{SS}) \cdot \Delta L/L$. The polarity of C_{SS} is positive and the polarity of C_{DS} is negative, and the magnitude of C_{SS} is larger than that of C_{DS}, and C_{DS} and C_{SS} are almost independent of V_{DS} in saturation region, and ΔL is larger for larger V_{DS}. From the above mentioned reasonings, we can see that the polarity of ΔC_{GS} due to the channel length modulation is positive and that its magnitude is larger for larger V_{DS}. Since the polarity of C_{GS} is negative, the magnitude of C_{GS} becomes smaller for larger V_{DS} in saturation region.

In the linear region, the short channel C_{GS} becomes much larger than the long channel C_{GS} due to the velocity saturation effect. This effect is evident for (V_{DS}=4V) curve. Due to the velocity saturation effect, more inversion charge resides in the channel for the short channel MOSFET than for the long channel case. Therefore, the sensitivity of the total channel charge and hence the sensitivity of the total gate charge with respect to V_S becomes larger for the short channel MOSFET than that for the long channel case. Hence, the short channel C_{GS} is larger than the long channel C_{GS} in the linear region, due to the velocity saturation effect.

$< C_{GB} >$

In the C_{GB} curve shown in Fig.2.16.(d), the splitting of C_{GB} for different V_{DS} can be observed in the cutoff region ($V_{GS} < 0.5V$) due to the channel side fringing field capacitance components. Larger V_{DS} gives smaller C_{GB} in the cutoff region. This is due to the fact that more bulk charge is under the control of the electric field which is originated from the drain junction and is terminated at the bulk electrode, as V_{DS} increases. Therefore, less bulk charge is under the control of the electric field which is originated from the gate electrode and is terminated in the bulk electrode, as V_{DS} increases. Hence C_{GB} becomes smaller for larger V_{DS} in the cutoff region.

In the strong inversion region, the short channel C_{GB} is much smaller than the long channel C_{GB}.

2.6. Simulation results and performance comparison

2.6.1. Transient gate current of a NMOSFET

Fig.2.17.(a) and Fig.2.17.(b) show the simulation results for the turn-on and the turnoff transients of a NMOSFET using this work (Charge Sheet MOSFET model) and the SPICE level 2 charge based model (XQC=0.4) [2.16] [2.23]. Since there is a big discontinuity of gate capacitance C_{GG} at the threshold voltage in the SPICE level 2 charge based model [2.16], as shown in Fig.2.15.(a), its simulated current oscillates whenever the gate voltage crosses the threshold voltage (V_{TH}) as shown in Fig.2.17.(a) and Fig.2.17.(b). The period of oscillation corresponds to the internal time step used in SPICE. On the contrary, the charge sheet model predicts the stable gate current waveform because of the continuity of the gate capacitance C_{GG} all over the operating regions, as shown in Fig.2.15.(a). The trapezoidal integration is used and the same convergence criterion is used for both models.

2.6.2. Performance comparison between models in SPICE3

Comparisons of run statistics have been made for some circuits using this work(CSM), the level-2 Meyer model and BSIM 40/60 (xpart=0) and BSIM 0/100 (xpart=1) models

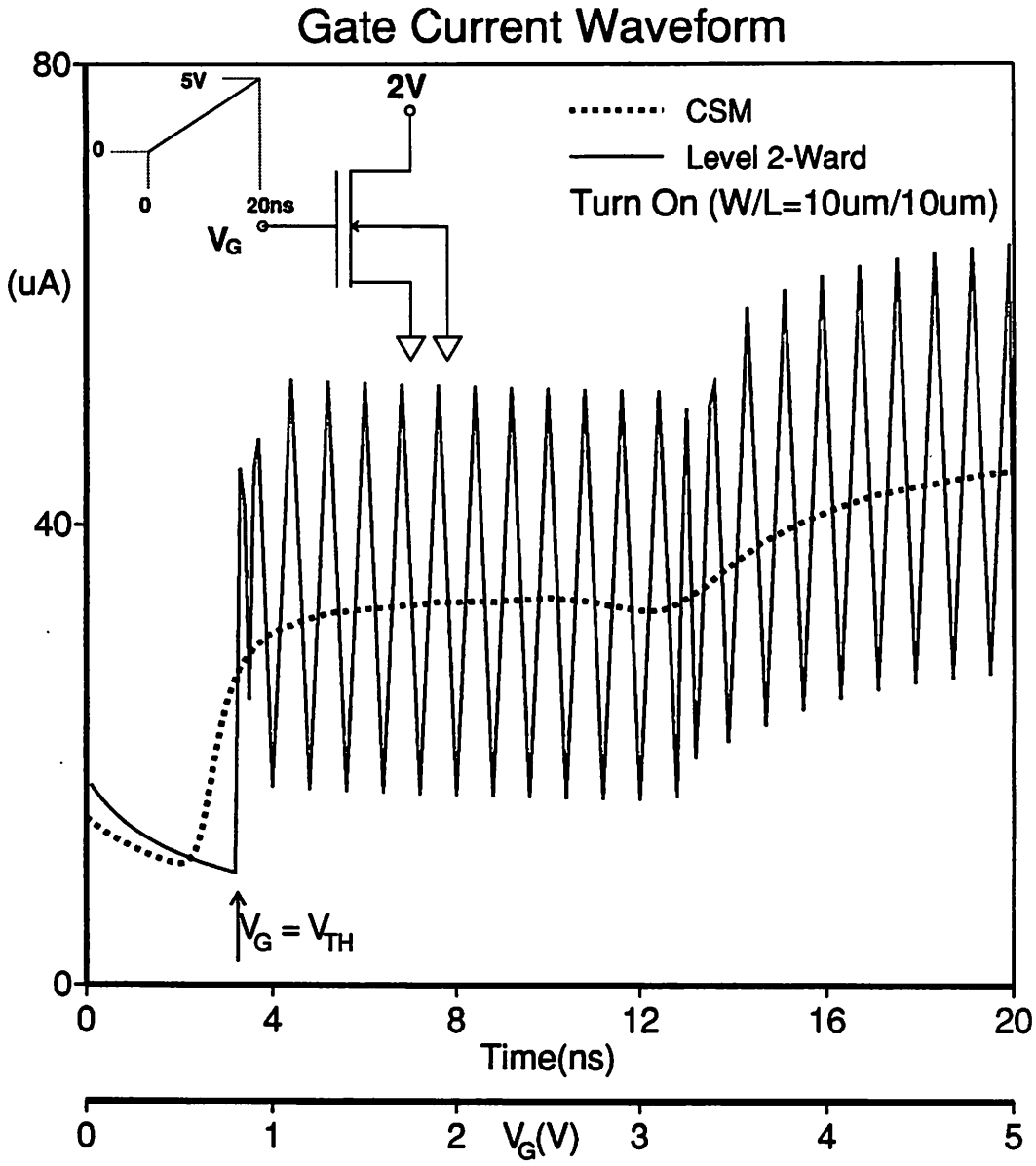


Fig.2.17.(a). Comparison of the gate current waveforms during the turn-on transient of a NMOSFET, simulated using this work(CSM) and the SPICE level 2 charge based model (Level 2-Ward).

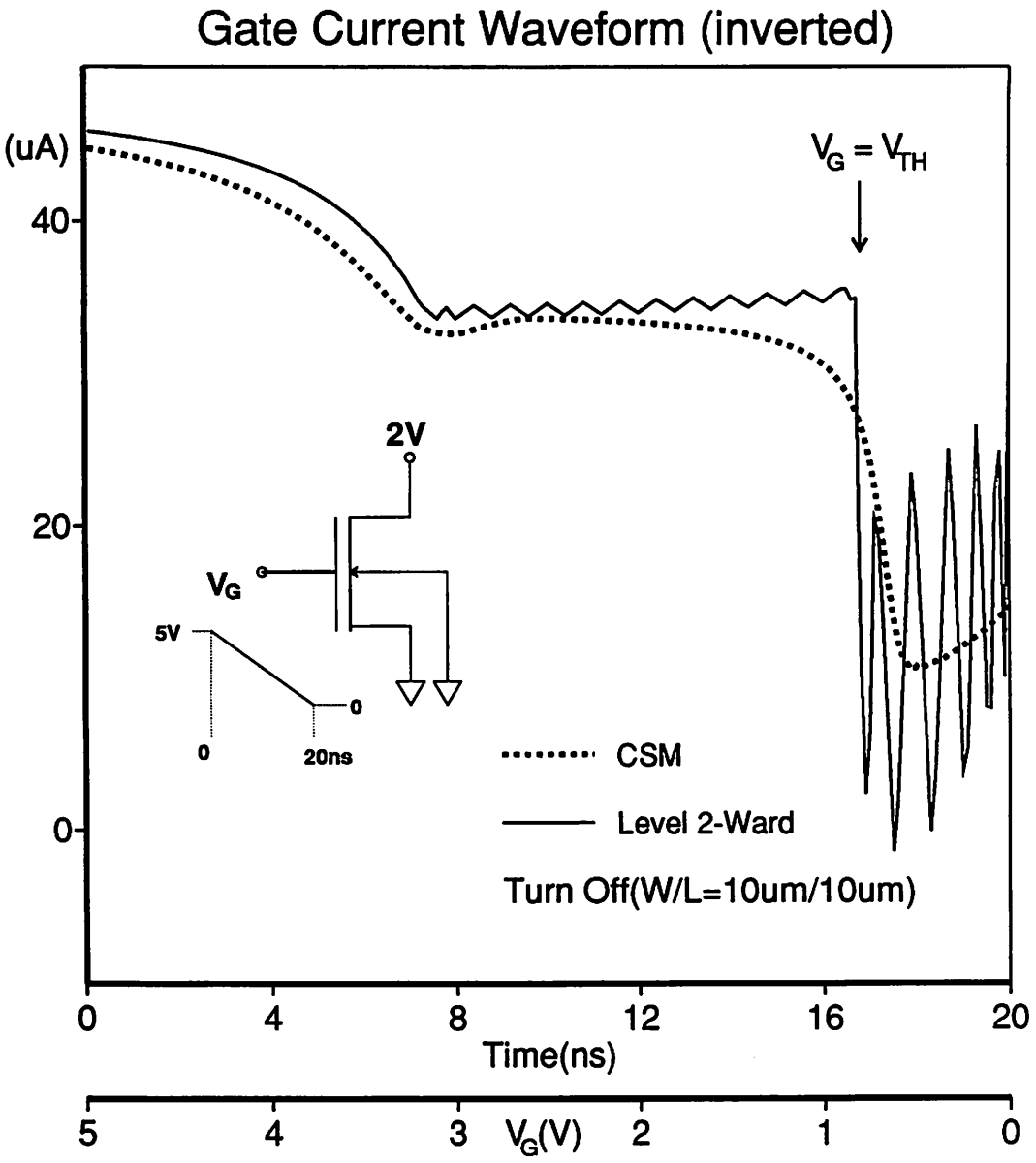


Fig.2.17.(b). Comparison of the gate current waveforms during the turn-off transient of a NMOSFET, simulated using this work(CSM) and the SPICE level 2 charge based model (Level 2-Ward).

[2.14] available in SPICE3. 40/60 and 0/100 of BSIM models refer to the ratios of drain and source charges in saturation region respectively. Table 2.1. shows the comparison of the CPU time per model computation. This work takes about twice longer than the conventional MOSFET models in SPICE3.

model	CSM	Meyer	BSIM
CPU time per model	1.6ms	0.7ms	1.0ms

Table 2.1. Comparison of CPU time per model computation using the VAX 8800 running Ultrix. V.2.2. 'CSM' is the charge sheet MOSFET model(this work) and 'Meyer' is the SPICE level 2 with Meyer capacitance model.

Table 2.2, Table 2.3 and Table 2.4 show the comparison of run statistics of the transient analysis of an 11 stage E-D NMOS ring oscillator, an NMOS OP Amp. and the same NMOS OP Amp with node numbers of source and drain interchanged in SPICE input, respectively. 'CSM.V' refers to the CSM model combined with the conventional voltage limiting routine and 'CSM.C' refers to the CSM model combined with the current limiting routine whose details are shown in Appendix 5. 'BSIM0' and 'BSIM1' refer to BSIM 40/60 (xpart=0) and BSIM 0/100 (xpart=1) models respectively.

In the NMOS ring oscillator example as shown in Table 2.2, the Meyer model gives too many transient iterations compared to other models.

In the NMOS OP Amp example as shown in Table 2.3, 'CSM.C', the CSM model using the current limiting routine, gives the least DC iterations. Both 'CSM.V' and 'CSM.C' give the number of iterations which is much less than that of 'Meyer' or 'BSIM0' model and is slightly larger than that of 'BSIM1' model.

When the node numbers of source and drain of NMOS OP Amp in the SPICE input, are reversed, the charge sheet model ('CSM.V', 'CSM.C') give much less transient iterations

than other models, as shown in Table 2.4.

From the preceding observations, we can see that this work(CSM) gives less iterations and hence better convergence property than other models due to its continuity property of current and charges and their derivatives with respect to biases, and also that the total CPU time of this work is comparable. to or even less than the other models. Also we can observe that the current limiting routine ('CSM.C') combined with the CSM model can be practically used as a complementary method for the conventional voltage limiting routine.

Table 2.5 shows the run statistics of the AC analysis of an NMOS OP Amp. 'CSM.V' gives less DC OP iterations than other models.

The SPICE input listings for the examples in Tables 2.2, 2.3, and 2.5 are shown in Appendix 6.

model	CSM.V	CSM.C	Meyer	BSIM0	BSIM1
Total CPU Time(s)	81	92	105	72	59
DC Iterations	9	31	11	17	17
Tran. Iterations	2160	2217	5530	2831	2437
Total Time Pts.	527	533	1265	750	645
Reject. Time Pts.	162	164	392	254	185
Load Time(s)	75	84	82	63	50

Table 2.2. Run statistics of the transient analysis of an 11 stage E-D NMOS ring oscillator.

model	CSM.V	CSM.C	Meyer	BSIM0	BSIM1
Total CPU Time(s)	24	23	15	16	15
DC Iterations	7	4	7	7	7
Tran. Iterations	401	401	432	409	396
Total Time Pts.	145	145	150	147	146
Reject. Time Pts.	16	16	17	18	15
Load Time(sec)	20	20	10	13	12

Table 2.3. The run statistics of the transient analysis of a NMOS OP Amp.

model	CSM.V	CSM.C	Meyer	BSIM0	BSIM1
Total CPU Time(s)	24	24	16	18	18
DC Iterations	7	4	7	7	7
Tran. Iterations	428	428	525	484	486
Total Time Pts.	137	137	150	150	151
Reject. Time Pts.	16	16	14	18	20
Load Time(sec)	20	20	11	14	14

Table 2.4. The run statistics of the transient analysis of the NMOS OP Amp, which is the same circuit as that in Table 2.3 but node numbers of source and drain in the SPICE input are interchanged.

model	CSM.V	CSM.C	Meyer	BSIM0	BSIM1
Total CPU Time(s)	2.6	5.2	2.1	3.4	3.4
DC OP Iterations	24	93	25	97	97
Load Time(sec)	0.9	3.0	0.4	1.3	1.3

Table 2.5. The run statistics of the AC analysis of the NMOS OP Amp., which is basically the same circuit as that in Table 2.3 but it is an open-loop configuration in this Table while it is a closed-loop unity gain configuration in Table 2.3.

2.7. Conclusion

An analytic charge sheet capacitance model of short channel MOSFETs has been derived and implemented in SPICE3. No iterations are needed to find the surface potential. The model equations are charge-based and includes the drift velocity saturation effect, the diffusion current, the effect of the bulk charge in analytic derivation steps. Also the channel length modulation effect on charges, and the channel side fringing field capacitances are included in the model as semi-empirical terms. The DC current, charge and their first and second derivatives are continuous under all operating regions. This continuity property improves the convergence property in circuit simulations.

An automatic gate capacitance measurement system has been set up to extract model parameters. Comparison of this model with the measured gate capacitances of short channel MOSFETs shows excellent agreements.

Comparison of the run statistics between this model implemented in SPICE3 and other models in SPICE3 shows that this model gives much less iterations than other models, due to the continuity property of this model.

Implementation details, the algorithm to find surface potentials for MOSFETs with non-uniform substrate doping concentration are shown in Appendices 1, 2, 3, 4, 5, 6, and 7.

References

[2.1] H.J.Park, P.K.Ko, C.Hu, "A Measurement Based Charge Sheet Capacitance Model of Short Channel MOSFETs for SPICE", IEEE IEDM 86 Technical Digest, pp.40-43, Dec. 1986

[2.2] B.J.Sheu, P.K.Ko, F.C.Hsu, "Characterization of Intrinsic Capacitances of Small-Geometry MOSFET's", IEEE 1984 Symposium on VLSI Technology Tech. Digest, pp.80-81, Sept. 1984

[2.3] H.Iwai, J.E.Oristian, J.T.Walker, R.W.Dutton, "A Scalable Technique for the Measurement of Intrinsic MOS Capacitances with Atto-Farad Resolution", IEEE J. of Solid State Circuits, Vol.SC-20, No.1, pp.264-276, Feb. 1985

[2.4] J.R.Brews, "A Charge-Sheet Model of the MOSFET", Solid State Electronics, Vol.21, pp.345-355, 1978

[2.5] F. Van de Wiele, "A Long-Channel MOSFET Model", Solid State Electronics, Vol.22, pp.991-997, 1979

[2.6] L.O.Chua, P.M.Lin, "Computer Aided Analysis of Electronic Circuits", Prentice Hall, 1975

[2.7] H.C.Pao, C.T.Sah, "Effects of Diffusion Current on Characteristics of Metal - Oxide (Insulator) - Semiconductor Transistors", Solid State Electronics, Vol.9, No.10, pp.927-937, 1966

[2.8] C.Turchetti, G.Masetti, Y.Tsividis, "On the Small-Signal Behavior of the MOS Transistor in Quasistatic Operation", Solid State Electronics, Vol.26, No.10, pp.941-949, 1983

[2.9] G.Merckel, J.Borel, N.Z.Cupcea, "An Accurate Large-Signal MOS Transistor Model for Use in Computer-Aided-Design", IEEE Trans. Electron Devices, Vol.ED-19,

No.5, pp.681-690, May 1972

[2.10] H.C.Poon, L.D.Yau, R.L.Johnston, D.Beecham, "DC Model for Short Channel IGFETs", IEEE IEDM 73 Technical Digest, pp.156-159, 1973

[2.11] S.Sze, "Physics of Semiconductor Devices", Ch.7, 2nd Ed., 1981

[2.12] C.F.Gerald, "Applied Numerical Analysis", 2nd Ed., Addison-Wesley, 1978

[2.13] R.H.Bartels, J.C.Beatty, B.A.Barsky, "An Introduction to Splines for use in Computer Graphics and Geometric Modeling", Ch.3, Morgan Kaufmann Publishers, Inc., 1987

[2.14] B.J.Sheu, D.L.Scharfetter, P.K.Ko, M.C.Jeng, "BSIM: Berkeley Short Channel IGFET Model for MOS Transistors", IEEE J. of Solid State Circuits, Vol.SC-22, No.4, pp.558-566, Aug. 1987

[2.15] A.G.Sabnis, J.T.Clemens, "Characterization of the Electron Mobility in the Inverted <100> Si Surface", IEEE International Electron Devices Meeting, Technical Digest, pp. 18-21, 1979

[2.16] A.Vladimirescu, S.Liu, "The Simulation of MOS Integrated Circuits using SPICE2", Electronics Research Lab., University of California, Berkeley, ERL Memo M80/7, 1980

[2.17] S.Y.Oh, D.E.Ward, R.W.Dutton, "Transient Analysis of MOS Transistors", IEEE Trans. Electron Devices, Vol.ED-27, No.8, pp.1571-1578, 1980

[2.18] R.Shrivastava, K.Fitzpatrick, "A Simple Model for the Overlap Capacitance of a VLSI MOS Device", IEEE Trans. Electron Devices, Vol.ED-29, No.12, pp.1870-1872, 1982

[2.19] R.S.Muller, T.I.Kamins, "Device Electronics for Integrated Circuits", John Wiley & Sons, 1977

[2.20] L.W.Nagel, "SPICE2: A Computer program to Simulate Semiconductor Circuits", ERL-Memo, M520, 9 May 1975, Electronics Research Lab., University of California, Berkeley, CA 94720

[2.21] R.S.C.Cobbold, "Theory and Application of FET's", New York: Wiley, 1970, pp.272-304

[2.22] J.E.Meyer, "MOS Models and Circuit Simulation", RCA Review, Vol.32, pp.42-63, 1971

[2.23] D.E.Ward, R.W.Dutton, "A Charge-Oriented Model for MOS Transistor Capacitances", IEEE J. of Solid State Circuits, Vol.SC-13, No.5, pp.703-708, 1978

[2.24] H.Iwai, M.R.Pinto, C.R.Rafferty, J.E.Oristian, R.W.Dutton, "Analysis of Velocity Saturation and Other Effects on Short-Channel MOS Transistor Capacitances", IEEE Trans. on Computer Aided Design, Vol.CAD-6, No.2, pp.173-184, Mar. 1987

[2.25] B.J.Sheu, P.K.Ko, "Measurement and Modeling of Short-Channel MOS Transistor Gate Capacitances", IEEE J. of Solid State Circuits, Vol.SC-22, No.3, pp.464-472, June 1987

Chapter 3

A SIMPLE NON-QUASISTATIC MOSFET MODEL FOR TRANSIENT ANALYSIS

3.1. Introduction

MOSFET has been widely used in the digital VLSI circuits as well as in the analog circuits[3.1]. And circuit simulation programs, such as SPICE, have been used extensively to predict the circuit performance. But the MOSFET device models available in the circuit simulation programs are not adequate in some applications, as pointed out by some researchers [3.2] [3.3] [3.4].

One of the problems is that the quasistatic small-signal conductance and capacitances are used for the large-signal transient analysis. This gives erroneous simulation results for signals whose rise or fall times are comparable to or smaller than the channel transit time. The small-signal frequency response has similar limitations at high frequencies.

Another problem of the present charge based MOSFET capacitance models is the uncertainty in the channel charge partitioning scheme [3.5] [3.6]. The gate, bulk and channel charge can be derived analytically. But to find the transient drain and source currents, the drain and the source charge components have to be evaluated separately. Oh et. al. derived the analytic channel charge partitioning scheme based on the quasistatic approximation [3.2]. This gives the 40/60 partitioning between the drain and the source charges respectively, in the saturation region. But this scheme produces anomalous results for the fast transient and the high frequency AC analyses. Other channel charge partitioning schemes, such as 0/100(TI [3.7],BSIM [3.8]) , 50/50(BSIM) and XQC(SPICE [3.9]), have been devised but none is valid under all operating conditions. Some examples of

their failings will be presented later. The ratios 0/100 and 50/50 refer to the ratio of the drain and source charges in the saturation region.

The need for a non-quasistatic MOSFET model has been pointed out by several researchers [3.2] [3.4]. Turchetti et. al.[3.10] solved the continuity equation for the transient analysis for an assumed QS(quasistatic) carrier density profile which is linear between source and drain plus a NQS(non-quasistatic) carrier density profile which is symmetrical with respect to source and drain. But the exact QS carrier density profile, which will be shown later, is not linear profile between source and drain. Due to this assumed linear profile the total QS channel charge becomes only 3/4 of the exact value for the slow signals in saturation region. Due to the assumed symmetrical NQS carrier density profile the NQS source charge component is the same as the NQS drain charge component under all conditions. This trivializes the NQS problem.

Bagheri et. al.[3.11] derived a NQS AC model based on the charge sheet model. In [3.11], iterations are used to find the surface potentials and also to solve the continuity equation. Also, the model in [3.11] cannot be directly used in the large signal transient analysis because it is based on the small signal AC analysis.

In this work, the current continuity equation is approximated to a diffusion equation with a time and position dependent diffusion coefficient. Simple analytic expressions have been derived for the drain and source currents of long channel MOSFETs without first computing the drain or the source charges. No channel charge partitioning scheme is needed at all. Similarly simple expressions have been derived for the small-signal AC analysis. The small-signal AC model is shown in Chapter 4.

This model has been implemented into SPICE3 based on the charge sheet formulation [3.12] and the required CPU time per iteration, i.e. per model computation, is two to three times that required for the conventional quasistatic MOSFET models in SPICE3. Many other works are reported on the charge sheet model [3.13] [3.14] [3.15], but [3.12] has been chosen to reduce the CPU time.

But it turns out that this model does not conserve charges. A charge conserving NQS model which is an improved version of this model is shown in Chapter 5.

Section 3.2 shows the formulation of model equations. A diffusion equation has been derived from the current continuity equation and the current relation. Analytic equations for node currents are derived from the solution of this diffusion equation.

Section 3.3 shows the simulation results and the comparison with other models, for the turn-on and the turn-off transients of a NMOSFET, the large signal channel current partitioning ratio, the turn-on transient of a NMOS inverter, the comparison with the measured data, and a NMOS ring oscillator.

Section 3.4 concludes this chapter.

3.2. Formulation of model equations

3.2.1. Derivation of diffusion equation

The channel current I_y can be written as the sum of the drift and the diffusion components following the charge sheet formulation described in [3.13].

$$I_y(y,t) = WC_{OX} \cdot \left\{ -\mu_n Q_n'(y,t) \frac{\partial \Psi_S(y,t)}{\partial y} + D_n \frac{\partial Q_n'(y,t)}{\partial y} \right\} \tag{3.1}$$

where W is the effective channel width, C_{OX} is the gate oxide capacitance per unit area. μ_n is the carrier mobility, Q_n' is the channel charge density normalized by $(-WC_{OX})$ and Ψ_S is the surface potential referenced to the bulk and D_n is the carrier diffusion constant. y is the lateral dimension from the source toward the drain. The direction of positive I_y is from the source toward the drain as shown in Fig.3.1.

The normalized channel charge density Q_n' can be written as. [3.11]

$$Q_n'(y,t) = V_{GB}(t) - V_{FB} - \Psi_S(y,t) - \gamma\sqrt{\Psi_S(y,t) - V_t} \tag{3.2}$$

where V_{GB} is the applied voltage between gate and bulk. V_{FB} is the flat band voltage,

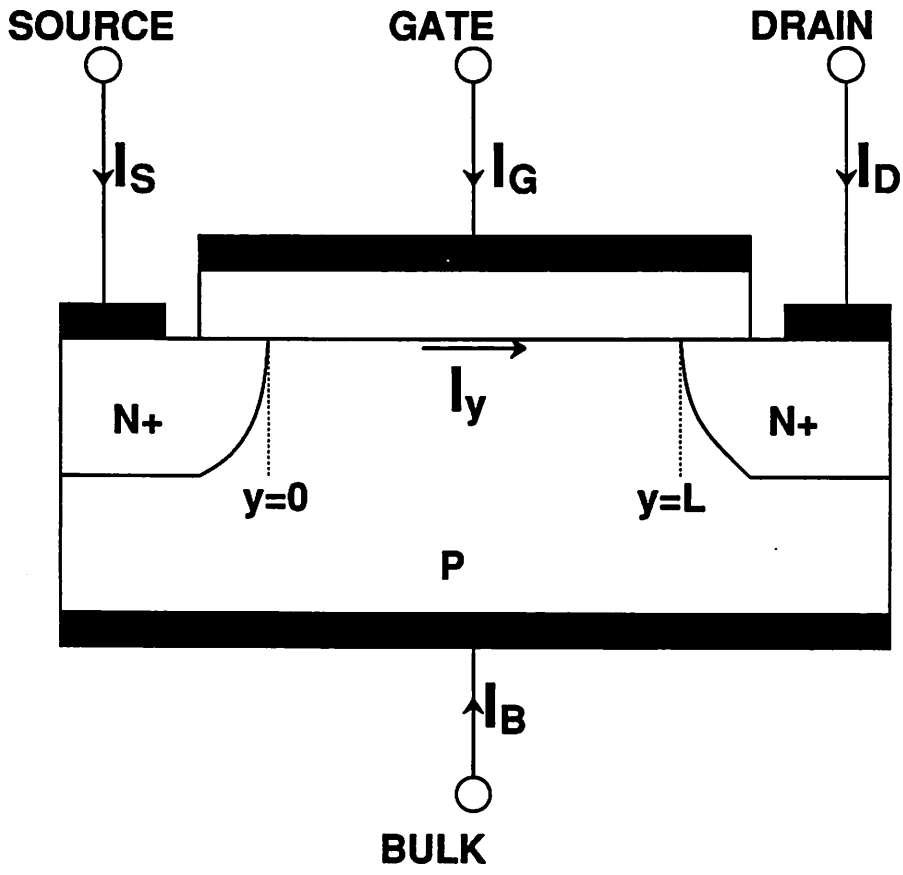


Fig.3.1. A cross section of NMOSFET in the channel length direction to show the polarities of currents. y is the lateral dimension from source toward drain.

V_t is the thermal voltage kT/q.

The square root term in (3.2) is commonly linearized with respect to Ψ_S for simplicity and efficiency of model equations. Hence Q_n' can be approximated as

$$Q_n'(y,t) = V_{GST}(t) - F_B \cdot (\Psi_S(y,t) - \Psi_{SO}(t)) \tag{3.3}$$

$$\textit{where} \quad V_{GST}(t) = V_{GB}(t) - V_{FB} - \Psi_{SO}(t) - \gamma\sqrt{\Psi_{SO}(t) - V_t} \tag{3.4}$$

$$F_B = 1 + \frac{0.5\gamma}{\sqrt{\Psi_{SO}(t) - V_t}} \cdot \left\{ 1 - \frac{1}{1.744 + 0.8364(\Psi_{SO}(t) - V_t)} \right\} \tag{3.5}$$

F_B is the correction factor for the bulk charge, Ψ_{SO} is the surface potential at the source end of the channel and γ is the conventional bulk effect coefficient. Although F_B has a slight time dependence, the argument (t) is not attached to F_B for simplicity. F_B has a weak dependence on applied voltages and its value is slightly larger than 1.0 and ranges between 1.0 and 1.2. Although there are several models available for the correction factor F_B ([3.16] [3.17] [3.9] [3.7] [3.8] [3.11] [3.10] [3.18]), the equation derived in [3.17] has been used in this work. The correction factor F_B in [3.17] was derived for the strong inversion region. The surface potential, $(2\Phi_F - V_{BS})$, in [3.17] is replaced by Ψ_{SO} in this work to extend the usage of F_B into the weak inversion region. Although this direct extension may give some errors, the effect of this error on the overall accuracy is expected to be small, because $(\Psi_S(y,t) - \Psi_{SO}(t))$ in (3.3) is small in the weak inversion region so that the accuracy of F_B is relatively unimportant in this region and that the weak inversion region contributes little to the channel charge and the transient current.

V_{GST} is equivalent to $(V_{GS} - V_{TH})$ in the conventional notation and has a non-negative value. Using (3.3) and the fact that F_B is independent of y, (3.1) can be rewritten as

$$I_y(y,t) = W\ C_{OX} \cdot \frac{\mu_n}{2F_B} \cdot \frac{\partial}{\partial y}(Q_n'(y,t) + F_B V_t)^2 \tag{3.6}$$

Substituting (3.6) into the current continuity equation (3.7) and assuming that μ_n has no y dependence, ie. neglecting the velocity saturation effect, we have (3.8).

$$\frac{\partial I_y(y,t)}{\partial y} = WC_{OX}\cdot\frac{\partial Q_n'(y,t)}{\partial t} \tag{3.7}$$

$$\frac{\partial Q_n'(y,t)}{\partial t} = \frac{\mu_n}{2F_B}\cdot\frac{\partial^2}{\partial y^2}(Q_n'(y,t)+F_B V_t)^2 \tag{3.8}$$

By further assuming (3.9) and multiplying both sides of the equation (3.8) by $(Q_n'(y,t)+F_B V_t)$, we get the diffusion equation (3.10) with the time and position dependent diffusion coefficient, $D(y,t)$. Eq. (3.9) is always valid except in the cutoff regime, where Q_n' becomes 0 but F_B changes with time. However the effect of this exception on the overall accuracy of the NQS model is expected to be small.

$$\left| \frac{\partial Q_n'(y,t)}{\partial t} \right| >> V_t \left| \frac{\partial F_B}{\partial t} \right| \tag{3.9}$$

$$\frac{\partial P(y,t)}{\partial t} = D(y,t)\cdot\frac{\partial^2 P(y,t)}{\partial y^2} \tag{3.10}$$

where $$P(y,t) = (Q_n'(y,t) + F_B\ V_t)^2 \tag{3.11}$$

$$D(y,t) = \frac{\mu_n}{F_B}\cdot(Q_n'(y,t) + F_B V_t) \tag{3.12}$$

3.2.2. Solution of the diffusion equation

To solve the diffusion equation (3.10), we can establish two boundary conditions for $P(y,t)$, $P_S(t)$ and $P_D(t)$ at the source and drain ends of the channel respectively, by assuming that the carrier densities at those positions respond instantaneously to the applied bias.

$$P_S(t) = (V_{GST}(t) + F_B V_t)^2 \tag{3.13}$$

$$P_D(t) = (V_{GST}(t) + F_B V_t - F_B(\Psi_{SL}(t)-\Psi_{SO}(t)))^2 \tag{3.14}$$

Ψ_{SO} and Ψ_{SL} are surface potentials at the source and drain ends of the channel respectively. Ψ_{SO} is computed using the charge sheet formulation [3.12]. The charge sheet formulation in [3.12] is derived from the original charge sheet approximation [3.13]. In [3.12], a small part of the surface potential curve versus $(V_{GB}-V_{FB})$ is stored in

computer memory in the form of a cubic spline function. Hence by referring the cubic spline function, no iteration is needed to find the surface potential. Ψ_{SL} is computed from

$$\Psi_{SL}(t) = \Psi_{SO}(t) + VDS(t) \tag{3.15}$$

where the effective drain to source voltage VDS is computed from the applied drain to source voltage V_{DS} and the drain saturation voltage, V_{DSSAT}, using the smoothing cubic spline function [3.12]. VDS is close to V_{DS} in linear region and goes asymptotically to V_{DSSAT} in saturation region. VDS changes smoothly between linear and saturation regions with the continuity of up to the second derivatives with respect to applied biases. V_{DSSAT} is derived from the asymptotic form of the short channel V_{DSSAT} with the velocity saturation effect [3.12], by assuming an infinite saturation velocity.

$$V_{DSSAT}(t) = \frac{V_{GST}(t)}{F_B} + V_t \tag{3.16}$$

This V_{DSSAT} includes the effect of the diffusion current through the thermal voltage term V_t [3.12]. $P(y,t)$ is decomposed into the QS and the NQS components.

$$P(y,t) = P_S(t) - (P_S(t) - P_D(t))\cdot\frac{y}{L} + \sum_{n=1}^{N} A_n(t)\cdot\sin(n\pi\frac{y}{L}) \tag{3.17}$$

The first two terms in the right hand side of (3.17) are QS components and the third term is the NQS component. The QS components can be derived from the diffusion equation (3.10) by setting the time derivative to be zero. Since the NQS component is 0 at drain and source ends, any NQS profile can be represented as a Fourier sine series expansion. From (3.11),(3.12) and (3.17), the diffusion coefficient $D(y,t)$ can be rewritten as

$$D(y,t) = \frac{\mu_n}{F_B}\cdot\sqrt{P_S(t) - (P_S(t) - P_D(t))\cdot\frac{y}{L} + \sum_{n=1}^{N} A_n(t)\cdot\sin(n\pi\frac{y}{L})} \tag{3.18}$$

$D(y,t)$ is approximated as a Fourier cosine series expansion with two terms.

$$D(y,t) = D_0(t) + D_1(t)\cdot\cos(\pi\frac{y}{L}) \tag{3.19}$$

The following steps deviate from an earlier version of this work described in [3.19]. First we decompose (3.18) into QS and NQS components. The NQS component is derived as a perturbation term by assuming that the coefficients $\{A_n\}$ are small compared to P_S. This assumption is introduced to simplify the analysis and is valid under all operating conditions except weak inversion and cutoff regions, which contributes little to channel charge and transient current. $D_0(t)$ is derived by matching the integration of $D(y,t)$ from y=0 to y=L in (3.18) and (3.19).

$$D_0(t) = \frac{\mu_n}{F_B} \cdot \left\{ \frac{2}{3} \cdot \frac{(P_S^{1.5} - P_D^{1.5})}{P_S - P_D} + \frac{1}{2\pi} \cdot \frac{5 - \frac{P_D}{P_S}}{\sqrt{P_S} + 0.5} \cdot \sum_{m=1}^{N/2} \frac{A_{2m-1}}{2m-1} \right\} \tag{3.20}$$

where the argument (t) is dropped from $P_S(t)$, $P_D(t)$ and $A_{2m-1}(t)$ in (3.20) for simplicity. The term, 0.5 (V), in the denominator of the NQS term in (3.20) is a heuristic factor to prevent a blow up at the cutoff where P_S is close to zero. $D_1(t)$ is derived by matching the source diffusion coefficient of (3.19) to the exact value, $D(0,t)$ in (3.18).

$$D_1(t) = \frac{\mu_n}{F_B} \cdot \sqrt{P_S(t)} - D_0(t) \tag{3.21}$$

$D_1(t)$ is limited to be less than or equal to $D_0(t)$, so that the diffusion coefficient $D(y,t)$ in (3.19) is non-negative all along the channel region. This guarantees the numerical stability of the solution method.

In this derivation, the diffusion coefficient approximation is chosen to be more accurate near the source side of the channel where the channel charge density is large than near the drain side of the channel where the channel charge density is smaller.

To show the validity of the approximations, (3.20) and (3.21), the approximate and the exact diffusion coefficients are shown in Fig.3.2.(a) for P_D/P_S= 0, 0.5 and 1.0 . In Fig.3.2.(a), only the QS component is considered and in Fig.3.2(b) the NQS component is included with A_1=0.1V^2. The error between the exact and the approximate ones is less than 10% of the source diffusion coefficient over 90% of the channel for all cases considered. The remaining 10% is the drain-side channel where the inversion carrier

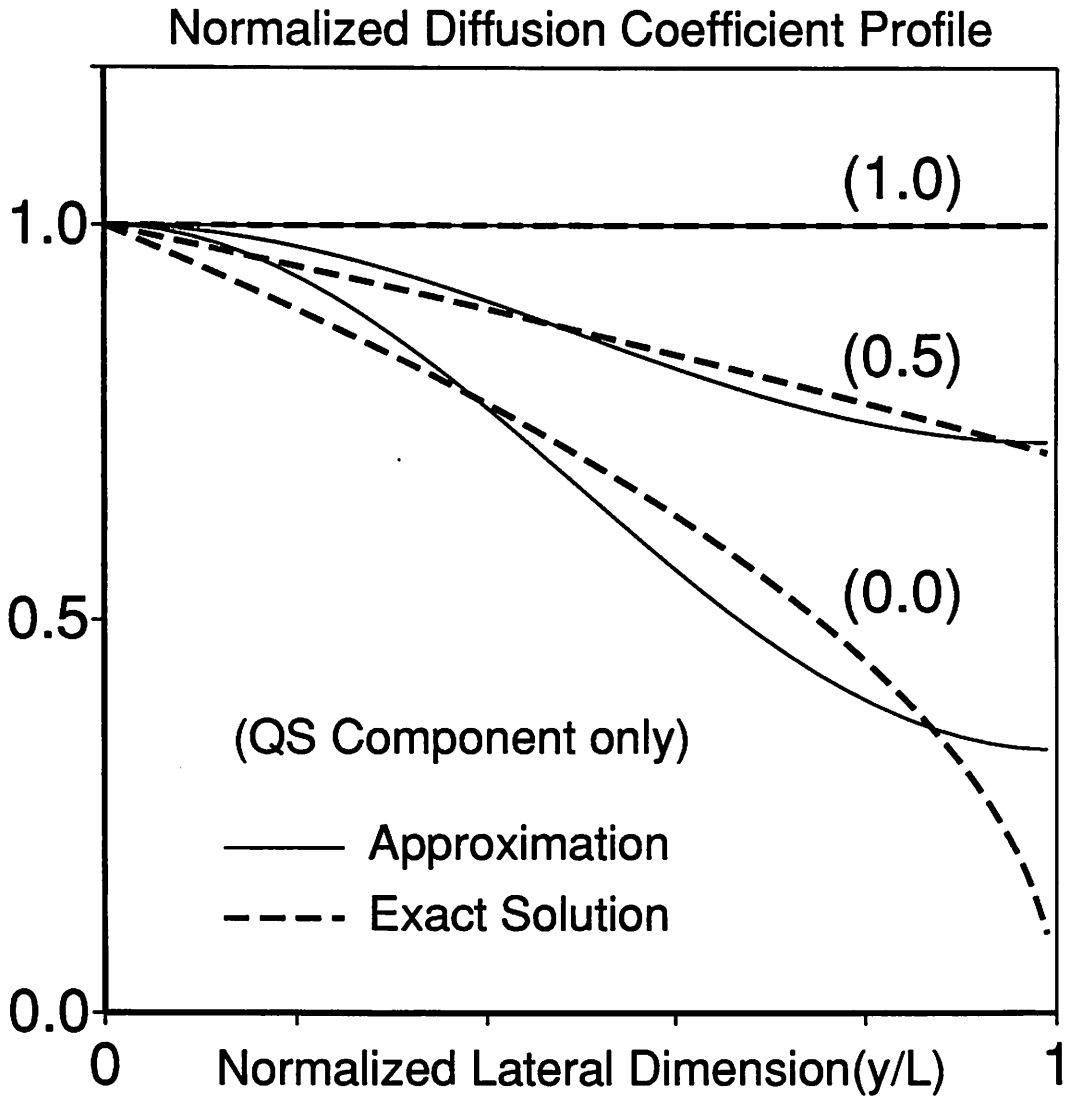


Fig.3.2. The approximation of the diffusion coefficient in Eq. (3.18) and (3.19). Diffusion Coefficient profiles are normalized with respect to the source diffusion coefficient.

(a). The approximate (3.19) and the exact values (3.18) of the diffusion coefficient for the QS case (A_i=0), with P_D/P_S = 0, 0.5 and 1.0. (0.0), (0.5) and (1.0) in the figures represent the ratio P_D/P_S.

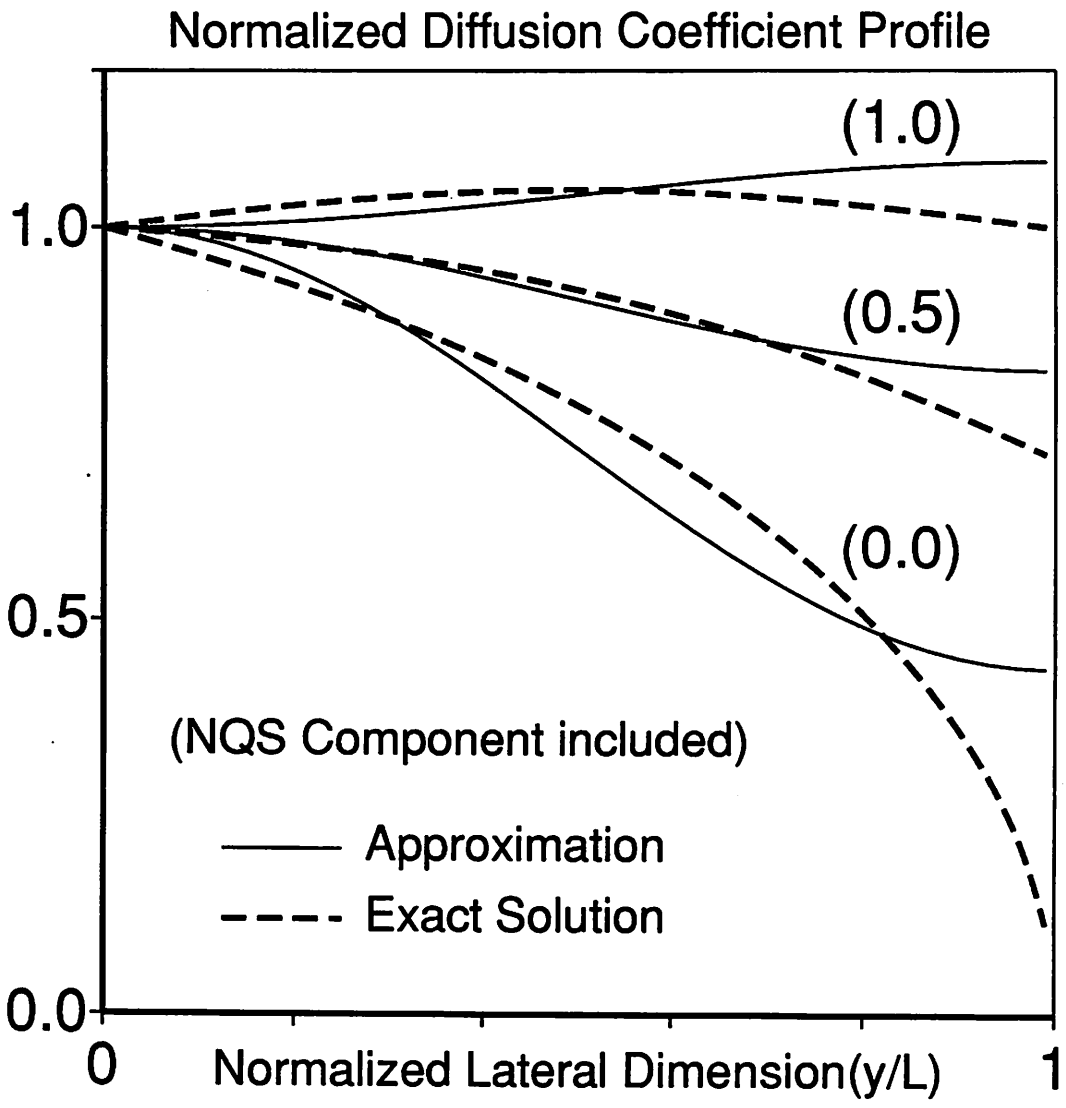


Fig.3.2.(b). The approximate (3.19) and the exact values (3.18) of the diffusion coefficient with the NQS component included, with P_D/P_S = 0, 0.5, 1.0 ($A_1 = 0.1V^2$ and all other A_i's are zero and $P_S = 1V^2$.)

density falls below one third of the source carrier density and is relatively less important than the other part.

The error can be reduced by increasing the number of terms in the Fourier cosine series expansion (3.19). But the number of terms in (3.19) is critical to the computational efficiency of the whole numerical scheme, which will be explained later. And so the number of terms in the Fourier cosine series expansion is limited to two as in (3.19) in this work. This assumption is further justified by comparing the results with PISCES

Substituting (3.17) and (3.19) into the diffusion equation (3.10), we have the state equation (3.22).

$$\frac{d\mathbf{A}(t)}{dt} = \mathbf{D}(t)\cdot\mathbf{A}(t) + \frac{dP_D(t)}{dt}\mathbf{C_D} + \frac{dP_S(t)}{dt}\mathbf{C_S} \tag{3.22}$$

$\mathbf{A}$ is a column matrix for the coefficients $\{A_n\}$, $\mathbf{D}$ is a tri-diagonal matrix and $\mathbf{C_D}$ and $\mathbf{C_S}$ are constant column matrices accounting for the excitations at drain and source ends respectively. $\mathbf{D}\cdot\mathbf{A}$ is the history term and the last two terms in the right hand side of (3.22) are excitation terms. This state equation (3.22) has been solved using the trapezoidal integration method. The coefficients $\{A_n\}$ and their derivatives with respect to the applied biases can be found simultaneously in one solution step. If the number of terms in the Fourier cosine series expansion (3.19) is larger than two, the matrix $\mathbf{D}$ is no longer a tri-diagonal matrix and the full Gaussian elimination method is needed to solve that. In this case, the CPU time of the whole model evaluation time is 2.2 times that of the tri-diagonal case (Vax 8800 with Ultrix V2.0). The detailed derivation and the solution steps of (3.22) are shown in Appendix 8.

3.2.3. Current equations

Using (3.11), the channel current equation (3.5) can be rewritten as

$$I_y(y,t) = WC_{OX}\frac{\mu_n}{2F_B}\cdot\frac{\partial P(y,t)}{\partial y} \tag{3.23}$$

Using (3.17) and (3.23), we can get the analytic equations for the transient drain and source currents.

$$I_D(t) = \frac{W}{L} C_{OX} \frac{\mu_n}{2F_B} \left\{ P_S(t) - P_D(t) - \sum_{n=1}^{N} (A_n(t) \cdot (-1)^n \cdot n\pi) \right\} \tag{3.24}$$

$$I_S(t) = -\frac{W}{L} C_{OX} \frac{\mu_n}{2F_B} \left\{ P_S(t) - P_D(t) - \sum_{n=1}^{N} (A_n(t) \cdot n\pi) \right\} \tag{3.25}$$

In the current SPICE implementation and the following examples, N is set to 10.

Including the channel length modulation effect, we can get the final equations (3.26) and (3.27) for I_D and I_S.

$$I_D(t) = I_{DC}(t) - \frac{W}{L} C_{OX} \frac{\mu_n}{2F_B} \cdot \sum_{n=1}^{N} (A_n(t) \cdot (-1)^n \cdot n\pi) \tag{3.26}$$

$$I_S(t) = -I_{DC}(t) + \frac{W}{L} C_{OX} \frac{\mu_n}{2F_B} \cdot \sum_{n=1}^{N} (A_n(t) \cdot n\pi) \tag{3.27}$$

$$\textit{where} \quad I_{DC}(t) = \frac{W}{L} C_{OX} \frac{\mu_n}{2F_B} (P_S(t) - P_D(t))(1 + \lambda V_{DS}(t)) \tag{3.28}$$

λ is the channel length modulation factor and V_{DS} is the applied drain to source voltage. A simple model is used for the channel length modulation in this work to concentrate on the NQS phenomenon. As shown in (3.26) and (3.27), I_D and I_S can be decomposed into the dc component, I_{DC}, and the transient component which is a linear combination of the state variables $\{A_n\}$.

The non-quasistatic behavior of the bulk current is neglected for simplicity and also because it is less important than the NQS drain or source currents. So the quasistatic bulk charge Q_B has been derived in (3.29) to get the bulk current I_B.

$$Q_B(t) = \int_0^L q_b(y,t)\,dy \tag{3.29}$$

$$q_b(y,t) = -WC_{OX} \cdot (\gamma\sqrt{\Psi_{SO}(t) - V_t} + (F_B - 1)(\Psi_S(y,t) - \Psi_{SO})) \tag{3.30}$$

where $q_b(y,t)$ is the bulk charge density per unit area. From (3.29) and (3.30) we can get the total bulk charge Q_B as

$$Q_B(t) = -WLC_{OX}\left\{\gamma\sqrt{\Psi_{SO}-V_t} + (1-\frac{1}{F_B})\cdot\left[V_{GST} + F_B V_t - \frac{2}{3}\cdot\frac{P_S{}^{1.5}-P_D{}^{1.5}}{P_S-P_D}\right]\right\} \quad (3.31)$$

For $(P_S-P_D)\approx 0$, the denominator becomes zero and Eq. (3.31) is converted to the asymptotic form. The bulk current $I_B(t)$ is computed from (3.32) using the trapezoidal integration scheme.

$$I_B(t) = \frac{dQ_B(t)}{dt} \quad (3.32)$$

The gate current $I_G(t)$ is derived from the other three current components as

$$I_G(t) = -(I_D(t) + I_S(t) + I_B(t)) \quad (3.33)$$

The derivatives of I_D, I_S, I_B and I_G with respect to V_G, V_D, V_S and V_B, which are needed for the Newton Raphson iterations in SPICE, have been derived analytically from the derivatives of P_S, P_D and $\{A_n\}$ with respect to biases.

3.2.4. Moving boundary condition

After the transistor is suddenly turned on, the carrier density is non-zero only near the source. 1-D numerical solution of (3.8) shows that the carrier density profile during this time interval is linear in y except for a small tail at the right end edge due to diffusion as shown in Fig.3.3.(a). The linear carrier density profile is a self-consistent solution because, given the linear carrier density profile, the surface potential becomes linear in y according to (3.2) and therefore the electric field is constant throughout the region of non- zero carrier density profile. Since drift transport is dominant over diffusion during this time period, the carriers move at the same velocity, assuming a constant mobility. Hence the carrier density profile remains linear in y at the next time step if V_{GST} changes linearly in time, as shown by a broken line in Fig.3.3.(b).

During this time period, I_D is of course zero. We call this period the moving boundary period and introduce a test condition to ensure $I_D=0$ until the end of the moving boundary period. Without such a test condition, there will be small but troublesome I_D's due to non-zero dQ_n/dy at $y=L$. This is due to the fact that we used a finite number of

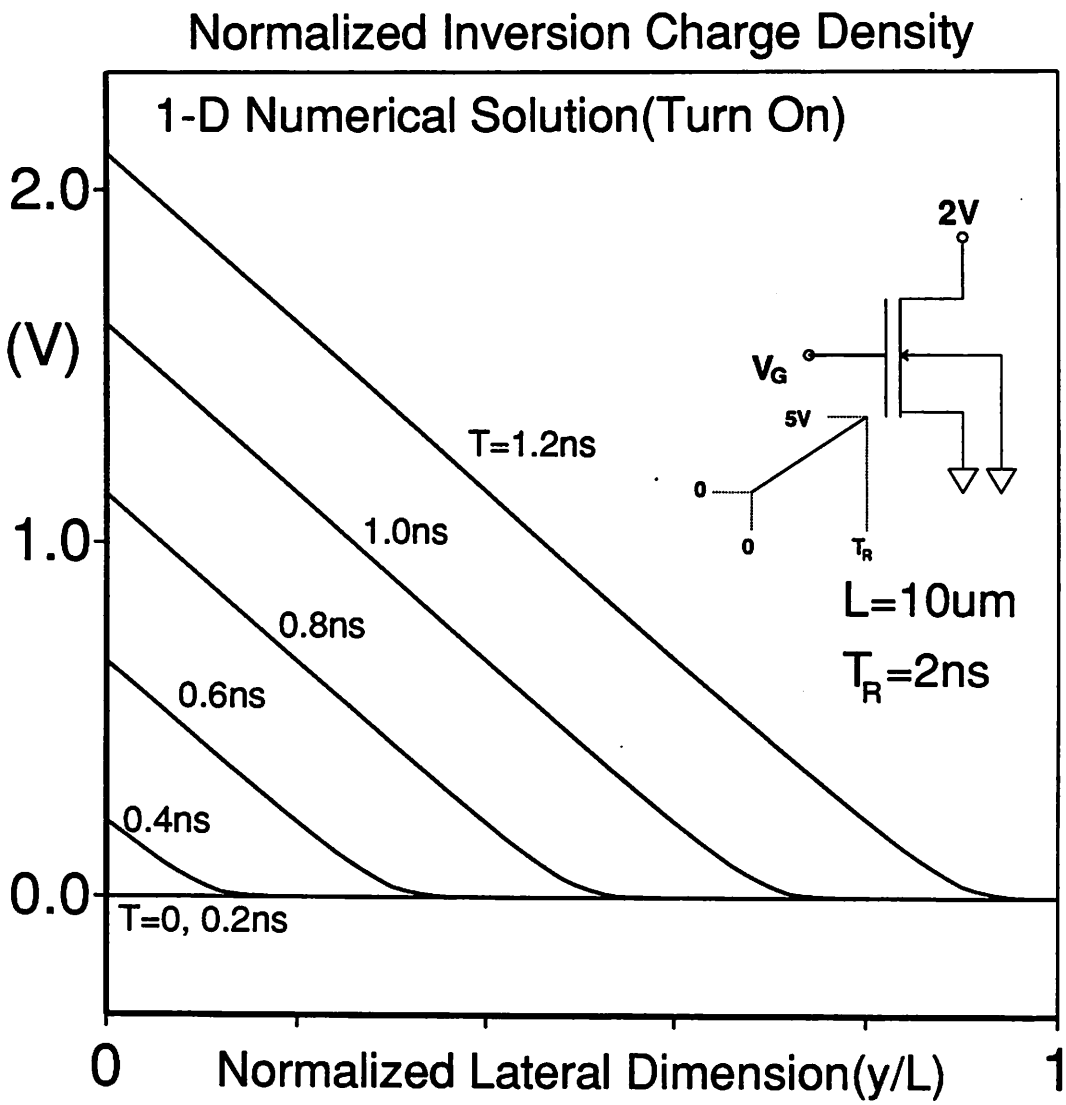


Fig.3.3. The approximation of the inversion charge density profile during the moving boundary period.

(a). The inversion charge density profile from the 1-D numerical solution of eq. (3.8). The profile is linear in y except in the small diffusion tail at the right hand edge.

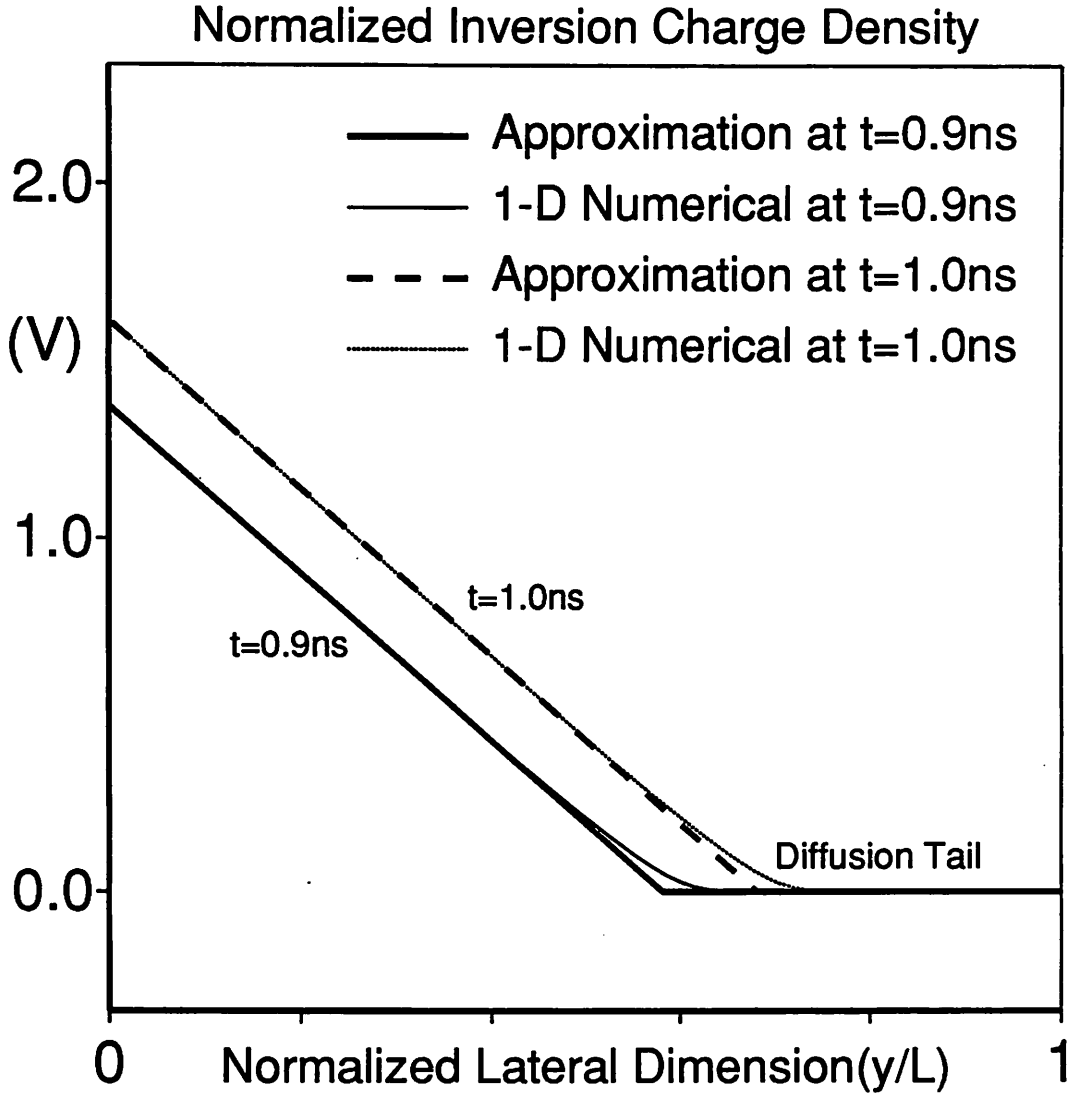


Fig.3.3.(b). Approximation of the inversion charge density profile into a linear profile, as shown by the thick and the broken lines. The profile from the 1-D numerical solution is also shown for comparison.

terms(10 in this work) in the Fourier sine series expansion (3.17) and it is difficult to represent the two section profiles shown in Fig.3.3.(a) accurately in terms of a Fourier series with finite terms.

To derive the moving boundary condition quantitatively, we assume that the normalized carrier density profile, $Q_n'(y,t)$, during this time period, is

$$Q_n'(y,t) = Q_0'(t) - Q_M'(t)\cdot y' \tag{3.34}$$

where y' is y/L and Q_M' is the slope of the profile. This assumption is suggested by the results of the 1-D numerical analysis and also by the reasonings discussed above.

Substituting (3.34) into (3.8), we have

$$\frac{dQ_0'(t)}{dt} - \frac{dQ_M'(t)}{dt}\cdot y' = \frac{\mu_n}{F_B L^2}\cdot(Q_M'(t))^2 \tag{3.35}$$

Assuming

$$\left| \frac{dQ_M'(t)}{dt} \right| \ll \left| \frac{dQ_0'(t)}{dt} \right| \tag{3.36}$$

(3.35) becomes

$$Q_M'(t) = \sqrt{\frac{F_B L^2}{\mu_n}\cdot\frac{dQ_0'(t)}{dt}} \tag{3.37}$$

Eq. (3.36) is equivalent to assuming $d^2V_{GST}/dt^2 \ll 2\cdot(dV_{GST}/dt)^{1.5}\cdot\sqrt{\mu_n/(F_B L^2)}$, i.e., that V_{GST} changes almost linearly with time during the moving boundary period. The moving boundary condition lasts, until the edge of the carrier density profile reaches the drain side , i.e.

$$Q_0'(t) < Q_M'(t) \tag{3.38}$$

From (3.37) and (3.38), we have

$$Q_0'(t) < \sqrt{\frac{F_B L^2}{\mu_n}\cdot\frac{dQ_0'}{dt}} \tag{3.39}$$

$$\textit{where} \quad Q_0'(t) = V_{GST}(t) \tag{3.40}$$

By substituting (3.40) into (3.39), we can get the moving boundary condition as follows.

If

$$\frac{dV_{GST}}{dt} > \frac{\mu_n \cdot V_{GST}^2}{F_B L^2} = \frac{V_{GST}}{\tau_T} \tag{3.41}$$

and

$$\Delta P_D < (F_B V_t)^2 \tag{3.42}$$

Then

$$I_D(t+\Delta t) = I_D(t) \tag{3.43}$$

where τ_T is the channel transit time. If the conditions (3.41) and (3.42) are met, the drain current is set to the value at the previous time step and a single correction term is multiplied to $\{A_n(t+\Delta t)\}$ for all n so that Eq. (3.43) can be satisfied.

Condition (3.42) is added to guarantee that the drain current does change with time when the carrier density at the drain edge changes with time in linear region even during the moving boundary period.

3.3. Results and comparison with other models

To test the accuracy of the model, the model results have been compared with the PISCES 2-D device simulation [3.21], the numerical solution of the 1-D current continuity equation, QS(quasistatic) models in SPICE and a limited amount of available data. Also a NMOS ring oscillator has been simulated and compared with other models. Although the PISCES 2-D device simulation and the 1-D numerical solution are also non-quasistatic models, only the analytic non-quasistatic model developed in this work is denoted by the term 'NQS model' in the following discussions.

The 1-D numerical analysis in this work is essentially the same as that described by P.Mancini et. al [3.22]. The only difference is that a numerical package was used in

[3.22] while a new program has been written in this work.

For QS models, BSIM (Berkeley Short Channel IGFET Model [3.8]) in SPICE3 is used.

3.3.1. Turn-on transient of a NMOSFET

In Fig.3.4.(a), the drain current waveforms are shown for the turn-on transient of a NMOSFET with the channel length of 10 μm. The gate-drain overlap capacitance, C_{GDO}, has been adjusted to fit the PISCES results at $V_{GS}<V_{TH}$. Good agreements are observed among the NQS(non-quasistatic), PISCES and the 1-D numerical solution of eq. (3.8). For T_R=2ns, the QS(quasistatic) 40/60 model shows an anomalous negative drain current at the initial stage of turn-on, when the device is in the subthreshold and the saturation region. The QS 0/100 model eliminated the anomalous and troublesome negative current but produced larger discrepancies in the turn-on time. Charge sheet model(CSM) is also shown for comparison. CSM is basically a 40/60 QS model[3.12] and is based on the channel charge partitioning scheme derived by Oh et. al[3.2]. The ratios 0/100 and 40/60 in the QS models refer to the ratios of the drain and source charges in saturation region. This device has a channel transit time of around 0.7 ns. For T_R=20ns, the discrepancy is much smaller than for T_R=2ns, as expected.

Fig.3.4.(b) shows the source current waveforms. Again good agreements are observed among NQS, PISCES and 1-D numerical solution. For T_R=2ns, the QS 0/100 and 40/60 model show a large negative current at the initial stage of turn-on and the magnitude of the initial negative current is larger for 0/100 model than for 40/60 model.

Fig.3.4.(c) shows the gate current waveform for T_R=2ns. Fair agreements are observed among NQS, PISCES and the numerical solution. At the initial stage of turn-on, there is a discrepancy between NQS and PISCES simulation and this is considered to be due to the neglect of hole current in the PISCES simulation with one-carrier method. The small discrepancy between NQS and PISCES simulation at around t/T_R=0.6 is due to neglect of the NQS bulk current.

The average gate current of the NQS, PISCES and the 1-D numerical solution during the turn-on time, is observed to be much less than that of the QS models. Since the gate

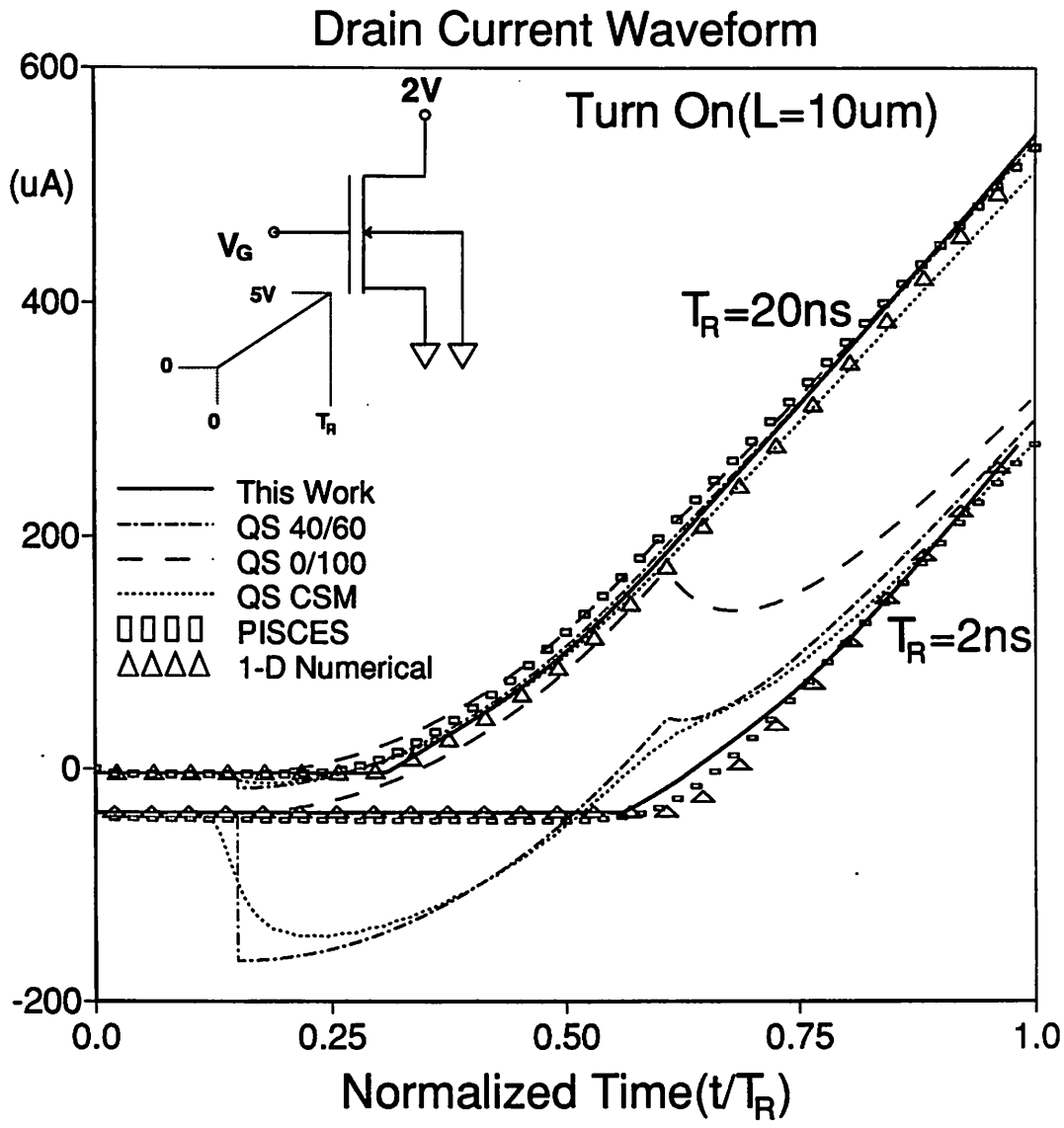


Fig.3.4. Turn-on transient of an NMOSFET with W/L=10μm/10μm, where V_{DS}=2V, V_{BS}=0 and V_{GS} is a rising ramp voltage. V_{GS} changes from 0 to 5V during the time interval of T_R, where T_R is 2 *ns* and 20 *ns* respectively. The NQS model('This Work' in the figures) has been compared with the PISCES 2-D device simulation program, 1-D numerical solution of eq.(3.8) and quasistatic(QS) 0/100 and 40/60 and CSM(Charge Sheet) models. The model parameters are T_{OX}=18 *nm*, N_{SUB}=2.1e16 cm^{-3}, V_{FB}=-0.5V,μ_0=500cm^2/V·sec, cgdo=1500 pF/m and cgso=600 pF/m.

(a). Drain current waveform

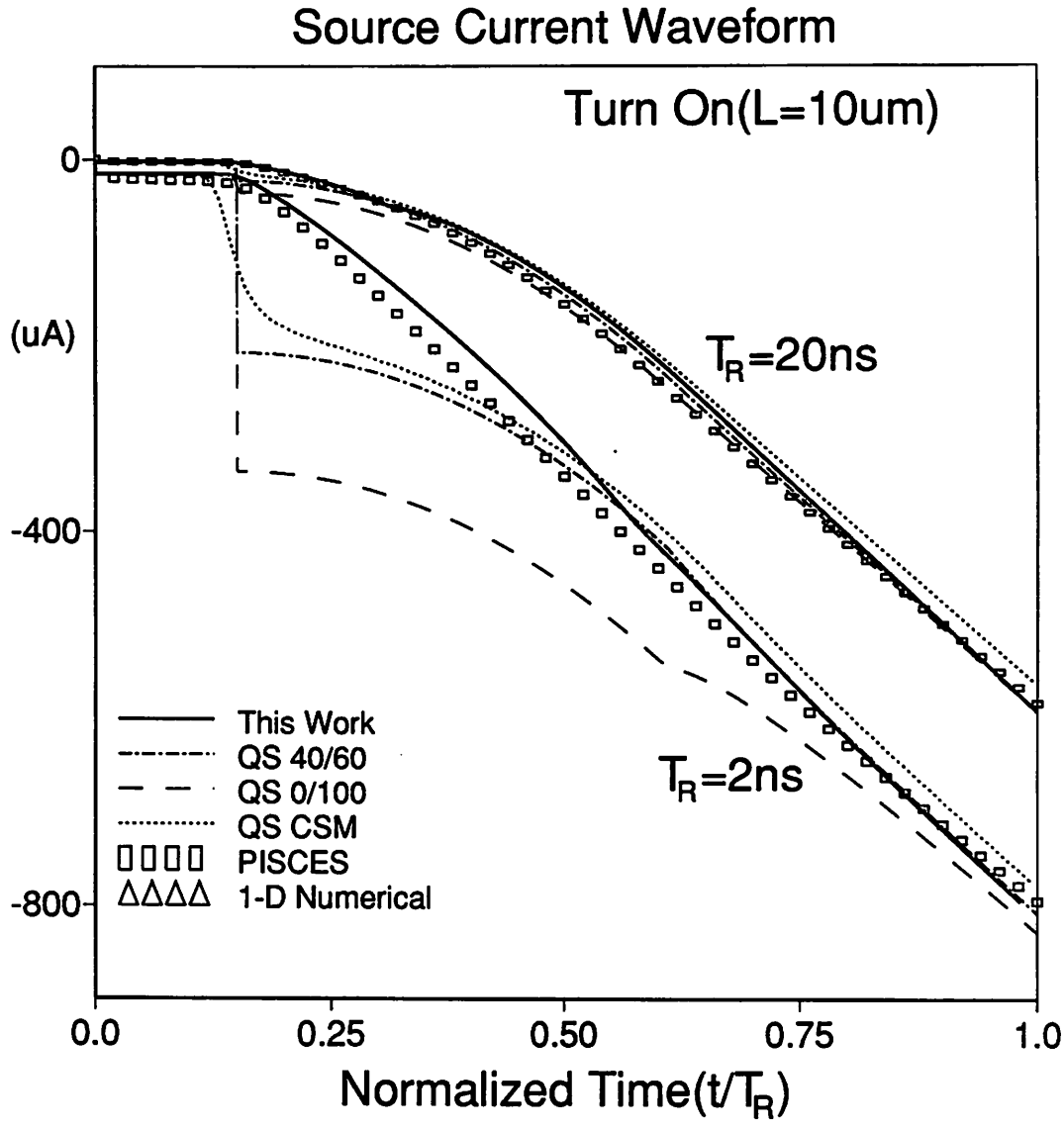


Fig.3.4.(b). Source current waveform

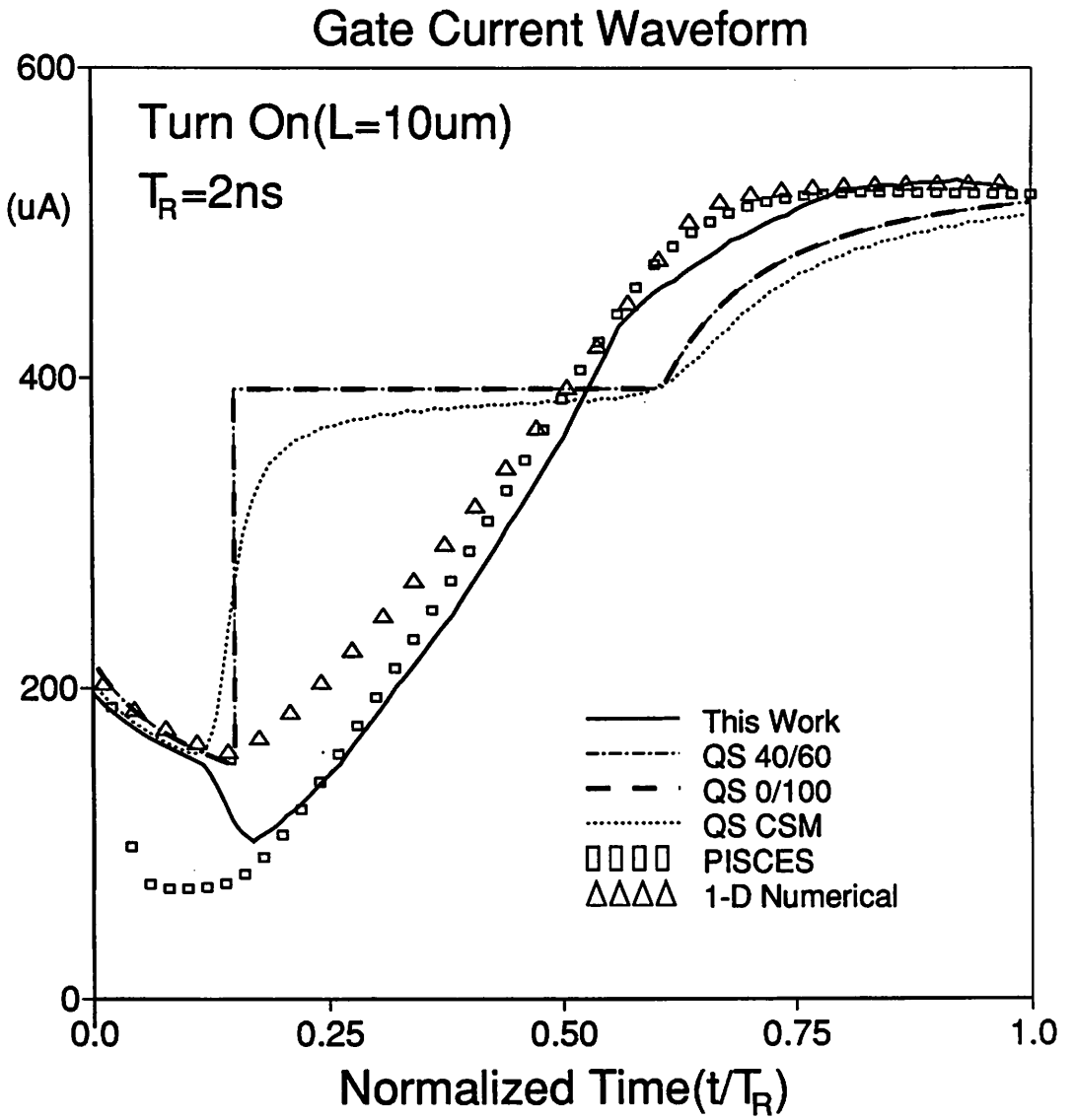


Fig.3.4.(c). Gate current waveform

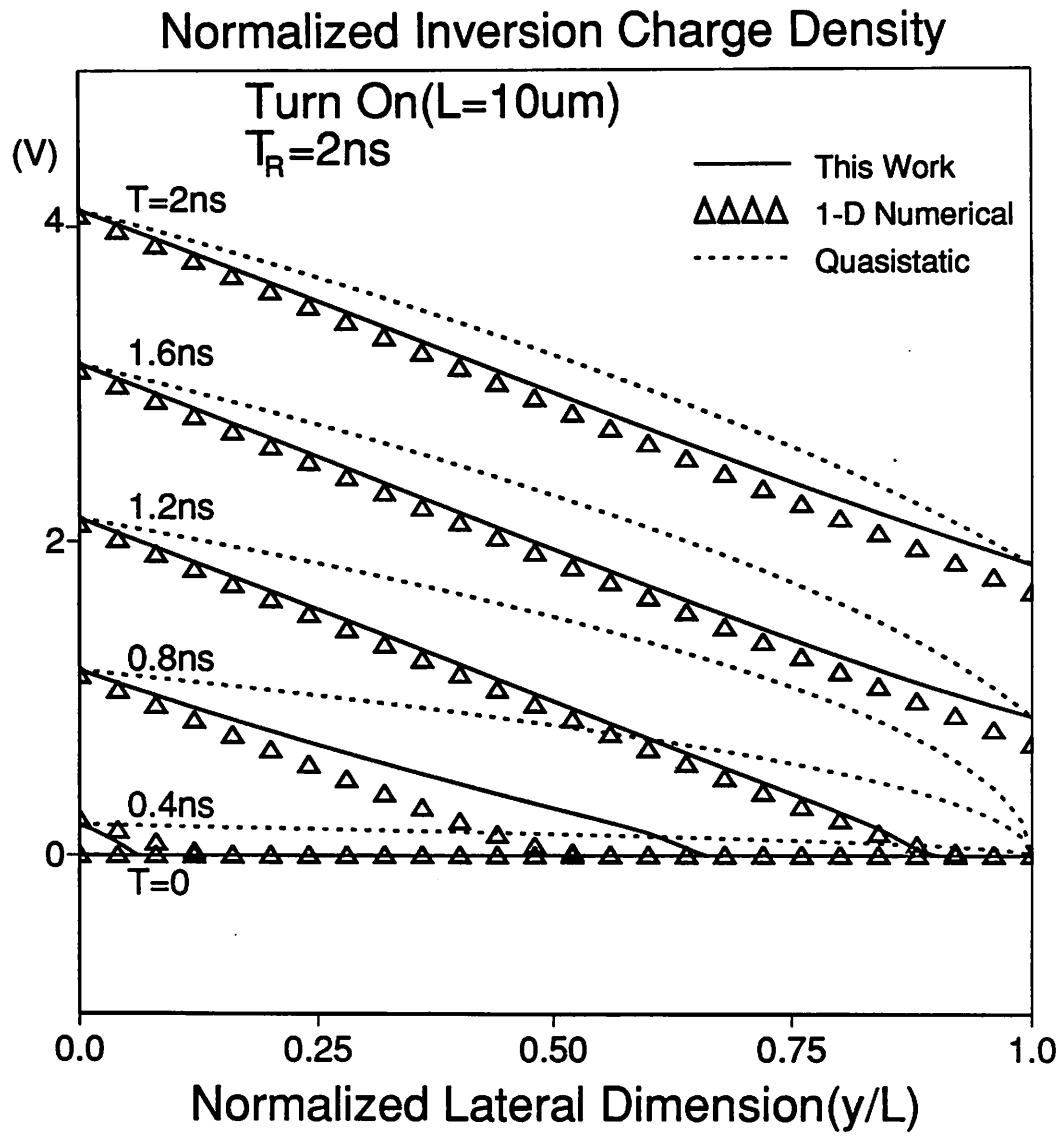


Fig.3.4.(d). Time evolution of the inversion charge density profile for T_R = 2 ns

voltage changes linearly with time, the gate current is directly proportional to the effective gate capacitance. So the effective gate capacitance is much smaller than that predicted by the QS models during the fast turn-on. In other words, during the fast turn-on of MOSFET's, the effective loading by the gate capacitance is much less than that predicted by the QS models.

Fig.3.4.(d) shows the time evolution of the inversion charge density profile for T_R=2ns. Good agreements are observed between NQS and numerical results but large discrepancies can be observed between NQS and QS results. During the initial stage of turn-on when the tail of the carrier density profile has not reached the drain side yet, the NQS model overestimates the carrier density in the channel near the drain and its drain current waveform may have some wiggles. To avoid this, the SPICE implementation checks for the moving boundary condition (3.41) and (3.42) and sets the drain current to the previous value if this condition is met.

The turn-on delay time of a MOSFET with a linearly ramped V_G and a fixed V_D,V_S and V_B, is derived from (3.41). The inversion charge density near the drain becomes non-zero at $\sqrt{F_B L^2/(\mu_n \cdot dV_G/dt)}$ seconds after V_G reaches V_{TH}, i.e., the turn-on delay time is

$$t_D = \frac{V_{TH}}{\frac{dV_G}{dt}} + \frac{L}{\sqrt{\frac{\mu_n}{F_B}\cdot\frac{dV_G}{dt}}} \tag{3.44}$$

where V_{TH} is the threshold voltage. Assuming that F_B=1 (i.e. neglecting the body effect), the second term in the right hand side of (3.44) agrees with the expression of the delay time ($\sqrt{T_R \cdot \tau_T}$) which has been derived by Swanson when the rise time T_R is larger than the channel transit time τ_T and the MOSFET is in saturation region [3.23] [3.2] [3.24].

The average CPU time per iteration, i.e. per model computation, for simulating a turn-on transient in SPICE3 is compared in Table.3.1, along with PISCES and the 1-D numerical solving of eq. (3.8). Meyer represents the SPICE level 2 Meyer capacitance model.

The NQS model is based on the charge sheet model(CSM) and takes about 1.6 times longer than the quasistatic CSM and about three times longer than the other QS SPICE models. For the PISCES simulation, the one carrier full Newton method is used with the number of grid points, 1769. And the uniform substrate doping concentration is used and the work function of the gate material is adjusted to fit the threshold voltage.

Analysis	NQS	CSM	BSIM	Meyer	PISCES	Numerical
TRAN	3.10	1.95	1.10	0.94	85000	2200

Table.3.1. Average model evaluation time(mS) per model computation in SPICE3 for the turn-on transient of a NMOSFET (Vax 8800 with Ultrix V2.0)

3.3.2. Turn-off transient of a NMOSFET

Fig.3.5.(a) shows the drain current waveform of the turn-off transient of a NMOSFET with the channel length of $10\mu m$. For T_F=$2ns$, the QS models predict an abrupt change of drain current at t=0, but NQS, PISCES and the numerical results show that the change of the drain current is continuous with time. Good agreements can be observed between NQS and 1-D numerical solution but there is a little discrepancy between NQS and PISCES results. This is believed to be due to the 2-D effect by the drain junction. In the PISCES simulation, the junction depth is set to be $0.33\mu m$, but in the NQS and 1-D numerical solution the junction depth is assumed to be 0.

Fig.3.5.(b) shows the source current waveform. Again, for T_F=$2ns$, the QS models predict an abrupt change of source current at t=0, but NQS, PISCES and the numerical results show that the source current changes continuously with time. The QS 0/100 model shows a larger discrepancy from the NQS model than the QS 40/60 model, which is the same tendency as in the turn-on transient as shown in Fig.3.4.(b).

Fig.3.5.(c) shows the gate current waveform for T_F=$2ns$. NQS, PISCES and the 1-D numerical solution show that the gate current changes continuously with time at the

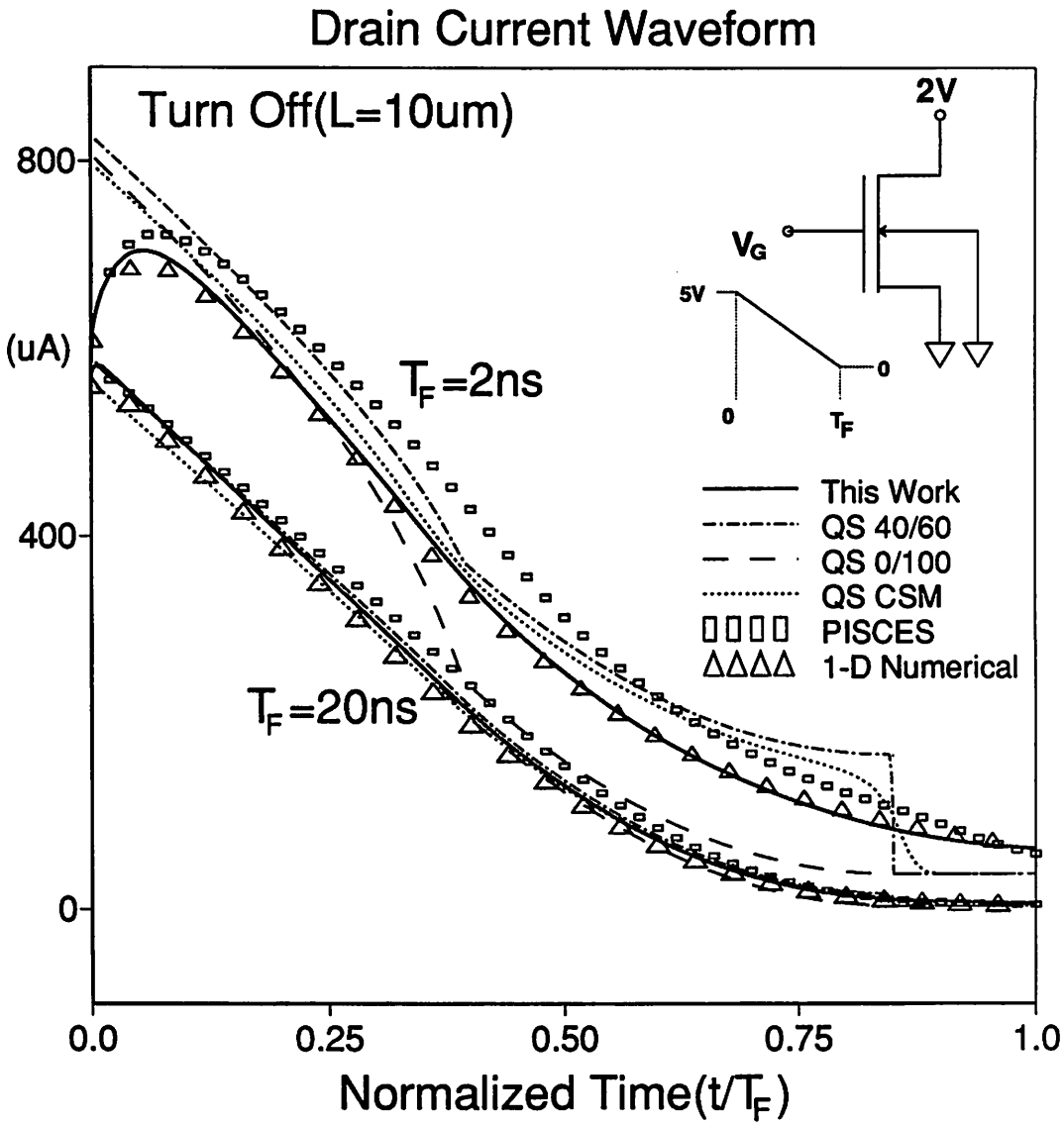


Fig.3.5. Turn-off transient of an NMOSFET with W/L=10μm/10μm, where V_{DS}=2V, V_{BS}=0 and V_{GS} is a falling ramp voltage. V_{GS} changes from 5V to 0 during the time interval of T_F, where T_F is 2 *ns* and 20 *ns* respectively. The NQS model('This Work' in the figures) has been compared with the PISCES 2-D device simulation program, 1-D numerical solution of eq.(3.8) and quasistatic(QS) 0/100 and 40/60 and CSM(Charge Sheet) models. The model parameters are the same as those in Fig.2.4.

(a). The drain current waveform

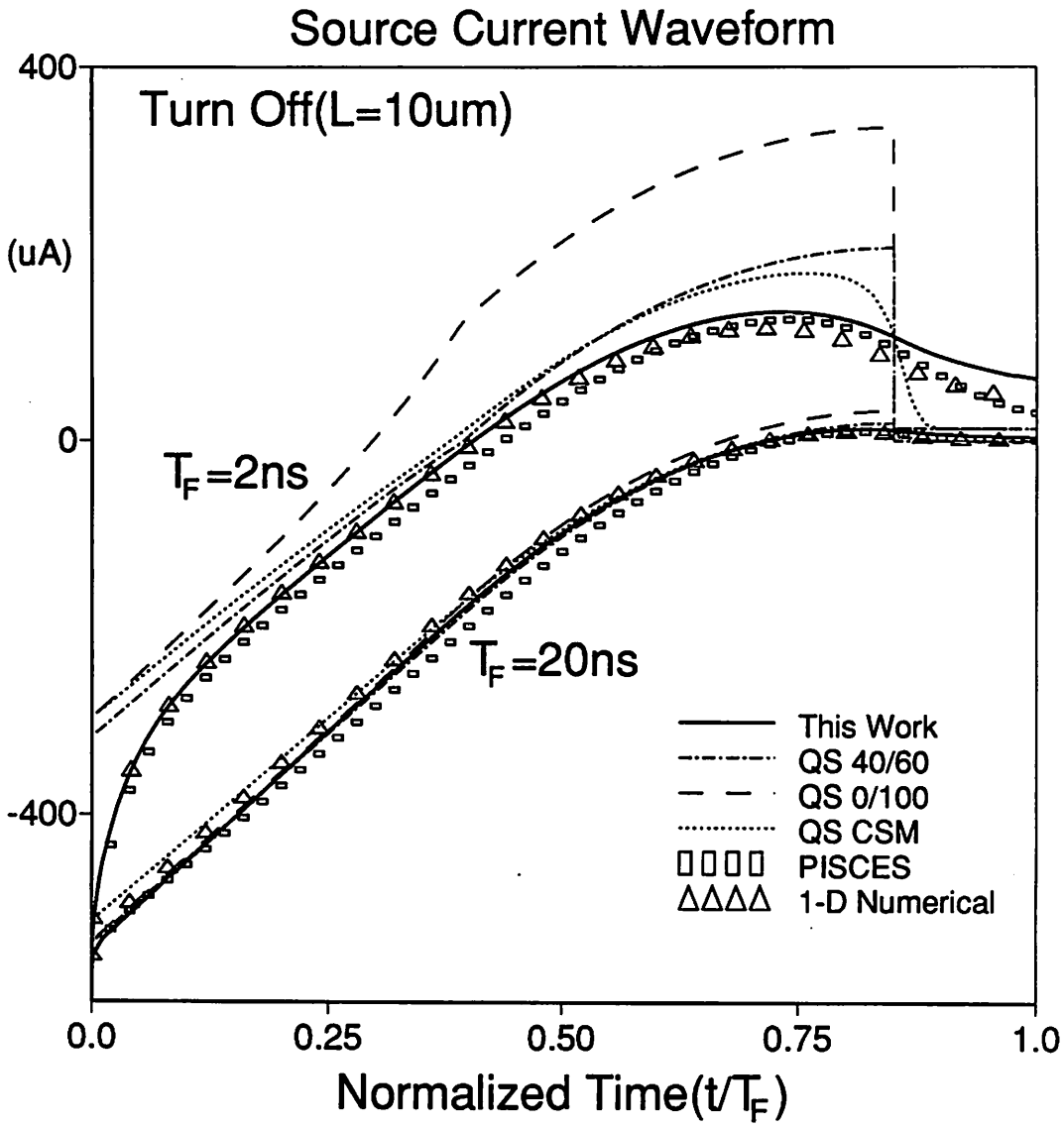


Fig.3.5.(b). Source current waveform

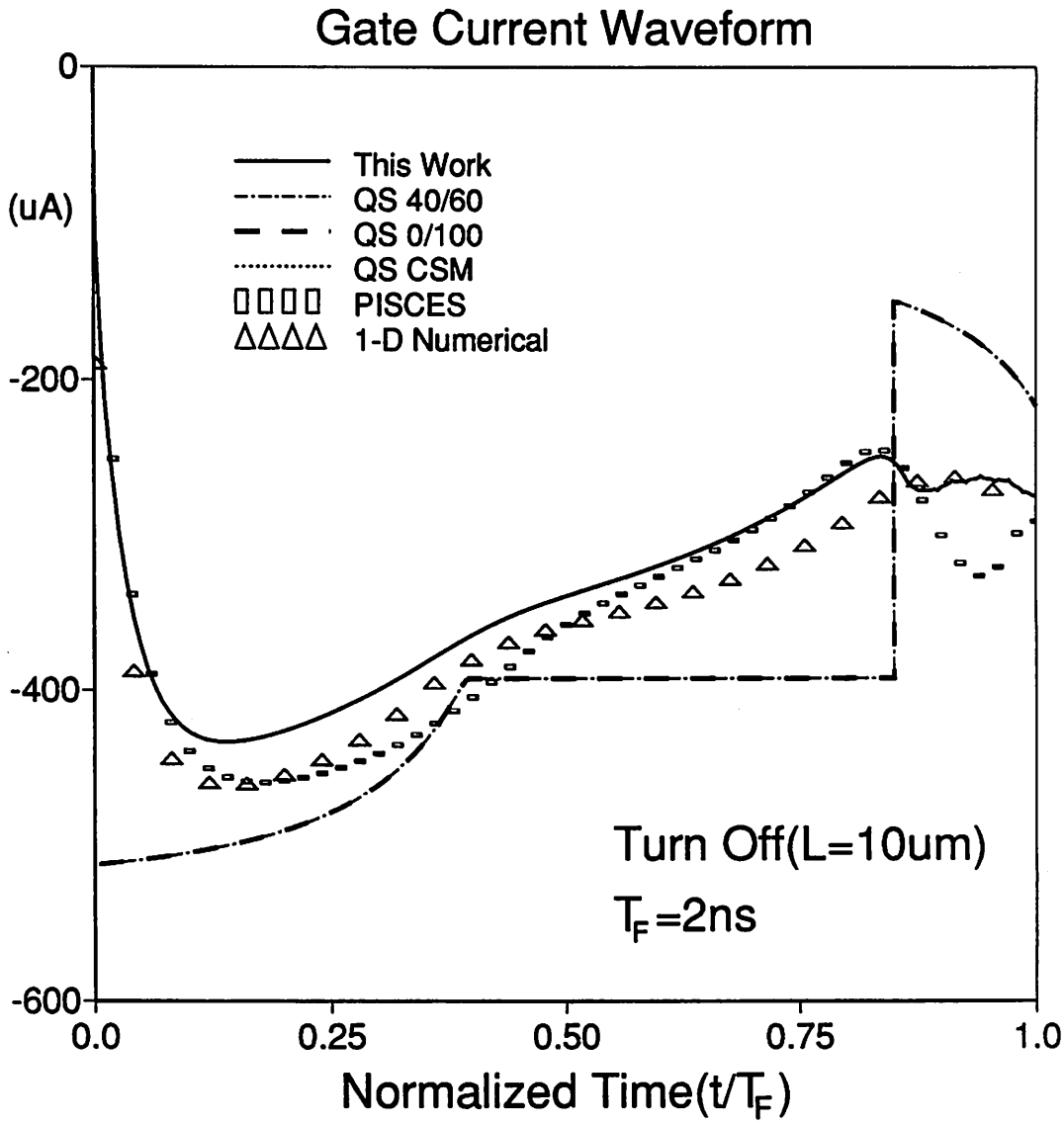


Fig.3.5.(c). Gate current waveform

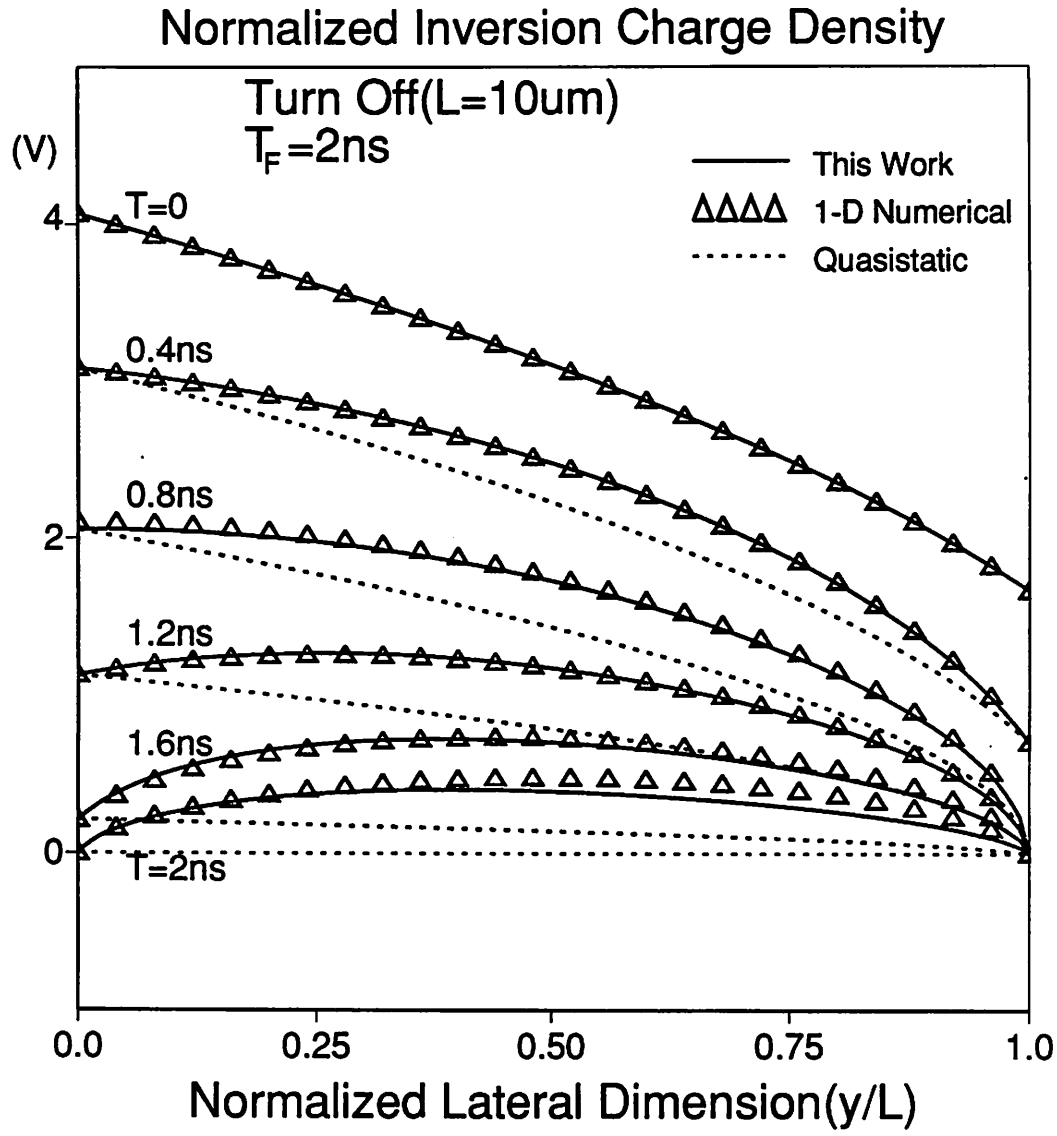


Fig.3.5.(d). Time evolution of the inversion charge density profile for T_F = 2 ns

turn-off transient but the QS models predict an abrupt change of gate current at t=0. The magnitude of the average gate current of NQS, PISCES and the numerical solution is less than that of QS models. From this we can see that, at the fast turn-off, the effective loading by the gate capacitance is less than that predicted by the QS models. This tendency is the same as that of the fast turn-on, as shown in Fig.3.4.(c). From Fig.3.4.(c) and Fig.3.5.(c), we can conclude that, during the fast transient, the effective loading by the gate capacitance is less than that predicted by the QS models due to the inertia of inversion carriers in the channel.

Fig.3.5.(d) shows the time evolution of the inversion charge density profile during the turn-off transient with T_F=2ns. Good agreements are observed between NQS and numerical results. The QS model curves are identical to those in Fig.3.4.(d), of course. The NQS model is reduced to the QS model at the steady state, therefore the two curves coincide at t=0.

3.3.3. Large-signal transient channel current partitioning ratio

For circuit analysis and CAD, it is necessary to determine the partitioning of the transient channel current (difference between the transient gate and bulk currents) between the transient source and drain currents. In charge based QS(quasistatic) models the transient source and drain currents can be computed by differentiating simple source and drain charge expressions. Therefore the concept of charge partitioning has been useful and popular in comparing different charge based QS models. [3.7] [3.8] [3.9]

However, since no simple charge expressions exist in NQS model, the ratio of transient currents is directly compared instead of the charge partitioning ratio. These two ratios are equal to each other for QS models in saturation region, where the ratios become constant independent of biases.

In this work, the transient current partitioning ratio, Δ, is defined to be the ratio of the transient drain and source currents, $I_{D.T}$ and $I_{S.T}$, where the DC drain to source current component is excluded.

$$\Delta = \frac{I_{D.T}}{I_{D.T}+I_{S.T}} \tag{3.45}$$

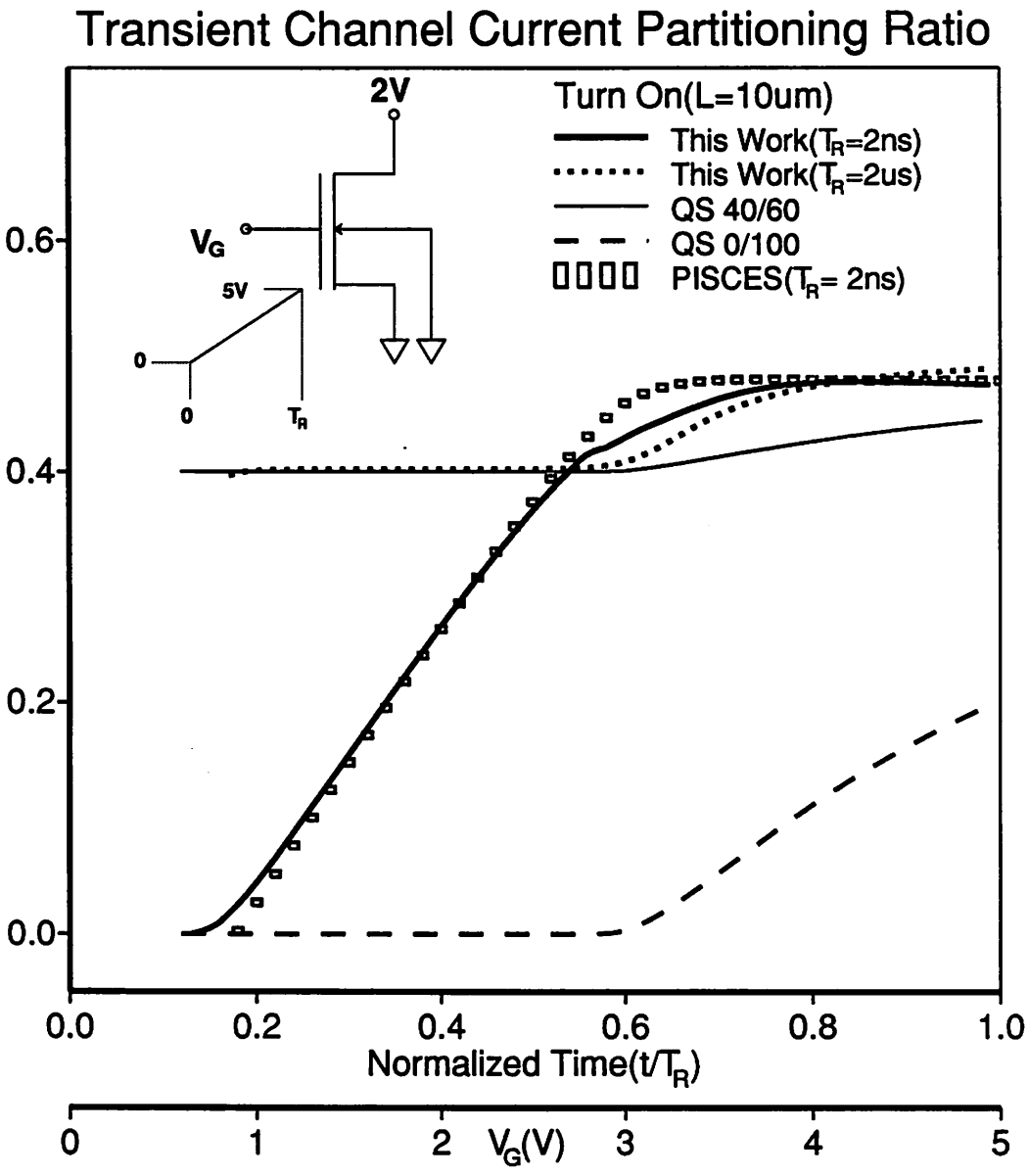


Fig.3.6. The transient channel current partitioning ratio for the NQS('This Work'), QS 40/60 and 0/100 models and PISCES simulation. Model parameters are the same as those in Fig.3.4. Partitioning ratio is not shown in portions of the cutoff region because division of the very small drain and source currents produces meaningless errors.

(a). Turn-on transient of an NMOSFET, for T_R of 2ns and 2μs respectively.

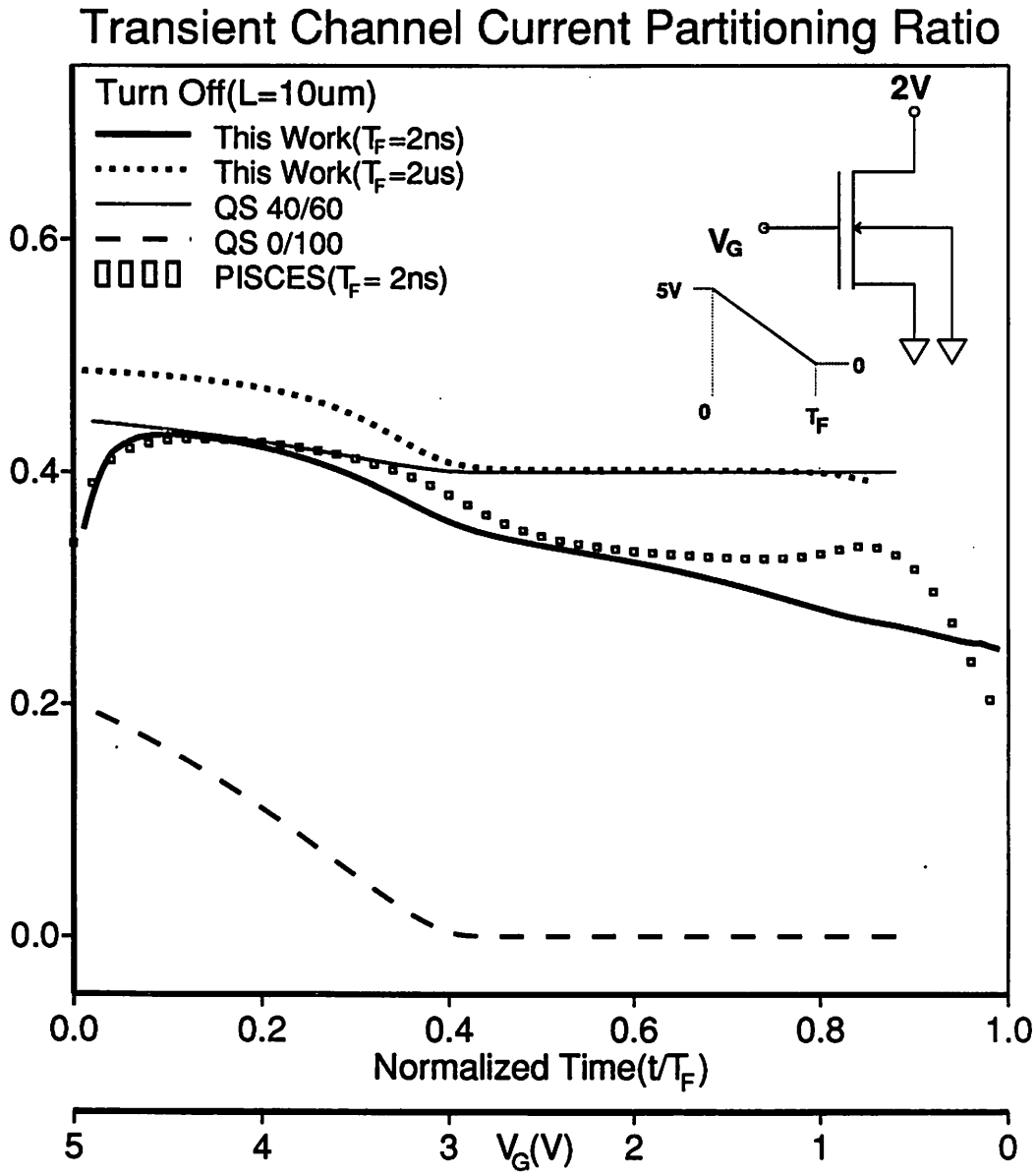


Fig.3.6.(b). Turn-off transient of a NMOSFET, for T_F of 2ns and 2μs respectively.

In saturation region, the charge based QS models give

$$I_{D.T} = \Delta_{QS} \cdot \frac{dQ_N}{dt} \tag{3.46}$$

$$I_{S.T} = (1-\Delta_{QS}) \cdot \frac{dQ_N}{dt} \tag{3.47}$$

where Q_N is the total channel charge and Δ_{QS} is the QS channel charge partitioning ratio and is equal to 0.4 for QS 40/60 model and 0 for QS 0/100 model. Substituting (3.46) and (3.47) into (3.45), we can see that Δ exactly matches Δ_{QS} in saturation region.

Fig.3.6.(a) and Fig.3.6.(b) show the transient current partitioning ratio during the turn-on and the turn-off transient, respectively, of a single NMOSFET with a ramp voltage at the gate and with V_{DS} and V_{BS} fixed at 2V and 0 respectively. for NQS, QS 40/60 and QS 0/100 models. Eq. (3.45), (3.24) and (3.25) are used to find the partitioning ratio of NQS model.

Fig.3.6.(a) shows the partitioning ratio during the turn-on transient. For the fast turn-on (rise time T_R = 2 $n$sec), the NQS model and PISCES show that the transient current partitioning ratio is close to 0 at the initial stage of turn-on, and it increases with time. For the slow turn-on (rise time T_R=2μsec), the NQS model gives almost the same result as that of QS 40/60 model in saturation region.

Fig.3.6.(b) shows the partitioning ratio during the turn-off transient of a NMOSFET. PISCES and NQS model show quite different partitioning ratio at the same gate voltage during the turn-on and the turn-off transients.

From Fig.3.6.(a) and Fig.3.6.(b), we can observe that the transient current partitioning ratio is not constant in saturation region, but is a function of the signal transition rate and also a function of the history of applied biases, while QS models predict a constant partitioning ratio in saturation region.

Also we can see that PISCES and NQS model are reduced to QS 40/60 model for the slow signals compared to the channel transit time. The QS 0/100 model was introduced rather arbitrarily as a convenient solution to solve some physical anomalies of QS 40/60

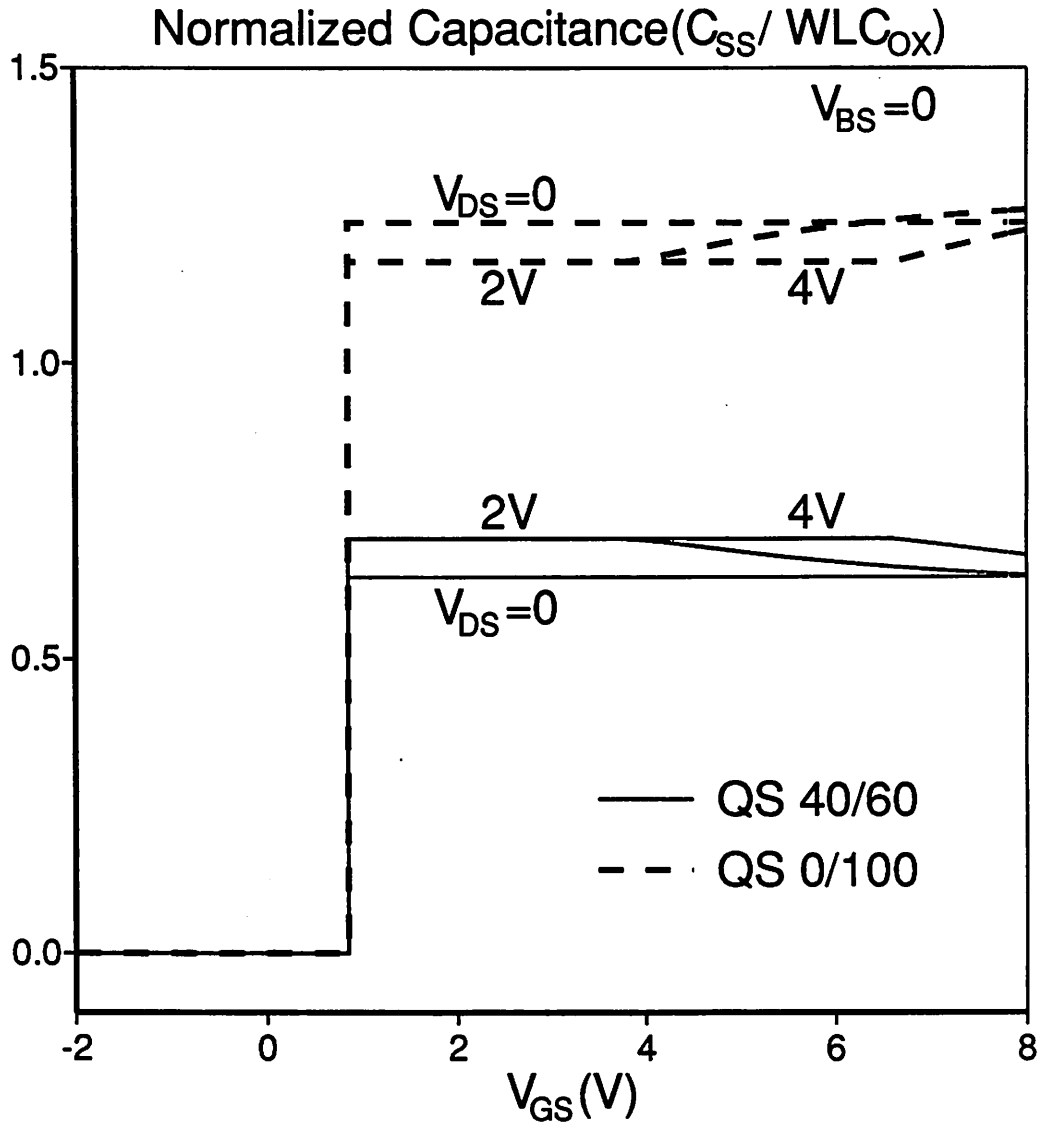


Fig.3.7. The normalized capacitance,C_{SS}/WLC_{OX}, of QS 0/100 and QS 40/60 model, to show the physical anomaly of the QS 0/100 model. C_{SS} is $\partial Q_S/\partial V_S$. In the QS 0/100 model, C_{SS} is larger than the total gate capacitance, WLC_{OX}, when the transistor is on. We can can see the same anomaly in C_{DD}.

model, for example, the negative I_D spike during the turn-on transient. However this 0/100 model shows another physical anomaly that some capacitance components are larger than the total gate oxide capacitance(WLC_{OX}). Fig.3.7 shows the capacitance C_{SS} ($\partial Q_S/\partial V_S$) of QS 0/100 model with respect to V_{GS} along with that of 40/60 model.

3.3.4. NMOS inverter transient

Fig.3.8 shows the turn-on transient of a NMOS inverter with a resistive load. The QS 40/60 model shows an output voltage overshoot at the initial stage of turn-on. This is related to the undershoot of drain current at the initial stage of turn-on as shown in Fig.3.4.(a). The QS 0/100 model does not show any voltage overshoots but shows large discrepancy in the turn-on time. The junction capacitance is included in this simulation but the overlap capacitance is not included to see the difference more clearly.

3.3.5. Comparison with measured data

In Fig.3.9.(a) and Fig.3.9.(b), the NQS model is compared with Oh's measured data for a long channel transistor[3.2] [3.24], along with the QS(quasistatic) simulation results. Although the measured data is available only up to t=90ns, the simulation results are presented up to t=120ns to show that the steady state values are the same for all the models.

Fig.3.9.(a) shows the measured data and simulation results without the drain junction capacitance. Overlap capacitance has been adjusted to fit the measured I_D at $V_{GS} < V_{TH}$. Good agreements are observed between the NQS results and the measured I_D. At t=60ns, there is an abrupt change of drain current in the NQS model due to the correctly abrupt cease of the overlap capacitance current. The QS 40/60 model shows a large negative current during the initial stage of turn-on, while the QS 0/100 model produces large discrepancies in the turn-on time. The MOSFET in this example stays in saturation region during all the time considered. In the QS 0/100 model, no channel charge is assigned to the drain node in saturation region and so the intrinsic drain current of 0/100 model in Fig.3.9.(a), consists of DC component only without any capacitive current. For t > 60ns, QS models predict a constant current with time but NQS and measurement still show a time varying drain current. This clearly illustrates the

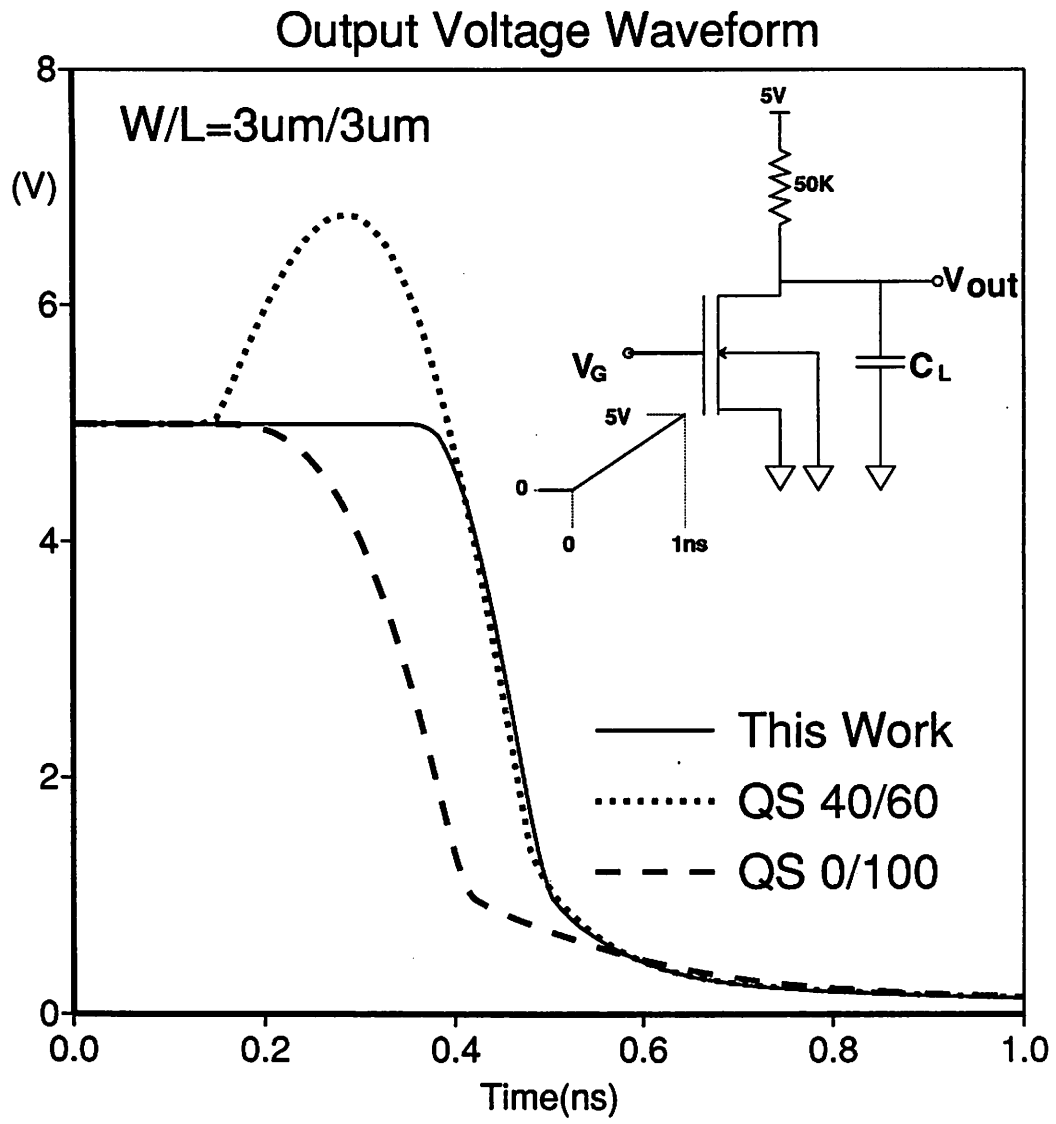


Fig.3.8. The turn-on transient of an NMOS inverter with a resistive load, for NQS('This Work'), QS 0/100 and 40/60 models. The junction capacitance is included but the overlap capacitance is not included to see the difference more clearly. The model parameters of the transistor are the same as those in Fig.3.4.

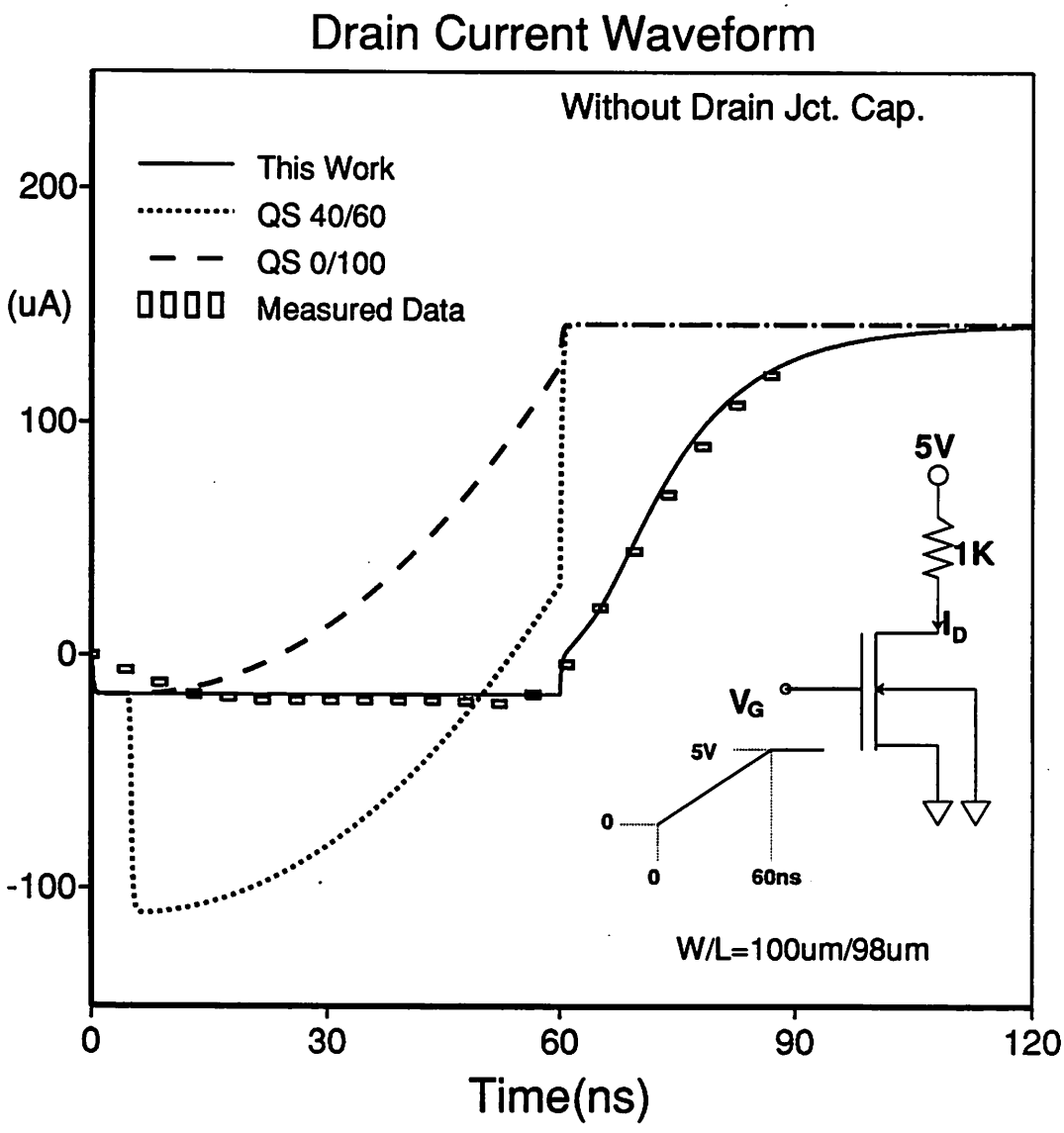


Fig.3.9. The comparison of the NQS model('This Work') with the measured data[3.2] for the turn-on transient of the long channel transistor. The results from the QS models are also shown for the comparison. where W/L=100μm/98μm, T_{OX}=80nm, V_{FB}=-1.25V, N_{SUB}=8·$10^{15}cm^{-3}$, μ_0=500cm^2/Vsec and cgdo=1500pF/m

(a). without the drain junction capacitance.

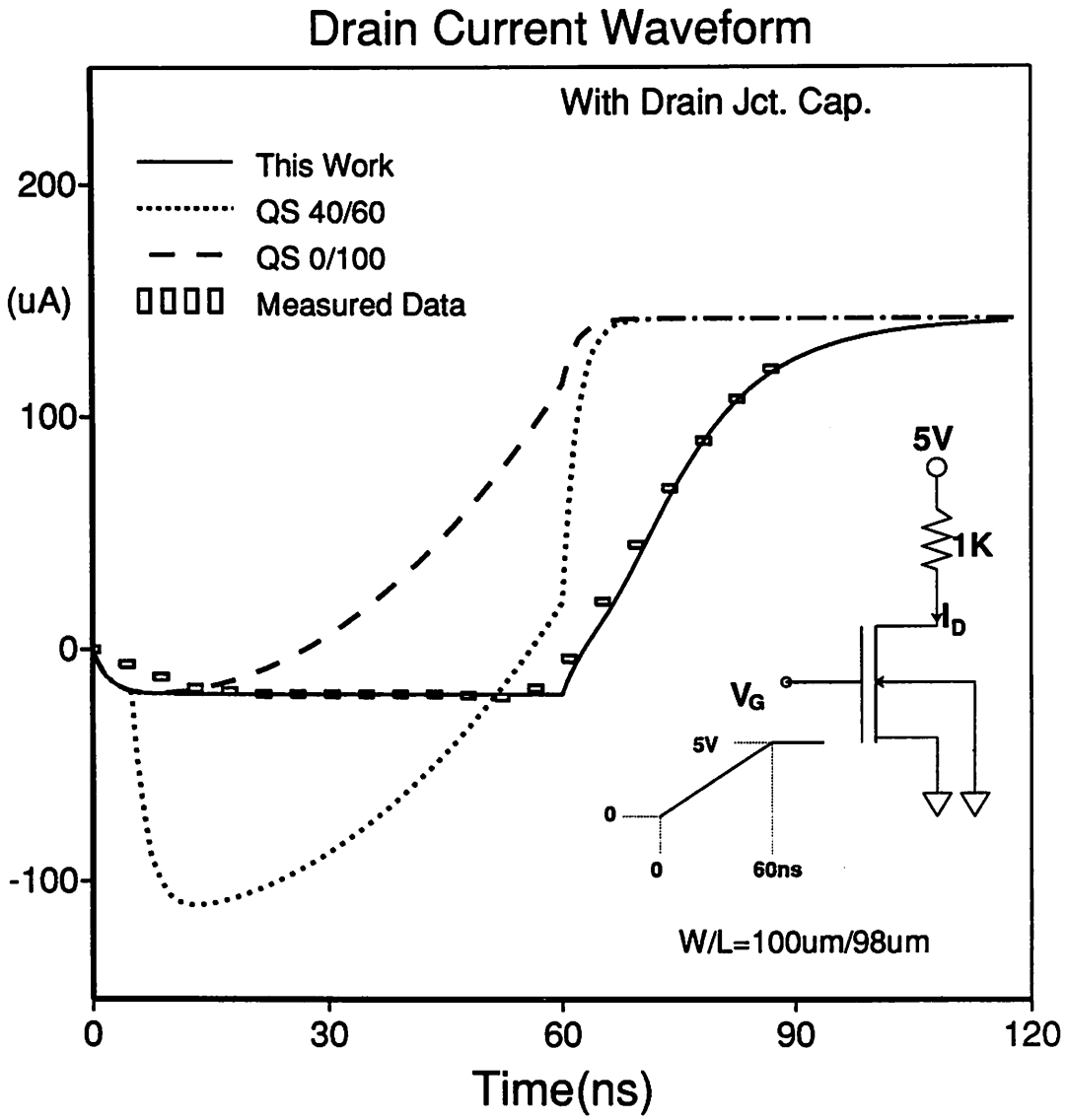


Fig.3.9.(b). The drain junction capacitance included in the simulation. The zero bias drain junction capacitance is 2pF.

limitation of the QS(quasistatic) models, which force the current and the charges to be functions of only the applied biases.

Fig.3.9.(b) shows the measured data and simulation results with the drain junction capacitance included. The current of NQS model is continuous at t=60ns.

3.3.6. Simulation of 11-stage NMOS ring oscillator

An 11-stage ring oscillator of depletion load NMOS inverters is simulated with NQS and QS models for four different channel length of 0.75, 3, 6 and 12 μm. The channel widths of the transistors are adjusted proportionately to guarantee the same DC current. Junction capacitance and overlap capacitance have been included in the simulation. The DC transfer curve and the current characteristics have been matched between each model by adjusting the model parameters, as shown in Fig.3.10.(a).

The junction capacitance, the overlap capacitance and the total intrinsic gate capacitance are shown in Table.3.2.

L(μm)	WLCox(fF)	CJ(fF)	Cov(fF)
0.75	1.3	20.4	0.4
3.0	20.6	20.4	1.5
6.0	82.6	20.4	3.0
12.0	330	20.4	6.0

Table.3.2. Comparison of intrinsic and extrinsic capacitances in the 1-stage of the ring oscillator. WLCox and Cov represent the total intrinsic gate capacitance and the drain overlap capacitance of the enhancement NMOSFET. CJ represents the total zero-bias junction capacitance at the output node of the inverter.

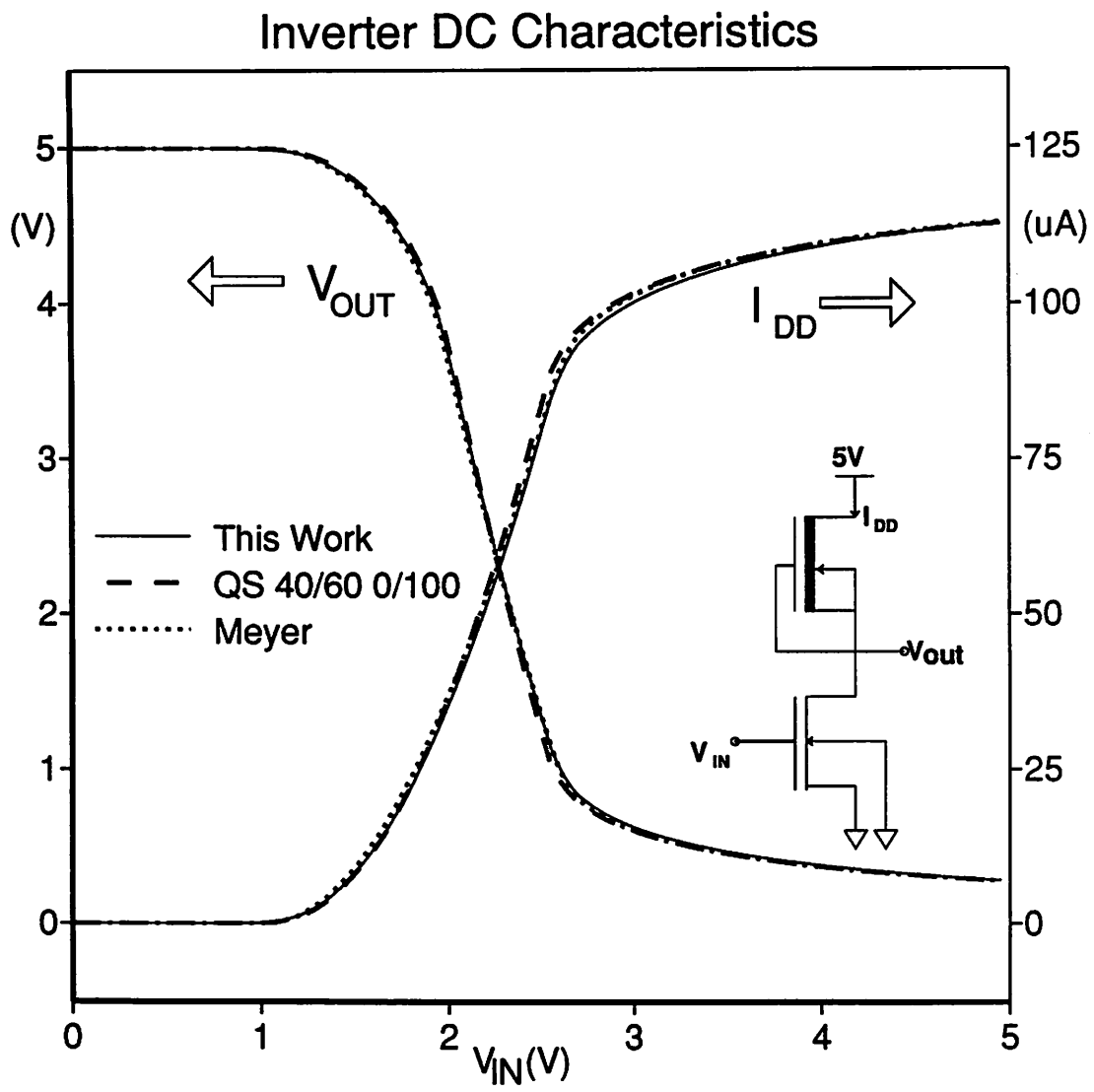


Fig.3.10. The simulation of 11-stage depletion load NMOS ring oscillators. The model parameters are V_{FB} = -0.87V for the enhancement driver and -5.0V for the depletion load, $N_{SUB}=2\cdot10^{16}cm^{-3}$, $\mu_0=500cm^2/V\cdot sec$, $\lambda=0.03V^{-1}$, cgdo=cgso=$500pF/m$, Cj=$3.0\cdot10^{-4}F/m^2$ and Cjsw=$8\cdot10^{-10}F/m$.

(a). The DC transfer curve and the DC supply current have been matched between each model to get the proper comparison.

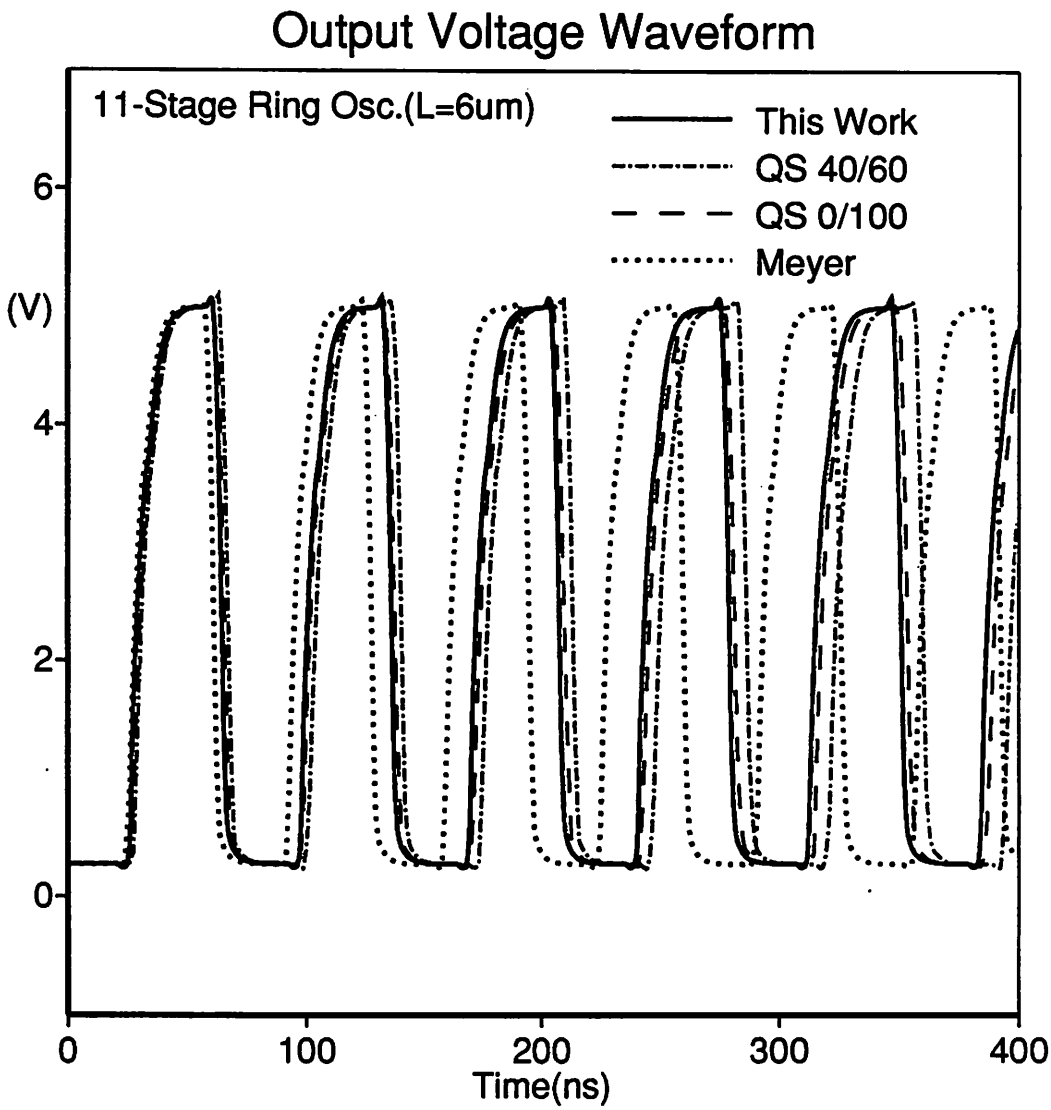


Fig.3.10.(b). The output voltage waveform of the ring oscillator for L=6μm.

The output voltage waveform for L=6μm is shown in Fig.3.10.(b).

The unit delay time of a single inverter is computed from the simulated output voltage waveform and is shown in Table.3.3. For L=0.75μm and 3μm, where the extrinsic capacitance is dominant over or comparable to the intrinsic component, the percent error of the unit delay time among models is within 1 %. For L=6 and 12μm, where the intrinsic component is dominant, the Meyer model predicts a delay time which is about 10 percent less than the NQS result and the charge based QS models show much less error.

L(μm)	NQS	QS 40/60	QS 0/100	Meyer
0.75	0.436(0%)	0.433(-0.7%)	0.432(-0.9%)	0.432(-0.9%)
3.0	1.29(0%)	1.29(0%)	1.27(-0.15%)	1.21(-0.58%)
6.0	3.27(0%)	3.30(0.9%)	3.25(-0.4%)	2.98(-8.7%)
12.0	10.0(0%)	10.3(3.8%)	10.0(0%)	9.1(-9.1%)

Table.3.3. Unit delay time(ns) of a depletion load NMOS inverter, computed from the 11-stage ring oscillator simulation. The numbers inside the brackets are the percent differences from the result of NQS model.

Table.3.4 shows the CPU time comparison between NQS and conventional QS models in SPICE3 for the simulation of the ring oscillator with L=6μm. The CPU time for the NQS model is around 3 times longer than that for the conventional QS models. A new time step control scheme in Appendix 9 is used for the NQS model and the same convergence criterion is used for all the models.

	NQS	QS 40/60	QS 0/100	Meyer
Total CPU Time(sec)	171	51	42	56
Total # of Iterations	2506	1901	1614	2638
Total Time Points	2495	521	436	628
Accepeted Time Points	875	347	305	425

Table.3.4. CPU time comparison between models in SPICE3, for the 11-stage depletion-load NMOS ring oscillator with four different sets of W/L's. One combination is $(W/L)_{ENH}$= 10μm/3μm and $(W/L)_{DEP}$= 3μm/3μm. The channel length has been varied to 0.75μm,6μm and 12μm and the channel width has been adjusted proportionately to keep the same DC characteristics. For the QS 0/100 and 40/60 models, BSIM in SPICE3 is used. For the NQS model, the time step control scheme in Appendix 9 is used. (Vax 8800 with Ultrix V2.0)

3.4. Conclusion

Based on the approximate solution of the current continuity equation, an analytic NQS(non-quasistatic) long channel MOSFET model for the transient analysis has been derived and implemented into SPICE3. This NQS model is independent of and can be implemented with any existing DC model, as illustrated in [3.20].

To check the accuracy of the model, this work has been compared with PISCES 2-D device simulation, 1-D numerical solution and the measured transient current waveform, and excellent agreements have been obtained. Also all the known anomalies in the transient analysis caused by the use of QS(quasistatic) models have been eliminated with this NQS model.

The channel charge partitioning ratio is shown to be not constant in saturation region but to change with time, the signal transition rate and the history of the applied biases.

The CPU time required for this NQS model, in SPICE3, is about two to three times that of the conventional QS models in SPICE3.

This NQS charge model has been combined with the SPICE level-2 DC model[3.9] and implemented in SPICE3 [3.20]. In [3.20], the drain saturation voltage V_{DSSAT} in eq. (3.16) is replaced by V_{DSSAT} computed from the SPICE level-2 DC model. In this way all the short channel effects such as the velocity saturation effect, the V_{GS} dependence of mobility and all other aspects considered by the level-2 DC model were also included in the implementation. It is true that while the DC model included all these short channel effects, the charge model is still based on the long channel theory. This compromise should be judged in the light of the fact that all the conventional charge(capacitance) models in SPICE (Meyer, Ward-Dutton, BSIM[3.8]) don't include any short channel effects either.

This model can be extended to AC analysis, which will be published elsewhere.[3.25]

Since this model is not a charge based model, it cannot guarantee the charge conservation.[3.26] A charge conserving NQS MOSFET model is shown in Chapter 5 and also is published in [3.27].

References

[3.1] D.A.Hodges, P.R.Gray, R.W.Brodersen, "Potential of MOS Technologies for Analog Integrated Circuits", IEEE JSSC, Vol. SC-13, pp.285-294, June 1978

[3.2] S.Y.Oh, D.E.Ward, R.W.Dutton, "Transient Analysis of MOS Transistors", IEEE JSSC Vol.SC-15, No.4, pp.636-643, 1980

[3.3] Y.Tsividis, G.Masetti, "Problems in the Precision Modeling of the MOS Transistor for Analog Applications", IEEE Trans. CAD Vol. CAD-3, No.1, pp. 72-79, Jan. 1984

[3.4] J.J.Paulous, D.A.Antoniadis, "Limitations of quasi-static capacitance models for the MOS Transistor", IEEE EDL Vol.EDL-4, pp.221-224, 1983

[3.5] J.G.Fossum, H.Jeong, S.Veeraraghavan, "Significance of the Channel-Charge Partition in the Transient MOSFET Model", IEEE Trans. Electron Devices, Vol.ED-33, No.10, pp. 1621-1623, Oct., 1986

[3.6] M.F.Sevat, "On the Channel Charge Division in MOSFET Modeling", Technical Digest, ICCAD, pp. 208-210, Nov., 1987,

[3.7] P.Yang, P.K.Chatterjee, "SPICE Modeling for Small Geometry MOSFET Circuits", IEEE Trans. CAD Vol.CAD-1, pp.169-182, 1982

[3.8] B.J.Sheu, D.L.Scharfetter, P.K.Ko, "SPICE2 Implementation of BSIM", ERL Memo, M85/42, U.C.Berkeley, May, 1985

[3.9] A.Vladimirescu, S.Liu, "The Simulation of MOS Integrated Circuits using SPICE2", ERL Memo, M80/7, U.C.Berkeley, Oct., 1980

[3.10] C.Turchetti, P.Mancini, G. Masetti, "A CAD-Oriented Non-Quasistatic Approach for the Transient Analysis of MOS IC's", IEEE JSSC Vol.SC-21, No.5, pp.827-836, 1986

[3.11] M.Bagheri, Y.Tsividis, "A Small Signal dc-to-High Frequency Nonquasistatic 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

[3.12] H.J.Park, P.K.Ko, C.Hu, "A Measurement Based Charge Sheet Capacitance Model of Short Channel MOSFET's for SPICE", IEDM 86 Technical Digest, pp.40-43,

Dec. 1986

[3.13] J.R.Brews, "A Charge-Sheet Model of MOSFET", Solid State Electronics, Vol.21, pp.345-355, 1978

[3.14] F. Van De Wiele, "A Long-Channel MOSFET Model", Solid State Electronics, Vol.22, pp.991-997, 1979

[3.15] P.P.Guebels, F. Van De Wiele, "A Charge Sheet Model for Small Geometry MOSFET's", IEDM 81 Technical Digest, pp.211-214, 1981

[3.16] G.Merckel, J.Borel, N.Z.Cupcea, "An Accurate Large-Signal MOS Transistor Model for Use in Computer-Aided Design", IEEE Trans. Electron Devices, Vol. ED-19, No.5, pp.681-690, May 1972

[3.17] H.C.Poon, L.D.Yau, R.L.Johnston and D.Beecham, "DC Model for Short Channel IGFETs", IEDM 73 Technical Digest, pp.156-159, 1973

[3.18] Y.P.Tsividis, "Operation and Modeling of the MOS Transistor", McGraw-Hill Book Company, pp.128-129, 1987

[3.19] H.J.Park, P.K.Ko, C.Hu, "A Non-Quasistatic MOSFET Model for SPICE", IEDM 87 Technical Digest, pp.652-655, Dec. 1987

[3.20] H.J.Park, P.K.Ko, C.Hu, "SPICE3 Implementation of NQS MOSFET Model with Level 2 DC Model", Internal Memo, Dept. of EECS, University of California, Berkeley, Mar. 1, 1988

[3.21] M.R.Pinto, C.S.Rafferty, R.W.Dutton, "PISCES-II Users's Manual", Stanford Electronics Lab., Stanford University, 1984

[3.22] P.Mancini, C.Turchetti, G.Masetti, "A Non-Quasi-Static Analysis of the Transient Behavior of the Long Channel MOST Valid in All Regions of Operation", IEEE Trans., Electron Devices, Vol.ED-34, No.2, pp.325-334, Feb. 1987

[3.23] R.M.Swanson, "Complementary MOS Transistors in Micropower Circuits", Technical Report No. 4963-1, Stanford Electronics Laboratories, Stanford University, 1974

[3.24] S.Y.Oh, "A Simplified Two-Dimensional Numerical Analysis of MOS Devices including Transient Phenomena", Technical Report No. G201-10, Stanford Electronics

Laboratories, Stanford University, June 1981

[3.25] H.J.Park, P.K.Ko, C.Hu, "A Non-Quasistatic MOSFET Model for SPICE - AC Analysis", was submitted to IEEE Trans. Electron Devices

[3.26] P.Yang, B.D.Epler, P.K.Chatterjee, "An Investigation of the Charge Conservation Problem for MOSFET Circuit Simulation", IEEE JSSC Vol.SC-18, No.1, pp.128-138, Feb. 1983

[3.27] H.J.Park, P.K.Ko, C.Hu, "A Charge-Conserving Non-Quasistatic MOSFET Model for SPICE Transient Analysis", to be published in IEDM 88 Technical Digest, Dec. 1988

[3.28] L.W.Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits", ERL Memo M520, University of California, Berkeley, May 1975

Chapter 4

A NON-QUASISTATIC MOSFET MODEL FOR AC ANALYSIS

4.1. Introduction

MOSFET has been widely used in the analog circuits [4.1] as well as in the digital VLSI circuits. And the circuit simulation programs, such as SPICE, have been used extensively to predict the circuit performance. But the MOSFET device models available in the circuit simulation programs are not adequate in some applications, as pointed out by some researchers [4.2] [4.3] [4.4].

One of the problems is that the quasistatic small-signal conductance and capacitance are used for the large-signal transient and the small signal AC analyses. This gives erroneous simulation results when the transient signal changes faster than the channel transit time. The small-signal frequency response has similar limitations at high frequencies.

Another problem of the present charge based MOSFET capacitance models is the uncertainty in the channel charge partitioning scheme [4.5] [4.6]. The gate, bulk and channel charge can be derived analytically. But to find the transient or AC drain and source currents, the channel charge must be partitioned into the drain and the source charges. Oh et. al. derived the analytic channel charge partitioning scheme based on the quasistatic approximation [4.2]. This gives the 40/60 partitioning between the drain and the source charges respectively, in saturation region. But this scheme produces anomalous results for the fast transient and the high frequency AC analyses. Other channel charge partitioning schemes, such as 0/100(TI [4.7], BSIM [4.8]) , 50/50(BSIM) and XQC(SPICE [4.9]), have been devised but none is valid under all operating conditions.

Some examples of their failings will be presented later. The ratio 0/100 and 50/50 refer to the ratio of the drain and source charges in saturation region.

The need for a non-quasistatic MOSFET model has been pointed out by several researchers [4.2] [4.4]. Turchetti et al.[4.10] solved the continuity equation for the transient analysis using an assumed quasistatic(QS) carrier density profile which is linear between source and drain plus a non-quasistatic(NQS) carrier density profile which is symmetrical between source and drain.

Bagheri et al.[4.11] and Tsividis[4.12] derived a NQS AC model based on the charge sheet model, but iterations are needed to find the surface potential and to solve the continuity equation. One may expect that these models would take rather long CPU times due to the iterations required, if they were implemented in circuit simulation programs.

In this work, the current continuity equation is approximated to a diffusion equation with a time and position dependent diffusion coefficient. Simple analytic expressions have been derived for the small-signal AC analyses as well as the large signal transient analysis. The model for the large-signal transient analysis is shown in Chapter 3.

This model has been implemented into SPICE3 based on the charge sheet formulation[4.13], and the required CPU time is about two to three times longer than that required for the conventional quasistatic(QS) MOSFET models in SPICE3. Many other works have been reported on the charge sheet model [4.14] [4.15] [4.16] but the model described in [4.13] has been chosen in this study to reduce the CPU time without sacrificing accuracy.

Section 4.2 shows the derivation of AC model equations along with the discussion on the validity of approximations and the 1-D numerical analysis for comparison with this work.

Section 4.3 shows model equations in saturation region along with the frequency response of AC drain and source currents, AC transconductances and AC channel current partitioning ratio in saturation region.

Section 4.4 shows simulation results and comparison with other models such as QS models, the 1-D numerical analysis and the multiple lumped model. Circuits simulated include an NMOS inverter, a CMOS inverter, a 2 stage CMOS OP Amp and a folded cascode CMOS OP Amp.

Section 4.5 concludes this chapter.

4.2. Formulation of model equations

4.2.1. NQS AC model equations

For AC analysis we start from the diffusion equation (3.10) which has been derived in Chapter 3 and it is repeated here for clarity.

$$\frac{\partial P(y,t)}{\partial t} = D(y,t)\cdot\frac{\partial^2 P(y,t)}{\partial y^2} \tag{4.1}$$

where $$P(y,t) = (Q_n'(y,t) + F_B\ V_t)^2 \tag{4.2}$$

$$D(y,t) = \frac{\mu_n}{F_B}\cdot(Q_n'(y,t) + F_B V_t) \tag{4.3}$$

The approximations used in the derivation of the diffusion equation (4.1) are the conventional linearization approximation of the bulk charge (eq.(2.18)) and the condition (4.4). The condition (4.4) is valid under all operating regions except for a narrow range of the weak inversion region, which will be shown in the next section (Fig.4.2.(a)).

$$\left| \frac{\partial Q_n'}{\partial t} \right| >> V_t \left| \frac{\partial F_B}{\partial t} \right| \tag{4.4}$$

In the same way as in Chapter 3, we can establish two boundary conditions for P, P_S and P_D at the source and drain ends of the channel respectively, by assuming that the carrier densities at those positions respond instantaneously to the applied bias.

$$P_S(t) = (V_{GST}(t) + F_B V_t)^2 \tag{4.5}$$

$$P_D(t) = (V_{GST}(t) + F_B V_t - F_B(\Psi_{SL}(t) - \Psi_{SO}(t)))^2 \tag{4.6}$$

Ψ_{SO} and Ψ_{SL} are surface potentials at the source and drain ends respectively and are computed in the same way as shown in Chapter 3.

For the small-signal AC analysis, the quasistatic value is used for the diffusion coefficient $D(y,t)$ in (4.1). This can be justified by assuming that the small-signal AC component in $(Q_n' + F_B V_t)$ of (4.3) is much smaller than the DC component. The DC component of $(Q_n' + F_B V_t)$ can be computed from (4.1) with the time derivative set to 0.

$$D(y,t) = D(0,t)\cdot\sqrt{1 - r\cdot\frac{y}{L}} \tag{4.7}$$

$$\text{where } D(0,t) = \frac{\mu_n}{F_B}\cdot(V_{GST} + F_B V_t) \tag{4.8}$$

$$r = 1 - \frac{P_D(t)}{P_S(t)} \tag{4.9}$$

For the small-signal sinusoidal excitation, we decompose $P(y,t)$ into the DC components and the small-signal component. The DC components are derived from (4.1) with the time derivative set to 0.

$$P(y,t) = P_S + (P_D - P_S)\cdot\frac{y}{L} + p(y,\omega)\cdot e^{j\omega t} \tag{4.10}$$

where P_S and P_D are DC values at the operating point and $p(y,\omega)$ is the phasor of the sinusoidal excitation. Substituting (4.10) into (4.1), we have

$$\alpha^2\cdot p(y',\omega) = \frac{\partial^2 p}{\partial y'^2}\cdot\sqrt{1 - r\cdot y'} \tag{4.11}$$

$$\text{where } \alpha = \sqrt{j\,\frac{\omega}{\omega_T}} = (1 + j)\cdot\sqrt{\frac{\omega}{2\omega_T}} \tag{4.12}$$

$$\omega_T = \frac{\mu_n(V_{GST} + F_B V_t)}{F_B L^2} = \frac{1}{\tau_T} \tag{4.13}$$

where $y' = y/L$ and τ_T is the channel transit time including the diffusion term $F_B V_t$. ω_T is determined by the DC operating point as shown in (4.14).

If r=0, the exact solution to (4.11) is $e^{\alpha y'}$ and $e^{-\alpha y'}$. Observing that the factor r lies between 0 and 1, we approximate the solution to (4.11), including the perturbation by non zero r, as

$$p(y',\omega) = G\cdot e^{\alpha y'}\cdot(1+b_1y'+b_2y'^2) + H\cdot e^{-\alpha y'}\cdot(1+d_1y'+d_2y'^2) \tag{4.14}$$

Eq. (4.14) is different from the early version of this work briefly described in [4.13] and y'^2 terms are newly added for good approximation. The coefficients b_1,b_2,d_1 and d_2 are asymmetry factors and can be found by substituting (4.14) into (4.11) and matching $p(0,\omega)$ and the integrated $p(y',\omega)$ from y'=0 to y'=1, in both sides of the equation. Hence,

$$b_1 = \frac{-B}{C - E\cdot\alpha} \tag{4.15}$$

$$b_2 = -\alpha\cdot b_1 \tag{4.16}$$

$$d_1 = \frac{-B}{C + E\cdot\alpha} \tag{4.17}$$

$$d_2 = \alpha\cdot d_1 \tag{4.18}$$

$$B = r^2\cdot\left\{ r - \frac{2}{3}(1-(1-r)^{1.5}) \right\} \tag{4.19}$$

$$C = 3r\cdot\left\{ \frac{2}{3}(1-(1-r)^{1.5}) - \frac{2}{5}(1-(1-r)^{2.5}) \right\} + 0.5r^3 \tag{4.20}$$

$$E = \frac{r^3}{3} - \frac{2}{3}(1-(1-r)^{1.5}) + \frac{4}{5}(1-(1-r)^{2.5}) - \frac{2}{7}(1-(1-r)^{3.5}) \tag{4.21}$$

Coefficients B,C and E are real positive numbers which increase monotonically with r, as shown in Fig.4.1. The magnitude is in the sequence of C > B > E for 0<r<1. Table 4.1 shows the values of B,C and E for r=0 and r=1. Coefficients b_1,b_2,d_1 and d_2 go to 0 asymptotically as r goes to 0, as implied in Table 4.1.

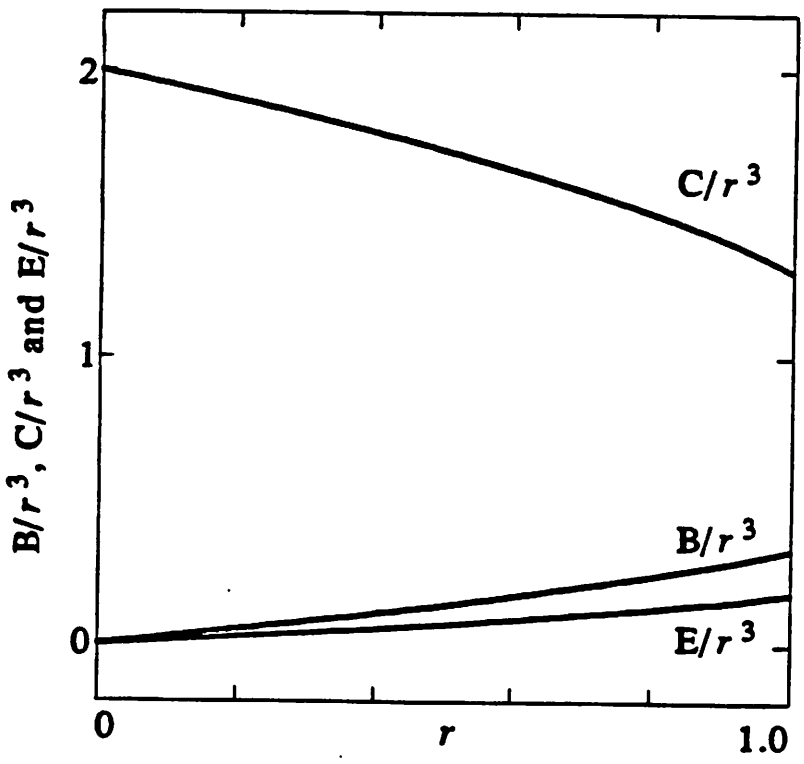


Fig.4.1. Coefficients B/r^3, C/r^3 and E/r^3 in eq. (4.19), (4.20) and (4.21) with respect to r

	$r \approx 0$	$r = 1$
B	$\frac{1}{4}\cdot r^4$	$\frac{1}{3}$
C	$2\cdot r^3$	1.3
E	$\frac{1}{8}\cdot r^4$	$\frac{19}{105}$

Table 4.1. Coefficients B, C and E in eq. (4.19), (4.20) and (4.21) for $r \approx 0$ and $r=1$ (*saturation*)

From the two boundary values of $p(y',\omega)$ at the source and the drain ends, p_S and p_D, we can find the coefficients G and H, as.

$$G = \frac{p_S(1+d_1+d_2)\cdot e^{-\alpha} - p_D}{(1+d_1+d_2)\cdot e^{-\alpha} - (1+b_1+b_2)\cdot e^{\alpha}} \tag{4.22}$$

$$H = \frac{p_D - p_S(1+b_1+b_2)\cdot e^{\alpha}}{(1+d_1+d_2)\cdot e^{-\alpha} - (1+b_1+b_2)\cdot e^{\alpha}} \tag{4.23}$$

Substituting eq.'s from (4.15) to (4.21) into (4.14), the phasor profile $p(y',\omega)$ can be rewritten as

$$p(y',\omega) = \tag{4.24}$$
$$\frac{\left\{ p_S\cdot\left[((C-B)(C-By')+(B-E)(E-By'^2)\alpha^2)\cdot\sinh(\alpha(1-y')) + B\cdot((C-E)-(B-E)y'-(C-B)y'^2)\cdot\alpha\cdot\cosh(\alpha(1-y'))\right] + p_D\cdot\left[(C(C-By')-E(E-By'^2)\alpha^2)\cdot\sinh(\alpha y') + B(-Ey'+Cy'^2)\cdot\alpha\cdot\cosh(\alpha y')\right]\right\}}{(C(C-B)+E(B-E)\alpha^2)\cdot\sinh\alpha + B(C-E)\cdot\alpha\cdot\cosh\alpha}$$

As shown in Table 4.2, the DC small-signal profile $p(y',0)$ is linear in y', and $p(y',\omega)$ at

very high frequencies ($\omega >> \omega_T$) is reduced to two exponential tails which are non-zero only near the drain and source ends respectively.

ω	$p(y',\omega)$
0	$p_S \cdot (1-y') + p_D \cdot y'$
$>> \omega_T$	$\approx (p_S \cdot e^{-\alpha y'} + p_D \cdot e^{-\alpha(1-y')})$

Table 4.2. The phasor profile $p(y',\omega)$ in (4.24) at DC (ω=0) and very high frequencies($\omega >> \omega_T$)

If r=0, that is P_S=P_D, (4.24) is reduced to (4.25), which is symmetric with respect to source and drain.

$$p(y',\omega)\,|_{r=0} = \frac{p_S \cdot \sinh(\alpha(1-y')) + p_D \cdot \sinh(\alpha y')}{\sinh\alpha} \tag{4.25}$$

The phasors p_S and p_D can be found using the derivatives of P_D and P_S in (4.5) and (4.6) with respect to applied biases.

$$p_S = \frac{\partial P_S}{\partial v_G} \cdot v_G + \frac{\partial P_S}{\partial v_B} \cdot v_B + \frac{\partial P_S}{\partial v_D} \cdot v_D + \frac{\partial P_S}{\partial v_S} \cdot v_S \tag{4.26}$$

$$p_D = \frac{\partial P_D}{\partial v_G} \cdot v_G + \frac{\partial P_D}{\partial v_B} \cdot v_B + \frac{\partial P_D}{\partial v_D} \cdot v_D + \frac{\partial P_D}{\partial v_S} \cdot v_S \tag{4.27}$$

where v_G, v_B, v_D and v_S are the small-signal phasors of each node voltage respectively. Eq. (4.26) and (4.27) are based on the assumption that the small-signal excitations at source and drain ends follow the applied biases instantaneously.

Using (3.6),(4.2) and (4.10), the small-signal channel current i_y can be written as

$$i_y = \frac{W}{L} \cdot C_{OX} \cdot \frac{\mu_n}{2F_B} \cdot \frac{\partial p(y',\omega)}{\partial y'} \tag{4.28}$$

From (4.24) and (4.28), the small-signal drain and source currents can be written as

$$i_D = \frac{W}{L}\cdot C_{OX}\cdot\frac{\mu_n}{2F_B}\cdot \tag{4.29}$$

$$\frac{p_S\cdot\left[((C^2-BE)-(B-E)^2\cdot\alpha^2)\cdot\alpha\right]-p_D\cdot\left[(-BC+B(3E-C)\cdot\alpha^2)\cdot\sinh\alpha+(C^2-BE+BC+E(B-E)\cdot\alpha^2)\cdot\alpha\cdot\cosh\alpha\right]}{(C(C-B)+E(B-E)\cdot\alpha^2)\cdot\sinh\alpha+B(C-E)\cdot\alpha\cdot\cosh\alpha}$$

$$i_S = -\frac{W}{L}\cdot C_{OX}\cdot\frac{\mu_n}{2F_B}\cdot \tag{4.30}$$

$$\frac{p_S\cdot\left[(B(C-B)+B(C-E)\cdot\alpha^2)\cdot\sinh\alpha+(B(B-E)+C(C-B)+E(B-E)\cdot\alpha^2)\cdot\alpha\cdot\cosh\alpha\right]-p_D\cdot\left[(C^2-BE-E^2\cdot\alpha^2)\cdot\alpha\right]}{(C(C-B)+E(B-E)\cdot\alpha^2)\cdot\sinh\alpha+B(C-E)\cdot\alpha\cdot\cosh\alpha}$$

For the AC analysis in SPICE, we need the derivatives of small-signal currents, i_D ,i_S ,i_G and i_B with respect to biases, which are used to construct the linearized small-signal equivalent circuit. From (4.29) and (4.30), we can find the derivatives of i_D and i_S with respect to p_D and p_S.

Including the channel length modulation effect, we can obtain the following derivatives of i_D and i_S with respect to biases

$$\frac{\partial i_D}{\partial v_G} = \frac{\partial i_D}{\partial p_S}\cdot\frac{\partial p_S}{\partial v_G} + \frac{\partial i_D}{\partial p_D}\cdot\frac{\partial p_D}{\partial v_G} + \frac{\partial i_D}{\partial \beta_p}\cdot\frac{\partial \beta_p}{\partial v_G} \tag{4.31}$$

$$\frac{\partial i_D}{\partial v_D} = \frac{\partial i_D}{\partial p_S}\cdot\frac{\partial p_S}{\partial v_D} + \frac{\partial i_D}{\partial p_D}\cdot\frac{\partial p_D}{\partial v_D} + \frac{\partial i_D}{\partial \beta_p}\cdot\frac{\partial \beta_p}{\partial v_D} + \lambda\cdot\beta_p(P_S-P_D) \tag{4.32}$$

$$\frac{\partial i_D}{\partial v_B} = \frac{\partial i_D}{\partial p_S}\cdot\frac{\partial p_S}{\partial v_B} + \frac{\partial i_D}{\partial p_D}\cdot\frac{\partial p_D}{\partial v_B} + \frac{\partial i_D}{\partial \beta_p}\cdot\frac{\partial \beta_p}{\partial v_B} \tag{4.33}$$

$$\frac{\partial i_D}{\partial v_S} = -\left[\frac{\partial i_D}{\partial v_G} + \frac{\partial i_D}{\partial v_D} + \frac{\partial i_D}{\partial v_B}\right] \tag{4.34}$$

$$\frac{\partial i_S}{\partial v_G} = \frac{\partial i_S}{\partial p_S}\cdot\frac{\partial p_S}{\partial v_G} + \frac{\partial i_S}{\partial p_D}\cdot\frac{\partial p_D}{\partial v_G} + \frac{\partial i_S}{\partial \beta_p}\cdot\frac{\partial \beta_p}{\partial v_G} \tag{4.35}$$

$$\frac{\partial i_S}{\partial v_D} = \frac{\partial i_S}{\partial p_S}\cdot\frac{\partial p_S}{\partial v_D} + \frac{\partial i_S}{\partial p_D}\cdot\frac{\partial p_D}{\partial v_D} + \frac{\partial i_S}{\partial \beta_p}\cdot\frac{\partial \beta_p}{\partial v_D} - \lambda\cdot\beta_p(P_S-P_D) \tag{4.36}$$

$$\frac{\partial i_S}{\partial v_B} = \frac{\partial i_S}{\partial p_S}\cdot\frac{\partial p_S}{\partial v_B} + \frac{\partial i_S}{\partial p_D}\cdot\frac{\partial p_D}{\partial v_B} + \frac{\partial i_S}{\partial \beta_p}\cdot\frac{\partial \beta_p}{\partial v_B} \tag{4.37}$$

$$\frac{\partial i_S}{\partial v_S} = -\left[\frac{\partial i_S}{\partial v_G} + \frac{\partial i_S}{\partial v_D} + \frac{\partial i_S}{\partial v_B}\right] \tag{4.38}$$

$$where \quad \beta_p = \frac{W}{L} C_{OX} \frac{\mu_n}{2F_B} \tag{4.39}$$

The DC drain current can be computed as shown in (3.28) of Chapter 3.

Derivatives of p_S and p_D with respect to biases can be found from (4.5) and (4.6) and λ is the channel length modulation factor and P_S and P_D in (4.32) and (4.36) represent DC values.

Also the derivatives of i_B with respect to biases can be found from the quasistatic capacitance values[4.12].

$$\frac{\partial i_B}{\partial v_G} = j\omega\, C_{BG} \tag{4.40}$$

$$\frac{\partial i_B}{\partial v_D} = j\omega\, C_{BD} \tag{4.41}$$

$$\frac{\partial i_B}{\partial v_B} = j\omega\, C_{BB} \tag{4.42}$$

$$\frac{\partial i_B}{\partial v_S} = -\left[\frac{\partial i_B}{\partial v_G} + \frac{\partial i_B}{\partial v_D} + \frac{\partial i_B}{\partial v_B}\right] \tag{4.43}$$

Since i_G is the negative sum of i_D,i_S and i_B, we have

$$\frac{\partial i_G}{\partial v_G} = -\left[\frac{\partial i_D}{\partial v_G} + \frac{\partial i_S}{\partial v_G} + \frac{\partial i_B}{\partial v_G}\right] \tag{4.44}$$

$$\frac{\partial i_G}{\partial v_D} = -\left[\frac{\partial i_D}{\partial v_D} + \frac{\partial i_S}{\partial v_D} + \frac{\partial i_B}{\partial v_D}\right] \tag{4.45}$$

$$\frac{\partial i_G}{\partial v_B} = -\left[\frac{\partial i_D}{\partial v_B} + \frac{\partial i_S}{\partial v_B} + \frac{\partial i_B}{\partial v_B}\right] \tag{4.46}$$

$$\frac{\partial i_G}{\partial v_S} = -\left[\frac{\partial i_D}{\partial v_S} + \frac{\partial i_S}{\partial v_S} + \frac{\partial i_B}{\partial v_S}\right] \tag{4.47}$$

4.2.2. Validity of approximations

In the derivation of the model equations, we used four approximations.

The first approximation is that the mobility is constant along the channel region, that is, the drift velocity saturation effect is neglected. This approximation has been used in the derivation of (3.8).

The second approximation is shown in eq. (4.4).

The third approximation is shown in eq. (4.14), whose validity will be checked by comparison with the 1-D numerical solution of (4.11).

The fourth approximation is the approximation of (4.3) into (4.7), which is valid because the small-signal amplitude is assumed to be much smaller than the DC component in AC analysis.

To check the validity of the second approximation (4.4), the V_{GS} and V_{BS} dependences of $Q_n'(y')$ (for y'=0, 0.99) and $F_B V_t$ are shown in Fig.4.2.(a), where V_{DS}=2V, V_{BS}=0 and V_{GS} changes from 0 to 5V. Except in the narrow range of the weak inversion region, the V_{GS} and V_{BS} dependences of $Q_n'(y')$ are larger than those of $F_B V_t$ by several orders of magnitude over the 99% of the channel region. The DC solution is used for $Q_n'(y')$. Since the V_{DS} dependence of $F_B \cdot V_t$ is 0, the approximation (4.4) is always valid for the variation of V_{DS}. So we can see that eq. (4.4) is a good approximation in the strong inversion region over the 99% of the channel region.

Fig.4.2.(b) shows the values of F_B for V_{BS}= 0, $-1V$, $-5V$ respectively with respect to V_{GS}. The normalized inversion charge density at source, $Q_n'(0)$, is also shown for comparison. Compared to $Q_n'(0)$, F_B has a very weak bias dependence and has an almost constant value which is slightly larger than 1.

4.2.3. 1-D numerical analysis

To check the validity of the approximation in (4.14), the results have been compared with the 1-D numerical solution of (4.11). Eq. (4.11) has been derived from the current

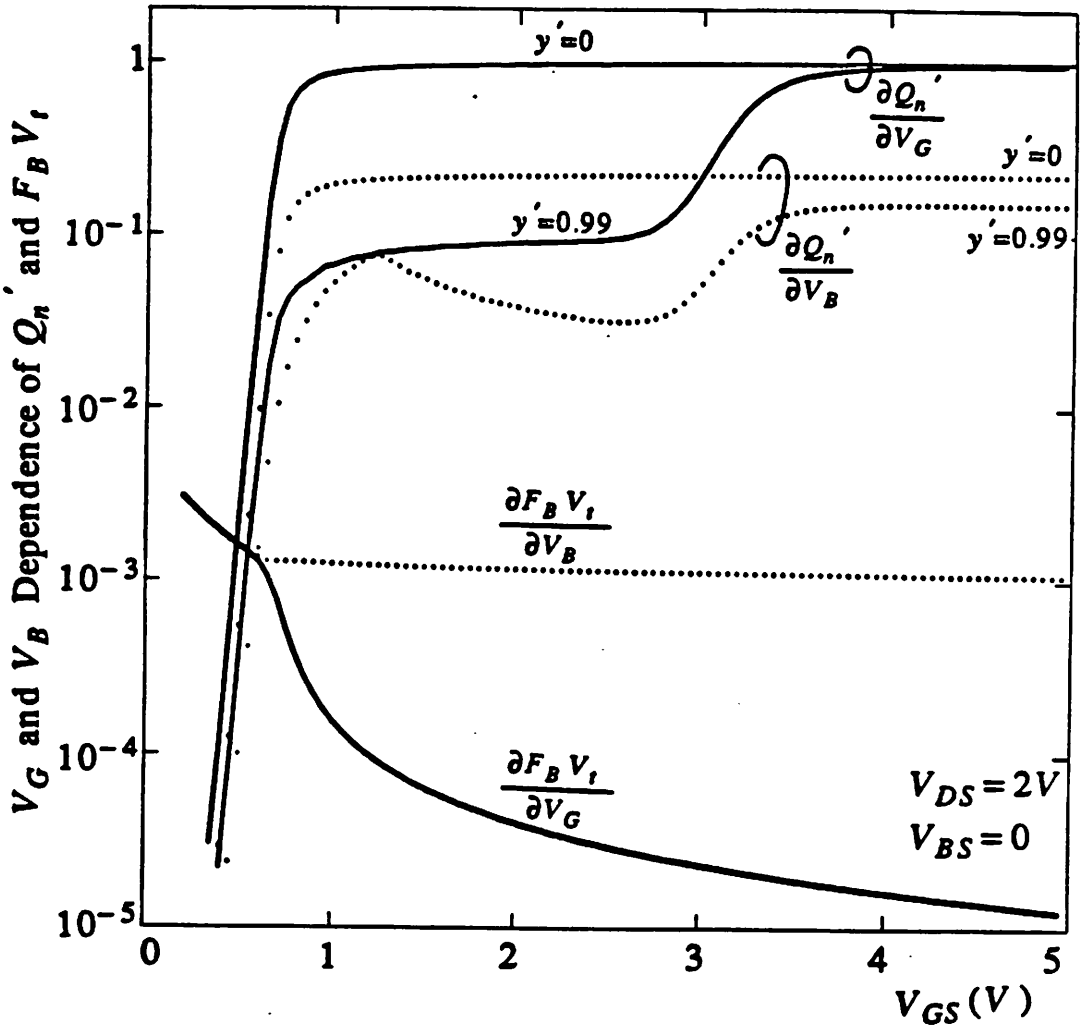


Fig.4.2. (a). Comparison of the V_{GS} and V_{BS} dependences of $Q_n'(y')$ (for $y' = 0$ and 0.99) and $F_B V_t$ to check the validity of the approximation (4.1).

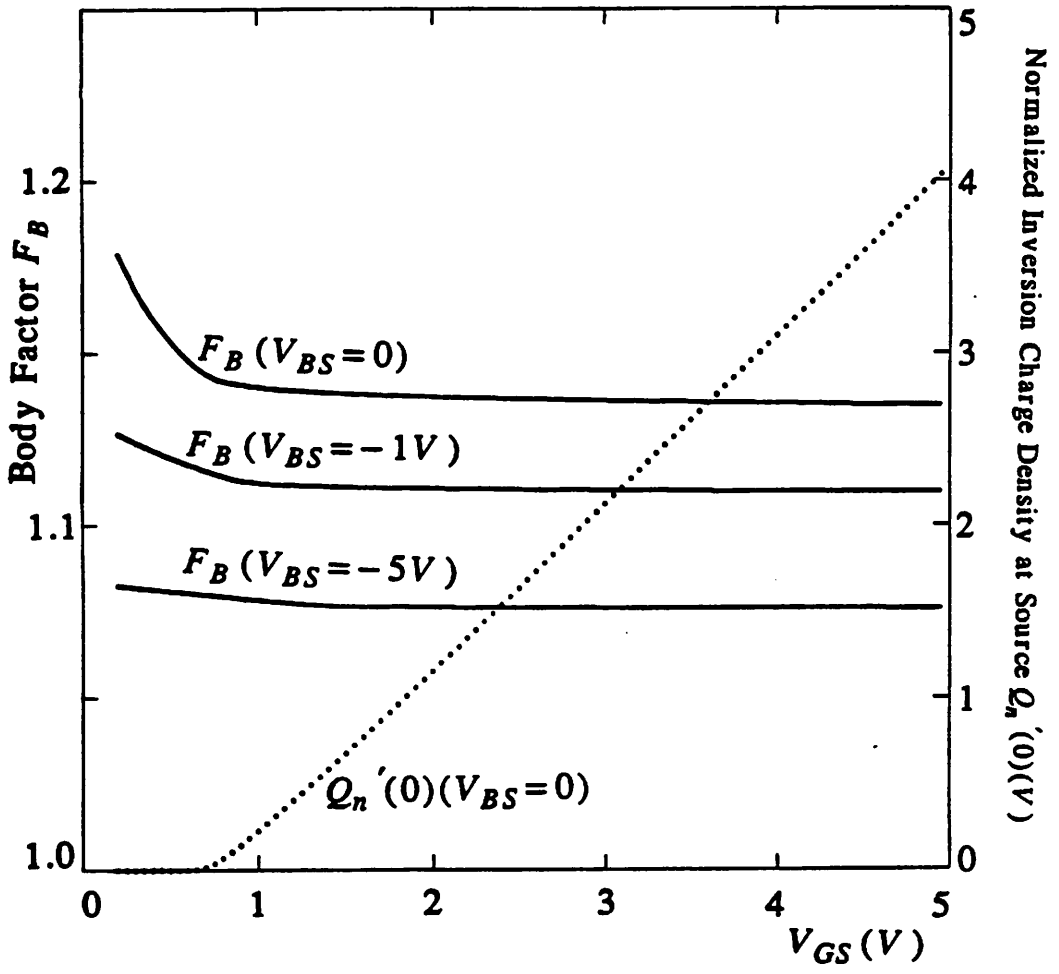


Fig.4.2.(b). Values of F_B for V_{BS} = 0, -1 and -5 V, with respect to V_{GS}. $Q_n'(0)$, the normalized inversion charge density at source, is also shown for comparison.

continuity equation using the approximations (4.4), (4.7) and assuming the constant mobility.

Discretizing (4.11) with respect to y', we have

$$p_i = \frac{\sqrt{1-r\cdot ih}}{2\sqrt{1-r\cdot ih}+j\cdot h^2\cdot\frac{\omega}{\omega_T}}\cdot(p_{i-1}+p_{i+1}) \tag{4.48}$$

where i is an integer for the grid number and ranges from 0 to NY. Uniform grids are used. The grid number 0 (i=0) represents the source end and the grid number NY (i=NY) represents the drain end of the channel. h is the interval of y' between grids (h=1/NY). j is the imaginary number notation. Decomposing $\{p_i\}$ in (4.48) into the in-phase (real) component $\{p_{re.i}\}$ and the out-of-phase (imaginary) component $\{p_{im.i}\}$ and matching the in-phase and out-of-phase components respectively in both sides of (4.48), we have

$$p_{re.i} = T_1\cdot(p_{re.(i-1)}+p_{re.(i+1)}) + T_2\cdot(p_{im.(i-1)}+p_{im.(i+1)}) \tag{4.49}$$

$$p_{im.i} = T_1\cdot(p_{im.(i-1)}+p_{im.(i+1)}) - T_2\cdot(p_{re.(i-1)}+p_{re.(i+1)}) \tag{4.50}$$

where T_1 and T_2 are real numbers and can be derived as $T1 = 4\cdot(1-r\cdot ih)/(4\cdot(1-r\cdot ih)+h^4\cdot(\omega/\omega_T)^2)$ and $T2 = \sqrt{1-r\cdot ih}\cdot h^2\cdot(\omega/\omega_T)/(4\cdot(1-r\cdot ih)+h^4\cdot(\omega/\omega_T)^2)$. The second subscript of p represents the grid number.

Assuming that $p(y',\omega)$ follows the applied small-signal voltages instantaneously at drain and source ends, we have two boundary conditions for each of $\{p_{re.i}\}$ and $\{p_{im.i}\}$.

In saturation region, there is no excitation at the drain end, that is, $p_{re.NY}$=$p_{im.NY}$=0. And the imaginary value of p at the source end ($p_{im.0}$) is set to 0 because only the relative phase is important. Hence, the boundary conditions of p in saturation region becomes

$$p_{re.0} = p_S \tag{4.51}$$

$$p_{re.NY} = 0 \tag{4.52}$$

$$p_{im.0} = 0 \tag{4.53}$$

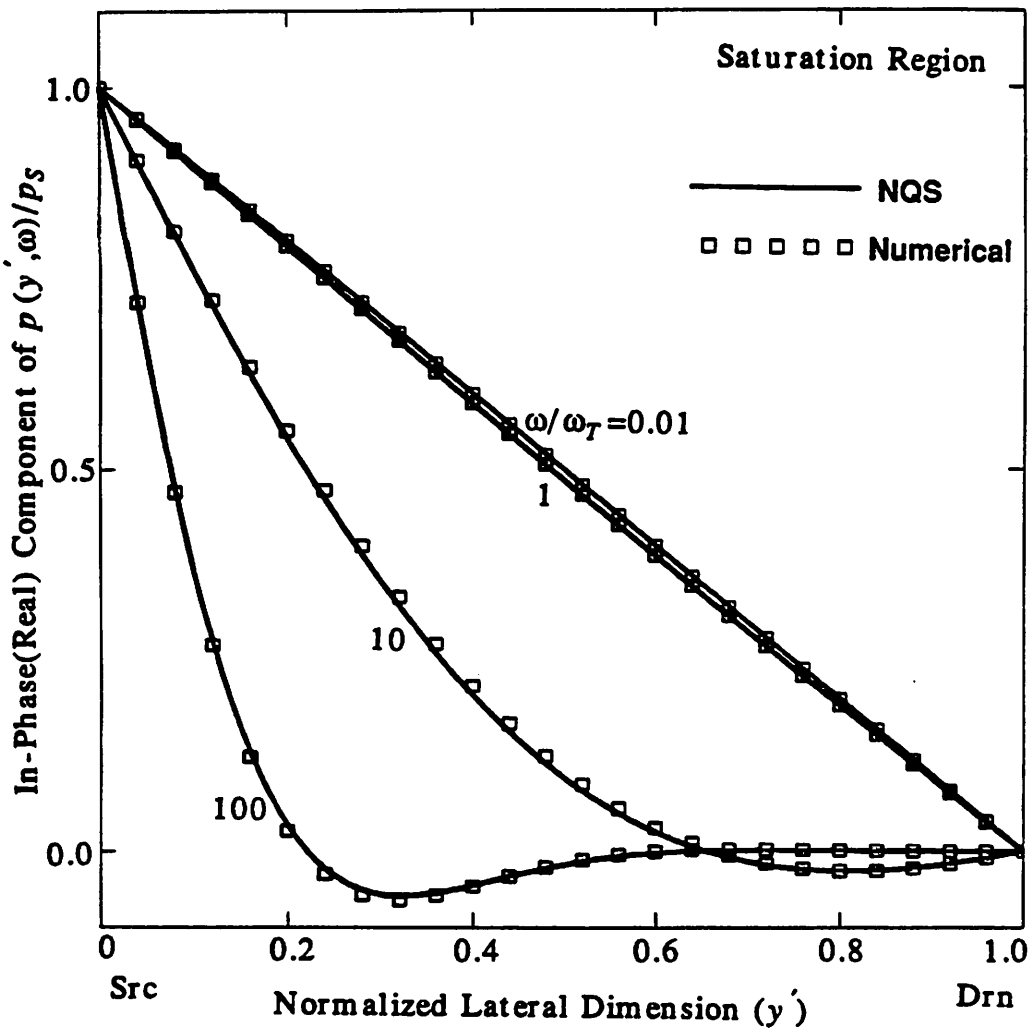


Fig.4.3. The profile $p(y',\omega)$ in saturation region for ω/ω_T= 0.01, 1, 10 and 100, from the approximation (4.14) (solid line) and the 1-D numerical solution of (4.11) (squares)
(a). In-phase (real) component

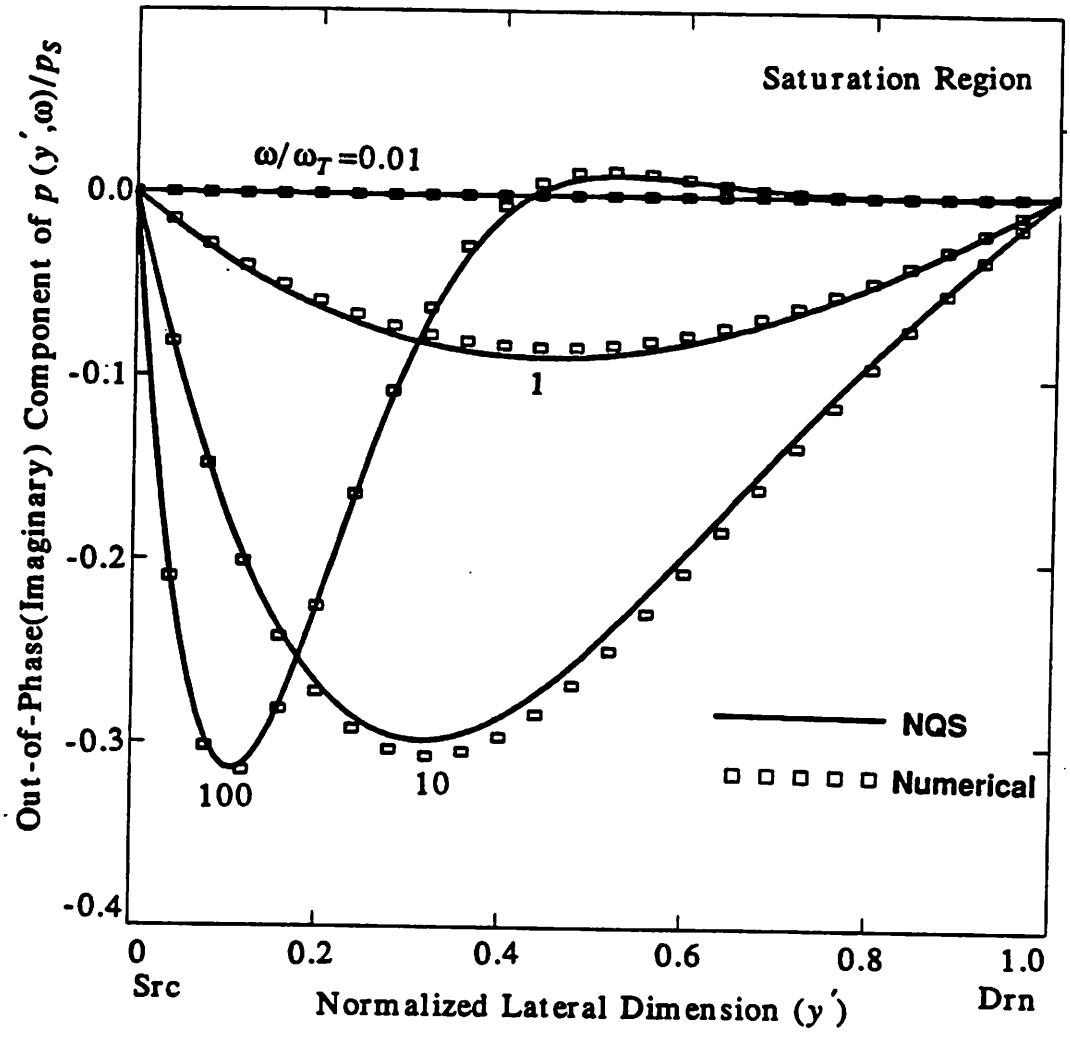


Fig.4.3.(b). Out-of-phase (imaginary) component

$$p_{im,NY} = 0 \qquad (4.54)$$

By iterations, we can find the in-phase profile $\{p_{re,i}\}$ and the out-of-phase profile $\{p_{im,i}\}$. From these profiles, we can get the in-phase and the out-of-phase components of i_D and i_S, using (4.28). NY=100 is used in the following examples.

Fig.4.3.(a) and 4.3.(b) show the in-phase and the out-of-phase components of $p(y',\omega)$ in saturation region which have been computed from the approximation (4.14) and also from the 1-D numerical solution of (4.11), for ω/ω_T= 0.01, 1, 10 and 100. Good agreements can be observed between the NQS model (this work) and the 1-D numerical analysis.

At low frequencies (ω/ω_T= 0.01), the in-phase component of $p(y',\omega)$ is almost linear in y'. At very high frequencies(ω/ω_T= 100), the in-phase component of $p(y',\omega)$ is proportional to $e^{-0.5\sqrt{\omega/\omega_T}\cdot y'}\cdot\cos(0.5\sqrt{\omega/\omega_T}\cdot y')$ and the out-of-phase component is proportional to $-e^{-0.5\sqrt{\omega/\omega_T}\cdot y'}\cdot\sin(0.5\sqrt{\omega/\omega_T}\cdot y')$. The peak point of the out-of-phase component moves toward the source as the frequency ω increases.

The profiles for the NQS model (this work) has been derived from (4.24). At ω = 0, (4.24) is reduced to $(p_S(1-y'))$ in saturation region. At $\omega >> \omega_T$, (4.24) is reduced to $(p_S\cdot e^{-\alpha y'} + p_D\cdot e^{-\alpha(1-y')})$ in saturation region. The profiles of the NQS model in saturation region are shown also in Table 4.2.

4.3. AC model in saturation region

4.3.1. NQS AC model in saturation region

For the AC analysis of MOSFET's we are mostly interested in the saturation region, where the voltage gain is high. For the high frequency MOS circuits, we are interested in the intrinsic behavior of MOSFET's at high frequencies[4.14]. To see the intrinsic behavior of MOSFET's, we neglect the extrinsic components such as the overlap

capacitance and the source, drain junction capacitance and junction current.

In saturation region, the factor r in (4.9) is 1 and p_D is 0 (no velocity saturation effect) and the small-signal drain and source current equations in (4.29) and (4.30) are reduced to

$$i_D = \beta_P p_S \cdot \frac{\alpha\cdot(1-0.014248\alpha^2)}{(0.77111+0.016920\alpha^2)\cdot\sinh\alpha + 0.22889\cdot\alpha\cdot\cosh\alpha} \tag{4.55}$$

$$i_S = -\beta_P p_S \cdot \frac{(0.19772+0.22889\alpha^2)\cdot\sinh\alpha + (0.80228+0.016920\alpha^2)\cdot\alpha\cdot\cosh\alpha}{(0.77111+0.016920\alpha^2)\cdot\sinh\alpha + 0.22889\cdot\alpha\cdot\cosh\alpha} \tag{4.56}$$

In this derivation the channel length modulation effect is not included. Table 4.3 shows i_D and i_S at DC(ω=0), at low frequencies ($\omega < \omega_T$) and at very high frequencies ($\omega \gg \omega_T$).

Expanding $\sinh\alpha$ and $\cosh\alpha$ in (4.55) and (4.56) as Taylor series in α, we can represent i_D and i_S as rational functions of s, where $s=j\omega=\alpha^2\cdot\omega_T$ and α is shown in (4.12).

$$i_D = \beta_P \cdot p_S \cdot \frac{1 - \frac{s}{70.185\omega_T}}{1 + \sum_{n=1}^{\infty}\left(\frac{s}{\omega_{p.n}}\right)^n} \tag{4.57}$$

$$i_S = -\beta_P \cdot p_S \cdot \frac{1 + \sum_{n=1}^{\infty}\left(\frac{s}{\omega_{z.n}}\right)^n}{1 + \sum_{n=1}^{\infty}\left(\frac{s}{\omega_{p.n}}\right)^n} \tag{4.58}$$

The coefficients $\{\omega_{p.n}\}$ and $\{\omega_{z.n}\}$ in (4.57) and (4.58) can be found from

$$\omega_{p.n} = \omega_T \cdot \left\{ \frac{0.77111}{(2n+1)!} + \frac{0.22889}{(2n)!} + \frac{0.016920}{(2n-1)!} \right\}^{-\frac{1}{n}} \tag{4.59}$$

$$\omega_{z.n} = \omega_T \cdot \left\{ \frac{0.19772}{(2n+1)!} + \frac{0.80228}{(2n)!} + \frac{0.22889}{(2n-1)!} + \frac{0.016920}{(2n-2)!} \right\}^{-\frac{1}{n}} \tag{4.60}$$

(for n=1,2,3,4,5,$\cdots$)

The first 10 values of $\{\omega_{p.n}/\omega_T\}$ and $\{\omega_{z.n}/\omega_T\}$ are shown in Table 4.4.

ω	$\frac{i_D}{\beta_P p_S}$	$\frac{i_S}{\beta_P p_S}$
0	1	−1
$< \omega_T$	$\approx 1 - \frac{j \cdot 0.27413 \cdot \frac{\omega}{\omega_T}}{1 + j \cdot 0.25988 \cdot \frac{\omega}{\omega_T}}$	$\approx -1 - \frac{j \cdot 0.42002 \cdot \frac{\omega}{\omega_T}}{1 + j \cdot 0.25988 \cdot \frac{\omega}{\omega_T}}$
$\gg \omega_T$	≈ 0	$\approx (-1-j)\sqrt{\frac{\omega}{2\omega_T}}$

Table.4.3. The small-signal drain and source currents i_D and i_S for NQS model (eq. (4.55) and (4.56)) in saturation region at DC(ω=0), low frequencies($\omega < \omega_T$) and high frequencies($\omega >> \omega_T$). i_D and i_S are normalized by the DC value $\beta_P p_S$

n	1	2	3	4	5
$\frac{\omega_{p.n}}{\omega_T}$	3.8479	7.2966	11.779	17.302	23.871
$\frac{\omega_{z.n}}{\omega_T}$	1.4708	3.4989	6.4275	10.285	15.094
n	6	7	8	9	10
$\frac{\omega_{p.n}}{\omega_T}$	31.490	40.163	49.892	60.680	72.529
$\frac{\omega_{z.n}}{\omega_T}$	20.870	27.626	35.372	44.118	53.871

Table.4.4. The first 10 values of coefficients $\{\omega_{p.n}\}$, $\{\omega_{z.n}\}$ in (4.57) and (4.58)

In (4.57), i_D has a single *zero* in the right-hand side of s-plane, which is almost two orders of magnitude larger than ω_T and can be neglected for the simple analysis. Since all the coefficients of s terms in the denominator of (4.57) are positive, i_D has infinitely many *poles* in the left-hand side of s-plane. Similarly from (4.58), we can see that i_S has infinitely many *zeroes* in the left-hand side of s-plane. The *poles* of i_S are the same as those of i_D.

$\beta_P \cdot p_S$ can be rewritten as

$$\beta_P \cdot p_S = g_M \cdot v_{GS} + g_{MBS} \cdot v_{BS} \tag{4.61}$$

where v_{GS} and v_{BS} are the phasors of the small-signal voltages and g_M and g_{MBS} are the DC transconductances and can be approximated in saturation region as.

$$g_M \approx \frac{W}{L} \cdot C_{OX} \cdot \frac{\mu_n}{F_B} \cdot (V_{GST} + F_B \cdot V_t) \tag{4.62}$$

$$g_{MBS} \approx g_M \cdot (-\frac{\partial V_{TH}}{\partial V_{BS}}) \tag{4.63}$$

where V_{TH} is the conventional threshold voltage.

Low frequency approximation

To see the low frequency behavior of the NQS model, we rewrite i_D and i_S as the sum of DC and capacitive components, using (4.57),(4.58) and (4.61).

$$i_D = (g_M \cdot v_{GS} + g_{MBS} \cdot v_{BS}) + s \cdot (C'_{DG} \cdot v_G + C'_{DB} \cdot v_B + C'_{DS} \cdot v_S + C'_{DD} \cdot v_D) \cdot \frac{1 + \frac{s}{14.595\omega_T} + \cdots}{1 + \frac{s}{3.8479\omega_T} + \cdots} \tag{4.64}$$

$$i_S = -(g_M \cdot v_{GS} + g_{MBS} \cdot v_{BS}) + s \cdot (C'_{SG} \cdot v_G + C'_{SB} \cdot v_B + C'_{SS} \cdot v_S + C'_{SD} \cdot v_D) \cdot \frac{1 + \frac{s}{2.8236\omega_T} + \cdots}{1 + \frac{s}{3.8479\omega_T} + \cdots} \tag{4.65}$$

For low frequencies such that $\omega < 0.38\omega_T$, the fraction of s terms at the right end of (4.64) and (4.65) can be neglected with error less than 10%.

	NQS	QS 40/60	QS 0/100	Meyer
$\frac{C_{DG}}{WLC_{OX}}$	–0.274	–0.267	0	0
$\frac{C_{DB}}{WLC_{OX}}$	–0.070*	–0.068*	0	0
$\frac{C_{DS}}{WLC_{OX}}$	+0.344*	+0.335*	0	0
$\frac{C_{DD}}{WLC_{OX}}$	0	0	0	0
$\frac{C_{SG}}{WLC_{OX}}$	–0.420	–0.400	–0.667	–0.667
$\frac{C_{SB}}{WLC_{OX}}$	–0.107*	–0.102*	–0.170*	0
$\frac{C_{SS}}{WLC_{OX}}$	+0.527*	+0.502*	+0.837*	+0.667
$\frac{C_{SD}}{WLC_{OX}}$	0	0	0	0

Table.4.5. The effective low-frequency capacitances of NQS model (eq. (4.64) and (4.65)) and the capacitances of QS models in saturation region. The parameters used are T_{OX}=18nm, N_{SUB}=2.1* $10^{16}cm^{-3}$ and V_{BS}=0. $C_{ij} = \partial q_i/\partial v_j$, where $i \in$ { D, S } and $j \in$ { G, B, S, D }. Values with superscript * are slightly dependent on N_{SUB}, T_{OX} and V_{BS}.

The effective capacitances in (4.64) and (4.65) are tabulated in Table 4.5. The capacitances from the QS models are also shown for comparison. Observing that the effective capacitances of NQS model are very close to those of QS 40/60 model with error less than 5% but are far from those of either QS 0/100 or Meyer model, we can see that the NQS AC model is reduced to the QS 40/60 model at low frequencies.

The capacitance values of the NQS model in Table 4.5 are not exactly the same as those of the QS 40/60 model because of the uncertainty in the truncation of s terms in the derivation steps of (4.64) and (4.65).

4.3.2. Frequency responses of small-signal currents and transconductances in saturation region

Frequency responses of small-signal drain and source currents i_D and i_S in saturation region are shown in Table 4.6 for each model. Frequency responses of QS models have been derived from the capacitance values in Table 4.5.

Frequency response of i_D in saturation region

The frequency response of the small-signal drain current i_D for the NQS model in saturation region is shown in (4.55) and (4.57). Fig.4.4.(a) and 4.4.(b) show the amplitude and the phase response of the normalized i_D with respect to the normalized frequency ω/ω_T, for the NQS model, the 1-D numerical solution of (4.11), QS 40/60, QS 0/100 and Meyer model. i_D is normalized by the DC value $\beta_P p_S$.

In the amplitude response as shown in Fig.4.4.(a), the NQS model and the 1-D numerical solution of (4.11) show that the amplitude of i_D decreases monotonically with frequency and goes asymptotically to 0 ($-\infty$ dB) at very high frequencies, while the QS 0/100 and Meyer model show that the amplitude is constant independent of frequency and the QS 40/60 model show that the amplitude increases monotonically with frequency. Since the carriers in the channel cannot respond to the very fast excitations at source and drain ends due to the inertia as implied in the diffusion equation (4.1), i_D in saturation region must be 0 at very high frequencies. So all the QS models give non-

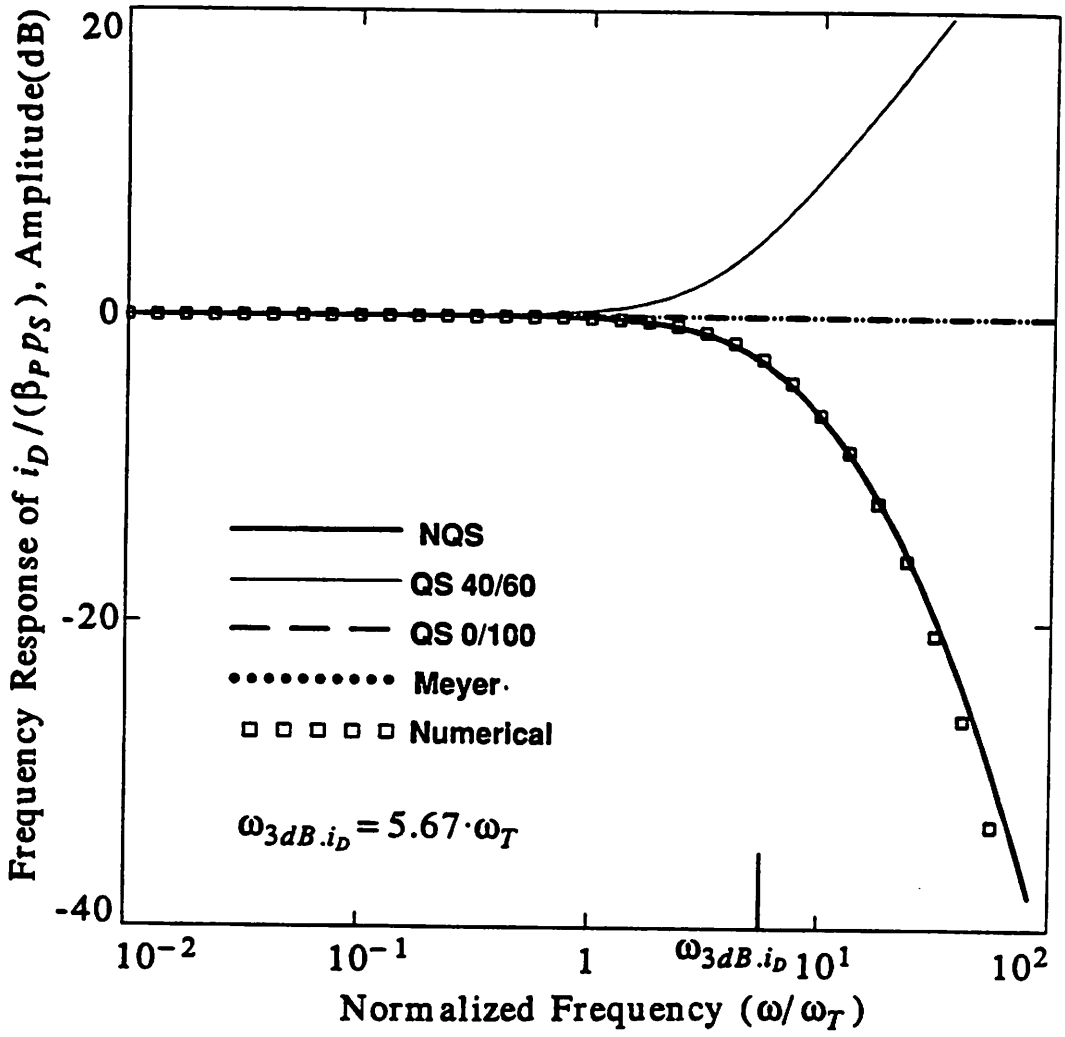


Fig.4.4. Frequency response of the small-signal drain current i_D in saturation region with respect to the normalized frequency ω/ω_T, for NQS, QS 40/60, QS 0/100, Meyer model and 1-D numerical solution of (4.11). Equations for the NQS model are shown in eq. (4.55) and (4.57)

(a). Amplitude response

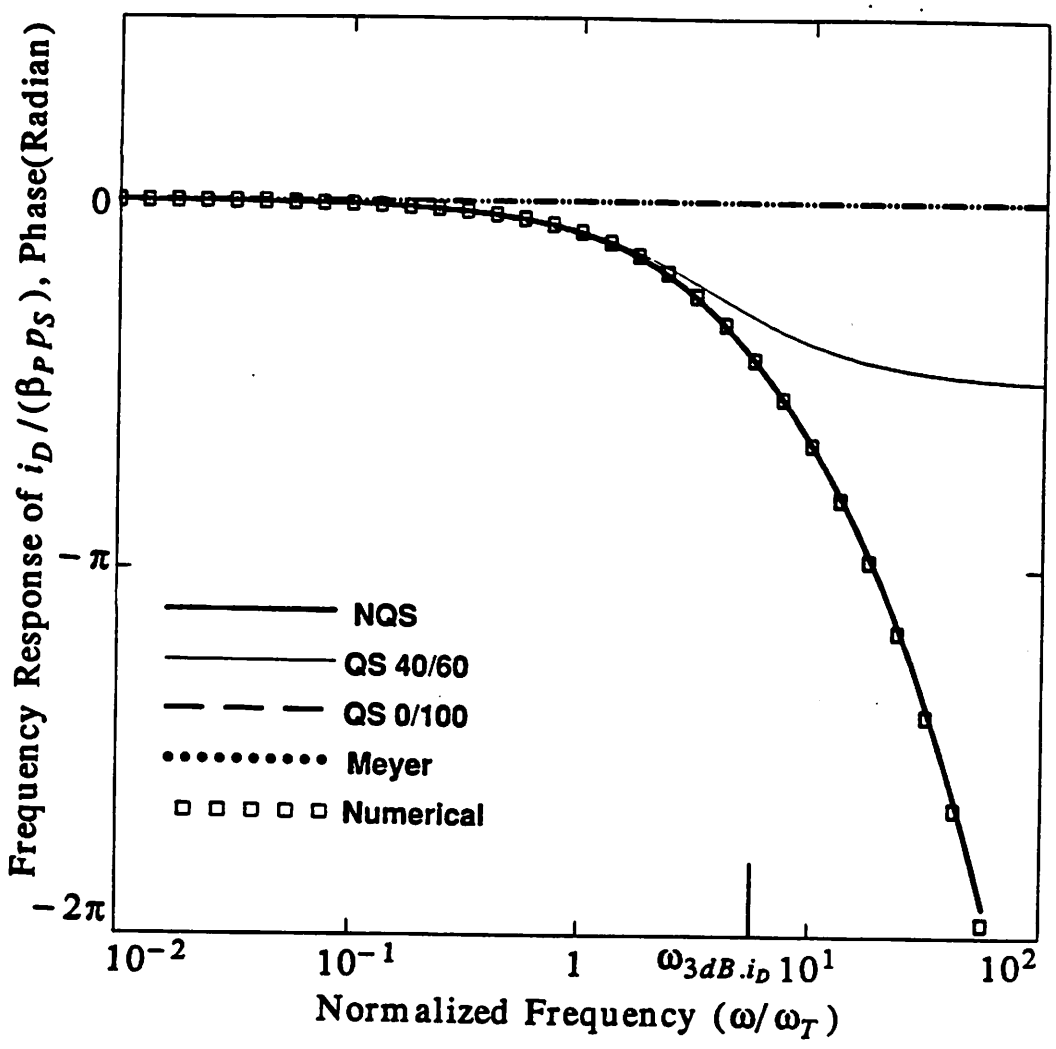


Fig.4.4.(b). Phase response

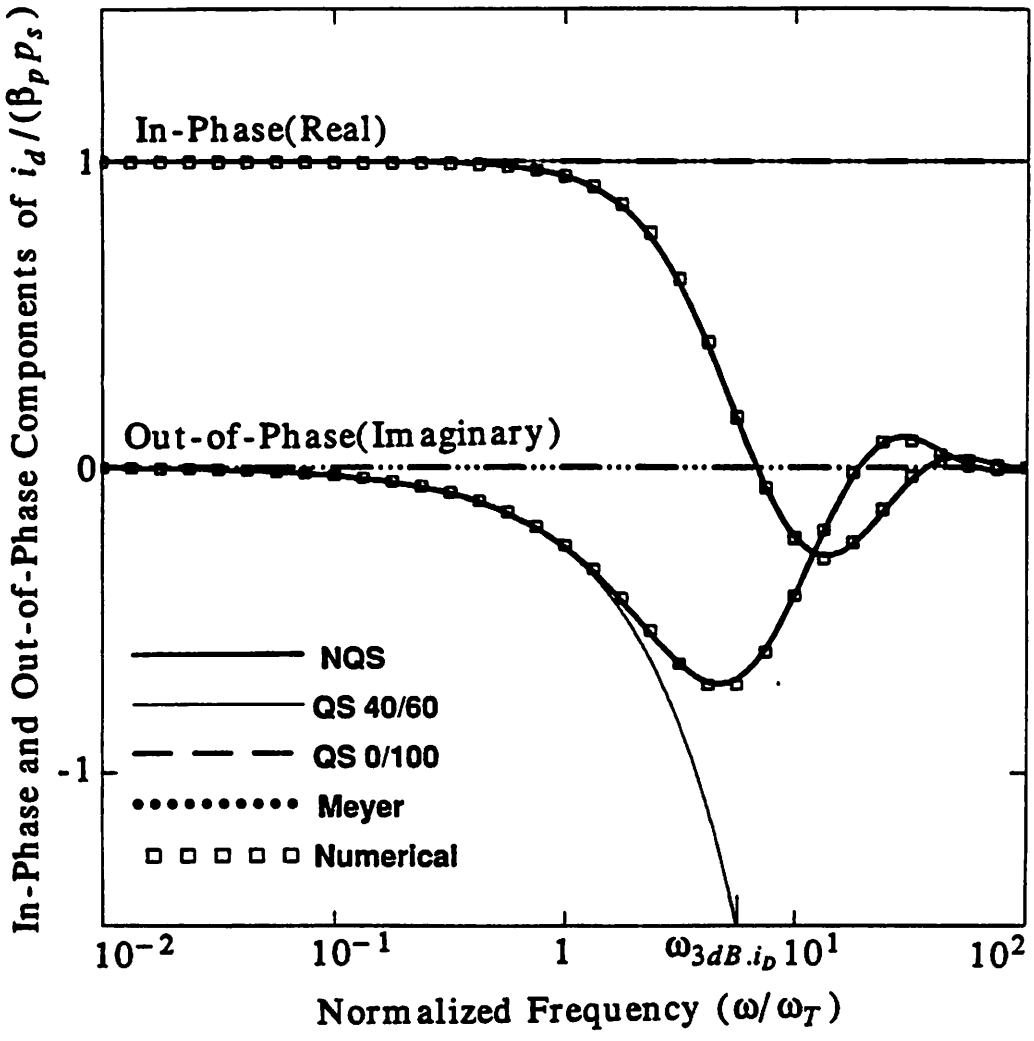


Fig.4.4.(c). In-phase (real) and out-of-phase (imaginary) components

physical results on i_D at very high frequencies.

In the phase response as shown in Fig.4.4.(b), the NQS model and the 1-D numerical solution of (4.11) show that the phase of i_D decreases unboundedly with frequency, while the QS 40/60 model shows that the phase decreases with frequency and goes asymptotically to $(-0.5)\cdot\pi$ radian at very high frequencies. The QS 0/100 and Meyer model show that the phase is 0 independent of frequency.

The 3dB frequency of NQS model, $\omega_{3dB.i_D}$, is defined to be the frequency where the amplitude of i_D drops to 3dB below the DC value. From Fig.4.4.(a) we can find $\omega_{3dB.i_D}$ as

$$\omega_{3dB.i_D} = 5.67\ \omega_T \tag{4.66}$$

And the phase shift at the 3-dB frequency in the NQS model is 75 degrees which is larger than 45 degrees in the single-pole roll-off characteristics. From this we can verify the multi-pole roll-off characteristics of i_D in the NQS model as shown in (4.57).

Fig 4.4.(c) shows the in-phase (real) and the out-of-phase (imaginary) components of i_D in saturation region with respect to ω/ω_T. Again i_D is normalized by the DC value $\beta_P p_S$. All the QS models show that the in-phase component of i_D is constant independent of frequency, but the NQS model and the 1-D numerical solution of (4.11) show that the in-phase component of i_D decreases with frequency and goes asymptotically to 0 at very high frequencies.

The QS 0/100 and Meyer model show that the out-of-phase component of i_D is 0 independent of frequency, but the QS 40/60 model shows that the out-of-phase component of i_D is proportional to ω as shown in Table 4.6.

But the NQS model and the 1-D numerical solution show that the magnitude of out-of-phase component increases with frequency for frequencies up to the 3-dB frequency $\omega_{3dB.i_D}$ and decreases with frequency beyond the 3-dB frequency and asymptotically goes to 0 at very high frequencies.

In both the in-phase and the out-of-phase components of i_D, the QS 40/60 model matches closely the NQS model with error less than 5% for $\omega < \omega_T$.

Frequency response of i_S in saturation region

The frequency response of the small-signal source current i_S in saturation region is shown in (4.56) and (4.58). Fig.4.5.(a) and 4.5.(b) show the amplitude and the phase response of the normalized i_S with respect to ω/ω_T. i_S is normalized by the DC magnitude $\beta_P p_S$.

In the amplitude response as shown in Fig.4.5.(a), all the models show that the amplitude of i_S increases monotonically with frequency. QS models predict that the magnitude is proportional to the signal frequency ω (20 dB per decade) at high frequencies, but the NQS model and the 1-D numerical analysis predict that the amplitude is proportional to $\sqrt{\omega}$ (10 dB per decade) at high frequencies.

In the phase response as shown in Fig.4.5.(b), all the models show that the phase is $-\pi$ radian at low frequencies. At very high frequencies, the QS models predict that the phase goes asymptotically to -0.5π radian, which implies that the out-of-phase component becomes much larger than the in-phase component. But the NQS model and the 1-D numerical analysis show that the phase goes asymptotically to $-(3/4)\pi$ radian at very high frequencies, which implies that the out-of-phase component becomes almost the same as the in-phase component.

Fig.4.5.(c) shows the in-phase(real) and the out-of-phase (imaginary) components of i_S with respect to ω/ω_T. All the QS models predict that the in-phase component is constant independent of frequency but the NQS model and the 1-D numerical analysis predict that the in-phase component increases in magnitude with frequency.

As shown in Table 4.6, the QS 40/60 model shows that the out-of-phase component of i_S is $-(2/5)\cdot WLC_{OX}\cdot\omega$ and the QS 0/100 and Meyer model show that it is $-(2/3)\cdot WLC_{OX}\cdot\omega$. The out-of-phase components of the NQS model and the 1-D numerical solution are almost the same as that of QS 40/60 model for $\omega < \omega_T$ and it becomes

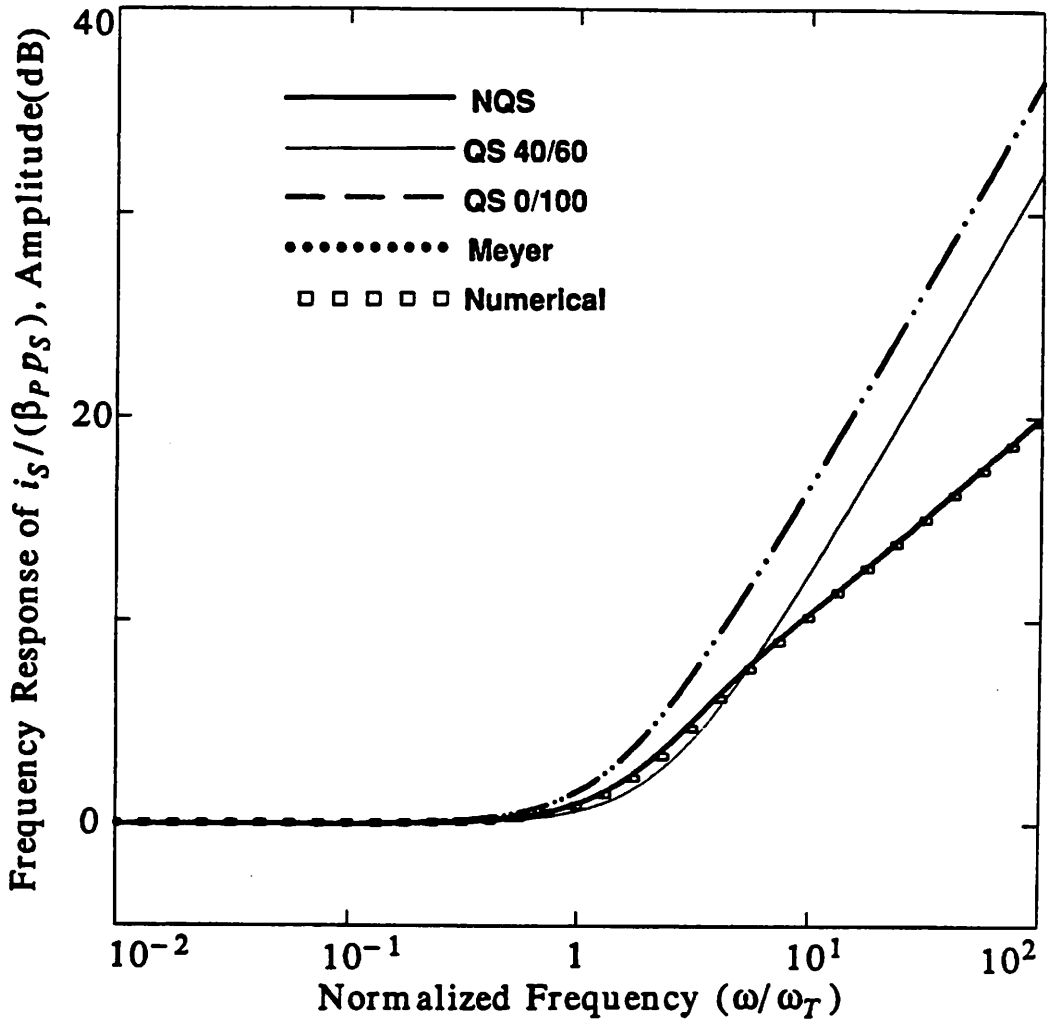


Fig.4.5. Frequency response of the small-signal source current i_S in saturation region with respect to the normalized frequency ω/ω_T, for NQS, QS 40/60, QS 0/100, Meyer model, and 1-D numerical solution of (4.11). Equations for the NQS model are shown in eq. (4.56) and (4.58)

(a). Amplitude response

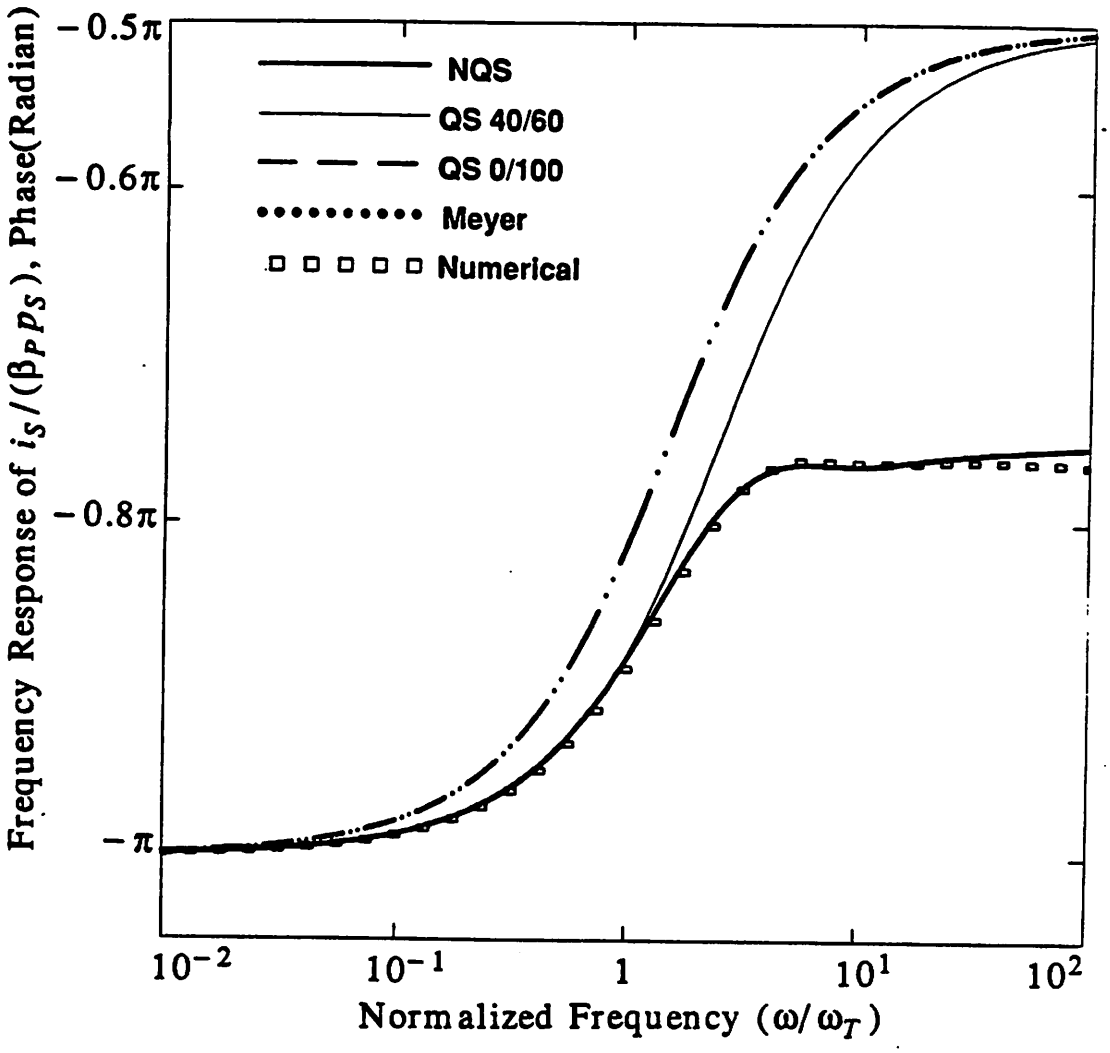


Fig.4.5.(b). Phase response

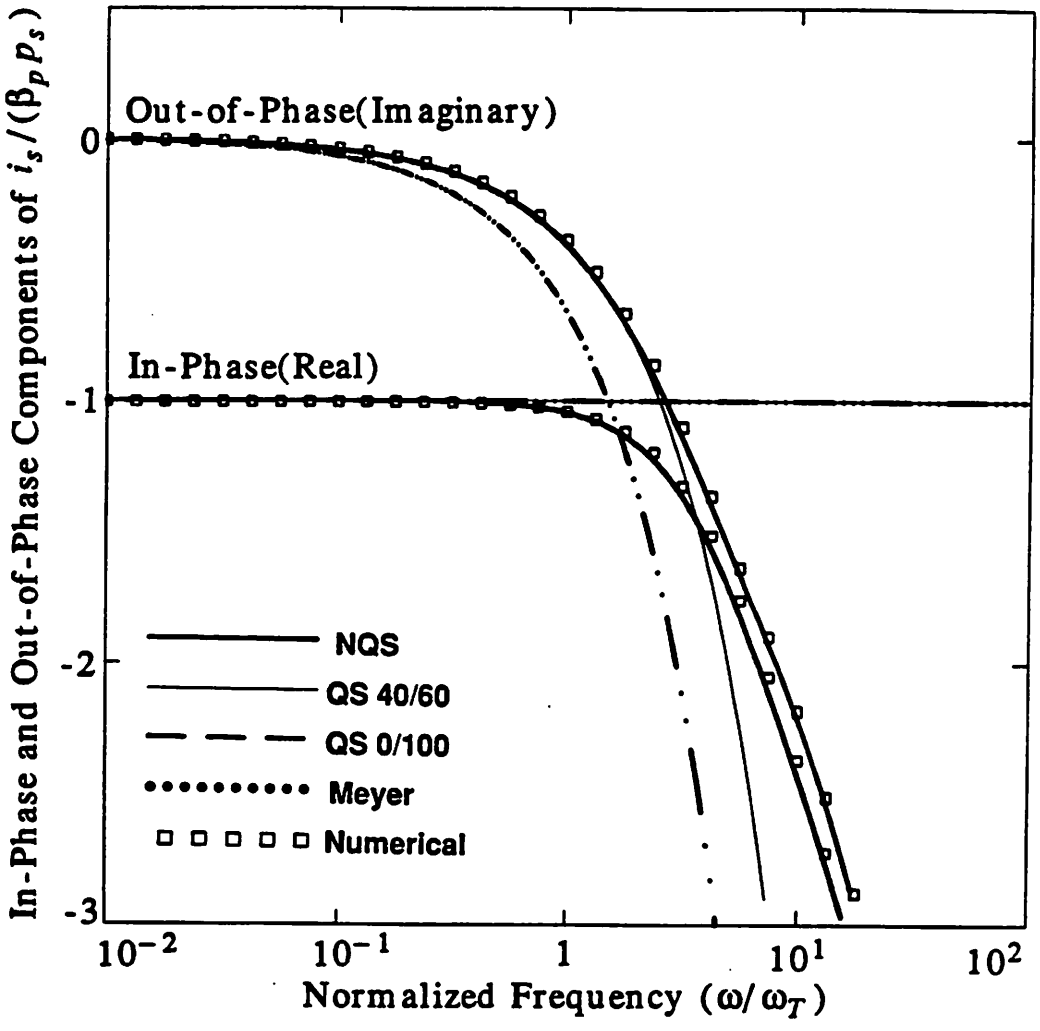


Fig.4.5.(c). In-phase (real) and out-of-phase (imaginary) components

	$\frac{i_D}{\beta_p p_s}$	$\frac{i_S}{\beta_p p_s}$
NQS	$\frac{1-\frac{s}{70.185\omega_T}}{1+\sum_{n=1}^{\infty}(\frac{s}{\omega_{p.n}})^n}$	$-\frac{1+\sum_{n=1}^{\infty}(\frac{s}{\omega_{z.n}})^n}{1+\sum_{n=1}^{\infty}(\frac{s}{\omega_{p.n}})^n}$
QS 40/60	$1-j\cdot\frac{4}{15}\cdot\frac{\omega}{\omega_T}$	$-1-j\cdot\frac{2}{5}\cdot\frac{\omega}{\omega_T}$
QS 0/100	1	$-1-j\cdot\frac{2}{3}\cdot\frac{\omega}{\omega_T}$
Meyer	1	$-1-j\cdot\frac{2}{3}\cdot\frac{\omega}{\omega_T}$

$$where \quad \omega_T(QS\ models) = \frac{\mu_n\cdot(V_{GS}-V_{TH})}{L^2}$$

Table.4.6. The small-signal drain and source currents i_D and i_S for each model in saturation region. i_D and i_S are normalized by the DC value $\beta_P p_S$. Coefficients $\{\omega_{p.n}\}$ and $\{\omega_{z.n}\}$ of NQS model can be found in Table.4.4.

much smaller than that of QS 40/60 model at high frequencies, as shown in Fig.4.5.(c).

From Table 4.3, we can see that i_S of NQS model in saturation region goes asymptotically to $-\beta_P p_S \cdot (1+j) \cdot \sqrt{\omega / (2\omega_T)}$ at very high frequencies ($\omega \gg \omega_T$) and the in-phase and the out-of-phase components become equal to each other and are proportional to $\sqrt{\omega}$. This behaviour can be verified in Fig.4.5.(c).

Frequency responses of transconductances in saturation region

Transconductances $\partial i_D/\partial v_G$, $\partial i_D/\partial v_S$ and $\partial i_S/\partial v_G$ are important in analog circuits because they determine the voltage gains of the common-source, the common-gate and the common-drain amplifiers respectively.

These transconductances of NQS model in saturation region can be derived from (4.57), (4.58) and (4.61) as.

$$\frac{\partial i_D}{\partial v_G} = \frac{g_M}{\beta_P p_S} \cdot i_D \tag{4.67}$$

$$\frac{\partial i_D}{\partial v_S} = -\frac{g_M + g_{MBS}}{\beta_P p_S} \cdot i_D \tag{4.68}$$

$$\frac{\partial i_S}{\partial v_G} = -\frac{g_M}{\beta_P p_S} \cdot i_S \tag{4.69}$$

where the small-signal voltage dependence of ω_T is not included in the derivation because ω_T is determined by the DC operating point as shown in (4.13). $\{\omega_{p,n}\}$ and $\{\omega_{z,n}\}$ are shown in Table 4.4 and g_M and g_{MBS} are shown in (4.62) and (4.63).

In saturation region, $\partial i_D/\partial v_G$ and $\partial i_D/\partial v_S$ have the same frequency response as i_D and $\partial i_S/\partial v_G$ has the same frequency response as i_S except for the real(in-phase) multiplication factors. The frequency response of $i_D/\beta_P p_S$ and $i_S/\beta_P p_S$ in saturation region are shown in (4.57) and (4.58) respectively and also are plotted in Fig.4.4 and Fig.4.5 respectively. Table 4.7 shows the frequency response of $\partial i_D/\partial v_G$, $\partial i_D/\partial v_S$ and $\partial i_S/\partial v_G$ in saturation region for each model. For the NQS model, asymptotic values at high frequencies are used.

	$\frac{\partial i_D}{\partial v_G}$	$\frac{\partial i_D}{\partial v_S}$	$\frac{\partial i_S}{\partial v_G}$
*NQS**	0	0	$-g_M \cdot (1+j) \cdot \sqrt{\frac{\omega}{2\omega_T}}$
QS 40 /60	$g_M \cdot (1-j \cdot \frac{4}{15} \cdot \frac{\omega}{\omega_T})$	$-(g_M+g_{MBS}) \cdot (1-j \cdot \frac{4}{15} \cdot \frac{\omega}{\omega_T})$	$-g_M \cdot (1+j \cdot \frac{2}{5} \cdot \frac{\omega}{\omega_T})$
QS 0 /100	g_M	$-(g_M+g_{MBS})$	$-g_M \cdot (1+j \cdot \frac{2}{3} \cdot \frac{\omega}{\omega_T})$
Meyer	g_M	$-(g_M+g_{MBS})$	$-g_M \cdot (1+j \cdot \frac{2}{3} \cdot \frac{\omega}{\omega_T})$

where $\omega_T(QS\ models) = \frac{\mu_n \cdot (V_{GS}-V_{TH})}{L^2}$

NQS : Asymptotic Values at High Frequencies*

Table.4.7 Values of $\partial i_D/\partial v_G$, $\partial i_D/\partial v_S$ and $\partial i_S/\partial v_G$ in saturation region. For the NQS model, the asymptotic values at high frequencies are shown.

The NQS model shows that $\partial i_D/\partial v_G$ and $\partial i_D/\partial v_S$, become 0 at high frequencies but the QS 40/60 model shows that they go to infinity at high frequencies and the QS 0/100 and Meyer model show that they are constant independent of frequency. So the QS models give qualitatively wrong results about $\partial i_D/\partial v_G$ and $\partial i_D/\partial v_S$ at high frequencies.

For $\partial i_S/\partial v_G$, all the models show the it goes to infinity at high frequencies. But the QS models predict that the out-of-phase component of $\partial i_S/\partial v_G$ is proportional to ω at high frequencies but the NQS model predicts that it is proportional to $\sqrt{\omega}$.

4.3.3. AC channel current partitioning ratio in saturation region

The AC channel current partitioning ratio, δ, is defined to be the ratio of imaginary (out-of-phase) components of the small-signal drain and source currents, i_D and i_S.

$$\delta = \frac{\mathrm{Im}(i_D)}{\mathrm{Im}(i_D)+\mathrm{Im}(i_S)} \tag{4.70}$$

The definition (4.70) matches the channel charge partitioning ratio of QS charge based models in saturation region as shown in the following discussions.

The imaginary parts of i_D and i_S for QS charge based models can be derived from

$$\mathrm{Im}(I_D) = \omega \cdot q_D \tag{4.71}$$

$$\mathrm{Im}(I_S) = \omega \cdot q_S \tag{4.72}$$

where ω is the signal angular frequency and q_D and q_S are AC small-signal phasors of drain and source charges respectively. Hence, the AC channel current partitioning ratio of QS charge based models, δ_{QS}, is defined to be

$$\delta_{QS} = \frac{\mathrm{Im}(i_D)}{\mathrm{Im}(i_D) + \mathrm{Im}(i_S)} = \frac{q_D}{q_D + q_S} \tag{4.73.a}$$

In saturation region, QS charge based models gives the total drain and source charges as, $Q_D = XQC*Q_N$ and $Q_S = (1 - XQC)*Q_N$, where Q_N is the total inversion charge and XQC is the channel charge partitioning ratio which is a constant. $XQC = 0.4$ for QS 40/60 model and $XQC = 0$ for QS 0/100 model. Since XQC is a bias independent constant, the AC small-signal phasors of drain and source charges become $q_D = XQC*q_N$

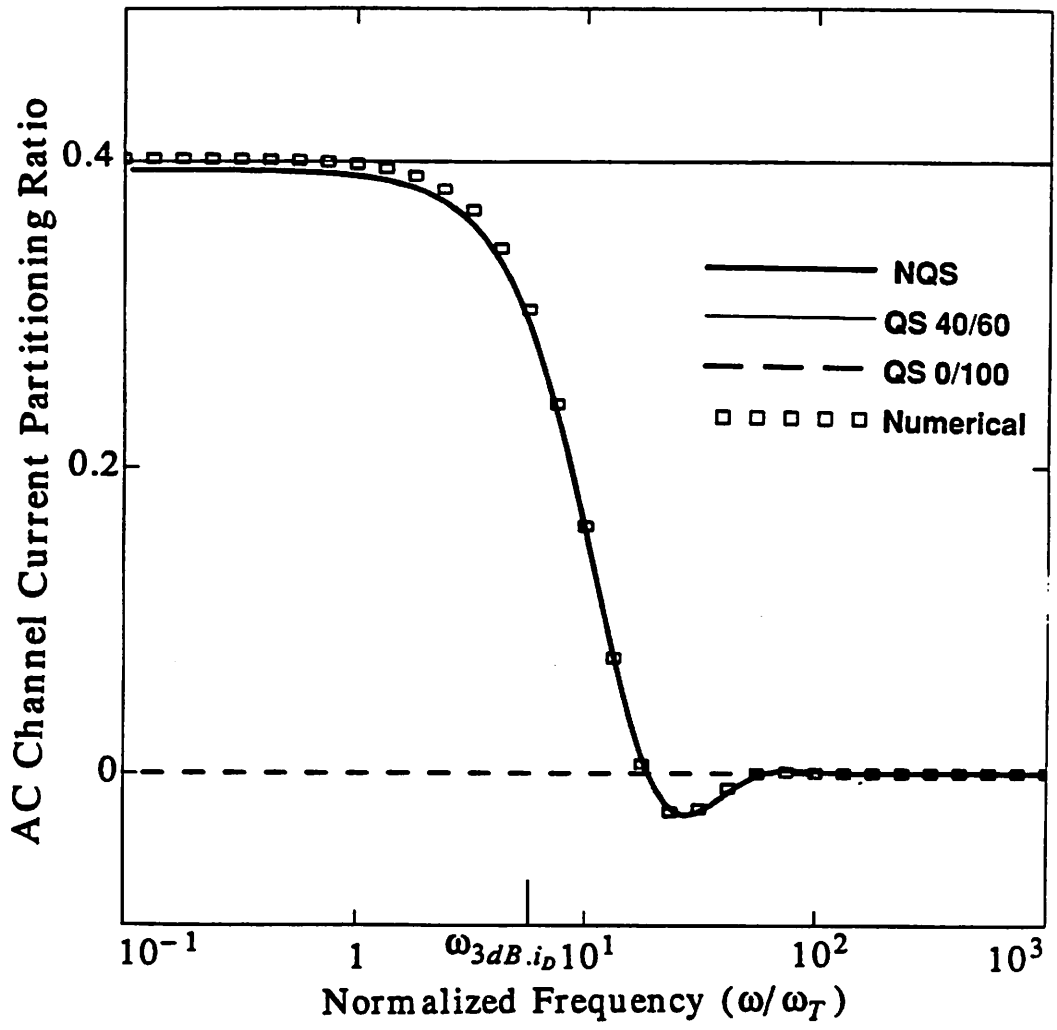


Fig.4.6. AC small-signal channel current partitioning ratio with respect to the normalized frequency (ω/ω_T), for NQS, QS 40/60, QS 0/100 model and 1-D numerical solution of (4.11).

and $q_S = (1 - XQC)*q_N$, where q_N is the AC small-signal phasor of the total inversion charge Q_N. Substituting these relations into (4.73.a), we have

$$\delta_{QS} = XQC \tag{4.73.b}$$

that is, the AC small-signal channel current partitioning ratio of QS charge based models is equal to the channel charge partitioning ratio in saturation region.

Fig.4.6 shows the AC small-signal channel current partitioning ratio with respect to the normalized frequency for NQS, QS 40/60, QS 0/100 model and 1-D numerical analysis. Eq. (4.55), (4.56) and (4.70) are used to find the partitioning ratio of NQS model. For the partitioning ratio of the 1-D numerical analysis, the slopes of the out-of-phase profile $\{p_{im,i}\}$ at source and drain ends are used. Good agreements can be observed between the NQS model and the 1-D numerical solution. The partitioning ratios of NQS model and 1-D numerical analysis are almost the same as that of QS 40/60 model at low frequencies and goes asymptotically to that of QS 0/100 model at very high frequencies. For the QS models, the partitioning ratio is constant independent of frequency.

4.4. Simulation results and comparison with other models

4.4.1. Frequency response of NMOS inverter with a resistive load

To see the frequency response of an NMOSFET, an NMOS inverter with a resistive load shown in Fig.4.7.(a), is simulated with the NQS, QS 40/60, QS 0/100 and SPICE level 2 Meyer model. This circuit is a common source amplifier. For the QS 0/100 and 40/60 model, BSIM in SPICE3 is used. The small-signal equivalent circuits of each model for the drain node are shown in Fig.4.7.(b), 4.7.(c) and 4.7.(d). The voltage gain can be found by applying K.C.L.(Kirchhoff Current Law) at the drain node and is shown in Table 4.8.

For the intrinsic MOSFET's, where the extrinsic capacitances such as overlap capacitance (C_{OV}) and junction and load capacitances (C_L) are 0, the voltage gain in saturation

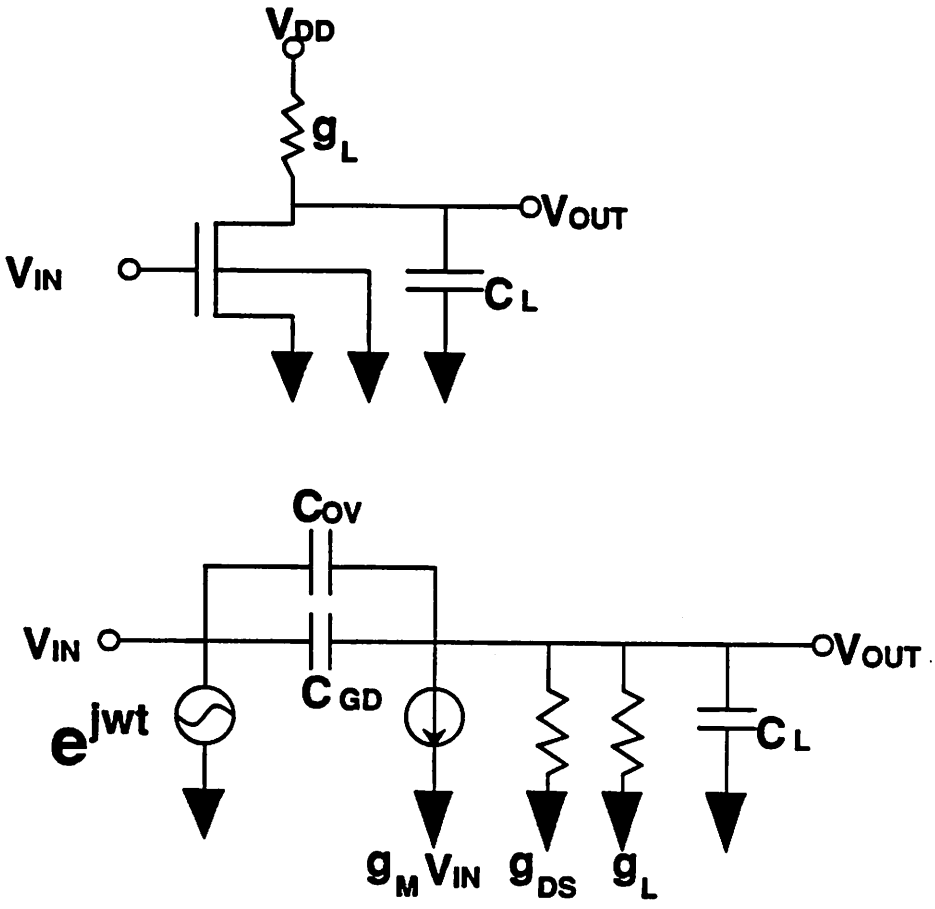
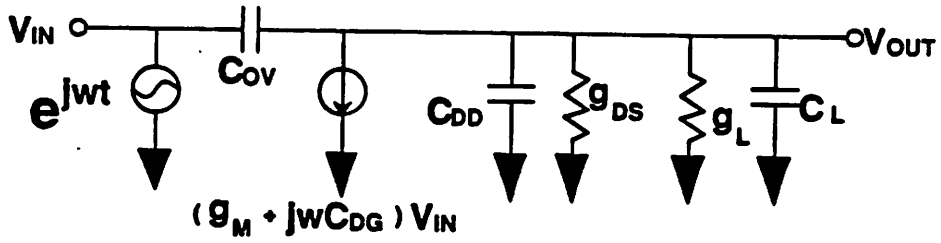


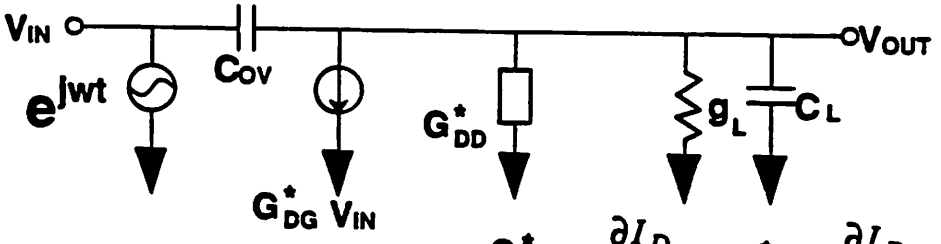
Fig.4.7. The circuit diagram and the equivalent circuits at the drain node for the AC analysis of an NMOS inverter (common-source amplifier) with a resistive load

(a). Circuit diagram

Fig.4.7.(b). Equivalent circuit from the Meyer model



(c).QS 40/60, 0/100



(d).NQS(this work)

$$G^*_{DG} = \frac{\partial I_D}{\partial V_G} \qquad G^*_{DD} = \frac{\partial I_D}{\partial V_D}$$

Fig.4.7.(c). Equivalent circuit from the QS charge based models(QS 40/60 and 0/100)

Fig.4.7.(d). Equivalent circuit from the NQS model

Model	$\frac{v_{OUT}}{v_{IN}}$
NQS	$\frac{-G_{DG}^{*} + sC_{ov}}{G_{DD}^{*} + g_L + s(C_{ov}+C_L)}$
QS 0/100,40/60	$\frac{-g_M + s(C_{ov}-C_{DG})}{g_{DS} + g_L + s(C_{DD}+C_{ov}+C_L)}$
Meyer	$\frac{-g_M + s(C_{ov}+C_{GD})}{g_{DS} + g_L + s(C_{GD}+C_{ov}+C_L)}$

where $G_{DG}^{*} = \frac{\partial i_D}{\partial v_G}$ $G_{DD}^{*} = \frac{\partial i_D}{\partial v_D}$

Table.4.8 AC voltage gain of an NMOS inverter (common source amplifier) with a resistive load shown in Fig.4.7.

	$\frac{v_{out}}{v_{in}}$	$\frac{v_{out}}{v_{in}}$(high freq.)
NQS	$-\frac{G_{DG}^{*}}{g_{DS}+g_L}$	0
QS 40/60	$-\frac{g_M - s\frac{4}{15}WLC_{OX}}{g_{DS}+g_L}$	$\frac{s\frac{4}{15}WLC_{OX}}{g_{DS}+g_L}$
QS 0/100	$-\frac{g_M}{g_{DS}+g_L}$	$-\frac{g_M}{g_{DS}+g_L}$
Meyer	$-\frac{g_M}{g_{DS}+g_L}$	$-\frac{g_M}{g_{DS}+g_L}$

Table.4.9. AC voltage gain of an NMOS inverter (common source amplifier) in saturation region without the extrinsic capacitances (intrinsic only)

	$\frac{v_{out}}{v_{in}}$	$\frac{v_{out}}{v_{in}}$(high freq.)
NQS	$-\frac{G_{DG}^{*} - sC_{ov}}{g_{DS}+g_L+s(C_{ov}+C_L)}$	$\frac{C_{ov}}{C_{ov}+C_L}$
QS 40/60	$-\frac{g_M - s(C_{ov}+\frac{4}{15}WLC_{OX})}{g_{DS}+g_L+s(C_{ov}+C_L)}$	$\frac{C_{ov}+\frac{4}{15}WLC_{OX}}{C_{ov}+C_L}$
QS 0/100	$-\frac{g_M - sC_{ov}}{g_{DS}+g_L+s(C_{ov}+C_L)}$	$\frac{C_{ov}}{C_{ov}+C_L}$
Meyer	$-\frac{g_M - sC_{ov}}{g_{DS}+g_L+s(C_{ov}+C_L)}$	$\frac{C_{ov}}{C_{ov}+C_L}$

Table 4.10. AC voltage gain of an NMOS inverter (common source amplifier) in saturation region with the extrinsic capacitances C_{ov} and C_L included

region can be simplified as in Table 4.9.

Fig.4.8.(a) and 4.8.(b) show the frequency response of an intrinsic NMOS inverter with a resistive load, for W/L=3μm/3μm and 30μm/30μm respectively.

For NQS model, the voltage gain is the product of the transconductance in (4.67) and the output impedance ($1/(g_{DS}+g_L)$). The amplitude decreases with frequency and goes to 0 at very high frequencies. The phase is π radian at DC and decreases unboundedly with frequency.

For the QS 0/100 and Meyer models, since the voltage gain has only a real (in-phase) part, as shown in Table 4.9, the amplitude and the phase are constant independent of frequency.

For the QS 40/60 model, since the voltage gain has a single zero and no poles, as shown in Table 4.9, the amplitude increases unboundedly with frequency. The phase is π radian at DC and goes asymptotically to 0.5 π radian at high frequencies. This is due to the fact that the imaginary (out-of-phase) part is proportional to the frequency and becomes much larger than the real(in-phase) part at high frequencies. At low frequencies, the phase response of QS 40/60 model agrees with that of NQS model.

Table 4.10 shows the equations for the voltage gain of a resistive load NMOS inverter in saturation region and with the extrinsic capacitances (C_{OV} and C_L) included. C_L is the drain junction capacitance in this example. All the models give the voltage gain with one pole and one zero.

For the Meyer and the QS 0/100 models, since the pole frequency is always lower than the zero frequency, the high frequency gain becomes a constant value which is always lower than the DC gain.

For the QS 40/60 model, since the pole frequency may be lower or higher than the zero frequency depending on the ratios of C_{ov}, C_L and WLC_{OX}, the high frequency gain may be lower or higher than the DC gain.

For the NQS model, since $G_{DG}{}^*$, that is, $\partial i_D/\partial v_G$, goes to 0 at very high frequencies as shown in (4.67), the high frequency gain becomes a constant value which is the same as

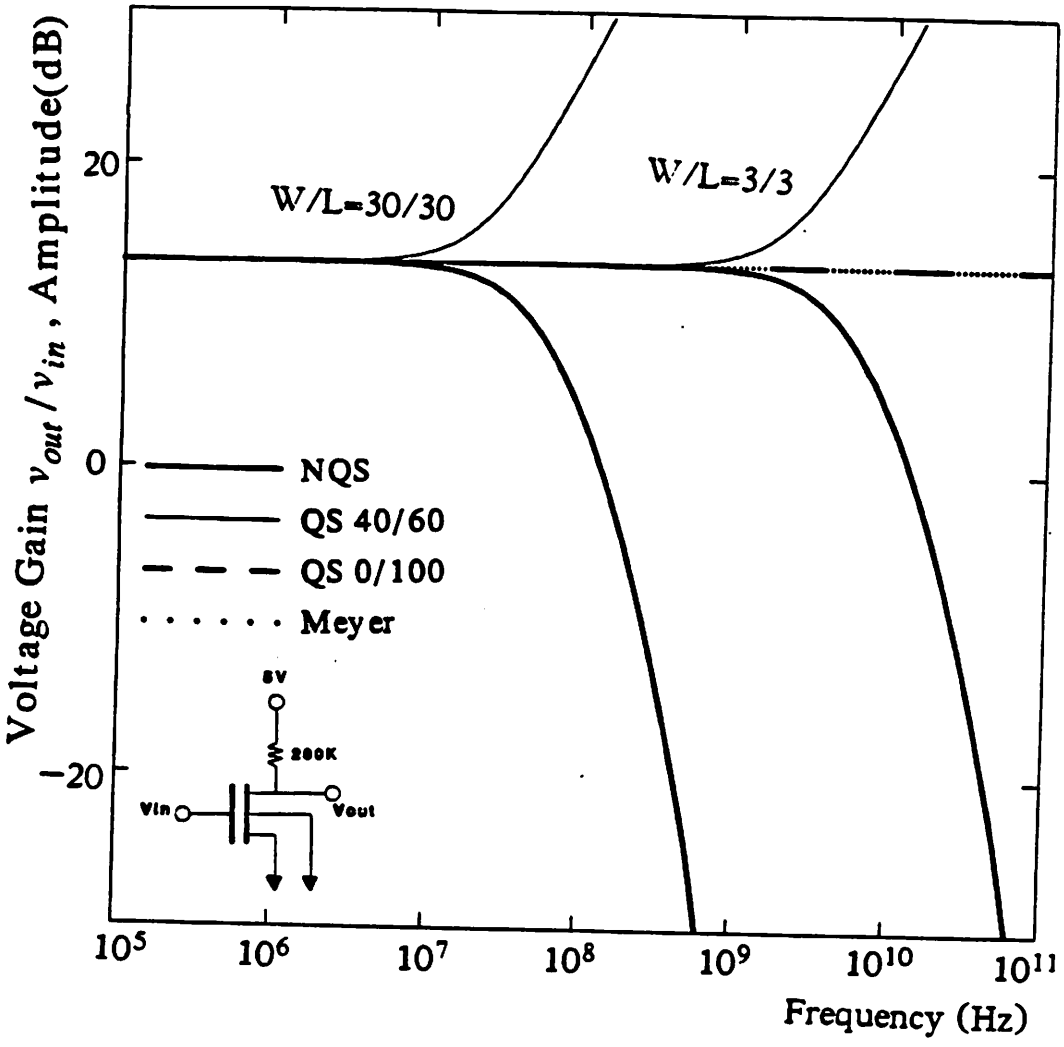


Fig.4.8. The frequency response of an intrinsic NMOS inverter with a resitive load. The extrinsic capacitances (overlap and junction capacitance) are not included. The model parameters are V_{FB}= –0.87V, N_{SUB}=2* $10^{16}cm^{-3}$, μ_0=500cm^2/V·sec and T_{OX}=50nm. W/L of the inverter is 3μm/3μm and 30μm/30μm respectively.

(a). Amplitude response

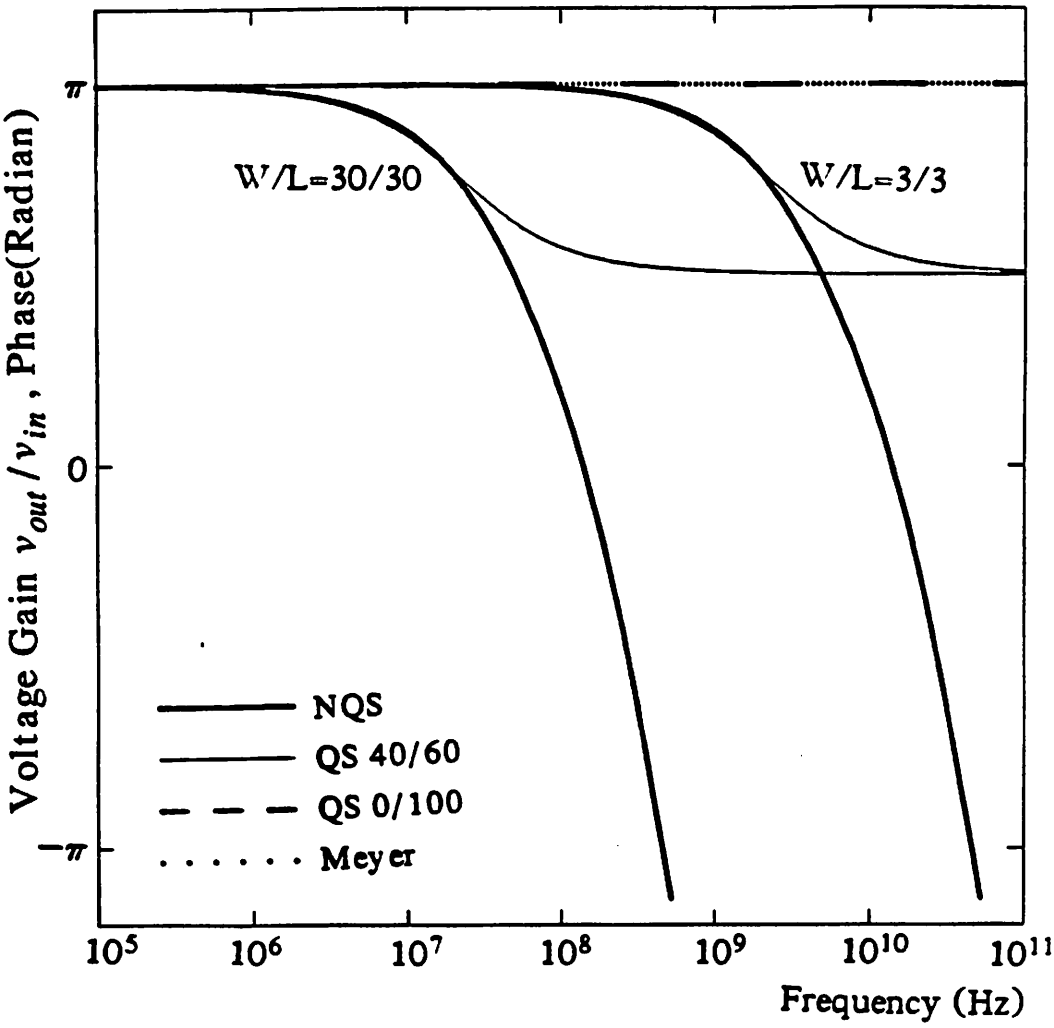


Fig.4.8.(b). Phase response

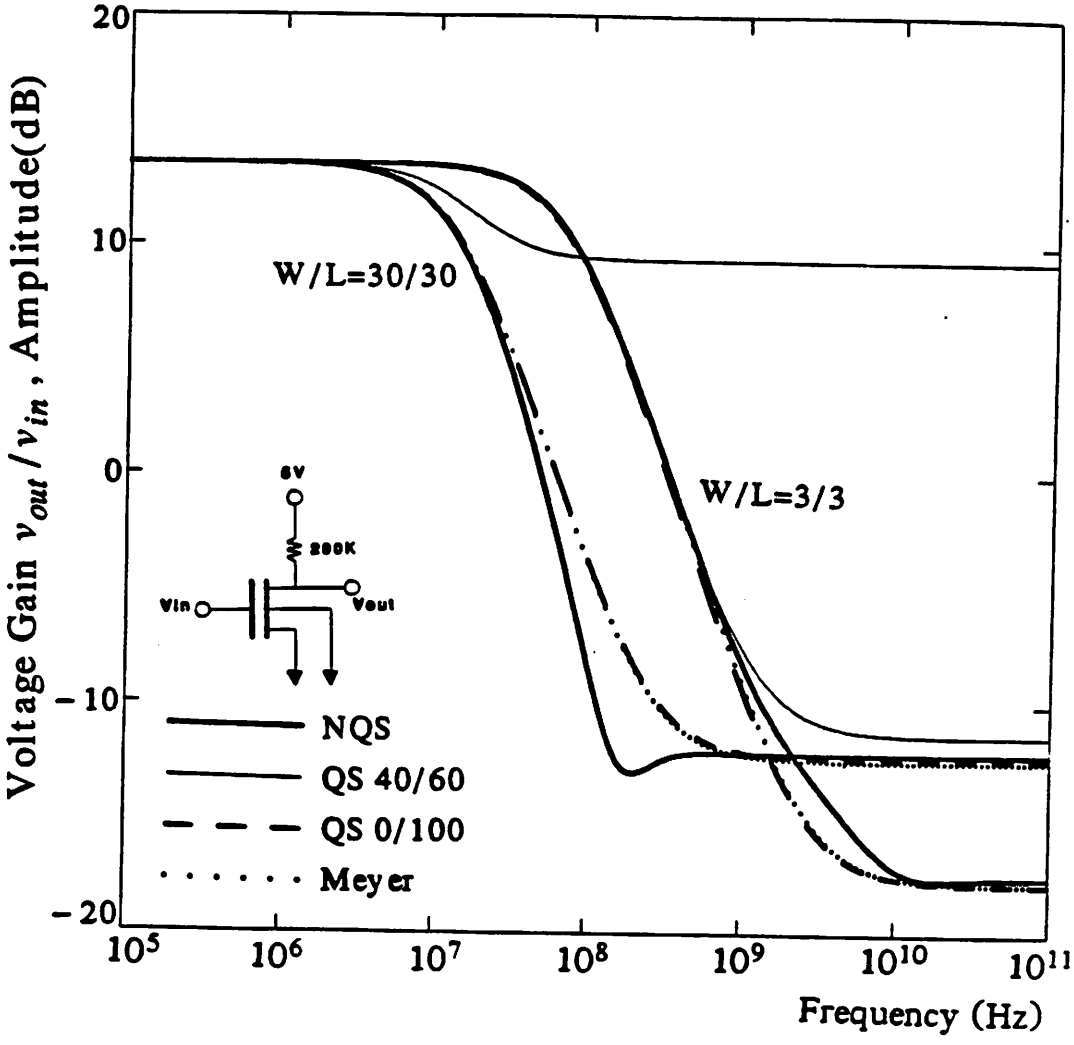


Fig.4.9. The frequency response of an NMOS inverter (common-source amplifier) with a resitive load and with extrinsic capacitances (overlap and junction capacitance). The model parameters are cgdo = cgso = $500pF/m$, Cj=3 $*10^{-4}F/m^2$, Cjsw = $8*10^{-10}F/m$ and the other parameters are the same as those in Fig.4.8. The zero bias junction capacitance at the output node is 17.4 fF for W/L=3μm/3μm and 87.6 fF for W/L=30μm/30μm, respectively.

(a). Amplitude response

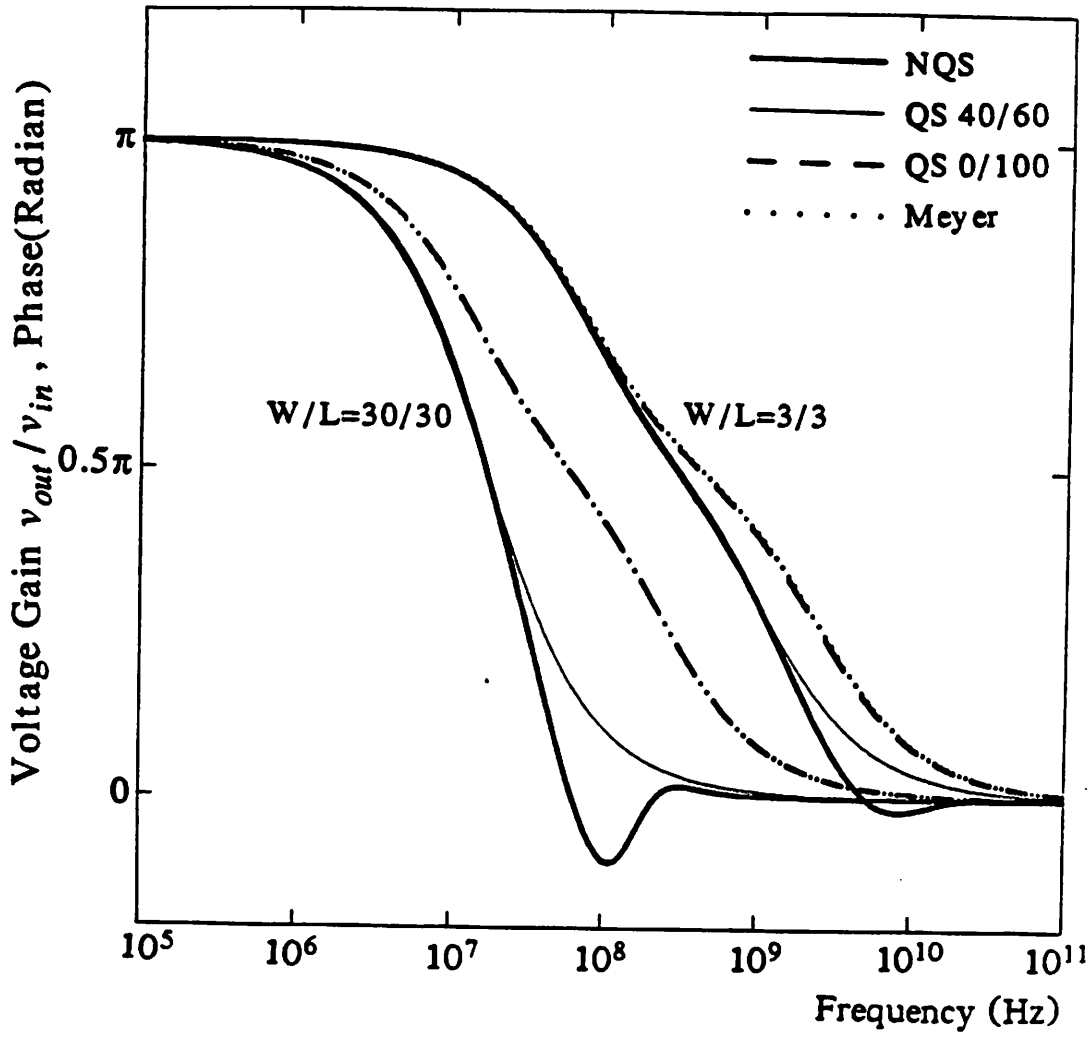


Fig.4.9.(b). Phase response

those of QS 0/100 and Meyer model.

Fig.4.9.(a) and 4.9.(b) show the frequency response of this inverter for W/L=30μ/30μ*m* and 3μ*m*/3μ*m* respectively. At high frequencies, QS 40/60 model gives too large amplitude error and QS 0/100 and Meyer model give too large phase error. From this, we can see that any of these three QS models is not satisfactory for AC analysis, while the NQS model solved all the physical anomalies of QS models in AC analysis.

4.4.2. Comparison with the multiple-lumped model

To compare this NQS model with the multiple-lumped model, an intrinsic part of MOSFET has been decomposed into N-lumped short-channel MOSFETs with the same channel width and channel length of one N-th the original channel length. And then the decomposed circuit is simulated with SPICE using the QS models. Values of N used in this example are 1, 2, 4, 8, 100 and 1000.

Fig.4.10.(a) and 4.10.(b) show the simulated frequency response of the N-lumped circuits derived from the NMOS inverter circuit in Fig.4.8.(a) with W/L=30μ*m*/30μ*m*. Simulation results using the NQS model are also shown for comparison. Extrinsic capacitances are not included in the simulation.

The N-lumped model goes asymptotically to the NQS model as N increases. In the amplitude response as shown in Fig.4.10.(a), DC gain of N-lumped model decreases slightly as N increases. This is considered to be due to the fact that the DC characteristics of MOSFET are not exactly proportional to the channel length.

Since ω_T is inversely proportional to L^2 as shown in (4.13), the new ω_T in N-lumped model is N^2 times the original ω_T. So in the N-lumped model, the valid frequency range increases in proportion to N^2, but beyond this valid range it gives anomalous results again.

For the 1000-lumped model, we can observe fair agreements with the NQS model for $\omega < 100 \cdot \omega_T$, both in the amplitude and in the phase response.

Table 4.11 shows the average CPU time (per frequency point in SPICE3) of the N-lumped model and the NQS model for the simulation of the circuit in Fig.4.8.(a). CPU

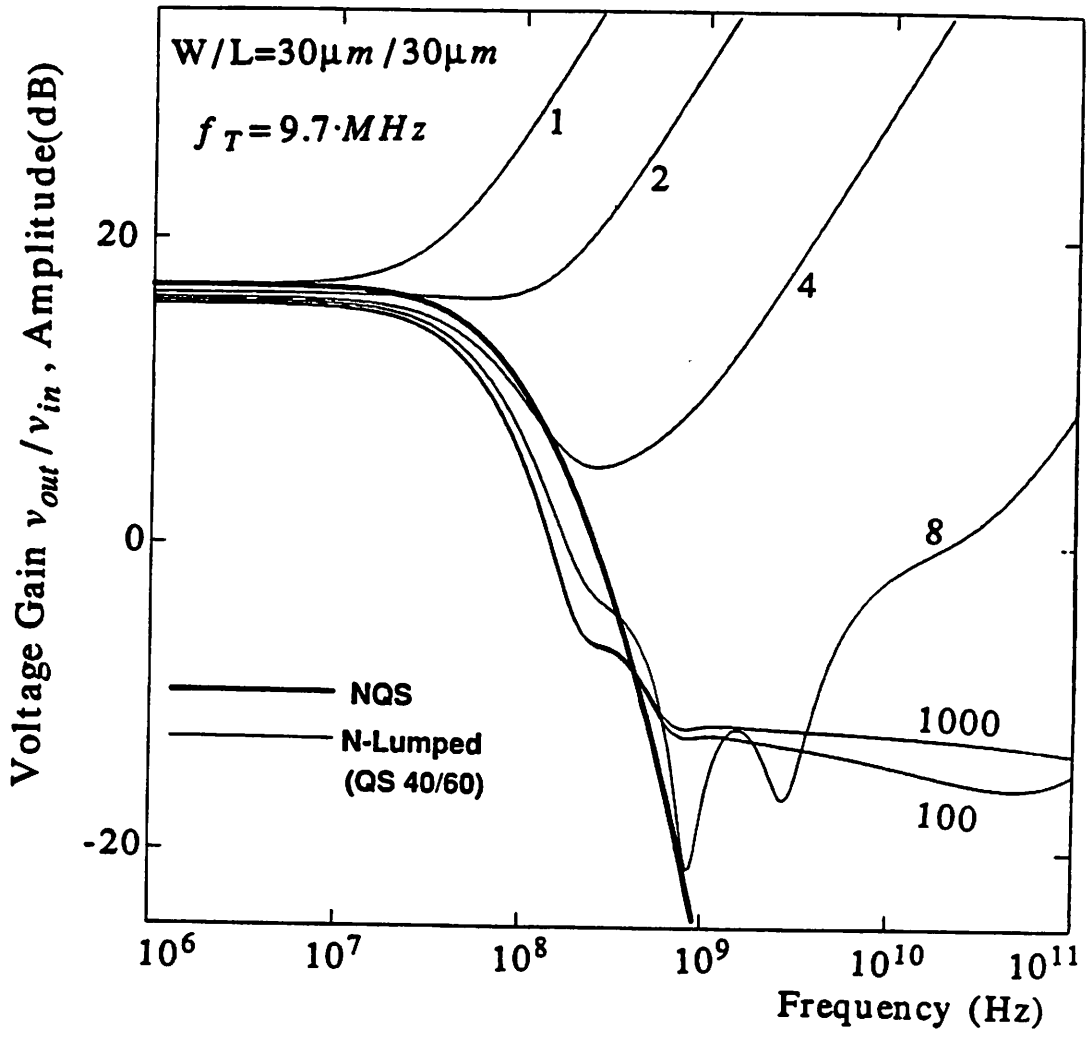


Fig.4.10. Frequency response of the NMOS inverter (W/L=30μm/μm) with a resistive load shown in Fig.4.8.(a), for the N-lumped models and the NQS model, where N = 1, 2, 4, 8, 100 and 1000. QS 40/60 model is used for the simulation of the N-lumped models. Model parameters of the MOSFET are the same as those in Fig.4.8.

(a). Amplitude response

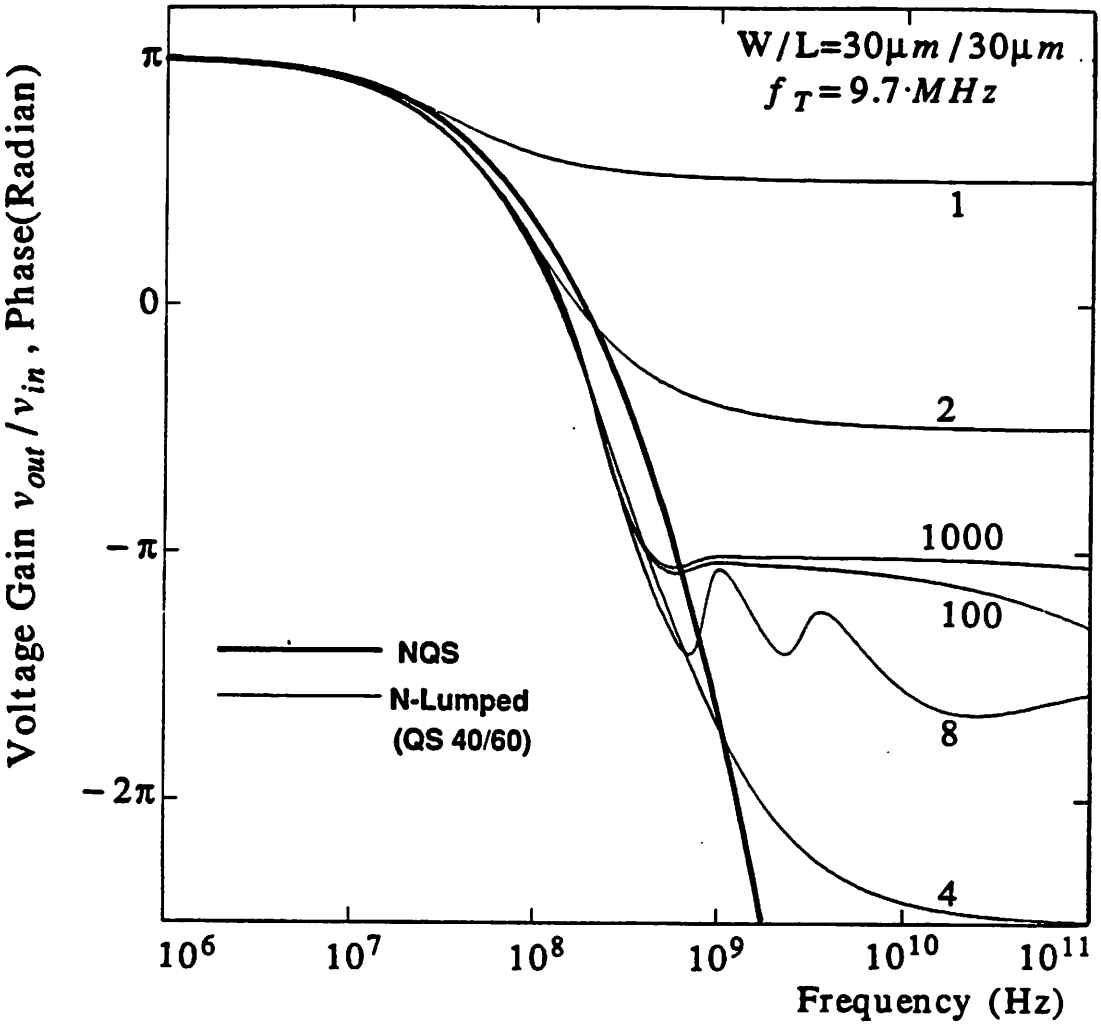


Fig.4.10.(b). Phase response

time of NQS model is comparable to that of 4-lumped model.

Model	1-Lump	2-Lump	4-Lump	8-Lump	100-Lump	1000-Lump	NQS
CPU Time	0.97	1.22	1.82	2.97	38.4	1420	1.90

Table.4.11. The average CPU time in [*msec*] per frequency point of the N-lumped models and the NQS model for the AC analysis of the NMOS inverter circuit in Fig.4.8.(a). VAX 8800 running Ultrix V2.0 is used.

small signal channel current partitioning ratio

Fig.4.11.(a) and 4.11.(b) show the AC small-signal channel current partitioning ratio of the N-lumped model and the NQS model in saturation region. The partitioning ratios have been computed using (4.70) and the ratios of the small-signal drain and source currents computed from SPICE simulation. For the simulation of the N-lumped model, QS 40/60 model is used in Fig.4.11.(a) and QS 0/100 model is used in Fig.4.11.(b).

For the 1-lumped model, the partitioning ratio is constant independent of frequency, as shown in Fig.4.11.(a) and 4.11.(b). As N becomes larger, the partitioning ratio of the N-lumped model goes asymptotically to that of NQS model irrespective of what QS model is used for the simulation. For N=100, the partitioning ratio becomes almost the same as that of NQS model except for a small mismatch in ω_T both in Fig.4.11.(a) and 4.11.(b).

4.4.3. Frequency response of CMOS inverter

Fig.4.12.(a) and 4.12.(b) show the frequency response of a CMOS inverter with W/L=3μm/3μm and 30μm/30μm respectively. The overlap capacitance and the junction capacitance have been included. We can observe that the Meyer and QS 0/100 model give almost the same results both in the amplitude and phase responses. This is due to the fact that the QS 0/100 model is reduced to the simple Meyer model in

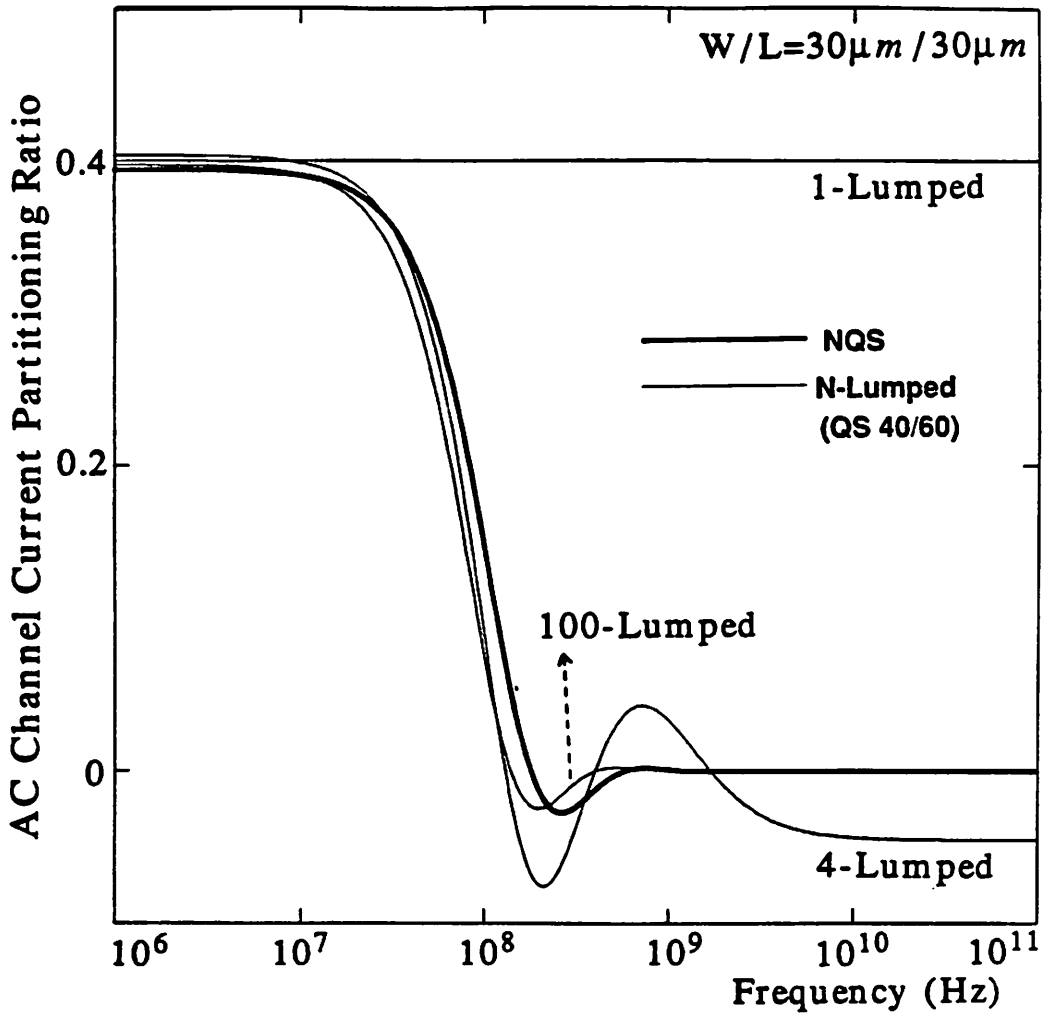


Fig.4.11. AC small-signal channel current partitioning ratios with respect to frequency for the N-lumped model and the NQS model. N = 1, 2, 4 and 100. The partitioning ratio of N=1000 is almost the same as that of N=100, and is not shown for clarity.

(a). QS 40/60 model is used for the simulation of N-lump models. N = 2 is not shown for clarity. For N=2, the partitioning ratio is 0.406 at low frequencies and is -0.402 at very high frequencies.

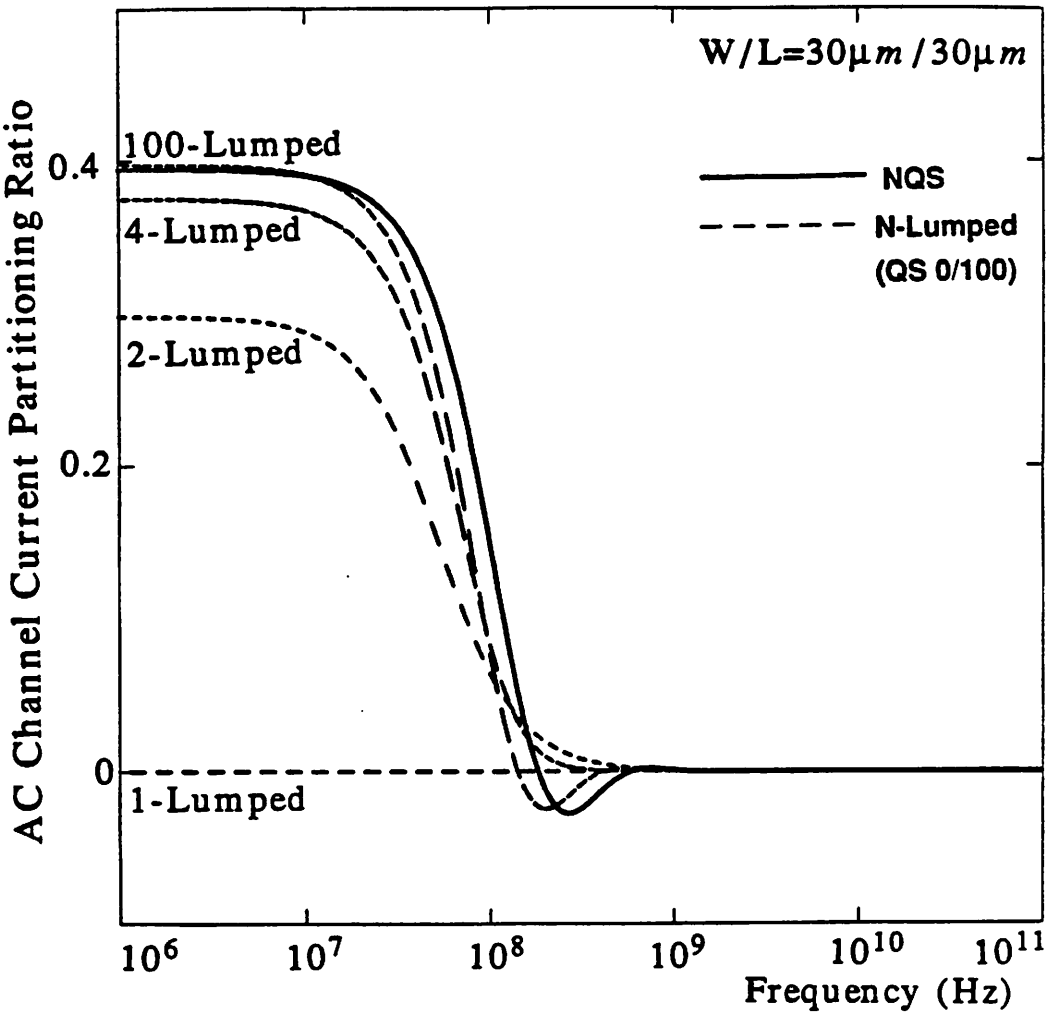


Fig.4.11.(b). AC small-signal channel current partitioning ratios with respect to frequency for the N-lumped model and the NQS model. QS 0/100 model is used for the simulation of N-lump models.

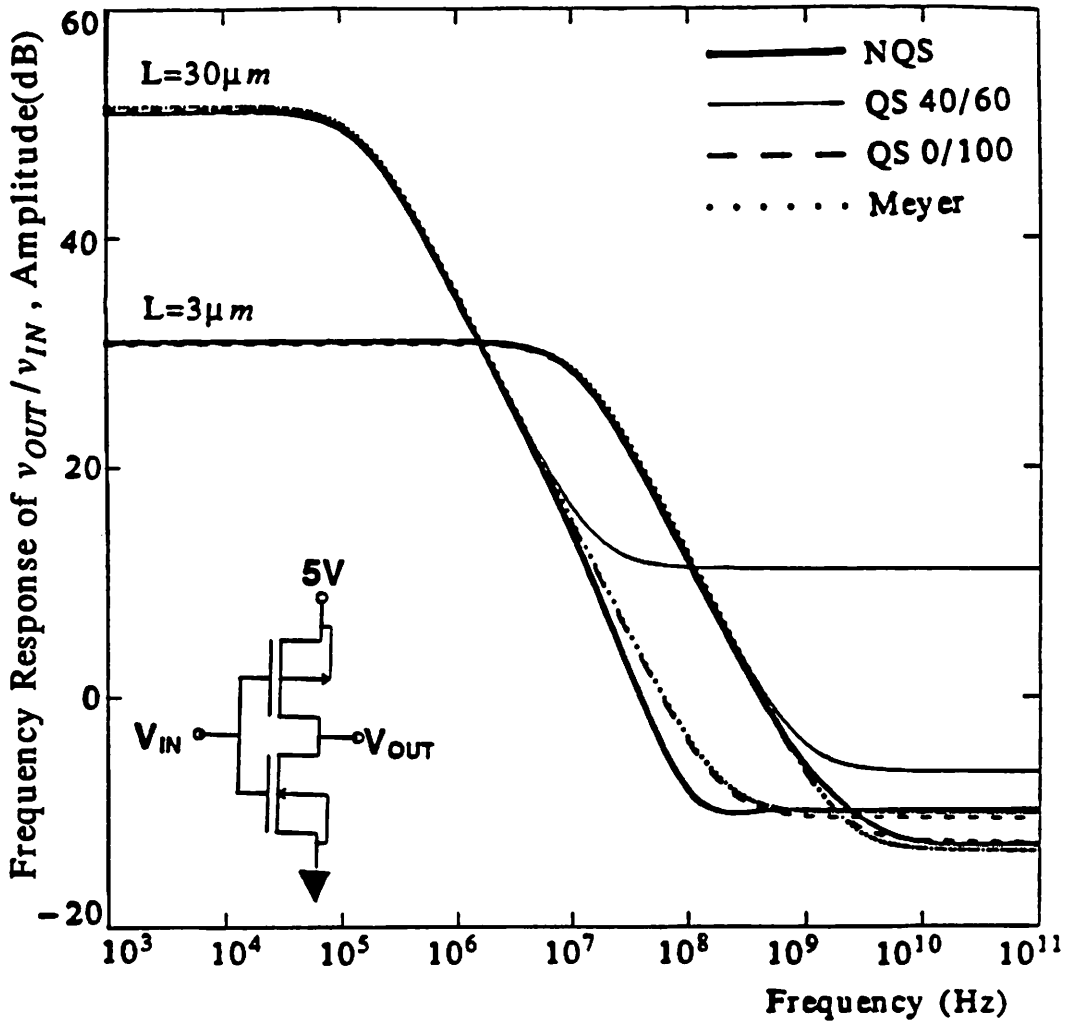


Fig.4.12. The frequency response of a CMOS inverter for L=3μm and 30μm respectively. For L=3μm, W/L of NMOSFET is 3μm/3μm and W/L of PMOSFET is 10μm/3μm. For L=30μm, W/L of NMOSFET is 30μm/30μm and W/L of PMOSFET is 100μm/30μm.

The model parameters for NMOSFET are the same as those in Fig.4.9 and the model parameters for PMOSFET are V_{FB}=–0.32V, N_{SUB}=6* $10^{15}cm^{-3}$, T_{OX}=50nm, μ_0=180cm^2/V·sec, Cj=2* $10^{-4}F/m^2$ and Cjsw=5* $10^{-10}F/m$. For L=3μm, λ=0.03V^{-1} and 0.05V^{-1} for NMOSFET and PMOSFET respectively. And for L=30μm, λ=0.003V^{-1} and 0.005V^{-1} for NMOSFET and PMOSFET respectively.

The zero bias junction capacitance at the output node is 40.4 fF for L=3μm and 264 fF for L=30μm.

(a). Amplitude response

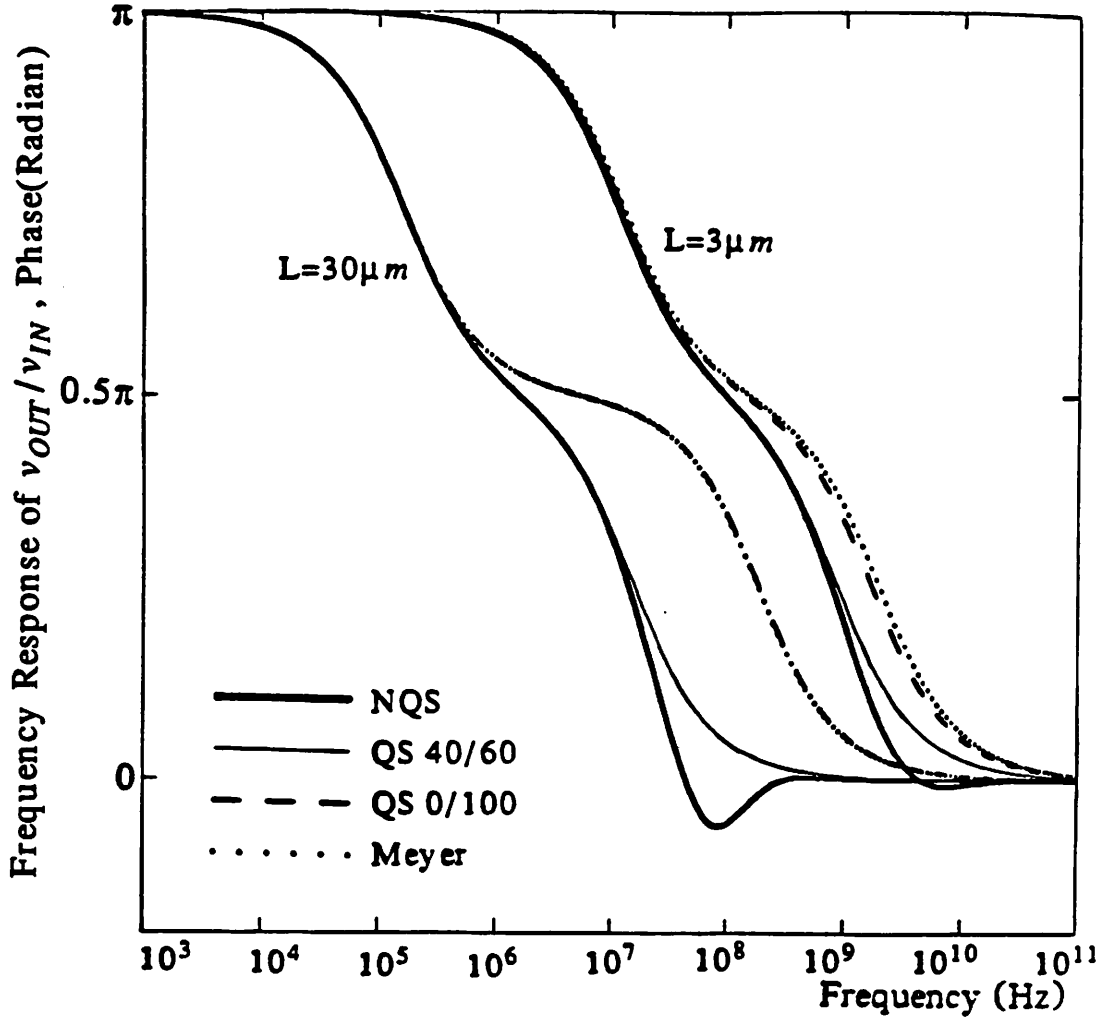


Fig.4.12.(b). Phase response

saturation region as far as the drain node is concerned. This tendency can also be observed in the resistive load NMOS inverter as shown in Fig.4.8.(a), 4.8.(b), 4.9.(a) and 4.9.(b). Again QS 40/60 model gives too large amplitude at high frequencies, due to the capacitive feed-through from gate to drain. QS 0/100 and Meyer model give too large phase error at high frequencies. For L=30μm, the differences between models are much larger than those of L=3μm.

4.4.4. Frequency response of 2-stage CMOS OP Amp

The circuit diagram of a 2-stage CMOS OP Amp is shown in Fig.4.13.(a) [4.15]. The channel length of each transistor is 3μm. The junction capacitance and the overlap capacitance have been included to simulate the real circuit and the areas and sidewall lengths of each junction are determined based on the 2μm design rule.

This circuit is a series connection of two common source amplifiers and QS models give two *poles* in the left-half s-plane and two *zeroes* in the right-half s-plane, as implied in Table 4.10 for a common source amplifier. The dominant *pole* frequency is determined by the impedance and capacitance at the output node of input PMOSFET and the non-dominant *pole* frequency is determined by the output node. The *zero* frequencies are determined by the gate to drain capacitance of the input PMOSFET and the output NMOSFET respectively.

To see the difference of the QS and the NQS model more clearly, the frequency response before compensation (C_C=0, C_L=0) is shown in Fig.4.13.(b) and 4.13.(c). The *pole* and *zero* frequencies for the QS models are computed using the *pole–zero* analysis routine in SPICE3 [4.16], and are shown in Table 4.12. The *pole* frequencies are almost the same for all three models. The QS 0/100 and Meyer model show almost the same *zero* frequencies. The slight difference between these two is due to the slight mismatch in model parameters. QS 40/60 model shows *zero* frequencies which are always smaller than those of QS 0/100 and Meyer model. This can be verified also from equations in Table 4.10.

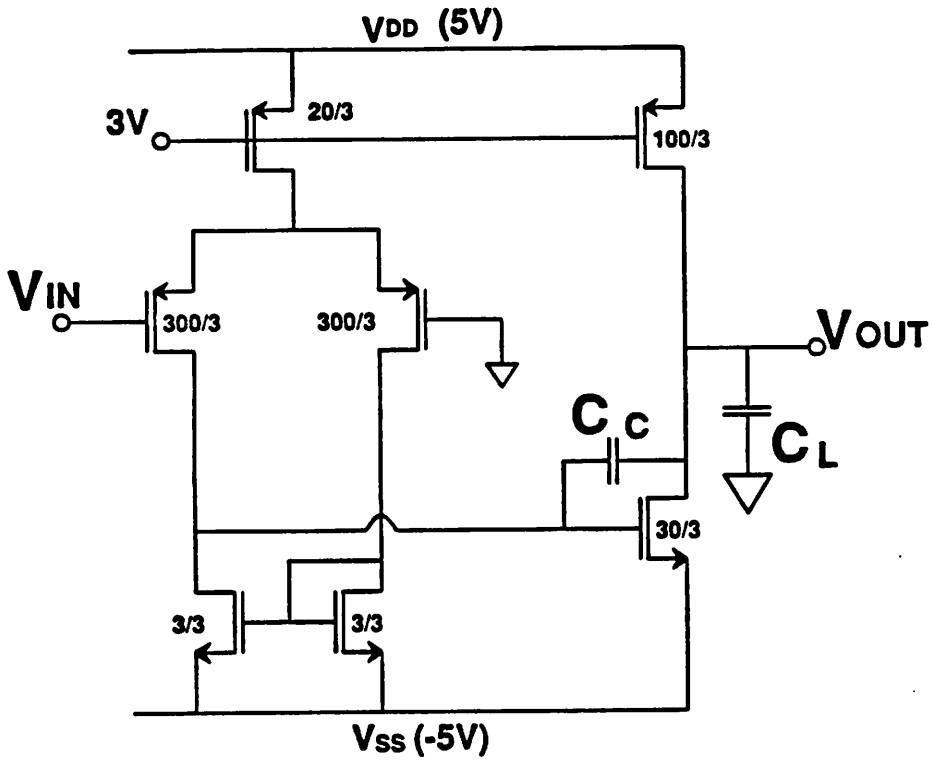


Fig.4.13. Frequency response of a 2-stage CMOS OP amp[4.15] before the frequency compensation. The channel length is 3μm and the compensation is not done deliberately ($C_C = C_L = 0$) to see the difference more clearly.

The model parameters for the NMOSFET is V_{FB}=–0.87V, T_{OX}=50 nm, N_{SUB}=2* $10^{16}cm^{-3}$, μ_0 = 500cm^2/V·sec, cgdo=cgso=150 pF/m and λ= 0.03V^{-1}.

And the model parameters for the PMOSFET are V_{FB} = -0.32V, T_{OX}=50 nm, N_{SUB}=6* $10^{15}cm^{-3}$, μ_0 = 180cm^2/V·sec, cgdo=cgso=150 pF/m and λ= 0.05V^{-1}.

The zero bias junction capacitance is 50.4 fF both at the first and the second stage output node.

(a). Circuit diagram

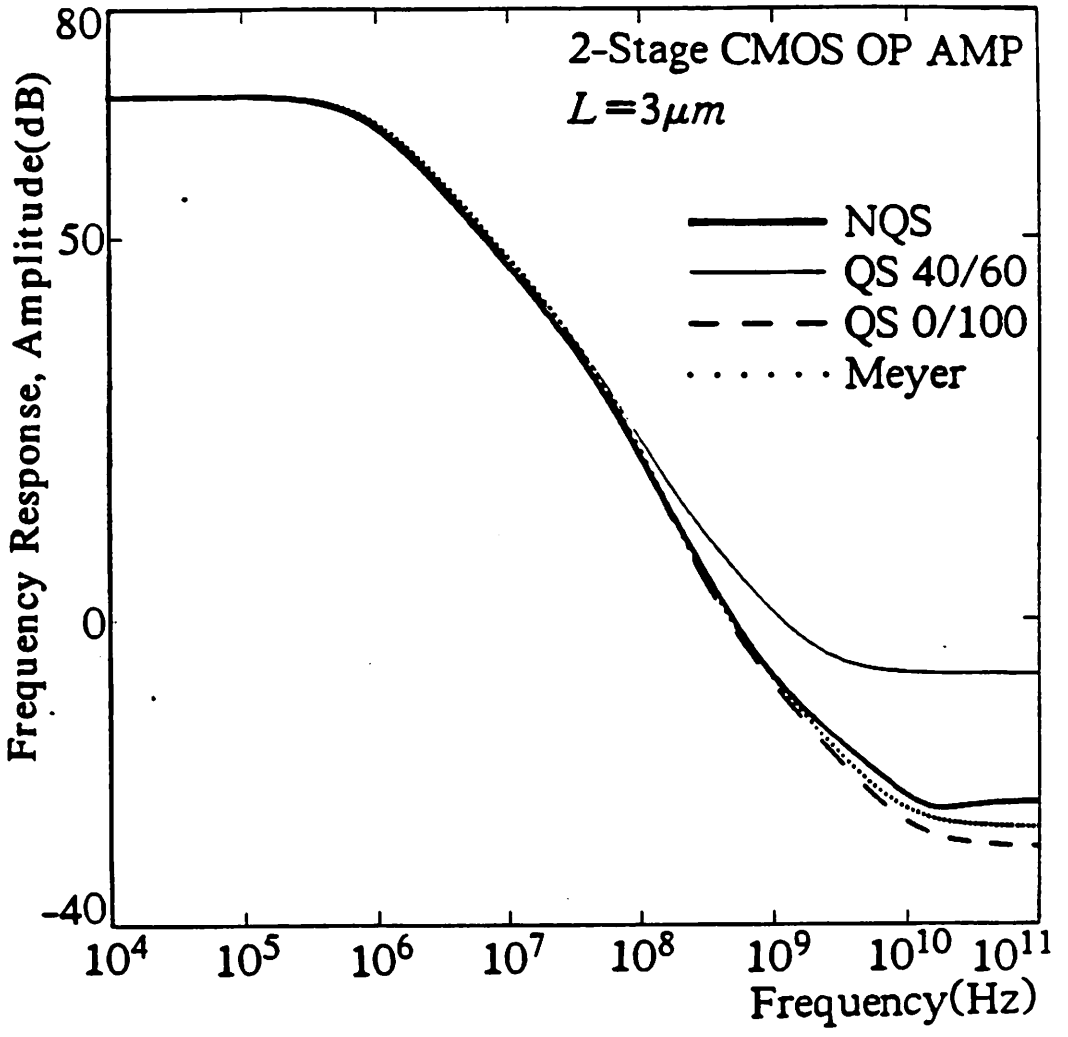


Fig.4.13.(b). Amplitude response

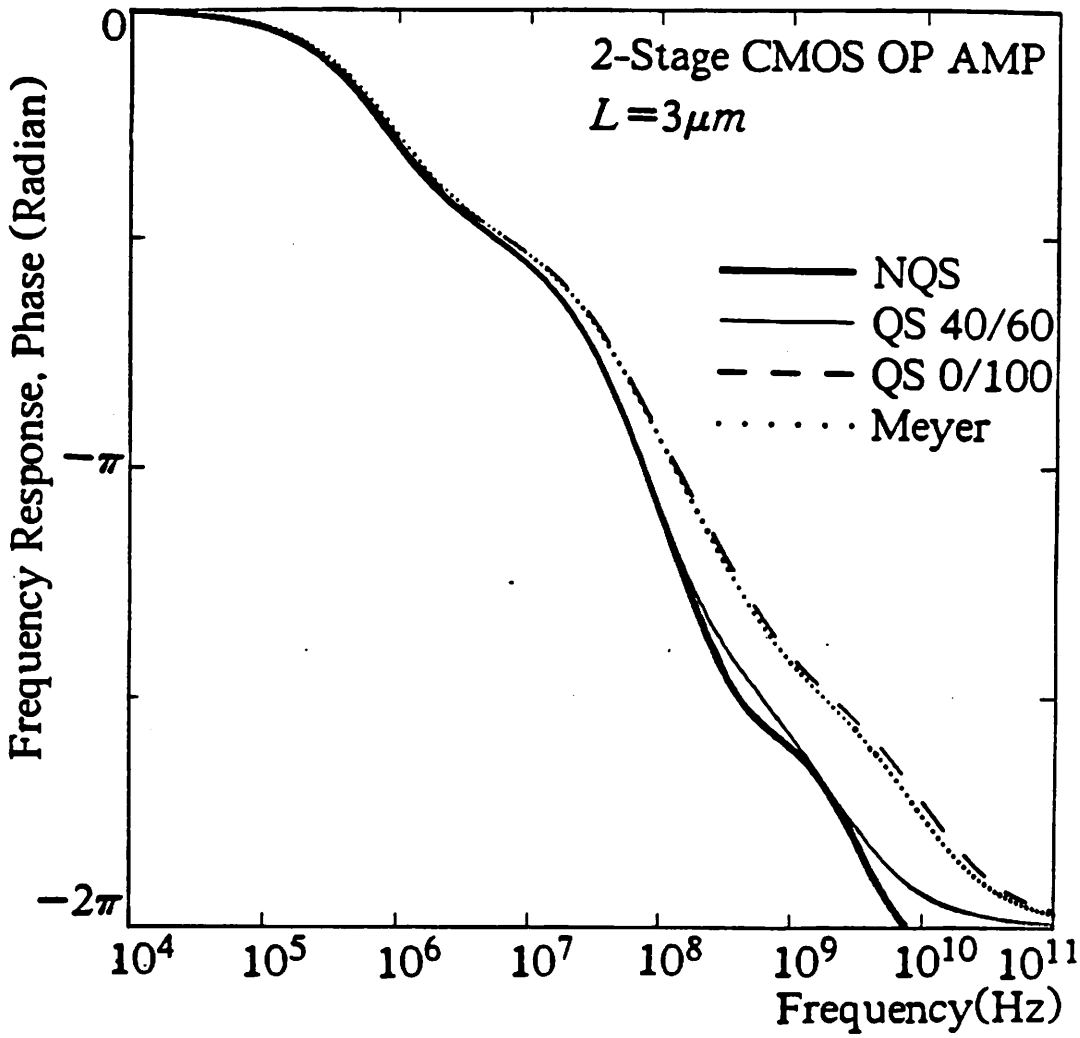


Fig.4.13.(c). Phase response

	pole 1	pole 2	zero 1	zero 2
Meyer	−830*KHz*	−92*MHz*	+320*MHz*	+8.9*GHz*
QS 0/100	−840*KHz*	−92*MHz*	+350*MHz*	+11*GHz*
QS 40/60	−840*KHz*	−90*MHz*	+120*MHz*	+2.3*GHz*

(+ : right-half s-plane - : left-half s-plane)

Table.4.12. *Pole* and *zero* frequencies of the 2-stage CMOS OP Amp (Fig.4.13) for the QS models. These *pole* and *zero* frequencies of QS models are computed using the pole-zero analysis routine in SPICE3 [4.16].

NQS model shows more complex behavior than the 2 *pole* and 2 *zero* characteristics. The first two *poles* of NQS model match those of QS models but, beyond the second *pole* frequency, the NQS model is closer to QS 0/100 and Meyer model in amplitude response and is closer to QS 40/60 model in phase response. This behavior is due to the multi-*pole* characteristics as shown in (4.67). $f_T(\omega_T/2\pi)$ is 45 MHz for the two input PMOSFET's and is 710 MHz for the output NMOSFET's.

4.4.5. Frequency response of folded-cascode CMOS OP Amp

Fig.4.14.(a) shows the circuit diagram of a folded-cascode CMOS OP amp, which is widely used for its good high frequency performance and good stability characteristics [4.15]. Each transistor has a channel length of $3\mu m$. Two input PMOSFET's(50/3) and two output NMOSFET's(9.1/3) form the cascode configuration and all other MOSFET's form the current sources. The differential input voltage is used to see the differential mode voltage gain.

The small-signal half circuit is shown in Fig.4.14.(b) and equivalent circuits for the half circuit in saturation region are shown in Fig.4.14.(c) and 4.14.(d), for QS and NQS models respectively. The voltage gain of QS models in saturation region can be derived from the equivalent circuit in Fig.4.14.(c), as

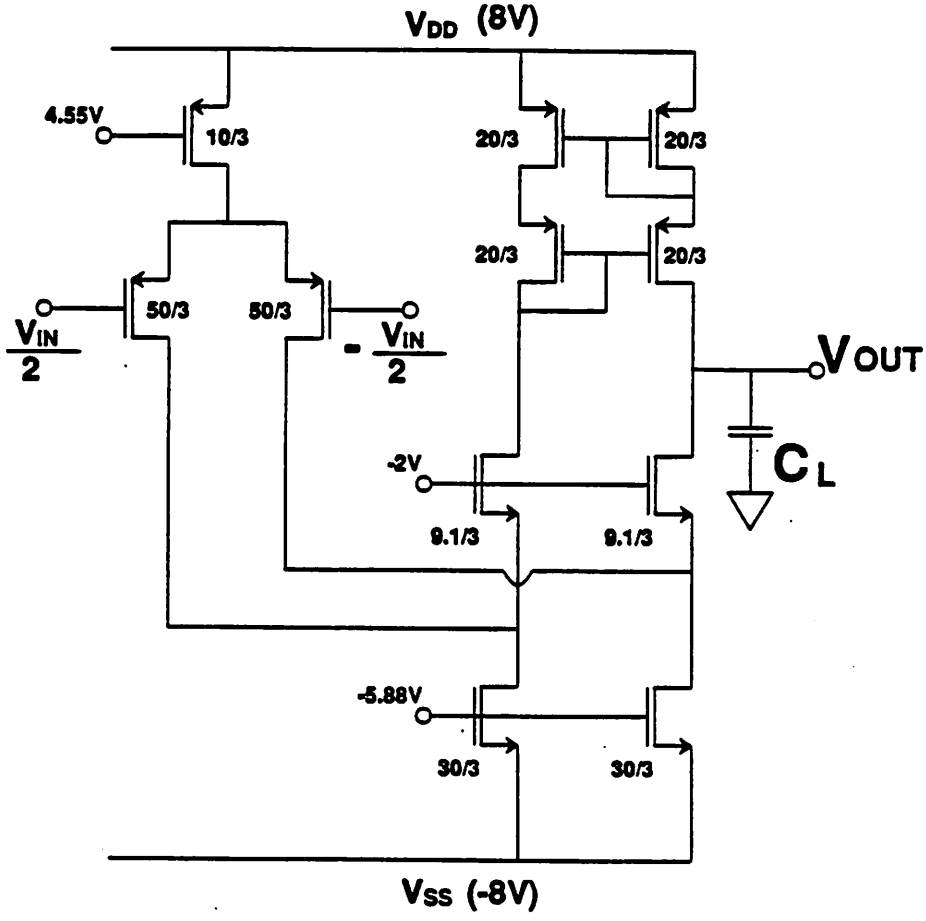


Fig.4.14. Frequency response of a folded-cascode CMOS OP amp[4.15] before the frequency compensation. Compensation is not done (C_L=0) to see the difference more clearly. The differential input voltage is applied to see the differential mode voltage gain.

Model parameters are the same as those in Fig.4.13. The zero-bias junction capacitance at the output node is 41.4 fF.

(a). Circuit diagram of a folded-cascode CMOS OP amp.

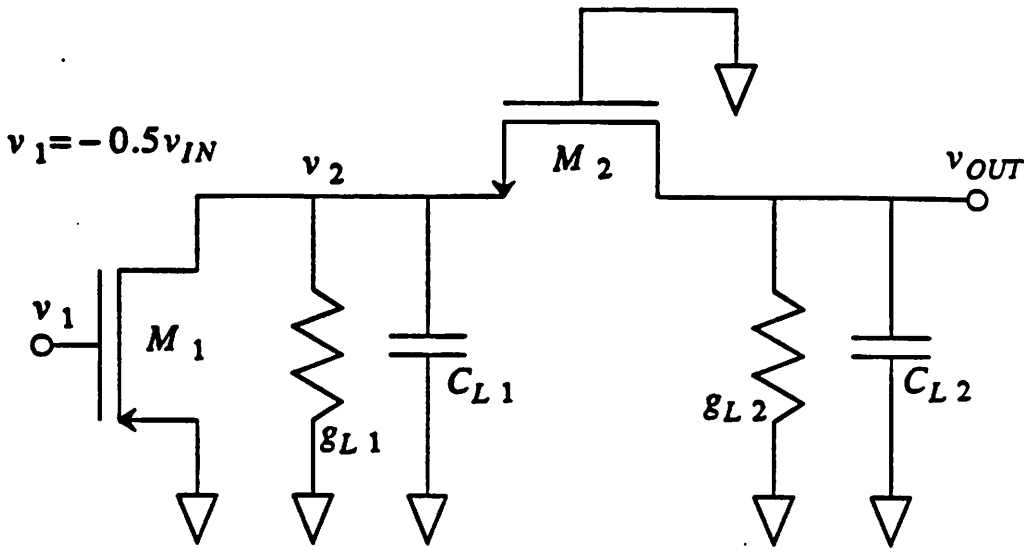


Fig.4.14.(b).

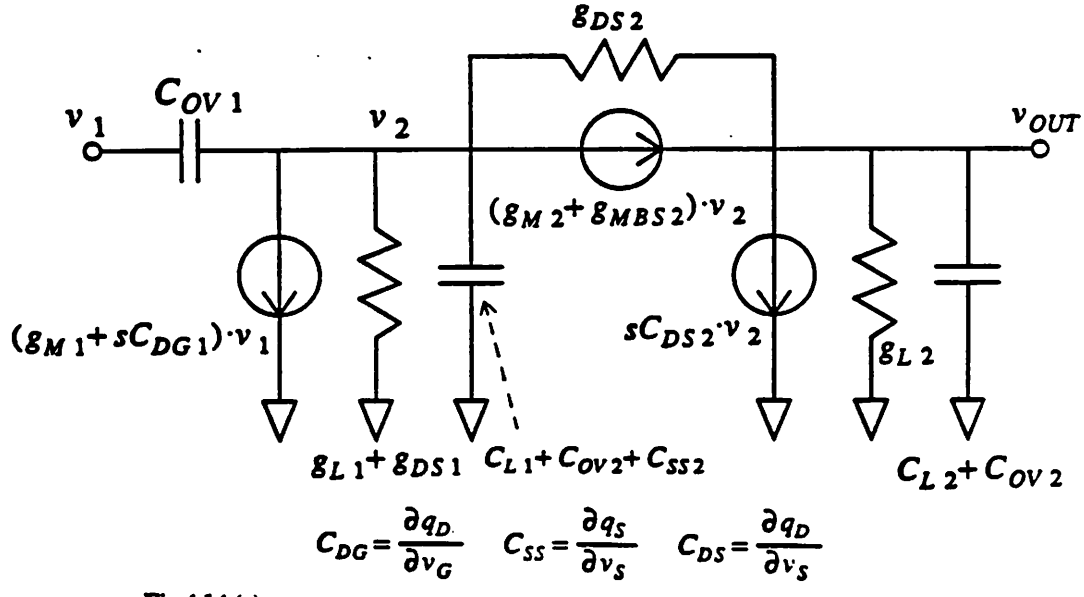


Fig.4.14.(c).

Fig.4.14.(b). Simplified cascode amplifier for AC analysis. This is the half circuit of the OP amp in Fig.4.14.(a).

Fig.4.14.(c). Equivalent circuit of Fig.4.14.(b) for QS charge based models

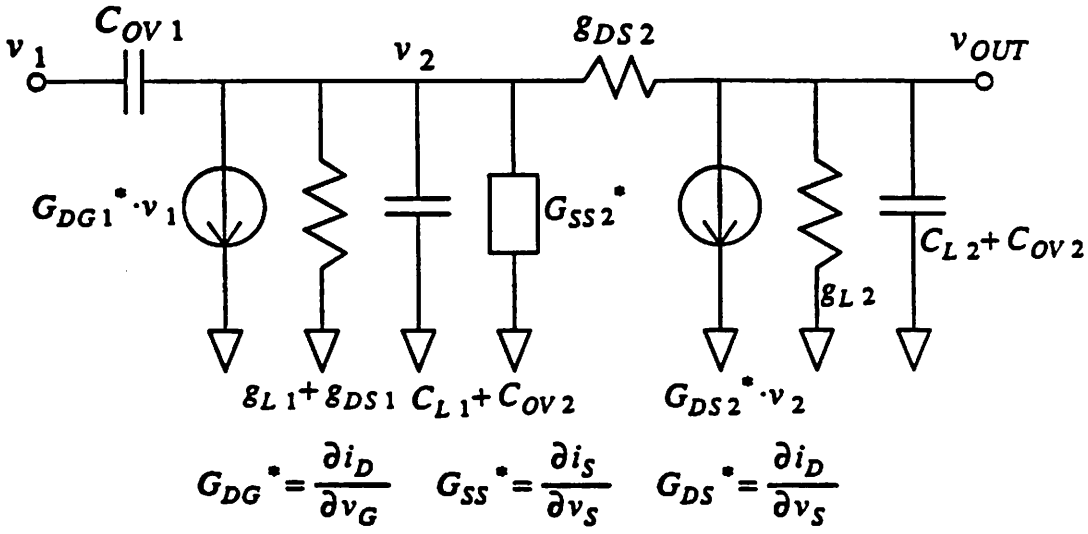


Fig.4.14.(d). Equivalent circuit of Fig.4.14.(b) for NQS model

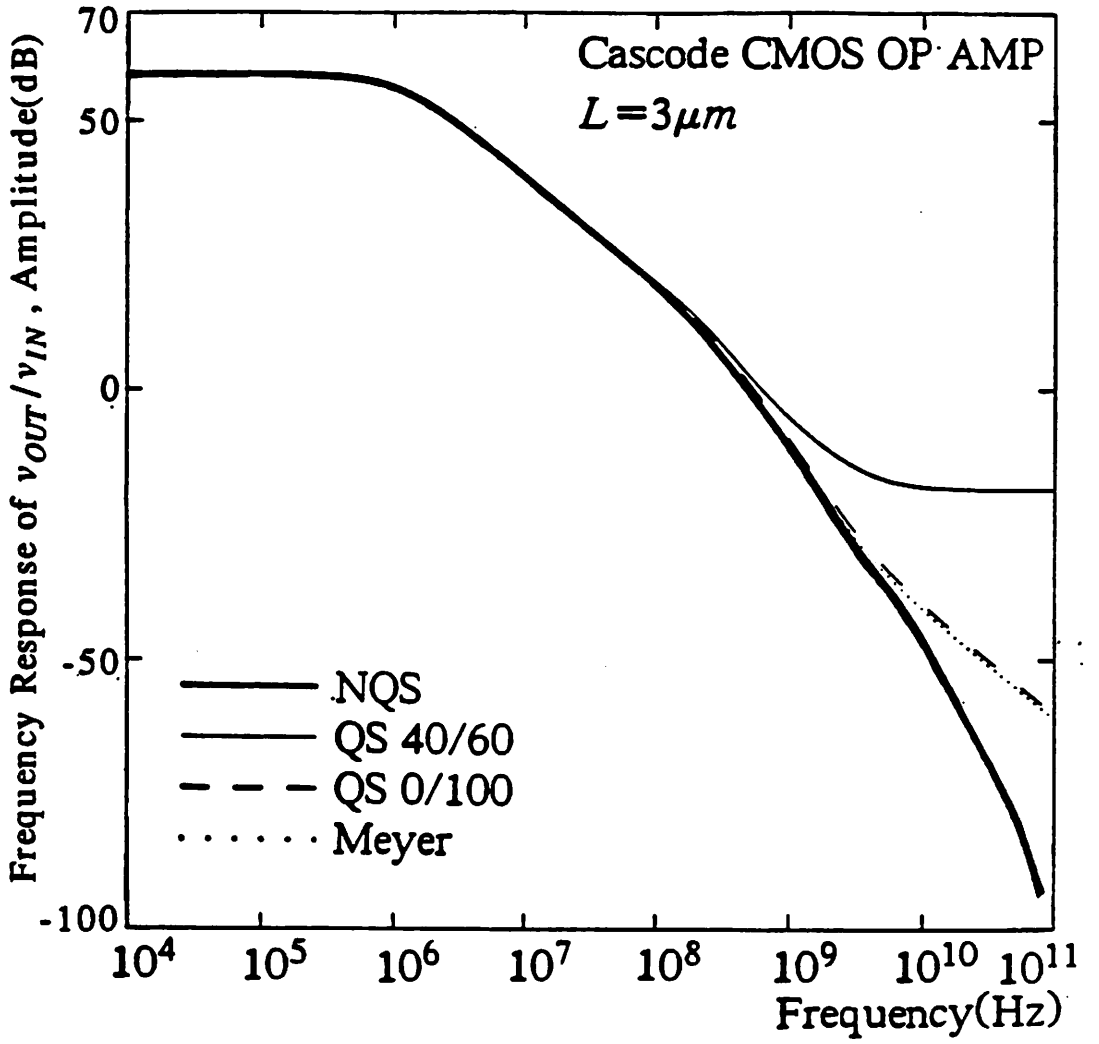


Fig.4.14.(e). Amplitude response of the folded-cascode CMOS OP amp in Fig.4.14.(a)

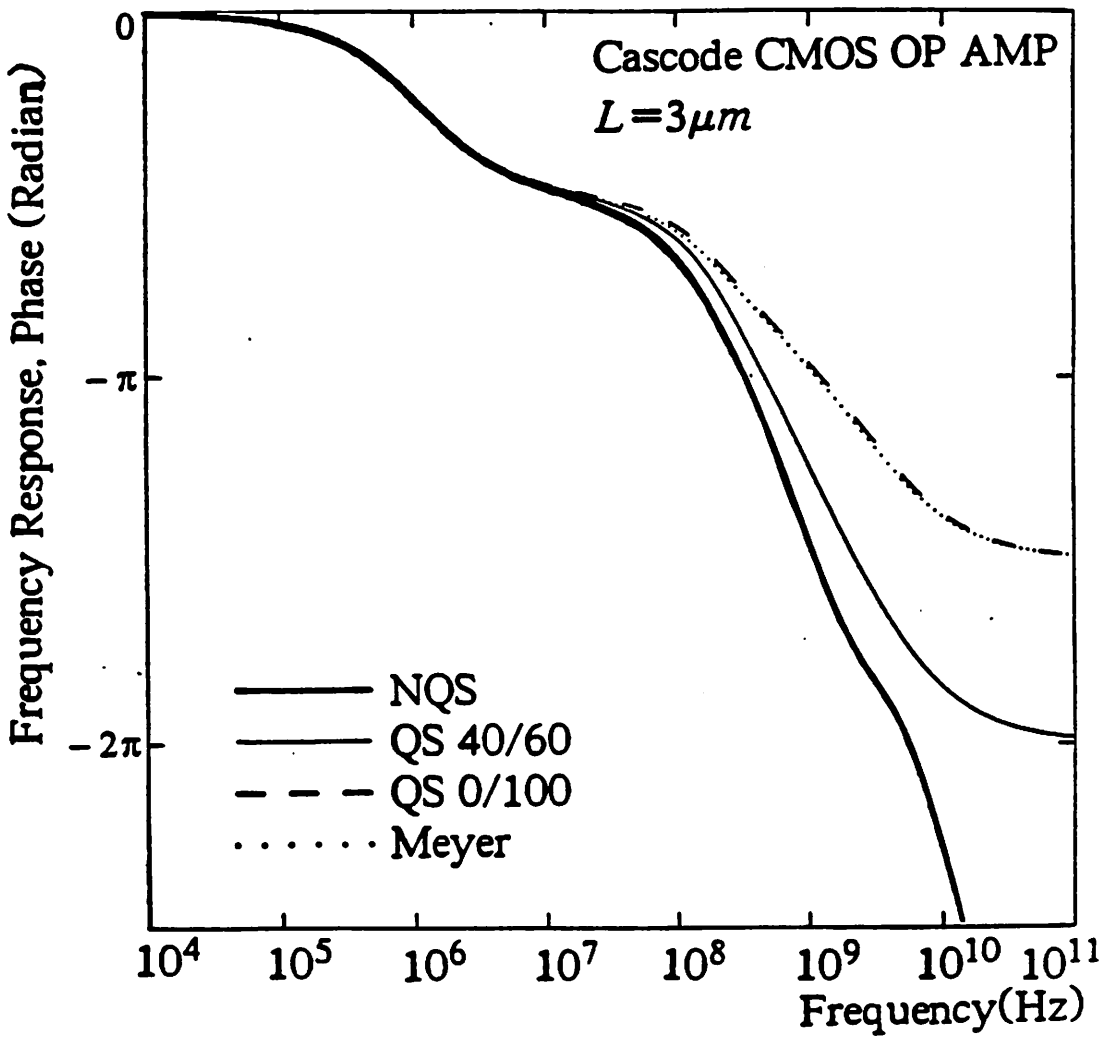


Fig.4.14.(f). Phase response of the folded-cascode CMOS OP amp in Fig.4.14.(a)

$$\frac{v_{OUT}}{v_{IN}} =$$

$$\frac{0.5\cdot\left[g_{M1}-s\,(C_{OV1}-C_{DG1})\right]\cdot\left[g_2-s\cdot C_{DS2}\right]}{GG_0+s\cdot\left[g_{L2}'\cdot(C_{SS2}+C_{L1}')+g_{DS2}\cdot C_{DS2}+(g_2+g_{L1}')\cdot C_{L2}'\right]+s^2\cdot C_{L2}'\cdot(C_{SS2}+C_{L1}')} \quad (4.74)$$

where $g_2 = g_{M2}+g_{MBS2}+g_{DS2}$ (4.75)

$g_{L1}' = g_{L1}+g_{DS1}$ (4.76)

$g_{L2}' = g_{L2}+g_{DS2}$ (4.77)

$C_{L1}' = C_{L1}+C_{OV1}+C_{OV2}$ (4.78)

$C_{L2}' = C_{L2}+C_{OV2}$ (4.79)

$GG_0 = g_{L2}\cdot(g_2+g_{L1}') + g_{DS2}\cdot g_{L1}'$ (4.80)

The subscript 1 refer to the transistor M1 and the subscript 2 refer to M2 in Fig.4.14.(b). The voltage gain in (4.74) has two *poles* in the left-half s-plane since all the coefficients of s terms in the denominator are positive. Neglecting the g_{DS2} term, we have the dominant pole frequency at $-(g_{L2}/C_{L2}')$ radian and the non-dominant pole frequency at $-(g_{M2}+g_{MBS2})/(C_{SS2}+C_{L1}')$ radian. The voltage gain of NQS model can be derived from Fig.4.14.(d).

Table 4.13 shows the voltage gains of the half circuit in Fig.4.14.(b) for Meyer, QS 0/100, 40/60 and NQS models respectively. The Meyer and QS 0/100 model show a single *zero* in the right-half s-plane (g_{M1}/C_{OV1} radian), while the QS 40/60 model shows two *zeroes* in the right-half s-plane ($g_{M1}/(4/15WLC_{OX1}+C_{OV1})$ and $g_2/0.45WLC_{OX2}$ radian).

The voltage gain equation of the cascode amplifier has also been derived in [4.17]. But eq. (16) in [4.17] shows two *poles* in the left-half s-plane and does not show any *zeroes*, since the capacitance model in [4.17] is basically the same as the Meyer model and the overlap capacitance is not included in the derivation.

model	$\frac{v_{OUT}}{v_{IN}}$
Meyer	$\frac{0.5\cdot(g_{M1}-sC_{OV1})\cdot g_2}{GG_0 + s\left[g_{L2}'\cdot(\frac{2}{3}WLC_{OX2}+C_{L1}')+(g_2+g_{L1}')\cdot C_{L2}'\right] + s^2\cdot C_{L2}'\cdot(\frac{2}{3}WLC_{OX2}+C_{L1}')}$
QS 0/100	$\frac{0.5\cdot(g_{M1}-sC_{OV1})\cdot g_2}{GG_0 + s\left[g_{L2}'\cdot(1.12WLC_{OX2}+C_{L1}')+(g_2+g_{L1}')\cdot C_{L2}'\right] + s^2\cdot C_{L2}'\cdot(1.12WLC_{OX2}+C_{L1}')}$
QS 40/60	$\frac{0.5\cdot\left[g_{M1}-s(\frac{4}{15}WLC_{OX1}+C_{OV1})\right]\cdot\left[g_2-s\cdot 0.45WLC_{OX2}\right]}{GG_0 + s\left[g_{L2}'\cdot(0.67WLC_{OX2}+C_{L1}')+g_{DS2}\cdot 0.45WLC_{OX2}+(g_2+g_{L1}')\cdot C_{L2}'\right] + s^2\cdot C_{L2}'\cdot(0.67WLC_{OX2}+C_{L1}')}$
NQS	$\frac{0.5(G_{DG1}^{\bullet}-sC_{OV1})\cdot(g_{DS2}-G_{DS2}^{\bullet})}{GG_1+(g_{L2}'\cdot G_{SS2}^{\bullet}+g_{DS2}\cdot G_{DS2}^{\bullet}) + s\cdot\left[C_{L1}'\cdot g_{L2}'+C_{L2}'(g_{L1}'+g_{DS2}+G_{SS2}^{\bullet})\right] + s^2\cdot C_{L2}'\cdot C_{L1}'}$

where $GG_1 = (g_{L1}'+g_{DS2})\cdot g_{L2}+g_{L1}'\cdot g_{DS2}$

Table.4.13. AC voltage gain equations of the half-circuit for the cascode OP Amp. in Fig.4.14.(b)

The frequency response of the original OP Amp in Fig.4.14.(a) before compensation (no external load capacitor) is shown in Fig.4.14.(e) and 4.14.(f). At low frequencies good agreements are observed among all four models but beyond the non-dominant *pole* frequency appreciable differences can be observed between each model. The DC gain is about 60 dB and the dominant *pole* frequency is 1.1 MHz and the non-dominant pole frequency is between 200 MHz and 300 MHz, as shown in Table 4.14. For the NQS model, $f_T(\omega_T/2\pi)$ is about 180 MHz for the input PMOSFET and 730 MHz for the cascode NMOSFET.

	pole 1	pole 2	zero 1	zero 2
Meyer	−1.1*MHz*	−220*MHz*	+3.6*GHz*	
QS 0/100	−1.1*MHz*	−240*MHz*	+3.8*GHz*	
QS 40/60	−1.1*MHz*	−310*MHz*	+820*MHz*	+3.9*GHz*

(+ : right-half s-plane - : left-half s-plane)

Table.4.14 *Pole* and *zero* frequencies of the folded cascode CMOS OP amp (Fig.4.14) for the QS models. These *pole*, *zero* frequencies are computed using the pole-zero analysis routine in SPICE3 [4.16].

The amplitudes of NQS, Meyer and QS 0/100 model go to 0 at very high frequencies but that of QS 40/60 model goes to a non-zero constant value, as shown in Fig.4.14.(e). This can be verified also from the equations in Table 4.13. $G_{DG}{}^{*}$ $(\partial i_D/\partial v_G)$ and $G_{DS}{}^{*}$ $(\partial i_D/\partial v_S)$ go to 0 at very high frequencies.

The phase shift of NQS model increases unboundedly with frequency, while that of Meyer and QS 0/100 model goes to -1.5π radian and that of QS 40/60 model goes to -2π radian due to the 2-*pole*-1-*zero* and 2-*pole*-2-*zero* characteristics respectively.

In comparison with the NQS model, the QS 40/60 model shows too large amplitude and the QS 0/100 and Meyer model show too large phase error at high frequencies.

CPU time for the AC analysis of CMOS OP Amp's using SPICE3 is shown in Table 4.15. NQS model takes about twice longer than the conventional SPICE models in the AC analysis.

Circuit	NQS	QS 40/60	QS 0/100	Meyer
2-Stage CMOS OP Amp	2.4(3.0)	1.2(1.9)	1.2(1.9)	1.1(1.7)
Cascode CMOS OP Amp	3.2(3.8)	1.6(2.2)	1.6(2.2)	1.5(2.0)

Table.4.15. CPU time comparison in SPICE3(VAX 8800 with Ultrix V2.0) for the AC analysis of the 2-stage and the folded-cascode CMOS OP Amp. The numbers represent the total analysis time in seconds and the numbers inside the brackets are the total run time in seconds. The total frequency points in the simulation is 226 and 181 for the 2-stage and the cascode OP Amp respectively.

4.5. Conclusion

Based on the approximate solution of the current continuity equation, an analytic NQS(non-quasistatic) MOSFET model for AC analysis has been derived and implemented in SPICE3. This model is based on the charge sheet formulation [4.12] and it includes both the drift and the diffusion current components. Mobility is assumed to be constant and the NQS behavior on the AC small-signal bulk current is not included in this model.

To check the validity of the model, this work has been compared with the 1-D numerical solution to the current continuity equation and also with the multiple lumped model where a MOSFET has been decomposed into many short channel MOSFETs and the resulting circuit has been simulated with SPICE using QS models. Excellent agreements

among them have been observed. Also all the known anomalies of QS(quasistatic) models in the AC analysis have been solved with this NQS model.

The frequency dependence of AC small-signal channel current partitioning ratio in saturation region is investigated. In saturation region, this ratio matches exactly the channel charge partitioning ratio of QS charge based models. For NQS, 1-D numerical and multiple-lumped models, this ratio is 40/60 at low frequencies and goes asymptotically to 0/100 at high frequencies, while , for QS charge based models, this ratio is fixed to be a constant value independent of frequency.

This model has been used for the AC analysis of example circuits, such as, NMOS and CMOS inverters and CMOS OP Amp's. And it shows results quite different from those of QS models. The QS 40/60 model gives too large amplitude and the QS 0/100 and Meyer model give too large phase error at high frequencies. So any of these QS models is not satisfactory for the high frequency AC analysis, while this NQS model is valid for all the frequency range.

The CPU time required for this NQS model, is about twice those of the conventional QS models in SPICE3.

References

[4.1] D.A.Hodges, P.R.Gray, R.W.Brodersen, "Potential of MOS Technologies for Analog Integrated Circuits", IEEE JSSC, Vol. SC-13, pp.285-294, June 1978

[4.2] S.Y.Oh, D.E.Ward, R.W.Dutton, "Transient Analysis of MOS Transistors", IEEE JSSC Vol.SC-15, No.4, pp.636-643, 1980

[4.3] Y.Tsividis, G.Masetti, "Problems in the Precision Modeling of the MOS Transistor for Analog Applications", IEEE Trans. CAD Vol. CAD-3, No.1, pp. 72-79, Jan. 1984

[4.4] J.J.Paulous, D.A.Antoniadis, "Limitations of quasi-static capacitance models for the MOS Transistor", IEEE EDL vol.EDL-4, pp.221-224, 1983

[4.5] J.G.Fossum, H.Jeong, S.Veeraraghavan, "Significance of the Channel-Charge Partition in the Transient MOSFET Model", IEEE Trans. Electron Devices, Vol.ED-33, No.10, pp. 1621-1623, Oct., 1986

[4.6] M.F.Sevat, "On the Channel Charge Division in MOSFET Modeling", Technical Digest, ICCAD, pp. 208-210, Nov., 1987,

[4.7] P.Yang, P.K.Chatterjee, "SPICE Modeling for Small Geometry MOSFET Circuits", IEEE Trans. CAD vol.CAD-1, pp.169-182, 1982

[4.8] B.J.Sheu, D.L.Scharfetter, P.K.Ko, "SPICE2 Implementation of BSIM", ERL Memo, M85/42, U.C.berkekley, May, 1985

[4.9] A.Vladmirescu, S.Liu, "The Simulation of MOS Integrated Circuits using SPICE2", ERL Memo, M80/7, U.C.Berkeley, Oct., 1980

[4.10] C.Turchetti, P.Mancini, G. Masetti, "A CAD-Oriented Non-Quasistatic Approach for the Transient Analysis of MOS IC's", IEEE JSSC vol.SC-21, No.5, pp.827-836, 1986

[4.11] M.Bagheri, Y.Tsividis, "A Small Signal DC-to-High Frequency Nonquasistatic Models for the Four-Terminal MOSFET Valid in All Regions of Operation", IEEE Trans. ED, vol. ED-32, No.11, pp. 2383-2391, 1985

[4.12] H.J.Park, P.K.Ko, C.Hu, "A Measurement Based Charge Sheet Capacitance Model of Short Channel MOSFET's for SPICE", IEDM 86 Technical Digest, pp.40-43,

Dec. 1986

[4.13] H.J.Park, P.K.Ko, C.Hu, "A Non-Quasistatic MOSFET Model for SPICE", IEDM 87 Technical Digest, pp.652-655, Dec. 1987

[4.14] H.Khorramabadi, P.R.Gray, "High-Frequency CMOS Continuous Time Filters", IEEE JSSC, Vol.SC-19, No.6, pp.939-948, Dec. 1984

[4.15] P.R.Gray, R.G.Meyer, "MOS Operational Amplifier Design - A Tutorial Overview", IEEE JSSC, Vol.SC-17, No.6, pp.969-982, Dec. 1982

[4.16] T.L.Quarles, A.R.Newton, D.O.Pederson, A.Sangiovanni-Vincentelli, "SPICE 3B1 User's Guide", Electronics Research Lab., U.C.Berkeley, Nov. 1987

[4.17] S.Liu, L.W.Nagel, "Small-Signal MOSFET Models for Analog Circuit Design", IEEE JSSC, Vol.SC-17, No.6, pp.983-998, Dec. 1982

Chapter 5

A CHARGE-CONSERVING NON-QUASISTATIC MOSFET MODEL FOR SPICE TRANSIENT ANALYSIS

5.1. Introduction

As MOSFET is widely used in charge storage applications such as switched analog circuits[5.1] and DRAMs, it is important for the device model to conserve charge. Non-charge-conserving models give physically wrong simulation results for SOS[5.2] or SOI MOSFETs and for conventional MOSFETs where repetitive input waveforms are used [5.3]. It has been shown that charge based models which give explicit expressions for node charges are guaranteed to conserve charge [5.2] [5.3].

During the turn-off transient of MOSFET switches, the channel charge is injected into either the source or the drain node. This channel charge injection is known to be one of the major distortion sources in the low distortion switched capacitor filters [5.4] and also is one of the major bottle necks for the high-speed high-resolution MOS A/D converters [5.5]. Since the conventional QS SPICE models cannot predict this channel charge injection problem accurately, the need for a NQS MOSFET model becomes very important in the design of switched-analog circuits.

Turchetti et. al [5.6] reported a NQS transient MOSFET model for an assumed channel charge density profile which is linear for the QS component and is symmetrical for the NQS component. But since the assumed channel charge profile is over-simplified, the usefulness of the model [5.6] is limited. We reported a NQS MOSFET model [5.7]

which predicts the NQS behavior accurately both for the transient [5.8] and AC [5.9] analyses but does not conserve charge for the transient analysis. In this Chapter, we describe a charge conserving NQS MOSFET model for the transient analysis.

A comparison of various MOSFET capacitance models is shown in Table.5.1.

	MEYER	WARD	NQS1	NQS	PISCES
SLOW SIGNAL	+	+	+	+	+
FAST SIGNAL	—	—	+	+	+
CHARGE CONSERVATION	—	+	—	+	+
SIMULATION TIME	1	1	3	4	10000

Table.5.1. General comparison between models. 'MEYER' is the Meyer capacitance model [5.10], 'WARD' is the Ward-Dutton charge based model [5.2], 'NQS1' is the previous version of non-quasistatic model [5.7] described in Chapter 3, 'NQS' is this work, the charge conserving non-quasistatic model, and 'PISCES' is a 2-D device simulation program [5.12].

For the slow signals, every model works fine, but for the fast signals, the QS(quasistatic) models (Meyer, Ward) are not valid. The Meyer capacitance model (MEYER) and the previous version of NQS model (NQS1) do not conserve charge, while the QS charge based model (WARD), this work(charge conserving NQS model) and PISCES do. The relative simulation time is also shown in Table.5.1.

Section 5.2 describes the formulation steps of model equations.

Section 5.3 shows the simulation results and comparison with other models. Charge conservation property of this work is demonstrated, channel charge partitioning schemes of QS(quasistatic) models are shown to be incorrect. This work has been compared with

PISCES, 1-D numerical simulation, QS models, and multiple-lump models. A CMOS inverter chain, a CMOS ring osciliator, a feedback pass-tr., and a differential sample-hold circuit have been simulated using this work and the results have been compared with those using other models.

Section 5.4 describes the implementation of this work into SPICE3. More details of the implementation are shown in Appendix 11 and Appendix 12.

Section 5.5 concludes this Chapter.

5.2. Formulation of Model Equations

5.2.1. Derivation of inversion charge density profile $Q'_n(y,t)$

As will be shown in the next section, Section 5.2.2, once the normalized inversion charge density $Q'_n(y,t)$ is given, node charges and node currents can be found. To derive the equation for $Q'_n(y,t)$, we start from the current relation and the current continuity equation shown in Chapter 3, Section 3.2.1. From the current relation (3.1) and the current continuity equation (3.7), we can derive (3.8). Eq. (3.8) is repeated here for clarity.

$$\frac{\partial Q'_n(y,t)}{\partial t} = \frac{\mu_n}{2F_B}\cdot\frac{\partial^2}{\partial y^2}\left[(Q'_n(y,t) + F_B V_t)^2\right] \tag{5.1}$$

where μ_n is the mobility of the inversion charge carriers and F_B is the correction factor which is slightly larger than 1.0 and is shown in (3.5) and V_t is the thermal voltage kT/q. The only approximation used in the derivation of (5.1) is the linearization approximation of the bulk charge as shown in (3.2) and (3.3).

To solve the partial differential equation (5.1), we decompose $Q'_n(y,t)$ into the QS(quasistatic) component and the NQS(non-quasistatic) component. The QS component can be derived from (5.1) by setting the time derivative to be zero. We assume the quasistatic approximation for inversion charge densities at source and drain ends, that

is, we assume that the inversion charge densities at source and drain ends respond instantaneously to applied biases. According to this approximation, the non-quasistatic component of $Q'_n(y,t)$ is zero at source(y=0) and drain(y=L) ends respectively. Since any profile which is 0 at both ends can be represented by a Fourier sine series expansion, the NQS component of $Q'_n(y,t)$ is represented by a Fourier sine series expansion. Hence $Q'_n(y,t)$ is represented as

$$Q'_n(y,t) = \sqrt{P_S(t) - (P_S(t) - P_D(t))\cdot\frac{y}{L}} - F_B V_t + \sum_{n=1}^{10} A_n(t)\cdot\sin(n\pi\cdot\frac{y}{L}) \tag{5.2}$$

where $$P_S(t) = (Q'_n(0,t) + F_B V_t)^2 \tag{5.3}$$

$$P_D(t) = (Q'_n(L,t) + F_B V_t)^2 \tag{5.4}$$

$$Q'_n(0,t) = V_{GST}(t) \tag{5.5}$$

$$Q'_n(L,t) = V_{GST}(t) - F_B\cdot(\Psi_{SL}(t) - \Psi_{SO}(t)) \tag{5.6}$$

$$V_{GST}(t) = V_{GB}(t) - V_{FB} - \Psi_{SO}(t) - \gamma\sqrt{\Psi_{SO}(t) - V_t} \tag{5.7}$$

where Ψ_{SL} and Ψ_{SO} are surface potentials at drain and source ends respectively and are computed from the charge sheet formulation as shown in Chapter 2, Section 2.2.

The square root term and $F_B V_t$ term in (5.2) are quasistatic components. In this work, 10 terms are used for the Fourier sine series in (5.2).

Substituting (5.2) into (5.1), we have

$$\frac{\partial G(y,t)}{\partial t} + \sum_{n=1}^{10}\frac{dA_n(t)}{dt}\cdot\sin(n\pi\cdot\frac{y}{L}) = \frac{\mu_n}{2F_B}\cdot\frac{\partial^2}{\partial y^2}\left[D(y,t)\cdot\sum_{n=1}^{10} A_n(t)\cdot\sin(n\pi\cdot\frac{y}{L})\right] \tag{5.8}$$

where $$G(y,t) = \sqrt{P_S(t) - (P_S(t) - P_D(t))\cdot\frac{y}{L}} - F_B V_t \tag{5.9}$$

$$D(y,t) = 2\sqrt{P_S(t) - (P_S(t) - P_D(t))\cdot\frac{y}{L}} + \sum_{n=1}^{10} A_n(t)\cdot\sin(n\pi\cdot\frac{y}{L}) \tag{5.10}$$

$G(y,t)$ in (5.9) can be rewritten as

$$G(y,t) = \sqrt{P_S(t)}\cdot\sqrt{1 - a(t)\cdot\frac{y}{L}} - F_B V_t \tag{5.11}$$

$$\text{where} \quad a(t) = 1 - \frac{P_D(t)}{P_S(t)} \tag{5.12}$$

$a(t)$ has a value in the range between 0 and 1. $a(t)$=0 in cutoff or when V_{DS}=0, and $a(t)$=1 in saturation region. $G(y,t)$ is approximated as (5.13) by expanding the square root term in (5.11) as a Taylor series with three terms. Hence,

$$G(y,t) \approx \sqrt{P_S(t)}\cdot\left[1 - \frac{a(t)}{2}\cdot\frac{y}{L} - b(t)\cdot\left(\frac{y}{L}\right)^2\right] - F_B V_t \tag{5.13}$$

$$\text{where} \quad b(t) = 3 - \frac{3}{4}\cdot a(t) - 2\cdot\frac{1-(1-a(t))^{1.5}}{a(t)} \tag{5.14}$$

$$= \frac{(a(t))^2}{8} \qquad (for \quad a(t) \approx 0)$$

$b(t)$ has been chosen such that it guarantees that the integration of the right hand side of (5.11) with y from y=0 to y=L is equal to the integration of the right hand side of (5.13). $D(y,t)$ in (5.10) is approximated as a Fourier cosine series so that the multiplication of a cosine series and a sine series in the right hand side of (5.8) gives another sine series and it enables the matching of the sine series coefficients in both sides of (5.8). Hence,

$$D(y,t) \approx D_0(t) + 2\cdot D_1(t)\cdot\cos\left(\pi\cdot\frac{y}{L}\right) \tag{5.15}$$

Only two terms are used for the Fourier cosine series expansion in (5.15) for the computational efficiency which will be shown later. $D_0(t)$ in (5.15) is the average value of $D(y,t)$ between y=0 and y=L. Hence,

$$D_0(t) = \frac{1}{L}\cdot\int_0^L D(y,t)\,dy \tag{5.16}$$

Substituting (5.10) into (5.16), we have

$$D_0(t) = \frac{4}{3}\cdot\sqrt{P_S(t)}\cdot\frac{1-(1-a(t))^{1.5}}{a(t)} + \frac{2}{\pi}\cdot\sum_{m=1}^{5}\frac{A_{2m-1}(t)}{2m-1} \tag{5.17}$$

$$= 2\sqrt{P_S(t)}\cdot(1 - \frac{a(t)}{4}) + \frac{2}{\pi}\cdot\sum_{m=1}^{5}\frac{A_{2m-1}(t)}{2m-1} \qquad (for \quad a(t) \approx 0)$$

$D_1(t)$ is chosen such that the approximation (5.15) is exact at the source end. Hence,

$$D_1(t) = \frac{1}{2}\cdot(D(0,t) - D_0(t))\cdot G(a(t))$$

$$= (\sqrt{P_S(t)} - \frac{1}{2}\cdot D_0(t))\cdot G(a(t)) \qquad (5.18)$$

where $G(a) = 1.0$ (for $a > 0.1$) (5.19)

$$= 1 - (1-\frac{a}{0.1})^2 \qquad (for \quad 0 < a \le 0.1)$$

$$= 0 \qquad (for \quad a \le 0)$$

The approximation (5.18) is more accurate near the source end of the channel than near the drain end. The function $G(a)$ is added to set D_1=0 when a=0, that is, when the MOSFET is in cutoff region or when V_{DS}=0. The constant 0.1 in (5.19) is a heuristic factor. $D_1(t)$ is limited to have values within the range $[-0.5\cdot D_0(t), +0.5\cdot D_0(t)]$ to guarantee $D(y,t) \ge 0$ for all $y(0 \le y \le L)$ and hence the numerical stability of the solution method.

Substituting (5.15) into (5.8), we have

$$\frac{\partial G(y,t)}{\partial P_S}\cdot\frac{dP_S(t)}{dt} + \frac{\partial G(y,t)}{\partial P_D}\cdot\frac{dP_D(t)}{dt} - \frac{dF_B}{dt}\cdot V_t + \sum_{n=1}^{10}\frac{dA_n(t)}{dt}\cdot\sin(n\pi\cdot\frac{y}{L})$$

$$= -\frac{\mu_n}{2F_B}\cdot\sum_{n=1}^{10}A_n(t)\cdot \qquad (5.20)$$

$$\left[((n-1)\pi)^2\cdot\frac{D_1(t)}{L^2}\cdot\sin((n-1)\pi\cdot\frac{y}{L}) + (n\pi)^2\cdot\frac{D_0(t)}{L^2}\cdot\sin(n\pi\cdot\frac{y}{L}) + ((n+1)\pi)^2\cdot\frac{D_1(t)}{L^2}\cdot\sin((n+1)\pi\cdot\frac{y}{L})\right]$$

Expanding (y/L) and $(y/L)^2$ in (5.13) as Fourier sine series, we can represent $\partial G(y,t)/\partial P_S$ and $\partial G(y,t)/\partial P_D$ in (5.20) as Fourier sine series. Hence,

$$\frac{\partial G(y,t)}{\partial P_S} = -\sum_{n=1}^{10}G_{S,n}(t)\cdot\sin(n\pi\cdot\frac{y}{L}) \qquad (5.21)$$

$$\frac{\partial G(y,t)}{\partial P_D} \approx -\sum_{n=1}^{10} G_{D,n}(t)\cdot\sin(n\pi\cdot\frac{y}{L}) \tag{5.22}$$

where

$$G_{S,n}(t) = -\frac{1}{\sqrt{P_S(t)}}\cdot$$

$$\left[\frac{1-(-1)^n}{n\pi} - (1-\frac{a}{2})\cdot\frac{(-1)^{n+1}}{n\pi} - (b + 2(1-a)\cdot\frac{db}{da})\cdot(\frac{(-1)^{n+1}}{n\pi} - \frac{2}{(n\pi)^3}\cdot(1-(-1)^n))\right] \tag{5.23}$$

$$G_{D,n}(t) = -\frac{1}{\sqrt{P_S(t)}}\cdot\left[\frac{(-1)^{n+1}}{n\pi}\cdot(1 + 2\cdot\frac{db}{da}) - \frac{2}{(n\pi)^3}\cdot(1-(-1)^n)\right] \tag{5.24}$$

In (5.23) and (5.24), the argument (t) is dropped from $a(t)$ and $b(t)$ for simplicity, db/da is the derivative of $b(t)$ in (5.14) with respect to $a(t)$.

Substituting (5.21) and (5.22) into (5.20) and matching the coefficients of each Fourier sine series term in both sides of the resulting equation, we can derive (5.25) for each n.

$$\frac{dA_n(t)}{dt} = G_{S,n}(t)\cdot\frac{dP_S(t)}{dt} + G_{D,n}(t)\cdot\frac{dP_D(t)}{dt}$$

$$-\frac{\mu_n}{2F_B}\cdot\frac{(n\pi)^2}{L^2}\cdot(A_{n-1}(t)\cdot D_1(t) + A_n(t)\cdot D_0(t) + A_{n+1}(t)\cdot D_1(t)) \tag{5.25}$$

$$(\text{for } n = 1, 2, 3, ..., 10)$$

where $A_0(t)$ and $A_{11}(t)$ are set to 0 in (5.25).

Rewriting (5.25) in the matrix notation, we can get the state equation (5.26).

$$\frac{d\mathbf{A}(t)}{dt} = \mathbf{D}(t)\cdot\mathbf{A}(t) + \mathbf{G_S}(t)\cdot\frac{dP_S(t)}{dt} + \mathbf{G_D}(t)\cdot\frac{dP_D(t)}{dt} \tag{5.26}$$

where $\mathbf{A}(t)$ is a column matrix for the coefficients $\{A_n(t)\}$, $\mathbf{D}(t)$ is a tri-diagonal square matrix, $\mathbf{G_S}(t)$ and $\mathbf{G_D}(t)$ are column matrices accounting for excitations at source and drain ends respectively. The coefficients of the square matrix $\mathbf{D}(t)$ are

$$D_{i,(i-1)}(t) = -\frac{\mu_n}{2F_B}\cdot\frac{\pi^2}{L^2}\cdot i^2\cdot D_1 \tag{5.27}$$

$$D_{i,i}(t) = -\frac{\mu_n}{2F_B}\cdot\frac{\pi^2}{L^2}\cdot i^2\cdot D_0 \tag{5.28}$$

$$D_{i,(i+1)}(t) = -\frac{\mu_n}{2F_B}\cdot\frac{\pi^2}{L^2}\cdot i^2\cdot D_1 \tag{5.29}$$

(for i = 1, 2, 3, ..., 10, but $D_{1,0}(t) = D_{10,11}(t) = 0$)

All the other coefficients of the matrix $\mathbf{D}(t)$ except the diagonal term in (5.27) and the two off-diagonal terms in (5.28) and (5.29) are 0.

The number of terms in the Fourier cosine series expansion of $D(y,t)$ in (5.15) is important for the computational efficiency, since, if more than 2 terms are used in (5.15), the matrix $\mathbf{D}$ is no longer a tri-diagonal matrix and the total computation time is about twice longer than the case of tri-diagonal matrix.

The coefficients of matrices $\mathbf{G_S}(t)$ and $\mathbf{G_D}(t)$ are shown in (5.23) and (5.24) respectively. Although several integration methods are available to solve the state equation (5.26), the trapezoidal integration method is used in this work.

At t = 0 (initial time point for the transient analysis), all the coefficients $\{A_n(t)\}$ (for n=1,2,3,...,10) are set to 0 since NQS components are 0 at initial DC operating point. Old values (values at the previous time point) are used for matrices $\mathbf{D}$, $\mathbf{G_S}$ and $\mathbf{G_D}$ in (5.26) to make the computation of derivatives of node currents with respect to node voltages easier. So $\mathbf{D}$, $\mathbf{G_S}$ and $\mathbf{G_D}$ are treated as constant matrices during the small time interval considered.

Applying the trapezoidal integration scheme to the state equation (5.26), we have

$$\frac{\mathbf{A}(t) - \mathbf{A}(t_0)}{k} =$$

$$\frac{1}{2}\mathbf{D}(t_0)\cdot(\mathbf{A}(t) + \mathbf{A}(t_0)) + \mathbf{G_S}(t_0)\cdot\frac{\mathbf{P_S}(t) - \mathbf{P_S}(t_0)}{k} + \mathbf{G_D}(t_0)\cdot\frac{\mathbf{P_D}(t) - \mathbf{P_D}(t_0)}{k} \tag{5.30}$$

where t is the present time point, t_0 is the previous time point and k is the time step which is equal to $(t - t_0)$. Eq. (5.30) can be rewritten as

$$(\mathbf{I} - \frac{k}{2}\mathbf{D})\cdot(\mathbf{A}(t) - \mathbf{A}(t_0)) =$$

$$k\,\mathbf{D}(t_0)\cdot\mathbf{A}(t_0) + \mathbf{G_S}(t_0)\cdot(P_S(t) - P_S(t_0)) + \mathbf{G_D}(t_0)\cdot(P_D(t) - P_D(t_0)) \qquad (5.31)$$

where **I** is a unit matrix. Eq. (5.31) is decomposed into 3 sets of equations to make the computation of derivatives of node currents and node charges with respect to node voltages easier.

$$(\mathbf{I} - \frac{k}{2}\mathbf{D}(t_0)) \cdot \frac{\partial \mathbf{A}}{\partial t} = \mathbf{D}(t_0) \cdot \mathbf{A}(t_0) \qquad (5.32)$$

$$(\mathbf{I} - \frac{k}{2}\mathbf{D}(t_0)) \cdot \frac{\partial \mathbf{A}}{\partial P_S} = \mathbf{G_S}(t_0) \qquad (5.33)$$

$$(\mathbf{I} - \frac{k}{2}\mathbf{D}(t_0)) \cdot \frac{\partial \mathbf{A}}{\partial P_D} = \mathbf{G_D}(t_0) \qquad (5.34)$$

Using (5.27), (5.28) and (5.29) and the fact that $D_0(t)$ is positive and $-0.5\cdot D_0(t) \le D_1(t) \le +0.5\cdot D_0(t)$, we can see that the matrix $(\mathbf{I} - 0.5k\cdot\mathbf{D})$ is diagonally dominant and hence it guarantees the stable solution.

From the solutions of three matrix equations (5.32), (5.33) and (5.34), we can find $\mathbf{A}(t)$ as

$$\mathbf{A}(t) = \mathbf{A}(t_0) + k\cdot\frac{\partial \mathbf{A}}{\partial t} + \frac{\partial \mathbf{A}}{\partial P_S}\cdot(P_S(t) - P_S(t_0)) + \frac{\partial \mathbf{A}}{\partial P_D}\cdot(P_D(t) - P_D(t_0)) \qquad (5.35)$$

Intermediate solutions $\partial\mathbf{A}/\partial t$, $\partial\mathbf{A}/\partial P_S$ and $\partial\mathbf{A}/\partial P_D$ are used to compute the derivatives of node currents and node charges with respect to applied biases. These derivatives will be used to form the Jacobian matrix for the nonlinear Newton-Raphson iterations in solving circuit equations.

5.2.2. Derivation of equations for node currents and node charges

From the current continuity equation, we can derive analytic equations for drain and source currents, and also for drain and source charges. The detailed derivation steps are shown in Appendix.10. Hence, the drain and source currents, $I_D(t)$ and $I_S(t)$ can be written as

$$I_D(t) = I_{DC}(t) + \frac{dQ_D(t)}{dt} \qquad (5.36)$$

$$I_S(t) = -I_{DC}(t) + \frac{dQ_S(t)}{dt} \quad . \tag{5.37}$$

$$\text{where} \quad Q_D(t) = -WC_{OX}\cdot\int_0^L \frac{y}{L}\cdot Q'_n(y,t)\,dy \tag{5.38}$$

$$Q_S(t) = -WC_{OX}\cdot\int_0^L (1-\frac{y}{L})\cdot Q'_n(y,t)\,dy \tag{5.39}$$

where $Q'_n(y,t)$ is the inversion charge density normalized by $(-WC_{OX})$, $Q_D(t)$ and $Q_S(t)$ are drain and source charges respectively.

The DC current $I_{DC}(t)$ can be written as (5.40) including the empirical channel length modulation effect.

$$I_{DC}(t) = \frac{W}{L}\cdot C_{OX}\cdot\frac{\mu_n}{2F_B}\cdot(P_S(t)-P_D(t))\cdot(1+\lambda\cdot V_{DS}) \tag{5.40}$$

where F_B is a correction factor which is slightly larger than 1.0 and is shown in (3.5), $P_S(t)$ and $P_D(t)$ are shown in (5.3) and (5.4) respectively. λ is the empirical channel length modulation factor and V_{DS} is the applied drain to source voltage.

$I_{DC}(t)$ in (5.36) and (5.37) can be replaced by any DC model equation other than (5.40). In fact, the SPICE level-2 DC model with all the short channel effects included has been combined with this NQS charge model and has been implemented in SPICE3, as shown in Section 5.4 and Appendix 12.

Substituting the equation for $Q'_n(y,t)$ (5.2) into (5.38) and (5.39), we can find

$$Q_D(t) = -WLC_{OX}\cdot$$

$$\left[\sqrt{P_S(t)}\cdot(\frac{2}{3}\cdot\frac{1-(1-a)^{1.5}}{a^2} - \frac{2}{5}\cdot\frac{1-(1-a)^{2.5}}{a^2}) - \frac{F_B V_t}{2} + \sum_{n=1}^{10}(-1)^{(n-1)}\cdot\frac{A_n(t)}{n\pi}\right] \tag{5.41}$$

$$Q_S(t) = -WLC_{OX}\cdot$$

$$\left[\sqrt{P_S(t)}\cdot(-\frac{2}{3}\cdot(1-a)\cdot\frac{1-(1-a)^{1.5}}{a^2} + \frac{2}{5}\cdot\frac{1-(1-a)^{2.5}}{a^2}) - \frac{F_B V_t}{2} + \sum_{n=1}^{10}\frac{A_n(t)}{n\pi}\right] \tag{5.42}$$

where $a(t)$ is shown in (5.12) and the argument (t) in $a(t)$ is not shown in (5.41) and

(5.42) for clarity. Since $a(t)$ comes in the denominator of (5.41), and (5.42), (5.41) and (5.42) are converted into asymptotic forms for small $a(t)$. Hence

$$Q_D(t) = -WLC_{OX}\cdot\left[0.5\sqrt{P_S(t)}\cdot(1-a\cdot(\frac{1}{3}+\frac{a}{16}))+\sum_{n=1}^{10}(-1)^{(n-1)}.\frac{A_n(t)}{n\pi}\right] \tag{5.43}$$

$$Q_S(t) = -WLC_{OX}\cdot\left[0.5\sqrt{P_S(t)}\cdot(1-a\cdot(\frac{1}{6}+\frac{a}{48}))+\sum_{n=1}^{10}\frac{A_n(t)}{n\pi}\right] \tag{5.44}$$

(for $a(t) \approx 0$)

The gate current $I_G(t)$ can be represented as the time derivative of the gate charge $Q_G(t)$.

$$I_G(t) = \frac{dQ_G(t)}{dt} \tag{5.45}$$

$$\textit{where} \quad Q_G(t) = WC_{OX}\cdot\int_0^L V_{GB}(t) - V_{FB} - \Psi_S(t)\,dy \tag{5.46}$$

Using the relation between $Q'_n(y,t)$ and $\Psi_S(y,t)$ shown in (3.3), we can rewrite $Q_G(t)$ as

$$Q_G(t) = WC_{OX}\cdot\int_0^L V_{GB}(t) - V_{FB} - \Psi_{SO}(t) - \frac{V_{GST}(t)}{F_B} + \frac{Q'_n(y,t)}{F_B}\,dy \tag{5.47}$$

Substituting (5.38) and (5.39) into (5.47), we can represent $Q_G(t)$ in terms of $Q_D(t)$ and $Q_S(t)$.

$$Q_G(t) = WLC_{OX}\cdot\left[V_{GB}(t)-V_{FB}-\Psi_{SO}(t) - \frac{V_{GST}(t)}{F_B}\right] - \frac{Q_D(t)+Q_S(t)}{F_B} \tag{5.48}$$

Derivatives of node currents and node charges with respect to node voltages can be computed from (5.40), (5.41), (5.42) and (5.48). These derivatives are used to form the Jacobian matrix for the nonlinear Newton-Raphson iterations in solving circuit equations.

5.2.3. Moving boundary condition

Although this moving boundary condition was discussed in Chapter 3, Section 3.2.2 for the non-charge-conserving NQS model, it is discussed here again because slightly different formulation schemes are required for this model due to the fact that this model is a charge based model while the model described in Chapter 3 is not.

After the transistor is suddenly turned on, the carrier density is non-zero only near the source. The 1-D numerical solution of the current continuity equation shows that the carrier density profile during this time interval is linear in y except for a small tail at the right end edge due to diffusion as shown in Fig.3.3.(a). The linear carrier density profile is a self-consistent solution because, given the linear carrier density profile, the surface potential becomes linear in y according to (3.2) and therefore the electric field is constant throughout the region of non- zero carrier density profile. Since drift transport is dominant over diffusion during this time period, the carriers move at the same velocity, assuming a constant mobility. Hence the carrier density profile remains linear in y at the next time point, if V_{GST} changes linearly in time, as shown by a broken line in Fig.3.3.(b).

During this time period, $I_D(t)$ is of course zero. We call this period the moving boundary period and introduce a test condition to ensure $I_D(t)$=0 until the end of the moving boundary period. Without such a test condition, there will be small but troublesome I_D's. This is due to the fact that we used a finite number of terms(10 in this work) in the Fourier sine series expansion (5.2) and it is difficult to represent the two section profile shown in Fig.3.3.(a) accurately in terms of a Fourier series with finite terms.

To derive the moving boundary condition quantitatively, we assume that the normalized carrier density profile, $Q'_n(y,t)$, during this time period, is

$$Q'_n(y,t) = Q'_0(t) - Q'_M(t) \cdot y' \tag{5.49}$$

where y' is y/L and Q'_M is the slope of the profile. This assumption is suggested by the results of the 1-D numerical analysis and also by the reasonings discussed above.

Substituting (5.49) into (5.1), we have

$$\frac{dQ_0'(t)}{dt} - \frac{dQ_M'(t)}{dt}\cdot y' = \frac{\mu_n}{F_B L^2}\cdot(Q_M'(t))^2 \tag{5.50}$$

Assuming

$$\left| \frac{dQ_M'(t)}{dt} \right| << \left| \frac{dQ_0'(t)}{dt} \right| \tag{5.51}$$

From (5.50) one can derive (5.52).

$$Q_M'(t) = \sqrt{\frac{F_B L^2}{\mu_n} \cdot \frac{dQ_0'(t)}{dt}} \tag{5.52}$$

Eq. (5.51) is equivalent to assuming $d^2V_{GST}/dt^2 << 2\cdot(dV_{GST}/dt)^{1.5}\cdot\sqrt{\mu_n/(F_B L^2)}$, i.e., that V_{GST} changes almost linearly with time during the moving boundary period. The moving boundary condition lasts, until the edge of the carrier density profile reaches the drain side , i.e.

$$Q_0'(t) < Q_M'(t) \tag{5.53}$$

From (5.52) and (5.53), we have

$$Q_0'(t) < \sqrt{\frac{F_B L^2}{\mu_n} \cdot \frac{dQ_0'}{dt}} \tag{5.54}$$

$$\textit{where} \quad Q_0'(t) = V_{GST}(t) \tag{5.55}$$

By substituting (5.55) into (5.54), we can get the moving boundary condition as follows.

If

$$\frac{dV_{GST}}{dt} > \frac{\mu_n \cdot V_{GST}^2}{F_B L^2} = \frac{V_{GST}}{\tau_T} \tag{5.56}$$

and

$$(P_D(t) - P_D(t_0)) < (F_B V_t)^2 \tag{5.57}$$

Then

$$I_D(t) = I_D(t_0) \tag{5.58}$$

where τ_T is the channel transit time, t is the present time point and t_0 is the previous time point. If the conditions (5.56) and (5.57) are met, the drain charge $Q_D(t)$ is adjusted so that the drain current $I_D(t)$ is set to the value at the previous time point, $I_D(t_0)$.

Condition (5.57) is added to guarantee that the drain current does change with time when the carrier density at the drain edge changes with time in linear region even during the moving boundary period.

For charge based models, (5.58) can be rewritten as

$$I_D(t) = I_{DC}(t) + I_{DT}(t) = I_D(t_0) \tag{5.59}$$

$$\textit{where} \quad I_{DT}(t) = \frac{dQ_D(t)}{dt} \tag{5.60}$$

To compute the displacement current $I_{DT}(t)$ from (5.60), different orders of integration schemes can be used. Two most commonly used integration schemes, Backward Euler and Trapezoidal integration schemes, are described below. In SPICE transient simulations, the integration routine automatically switches orders of integration during one simulation. By default, SPICE uses both Backward Euler and Trapezoidal integration schemes during transient simulations.

<Backward Euler integration scheme>

If the Backward Euler integration (order 1) is used, (5.60) can be rewritten as

$$I_{DT}(t) = \frac{Q_D(t) - Q_D(t_0)}{k} \tag{5.61}$$

where k is the time step which is equal to $(t-t_0)$. Substituting (5.61) into (5.59), we can find the new drain charge $Q_D(t)$ during the moving boundary period, as

$$Q_D(t) = Q_D(t_0) + k\cdot(-I_{DC}(t) + I_D(t_0)) \tag{5.62}$$

New drain capacitances can be computed from (5.62) as

$$C_{DX}(t) = \frac{\partial Q_D(t)}{\partial V_X} = -k \cdot \frac{\partial I_{DC}(t)}{\partial V_X} \tag{5.63}$$

(where $V_X = V_G, V_D, V_S, V_B$)

$\partial I_{DC}(t)/\partial V_X$ in (5.63) can be computed from (5.40).

<Trapezoidal integration>

If the trapezoidal integration (order 2) is used, (5.60) can be rewritten as

$$I_{DT}(t) = -I_{DT}(t_0) + \frac{2}{k} \cdot (Q_D(t) - Q_D(t_0)) \tag{5.64}$$

Substituting (5.64) into (5.59), we have

$$Q_D(t) = Q_D(t_0) + \frac{k}{2} \cdot (-I_{DC}(t) + I_{DT}(t_0) + I_D(t_0)) \tag{5.65}$$

New drain capacitances can be derived from

$$C_{DX}(t) = \frac{\partial Q_D(t)}{\partial V_X} = -\frac{k}{2} \cdot \frac{\partial I_{DC}(t)}{\partial V_X} \tag{5.66}$$

where $V_X = V_G, V_D, V_S, V_B$

<Adjustment of coefficients $A_n(t)$>

Coefficients $\{A_n(t)\}$ are adjusted so that the drain charge $Q_D(t)$ computed from (5.41) or (5.43) is equal to $Q_D(t)$ computed from (5.62) or (5.65). This adjustment of $\{A_n(t)\}$ also guarantees the continuity of $Q_D(t)$ between the moving boundary period and the non-moving boundary period without changing $Q_S(t)$.

$$A_{2m-1}(t) = A'_{2m-1}(t) - \frac{0.5 \cdot \pi \cdot (Q_D(t) - Q'_D(t))}{(2m-1) \cdot WLC_{OX} \cdot \sum_{i=1}^{5} \frac{1}{(2i-1)^2}} \tag{5.67}$$

$$A_{2m}(t) = A'_{2m}(t) + \frac{0.5 \cdot \pi \cdot (Q_D(t) - Q_D'(t))}{(2m) \cdot WLC_{OX} \cdot \sum_{i=1}^{5} \frac{1}{(2i)^2}} \tag{5.68}$$

(for m = 1, 2, 3, 4, 5)

$A'_{2m}(t)$ and $A'_{2m-1}(t)$ are computed from (5.35), $Q'_D(t)$ is computed from (5.41) or (5.43) and $Q_D(t)$ is computed from (5.62) or (5.65). Although this scheme guarantees the continuity of Q_D between the moving boundary period and the non-moving boundary period, the displacement current dQ_D/dt is discontinuous between those two time periods. This discontinuity in drain current did not cause any serious convergence problem in circuit simulations as shown in Section 5.3. Related to the discontinuity of the displacement drain current, one must note that there is a big discontinuity at the turn-on time ($V_{GS} = V_{TH}$) for QS 40/60 and QS 50/50 models as shown in Section 5.3.4, and yet they work fine in real circuit simulations.

5.3. Simulation results and comparison with other models

This work has been implemented in SPICE3B.1 as will be shown in Section 5.4. Many example circuits are simulated using this work and the results are compared with those from other models. When a MOSFET is turned off, the inversion charge stored in the channel is injected out of the MOSFET either through drain or source node. This channel charge injection is important in switched analog circuits [5.1]. For example, the channel charge injection is one of the major distortion sources in low distortion switched capacitor filters [5.4] and it is one of the major bottle necks for high-speed high-resolution MOS A/D converters [5.5]. Hence emphasis has been placed in the channel charge injection at the turn-off transient in the following discussions.

In 5.3.1, the charge conservation property of this work is demonstrated. In 5.3.2, the channel charge partitioning schemes of QS models are discussed, and the comparison has been made on the channel charge partitioning ratios for the turn-on and the turn-off transients of a NMOSFET predicted by this work and QS models. In 5.3.3, this work has been compared with the multiple-lumped model. In 5.3.4 and 5.3.5, the waveforms of node currents and node charges for the turn-on and turn-off transients of a NMOSFET are simulated using this work, BSIM [5.11], PISCES [5.12] and 1-D numerical simulation and the results are compared. In 5.3.6, CMOS inverter chain and CMOS ring

oscillator are simulated and voltage waveforms are shown. In 5.3.7, the channel charge injection of a feedback pass tr. has been simulated. In 5.3.8, a differential sample-hold circuit of MOS A/D converter has been simulated and detailed operations have been studied.

5.3.1. Charge conservation

Since this work is a charge-based model, it guarantees the charge conservation [5.2] [5.3]. To demonstrate the charge conservation property of this work, the turn-off transient of a pass transistor has been simulated using this work, a QS charge based model, and the Meyer capacitance model [5.10]. A train of pulses is applied at gate. As shown in Fig.5.1.(a), charge based models such as this work and QS 0/100 model predict correctly that the output voltage returns to 0 after each clock cycle, while the Meyer capacitance model shows incorrect results due to charge-non-conservation. The SPICE3 input file for this simulation is shown in Fig.5.1.(b). For the QS 0/100 model, BSIM[5.11] in SPICE3 is used.

5.3.2. Channel charge partitioning ratio

In the QS(quasistatic) charge based models such as those shown in [5.2], [5.3] and [5.11], the total channel charge must be partitioned into drain and source charge components respectively, in order to find the transient or AC drain, source current components.

Many channel charge partitioning schemes are reported.

Oh, Ward and Dutton derived analytic equations for channel charge partitioning from the current continuity equation. The derivation steps for this scheme are shown in Appendix.10 and equations for the partitioning are shown in (A10.7) and (A10.12). But quasistatic profile is used for $Q'_n(y,t)$ in (A10.7) and (A10.12). If we neglect the diffusion current components, the quasistatic profile of $Q'_n(y,t)$ in saturation region becomes $Q'_n(0,t) \cdot \sqrt{1-y/L}$. Substituting the above profile into (A10.7) and (A10.12),

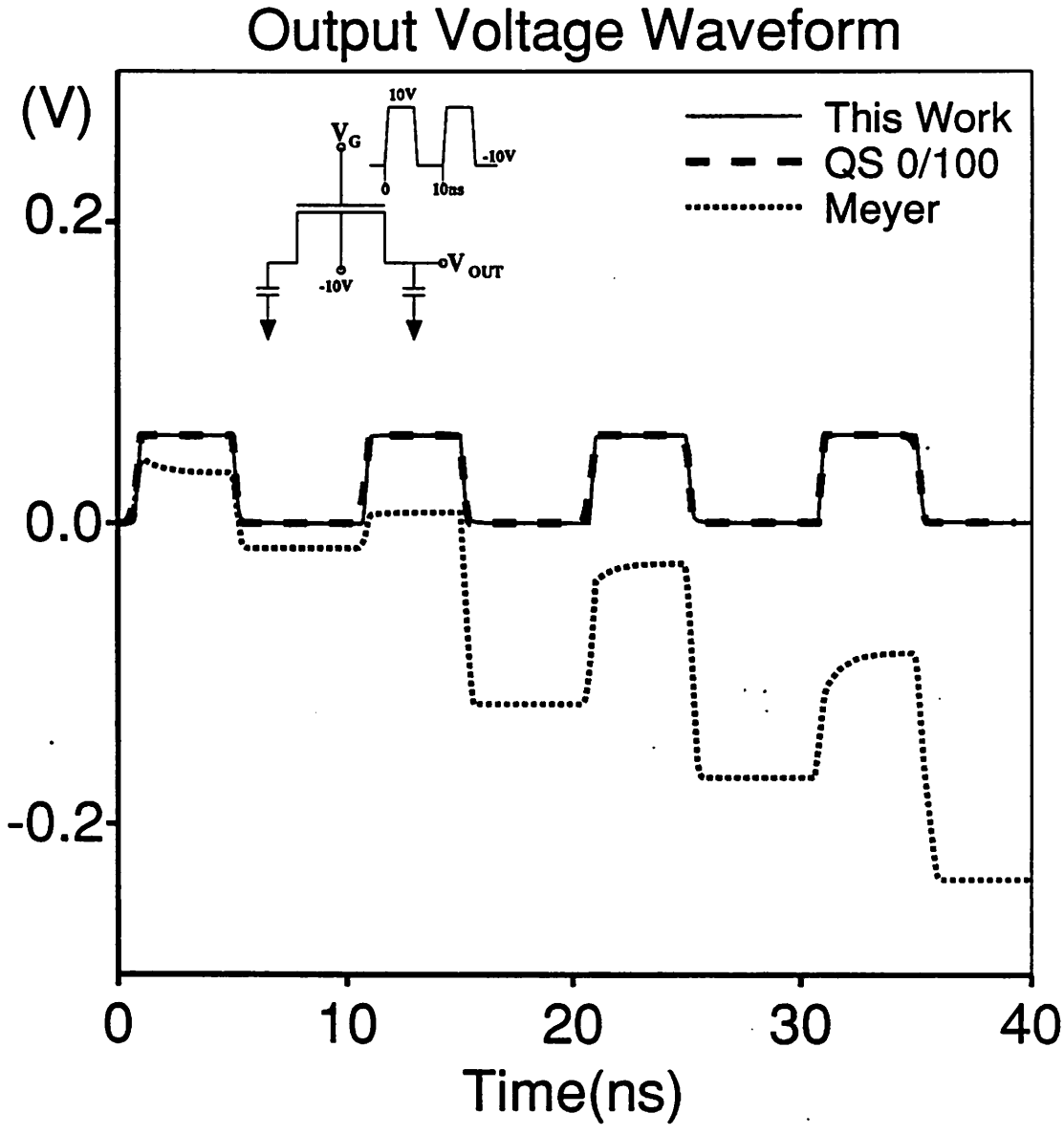


Fig.5.1.(a). The output voltage waveform for the turn-off transient of a NMOS pass-transistor, to demonstrate the charge conservation property of this work.

```
< SPICE3 Input File to show Charge Conservation >
the turnoff transient of pass transistor
m1 1 2 3 4 nch w=20u l=3u
Vg 2 0 dc -10 pulse -10 10 0 1n 1n 4n 10n
Vb 4 0 dc -10
Cload 3 0 2pF
Csrc  1 0 2pF
.tran 100p 40n
.print tran v(3)
.ic v(3)=0 v(1)=0
*
*      Model Parameters for This Work
************************************************
.model  nch  nmos  level=6 vfb=-1.17 gamma=1.36
+  phi=0.75  u0=541  js=1e-15  js=0  tox=0.05
+  nsub=2.5e16  cgdo=0  cgbo=0 cgso=0
+  qtrtol=10  mintol=1
*
*    Model Parameters for BSIM 0/100
************************************************
*.model nch nmos level=4 xpart=1 vfb=-1.17
*+ phi=0.96 k1=1.36 muz=541 mus=545 tox=0.050
*+ vdd=5 n0=1 cgdo=0 cgso=0 cgbo=0 cj=0 cjsw=0
*+ js=1e-15 pb=0.7 pbsw=0.7
*
* Model Parameters for Meyer capacitance model
************************************************
*.model nch nmos vto=0.9 uo=500 gamma=1.36
*+ phi=0.75 rsh=0 level=1 cj=0 cjsw=0 js=0
*+ tox=50n nsub=2.5e16 ld=0 cgdo=0 cgso=0 cgbo=0
.end
```

Fig.5.1.(b). The SPICE3 input file for the turn-off transient of a pass-tr., shown in Fig.5.1.(a).

we can get the 40/60 partitioning between the drain charge Q_D and the source charge Q_S in saturation region, as follows.

$$Q_D(40/60,\ saturation\ region) = -\frac{4}{15}\cdot WLC_{OX}\cdot(V_{GS} - V_{TH}) \tag{5.69}$$

$$Q_S(40/60,\ saturation\ region) = -\frac{2}{5}\cdot WLC_{OX}\cdot(V_{GS} - V_{TH}) \tag{5.70}$$

In linear region, we can get Q_D and Q_S from (A10.7) and (A10.12) by substituting (5.71) into (A10.7) and (A10.12).

$$Q'_n(y,t) = \sqrt{Q'_n(0,t)^2 - (Q'_n(0,t)^2 - Q'_n(L,t)^2)\cdot(y/L)} \tag{5.71}$$

In QS 0/100 model, which was reported by Yang and Chatterjee [5.3] and is available also in BSIM (Berkeley Short Channel IGFET Model) [5.11], all the channel charge is assigned to Q_S in saturation region. This assumption is rather arbitrary and has no physical basis. Q_D and Q_S in saturation region for QS 0/100 model can be written as

$$Q_D(0/100,\ saturation\ region) = 0 \tag{5.72}$$

$$Q_S(0/100,\ saturation\ region) = -\frac{2}{3}\cdot WLC_{OX}\cdot(V_{GS} - V_{TH}) \tag{5.73}$$

Since Q_D=0 in saturation region for QS 0/100 model, the drain current has no displacement current component in saturation region. In linear region, Q_D and Q_S are set to satisfy the following three conditions.

(1). $Q_D = Q_S = -0.5\cdot WLC_{OX}\cdot(V_{GS} - V_{TH})$ when V_{DS} =0

(2). $Q_D = 0$ and $Q_S = -2/3\cdot WLC_{OX}\cdot(V_{GS} - V_{TH})$ when $V_{DS} = V_{DSSAT}$

(3). Both $\partial Q_D/\partial V_{DS}$ and $\partial Q_S/\partial V_{DS}$ are continuous at $V_{DS}=V_{DSSAT}$

In QS 50/50 model which is useful for some charge injection problem, Q_D and Q_S have the same value under all operating conditions. In saturation region, Q_D and Q_S become

$$Q_D(50/50,\ saturation\ region) = -\frac{1}{3}\cdot WLC_{OX}\cdot(V_{GS} - V_{TH}) \tag{5.74}$$

$$Q_S(50/50,\ saturation\ region) = -\frac{1}{3}\cdot WLC_{OX}\cdot(V_{GS} - V_{TH}) \tag{5.75}$$

In linear region, Q_D and Q_S of QS 50/50 model can be found from $Q_D = Q_S = -0.5 * (Q_G + Q_B)$ [5.11].

In BSIM [5.11], all three partitioning schemes are available.

In SPICE level-2 charge based model [5.13], a model parameter 'XQC' determines the channel charge partitioning ratio in saturation region. Hence Q_D and Q_S in saturation region become

$$Q_D(level\text{–}2,\ saturation\ region) = -XQC \cdot \frac{2}{3} \cdot WLC_{OX} \cdot (V_{GS} - V_{TH}) \tag{5.76}$$

$$Q_S(level\text{–}2,\ saturation\ region) = -(1 - XQC) \cdot \frac{2}{3} \cdot WLC_{OX} \cdot (V_{GS} - V_{TH}) \tag{5.77}$$

Hence 'XQC' represents the ratio, $Q_D/(Q_D+Q_S)$ in saturation region. 'XQC' can be chosen to be any value between 0 and 0.5. In linear region, Q_D and Q_S are formulated such that it guarantees that the resulting equations satisfy (5.76) and (5.77) when $V_{DS}=V_{DSSAT}$ and that they satisfy $Q_D = Q_S = -0.5 \cdot WLC_{OX} \cdot (V_{GS} - V_{TH})$ when $V_{DS} = 0$. Hence, Q_D and Q_S in linear region for the level-2 charge based model, are written as follows [5.14].

$$Q_D = XQC \cdot Q_{N.SAT} + (1.5 - 2 \cdot XQC) \cdot (Q_N - Q_{N.SAT}) \tag{5.78}$$

$$Q_S = (1 - XQC) \cdot Q_{N.SAT} + (2 \cdot XQC - 0.5) \cdot (Q_N - Q_{N.SAT}) \tag{5.79}$$

$$where \quad Q_N = -(Q_G + Q_B) \tag{5.80}$$

$$Q_{N.SAT} = -\frac{2}{3} \cdot WLC_{OX} \cdot (V_{GS} - V_{TH}) \tag{5.81}$$

Q_N is the total channel charge and $Q_{N.SAT}$ is the value of Q_N in saturation region. In saturation region, since $Q_N = Q_{N.SAT}$, (5.78) and (5.79) are reduced to (5.76) and (5.77) respectively. When $V_{DS} = 0$, since $Q_N = - WLC_{OX} \cdot (V_{GS} - V_{TH})$, we can see that $Q_D = Q_S = -0.5 \cdot WLC_{OX} \cdot (V_{GS} - V_{TH})$. Hence (5.78) and (5.79) satisfy the above mentioned two requirements.

(XQC=0) is conceptually equivalent to QS 0/100 model, (XQC=0.4) is conceptually equivalent to QS 40/60 model, and (XQC=0.5) is conceptually equivalent to QS 50/50 model. But due to the differences in implementation details, they are not exactly the same.

As explained in the preceding discussions, there are many channel charge partitioning schemes, such as, QS 40/60, QS 0/100, QS 50/50 and XQC. But none of these partitioning schemes are valid under all operating conditions. A designer must choose the best partitioning scheme for each circuit. And this is tedious and error prone. In addition, the channel charge partitioning ratio is not constant in saturation region but changes with time, as will be shown in the following discussions.

Fig.5.2.(a) shows the circuit schematic to investigate the channel charge partitioning ratio at the turn-on and turn-off transient of a NMOSFET. The channel length of the MOSFET is 10 μm.

<Turn-On Transient>

For the turn-on transient, the gate voltage V_G changes from 0V(at t=0) to 5V(at t=Tr) and remains at 5V after t=Tr as shown in Fig.5.2.(a). The MOSFET is initially in cutoff region. Since the drain voltage VD is 2V, the MOSFET enters saturation region and then linear region, as time goes on.

Fig.5.2.(b) shows the channel charge partitioning ratio ($Q_D/(Q_D+Q_S)$) at the turn-on transient with respect to the normalized time t/Tr. All the QS(quasistatic) models (QS 40/60, QS 0/100, QS 50/50) give the channel charge partitioning ratio which are independent of the rise time Tr and are functions of node voltages only.

QS 0/100 model (dashed line) shows that the partitioning ratio is 0 in saturation region and increases in linear region. QS 40/60 model (dotted line) shows that the partitioning ratio is 0.4 in saturation region and increases in linear region. QS 50/50 model (dash dot dot dot) shows that the partitioning ratio is 0.5 throughout all the time.

But this work and 1-D numerical simulation show that the partitioning ratio depends on Tr(rise time) as well as on node voltages.

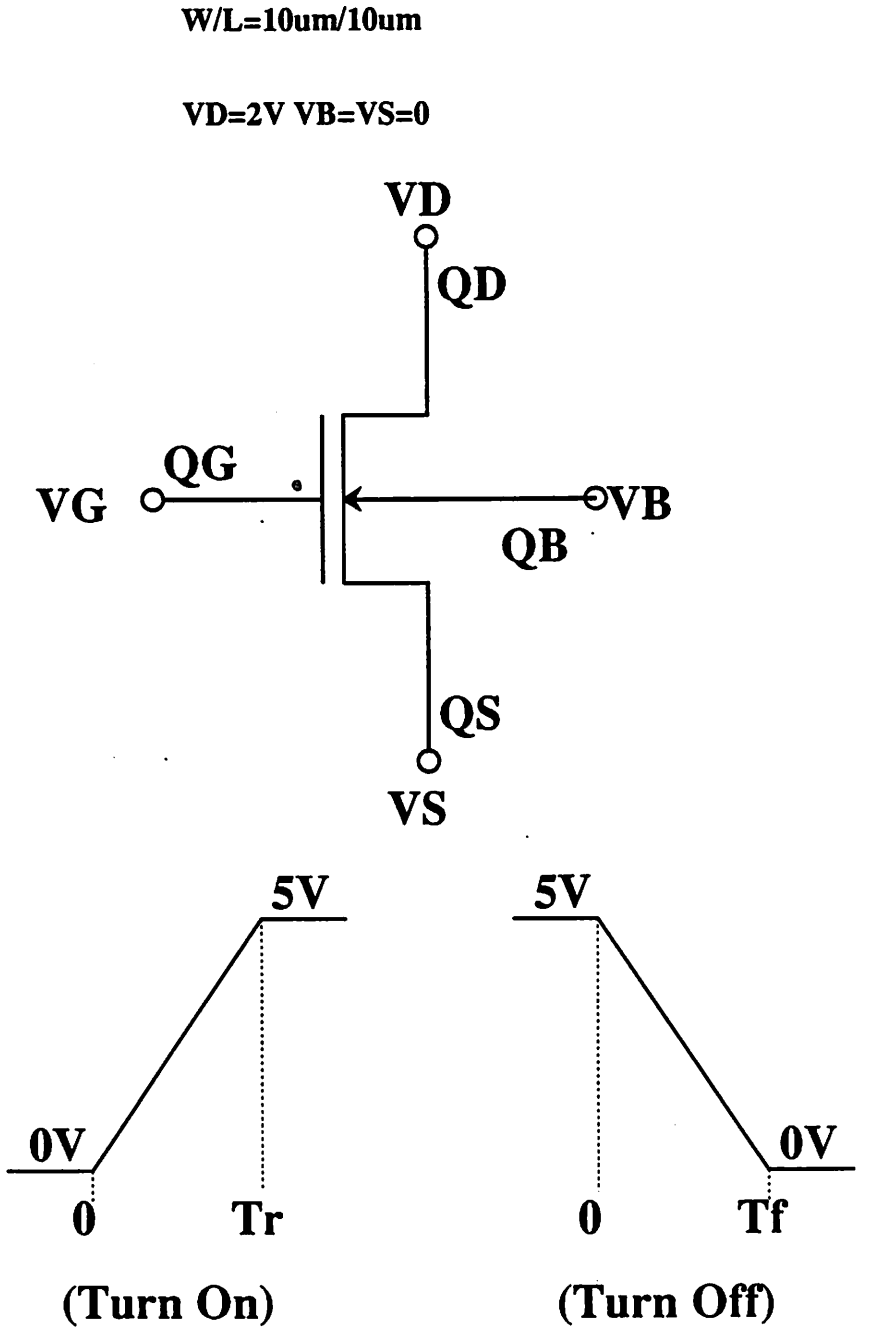


Fig.5.2.(a). The circuit diagram to investigate the channel charge partitioning ratio at the turn-on and the turn-off transients of a NMOSFET.

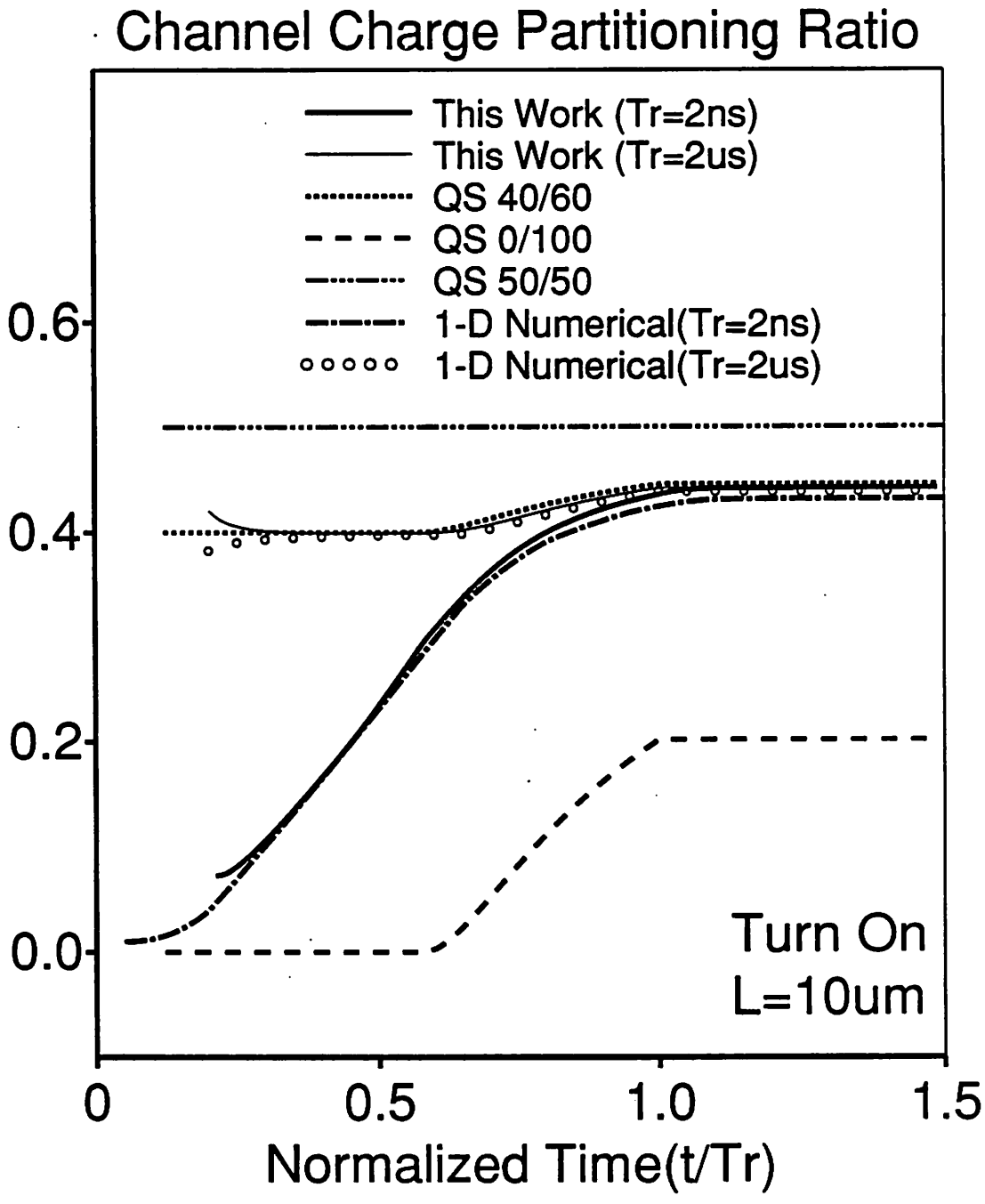


Fig.5.2.(b). The channel charge partitioning ratio at the turn-on transient of the NMOS-FET shown in Fig.5.2.(a).

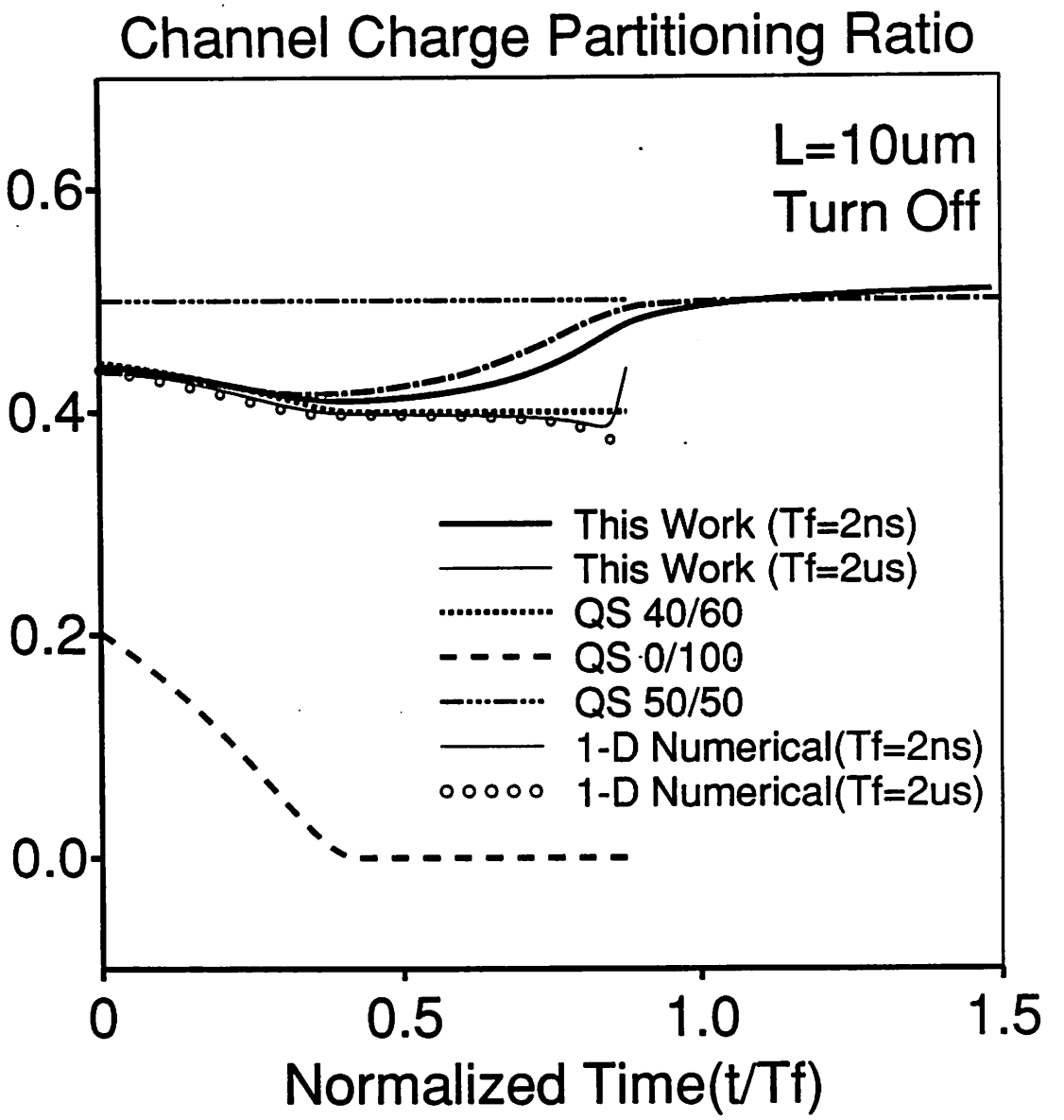


Fig.5.2.(c). The channel charge partitioning ratio at the turn-off transient of the NMOS-FET shown in Fig.5.2.(a).

For the fast turn-on (Tr=2*ns*), this work(thick solid line) and 1-D numerical simulation (dot dash) show that the partitioning ratio is close to 0/100 at the initial stage of turn-on and approaches 40/60 as the MOSFET approaches the steady state.

For the slow turn-on (Tr=2μs), this work (thin solid line) and 1-D numerical simulation (circles) give the partitioning ratios which are almost the same as 40/60. Hence we can see that this work and 1-D numerical model are reduced to QS 40/60 model for very slow turn-on transient compared to the channel transit time.

In cutoff region, the partitioning ratios are not shown, since the values of drain and source charges are so small that the division of too small quantities gives meaningless error.

<Turn-Off Transient>

For the turn-off transient, the gate voltage V_G ramps from 5V(at t=0) to 0V(at t=Tf) and remains at 0V after t = Tf as shown in Fig.5.2.(a). Since the drain voltage *VD* is 2V, the MOSFET is initially in linear region and enters saturation region and then cutoff region as time goes on. Fig.5.2.(c) shows the channel charge partitioning ratio at the turn-off transient with respect to the normalized time t/Tf. Again all the QS(quasistatic) models give the partitioning ratios which are independent of Tf and are functions of node voltages only. But this work and 1-D numerical simulation show that the partitioning ratio depends on the fall time Tf as well as on node voltages.

For the fast turn-off (Tf=2ns), this work(thick solid line) and 1-D numerical simulation (dot dash) show that the partitioning ratio is the same as 40/60 in linear region and that the ratio approaches 50/50 as the MOSFET enters saturation region and then cutoff region.

For the slow turn-off (Tf=2us), this work(thin solid line) and 1-D numerical simulation (circles) show that the partitioning ratio is almost the same as 40/60 throughout all the time. Again we can verify that this work and 1-D numerical simulation are reduced to QS 40/60 model for the slow turn-off transient compared to channel transit time.

The comparison of Fig.5.2.(b) and Fig.5.2.(c) shows that this work and 1-D numerical simulation give the partitioning ratios which are dependent on history in addition to signal transition rate (Tf, Tr), and node voltages. But all the QS models give the partitioning ratios which are functions of node voltages only.

5.3.3. Comparison with multiple-lump model

This work has been compared with the QS(quasi-static) multiple-lump model where each MOSFET is decomposed into N MOSFET's in series with channel length of 1/N times the original channel length, and the resulting circuit is simulated using QS(quasistatic) models in SPICE.

Fig.5.3.(a) shows the output voltage waveforms at the turn-off transient of a NMOS pass transistor, predicted by this work and the QS multiple-lump model. As N (Number of lumped elements) increases, this QS multiple-lump model approaches this work. The QS 8-lump model gives almost the same waveform as the QS 4-lump model. Overlap and junction capacitances are not included in the simulation to concentrate on the intrinsic phenomenon. Table.5.2 shows the comparison of run statistics for this example.

	This Work	1-Lump	2-Lump	4-Lump	8-Lump
Total Run Time(sec)	3.1	0.7	1.1	3.1	837
Number of Iterations	591	145	176	422	85546
Total Time Points	277	70	65	100	14617
Rejected Time Points	5	3	0	11	3645
Load Time(sec)	1.9	0.2	0.5	2.0	613

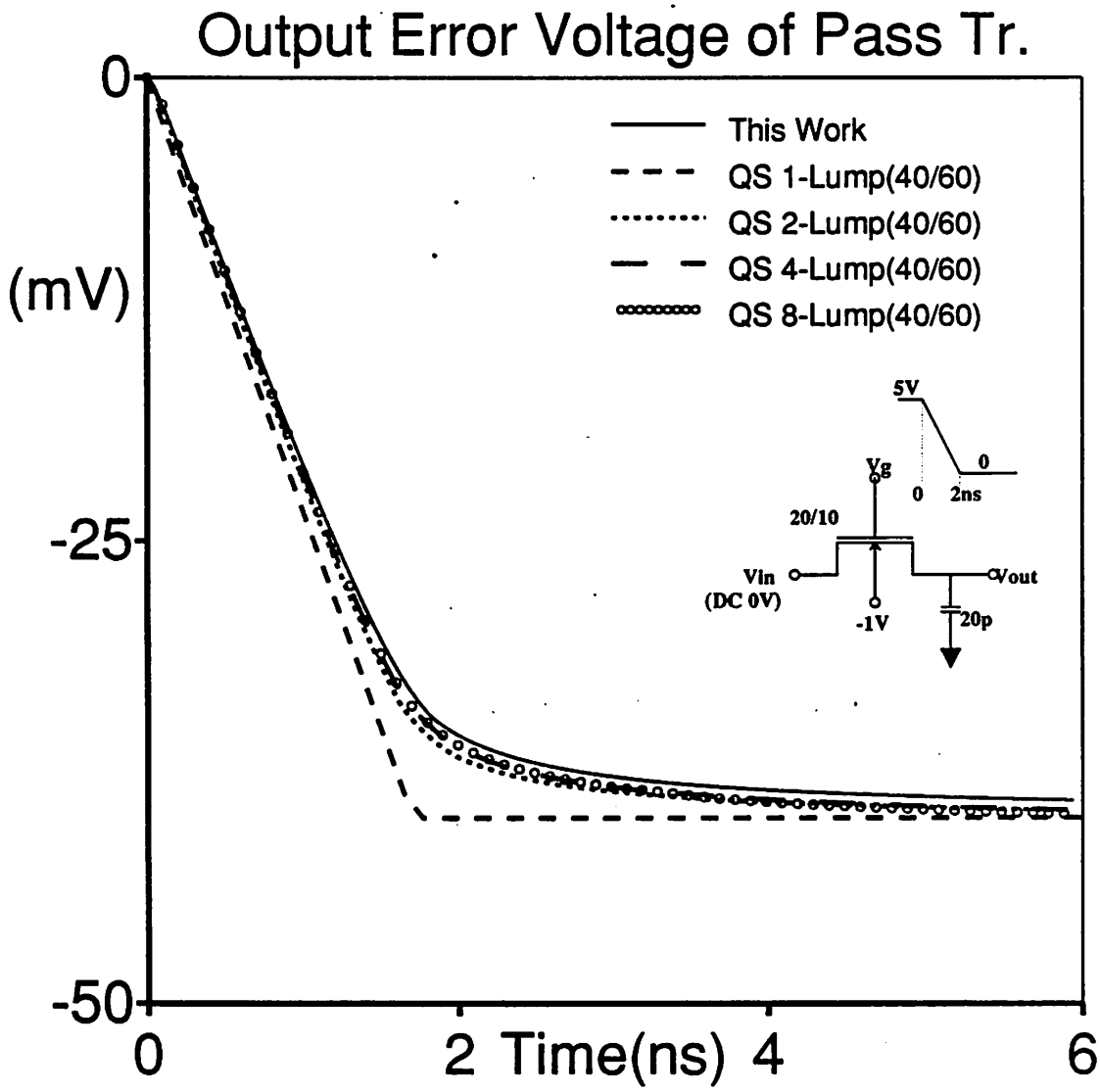


Fig.5.3.(a). The output voltage waveform for the turn-off transient of a NMOS pass-transistor, predicted by this work and QS multiple-lump models.

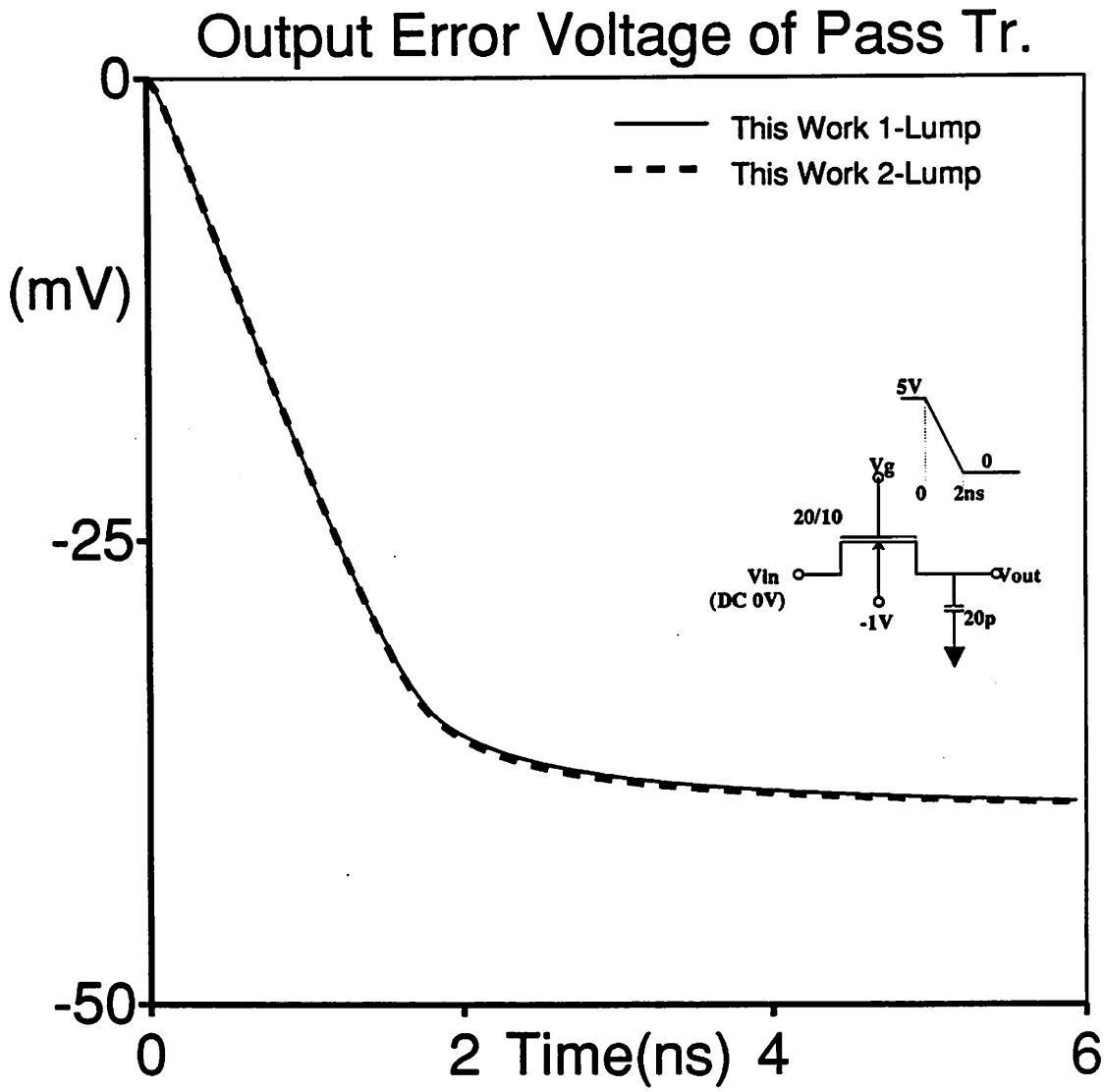


Fig.5.3.(b). The output voltage waveform for the turn-off transient of a NMOS pass-transistor, predicted by NQS multiple-lump models. The NQS 1-lump model refers to this work.

<Input file for this work>

```
turn off transient of NMOS pass tr.
m1 3 2 1 4 nch w=20u l=10u
vg  2 0 dc 5 pwl 0 5 2n 0 1 0
Vin 1 0 dc 0
Vb  4 0 dc -1
Cl 33 0 20pF
vc 33 3 dc 0
.ic v(3)=0 v(1)=0
.tran 0.01n 6n
.print tran v(3)
.model nch nmos level=6 vfb=-0.77
+ tox=0.018 nsub=2e16 u0=500
+ lambda=0.03 pb=0.7 pbsw=0.7
+ js=1e-4 cgdo=0 cgso=0
+ mintol=0.1 qtrtol=10
.end
```

<Input file for QS 2-lump model>

```
turn off transient of NMOS pass tr.
m1 5 2 1 4 nch w=20u l=5u
m2 3 2 5 4 nch w=20u l=5u
vg  2 0 dc 5 pwl 0 5 2n 0 1 0
Vin 1 0 dc 0
Vb  4 0 dc -1
Cl 33 0 20pF
vc 33 3 dc 0
.ic v(3)=0 v(5)=0
.tran 0.01n 6n
.print tran v(3)
.model nch nmos level=4
+ xpart=0 vfb=-0.77 phi=0.956
+ k1=0.425 muz=500 mus=520
+ tox=0.018 vdd=5 n0=1 cgdo=0 cgso=0
+ cj=0 cjsw=0 js=1e-15 pb=0.7 pbsw=0.7
.end
```

Fig.5.3.(c). The SPICE3 input files for the turn-off transient of a pass-tr. in Fig.5.3.(b) and Fig.5.3.(c), using this work and the QS 2-lump model.

Table.5.2. Comparison of run statistics for the turn-off transient of the pass-tr. shown in Fig.5.3.(a), between this work and QS multiple-lump models.

This work takes about the same time as the QS 4-lump model. As N (the number of lump elements) increases, the CPU time takes longer and it becomes more difficult to get convergence as partially reflected in the number of 'Rejected Time Points' in Table.5.2.

Fig.5.3.(b) shows the output voltage waveform of the NQS multiple-lump model where each MOSFET is decomposed into N MOSFETs in series, and the resulting circuit is simulated using this work(NQS model). The output voltage waveform of the NQS 2-lump model is almost the same as that of the NQS 1-lump model(this work). This clearly demonstrates that this work is non-quasi-static.

Fig.5.3.(c) shows the SPICE3 input file for this work and the QS 2-lump model. BSIM in SPICE3 is used for QS multiple-lump models.

5.3.4. Turn-on transient of a NMOSFET

Fig.5.4.(a) shows the circuit diagram to evaluate the turn-on transient of a NMOSFET. A rising ramp voltage is applied at gate and V_D=2V, V_B=V_S=0. Initially the MOSFET is in cutoff region and enters saturation region and then linear region as time goes on. This turn-on transient has been simulated using this work, QS charge based models, PISCES(2-D device simulation) [5.12] and 1-D numerical solution of current continuity equation. For QS charge based models, BSIM in SPICE3 is used. The ratios 40/60, 0/100 and 50/50 represent the ratios between drain charge and source charge in saturation region, as shown in Section 5.3.2. The scheme for 1-D numerical solution of current continuity equation is shown in Appendix.13

Fig.5.4.(b) shows the drain current waveform for the turn-on transient. At t=0, all the models predict the same drain current which is slightly negative due to the feed-through current through gate-drain overlap capacitance. At t=0.35ns when V_{GS} becomes V_{TH}, there are sudden negative jumps in drain currents for QS 50/50 and QS 40/60 models.

This sudden negative jump is caused by the assignment of fixed ratio (50% for QS 50/50 and 40% for QS 40/60) of channel charge to drain node without considering the channel transit time effect. But this work, PISCES and 1-D numerical simulation show that the drain current doesn't change with time until t=1.17ns when the inversion carriers which started from source node at the turn-on time (t=0.35ns) reach drain node. The channel transit time for this example is (1.17ns-0.35ns) = 0.82ns. The drain current component of QS 0/100 model in saturation region is purely DC (transport) current component since the drain charge is 0 in saturation region and the displacement current component is 0. This work, PISCES and 1-D numerical solution give good agreements among them throughout all the time considered except a small discontinuity in drain current at t=1.17ns. This small discontinuity of drain current between the moving boundary period and the non-moving boundary period is caused by the adjustment of coefficients $\{A_n\}$ as shown in Section 5.2.3 of this Chapter. By adjusting coefficients $\{A_n\}$, the drain charge is made to be continuous but the drain current cannot be made to be continuous between two time periods. This discontinuity in drain current did not cause any serious convergence problems in circuit simulations as shown in the subsequent simulation examples. Related to this discontinuity, one must note that there is a big discontinuity in drain current at the turn-on time(t=0.35ns) for QS 40/60 and QS 50/50 models as shown in Fig.5.4.(b), and yet it works fine in real circuit simulations.

For the time interval after t=2ns, during which the applied bias doesn't change with time but this work, PISCES and 1-D numerical simulation show that the drain current does change with time.

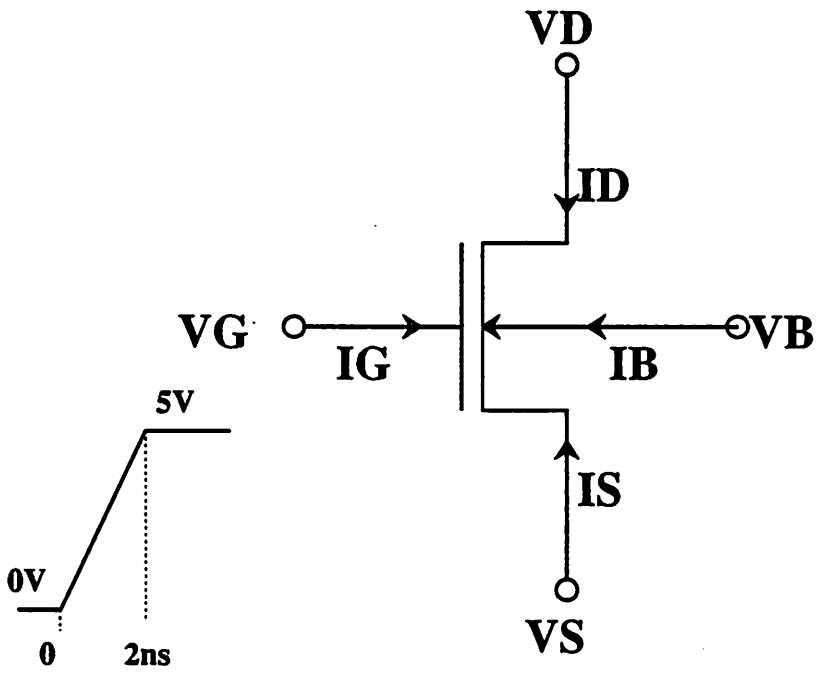
Fig.5.4.(c) shows the intrinsic drain charge waveform normalized by WLC_{OX}. It doesn't include any charge components due to overlap or junction capacitance. This work and 1-D numerical simulation show good agreements between each other throughout all the time considered. This work and QS 40/60 model gives the same drain charge at the steady state (t=3ns) but this work shows some delay in charge build-up during the time interval before t=2ns. At the steady state(t=3ns), the QS 0/100 model gives smaller drain charge than this work and the QS 50/50 model gives larger drain charge than this

work.

Fig.5.4.(d) shows the source current waveform. This work, PISCES and 1-D numerical simulation give good agreements throughout all the time considered. All the QS models give a sudden negative jump at the turn-on time(t=0.35ns). The QS 0/100 model gives the largest jump among the three QS models, because 100% of the channel charge is assigned to the source node in saturation region for QS 0/100 model.

Fig.5.4.(e) shows the intrinsic source charge waveform. Again this work and 1-D numerical solution give good agreements. The steady state source charge at t=3ns is the same for this work, 1-D numerical simulation and QS 40/60 model. But this work and 1-D numerical simulation show the delay in charge-buildup compared to QS 40/60 model. Among QS models, QS 0/100 model gives the largest magnitude of source charge and QS 50/50 model gives the smallest magnitude.

Fig.5.4.(f) shows the gate current waveform. All the QS models give exactly the same gate current waveform irrespective of channel charge partitioning scheme because channel charge partitioning scheme affects only drain and source charges not gate or bulk charge. Although this work and 1-D numerical simulation show good agreements throughout all the time, PISCES gives slightly different value from this work or 1-D simulation at the initial stage of turn-on. This discrepancy is considered to be due to the bulk resistance which is included in PISCES simulation but is not included in this work or 1-D numerical simulation. The effect of bulk resistance will be further discussed related to bulk current shown in Fig.5.4.(h). Since a linear ramp voltage is applied at gate and all other nodes of MOSFET are fixed at DC voltages, the gate current is directly proportional to gate capacitance for $t < 2ns$. Hence we can observe that, for the fast turn-on, the effective loading by the gate capacitance is much smaller than that predicted by QS models especially in saturation region. The discontinuity of the gate current of this work at t=1.17ns is due to the discontinuity of dQ_D/dt between the moving boundary period and the non-moving boundary period, which had been explained related to the drain current in Fig.5.4.(h) and also in Section 5.2.3.



W/L=10um/10um

VD=2V VB=VS=0

Fig.5.4.(a). The circuit schematic for the turn-on transient of a NMOSFET.

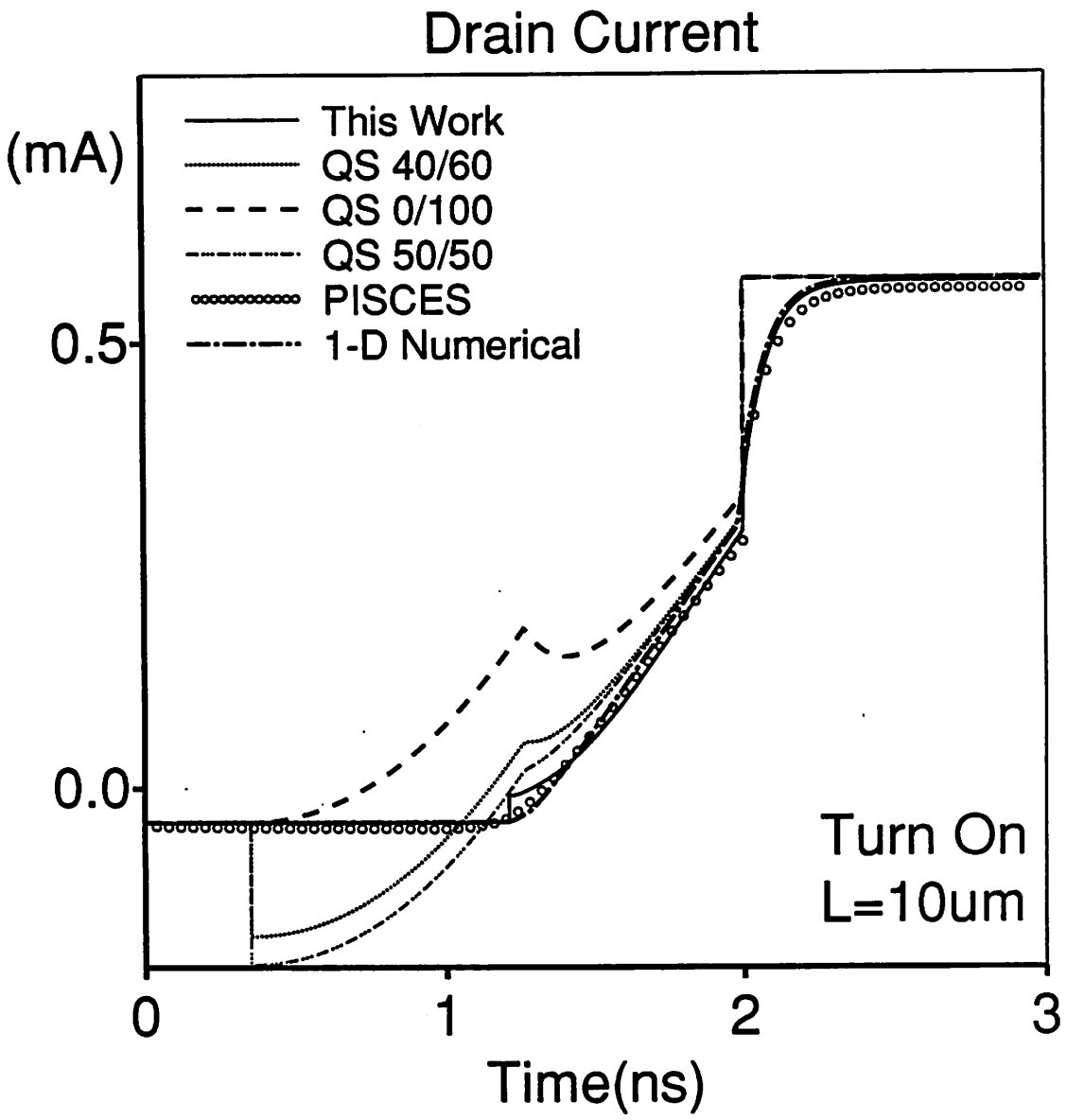


Fig.5.4.(b). The drain current waveform for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

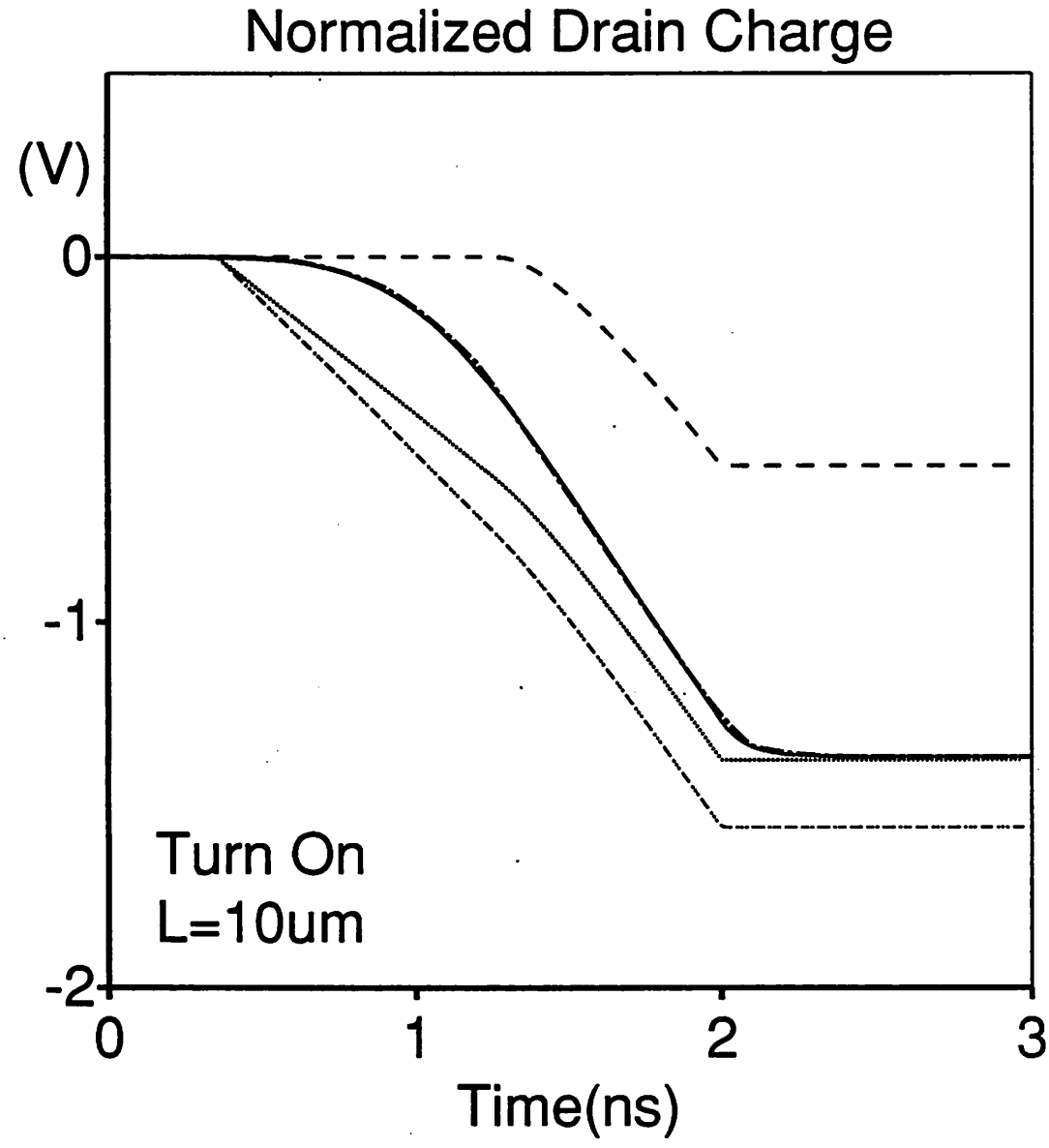


Fig.5.4.(c). The waveform of the intrinsic drain charge normalized by WLCox, for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

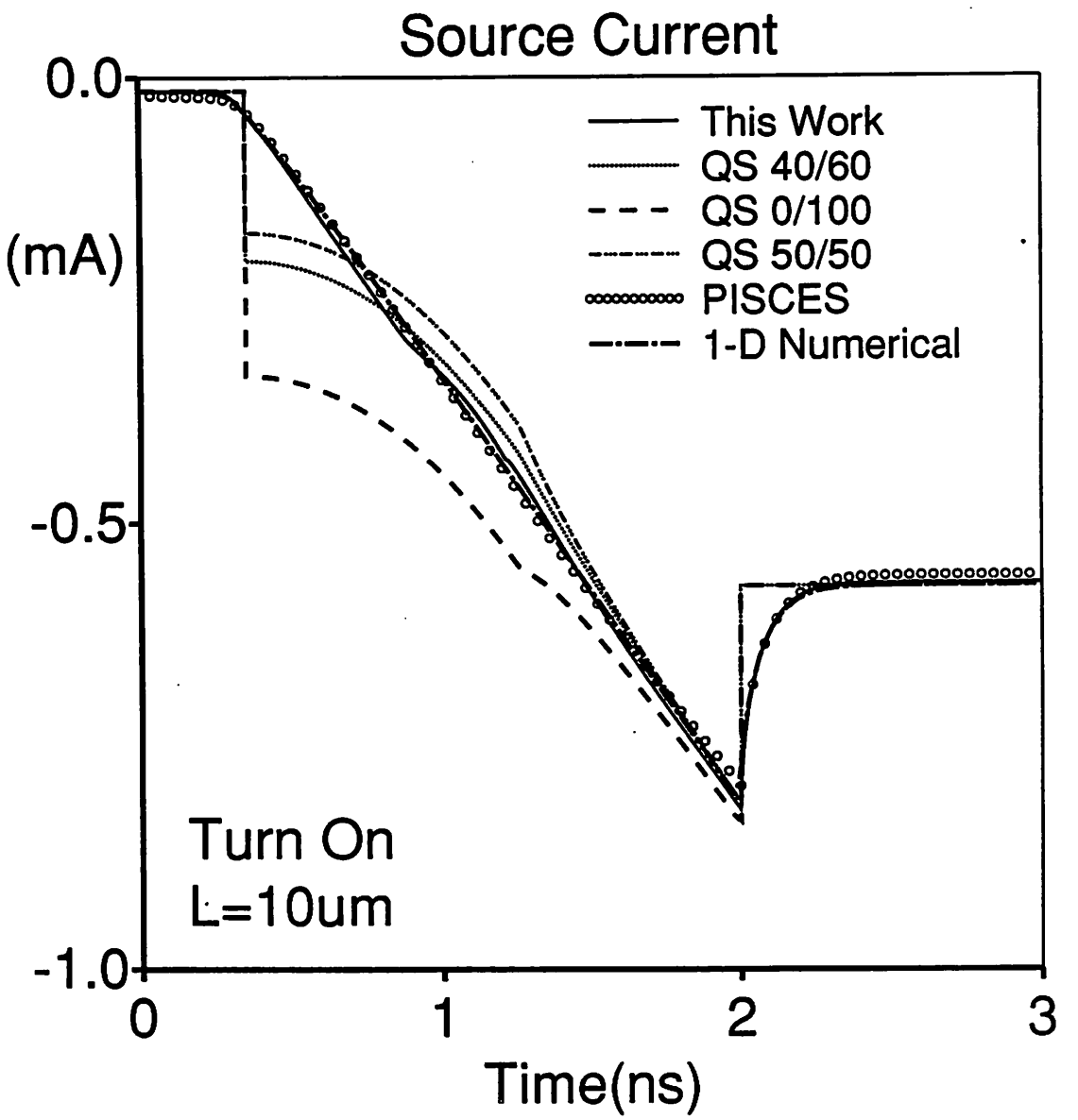


Fig.5.4.(d). The source current waveform for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

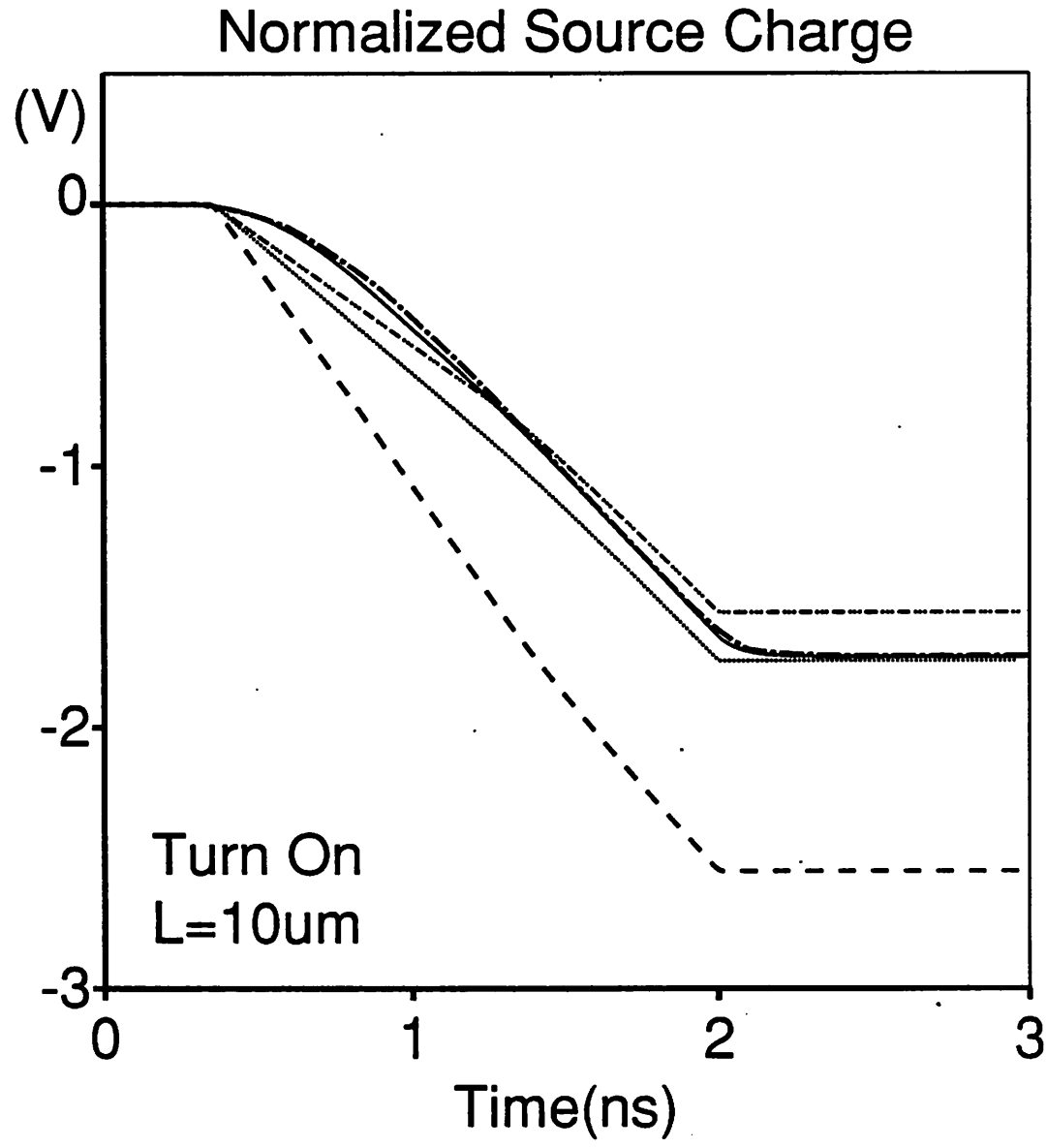


Fig.5.4.(e). The waveform of the intrinsic source charge normalized by WLCox, for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

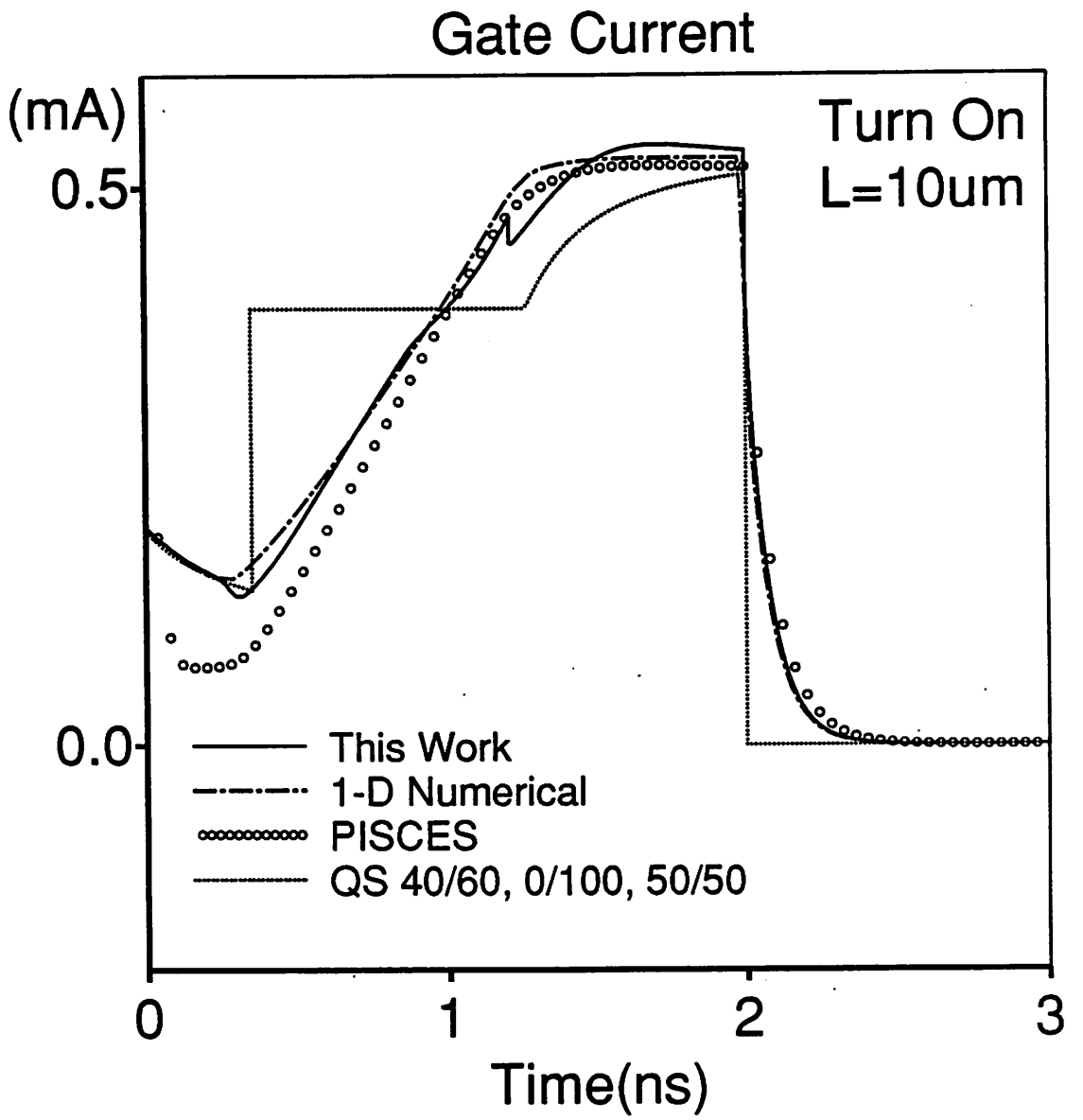


Fig.5.4.(f). The gate current waveform for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

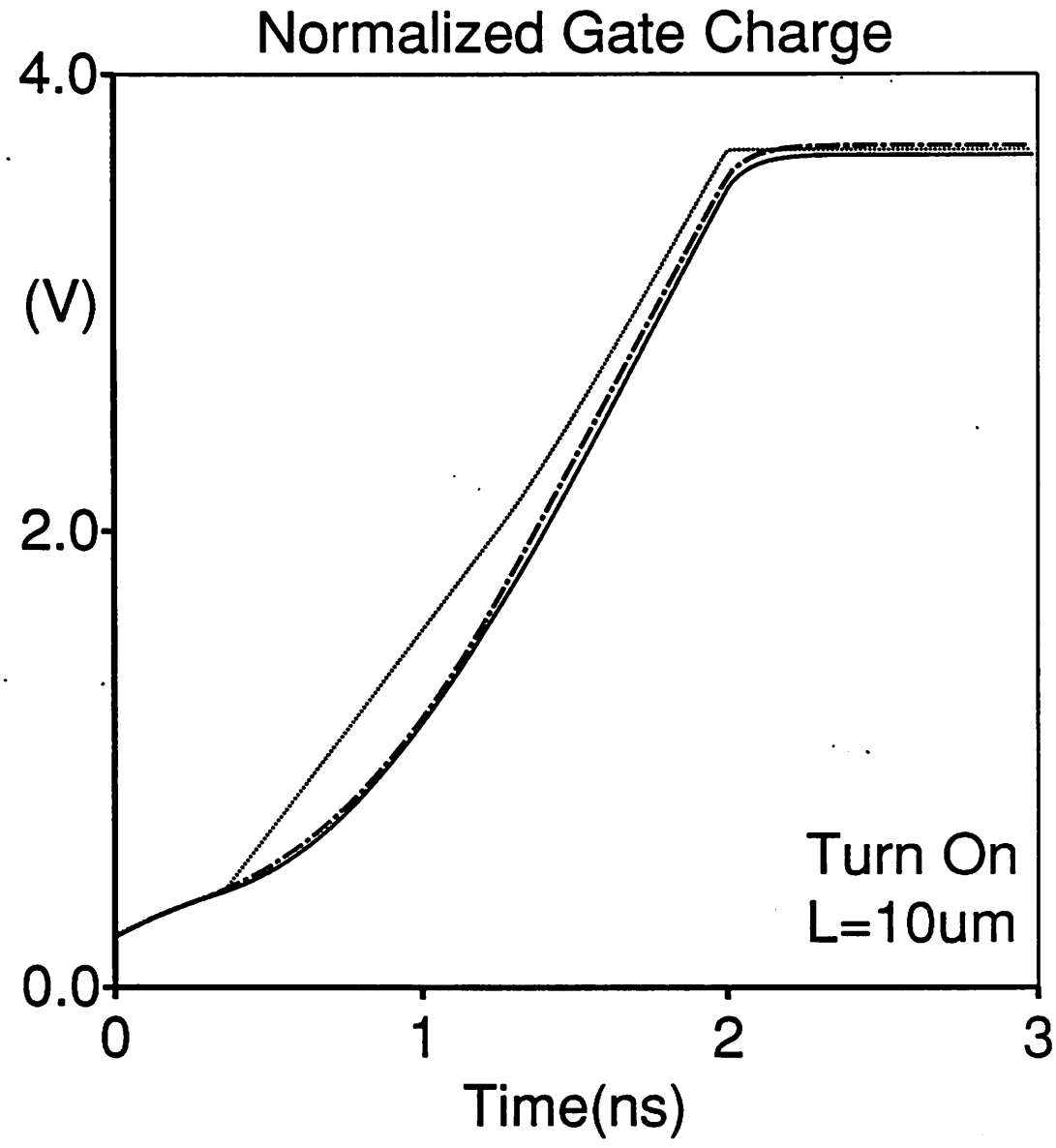


Fig.5.4.(g). The waveform of the intrinsic gate charge normalized by WLCox, for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

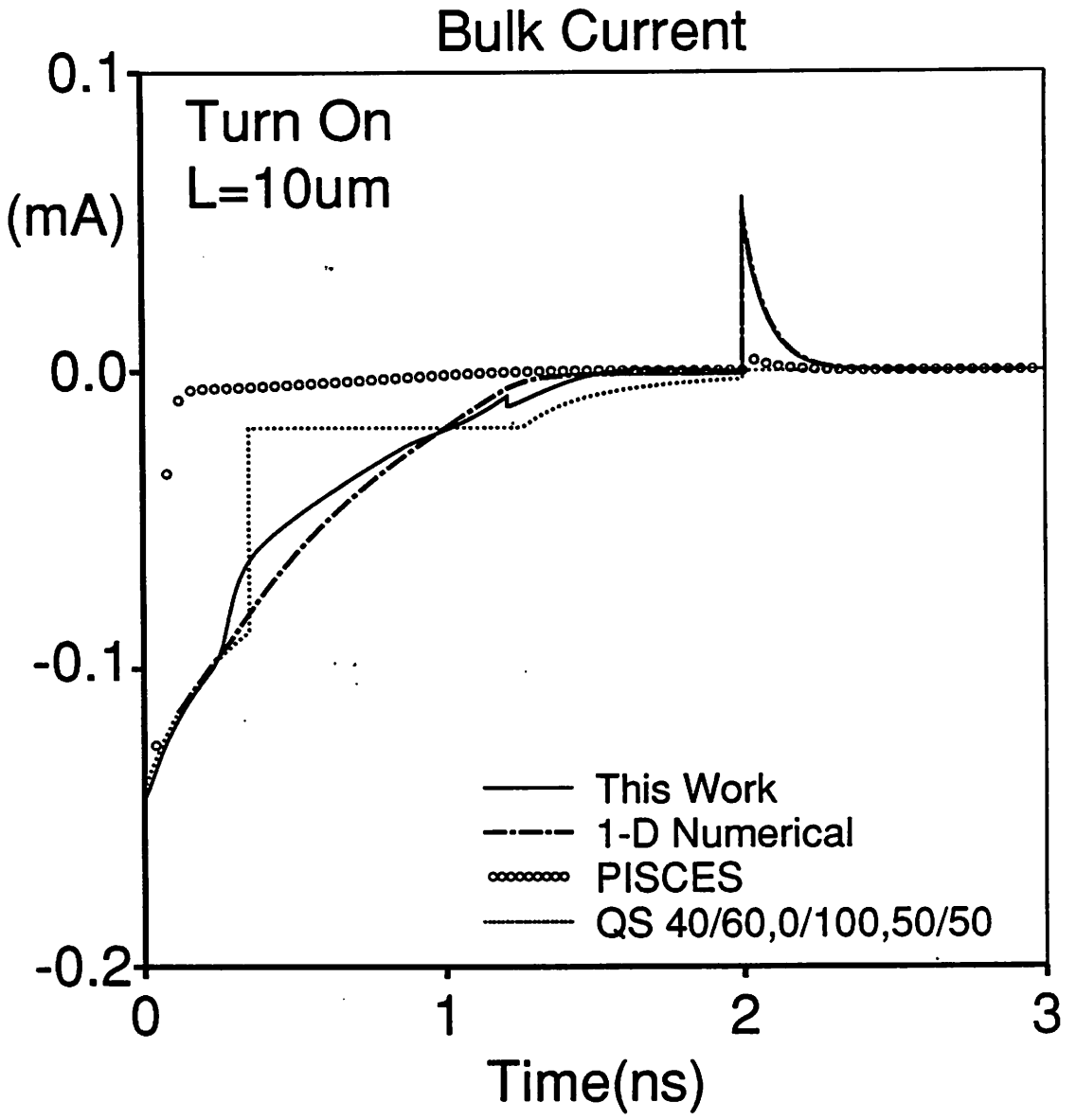


Fig.5.4.(h). The bulk current waveform for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

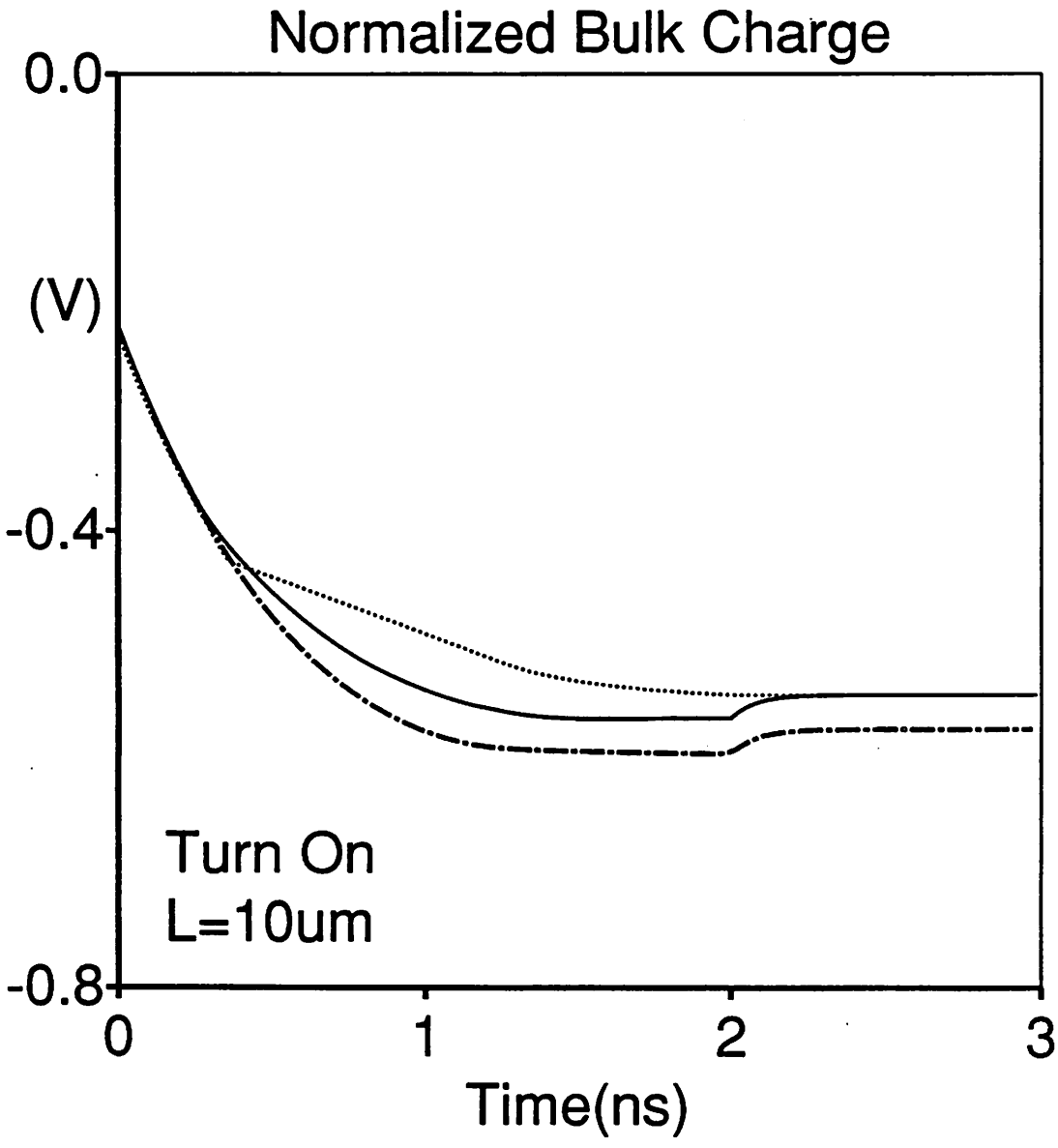


Fig.5.4.(i). The waveform of the intrinsic bulk charge normalized by WLCox, for the turn-on transient of the NMOSFET shown in Fig.5.4.(a).

SPICE3 Input Listings of
Turn-On Transient of an NMOSFET

```
turnon transient of a NMOSFET
m1 1 2 3 4 enh w=10u l=10u
VG 2 0  dc 0 PWL 0 0 2N 5 1 5
VD 1 0 2
VS 3 0 0
VB 4 0 0
.TRAN .01N 3N 0 0.01N
.PRINT TRAN I(VG) I(VD) I(VS) I(VB)
*
* model parameters for NQS model
.model enh nmos level=6 vfb=-0.5
+       tox=0.018 nsub=2.1e16 u0=500
+       lambda=0 js=1e-10 cgdo=1500p
+       cgso=600p mintol=0.1 qtrtol=1
*
* model parameters for BSIM
*    xpart=0 ----> QS 40/60
*    xpart=1 ----> QS 0/100
*    xpart=2 ----> QS 50/50
*.model enh nmos level=4 xpart=0
*+ vfb=-0.5 phi=0.96037 k1=0.435
*+ muz=500 mus=500 tox=0.018 vdd=5
*+ cgdo=1500p cgso=600p
* convergence paramters for BSIM SPICE3
*.OPT ACCT VNTOL=1N RELTOL=0.000002
*+ ABSTOL=1E-15 CHGTOL=1E-18
.end
```

Fig.5.4.(j). SPICE3 input file for the turn-on transient of a NMOSFET shown in Fig.5.4.(a).

Fig.5.4.(g) shows the gate charge waveform. This work and 1-D numerical simulation show the delay in charge build-up compared to QS model.

Fig.5.4.(h) shows the bulk current waveform. PISCES shows a large deviation from this work. This is considered to be due to the bulk resistance and bulk junction capacitance which are considered in PISCES but are not considered in this work, 1-D numerical simulation or the QS model. We can observe the NQS(non-quasistatic) behavior of this work in Fig.5.4.(h).

Fig.5.4.(i) shows the waveform of intrinsic bulk charge normalized by WLC_{OX}.

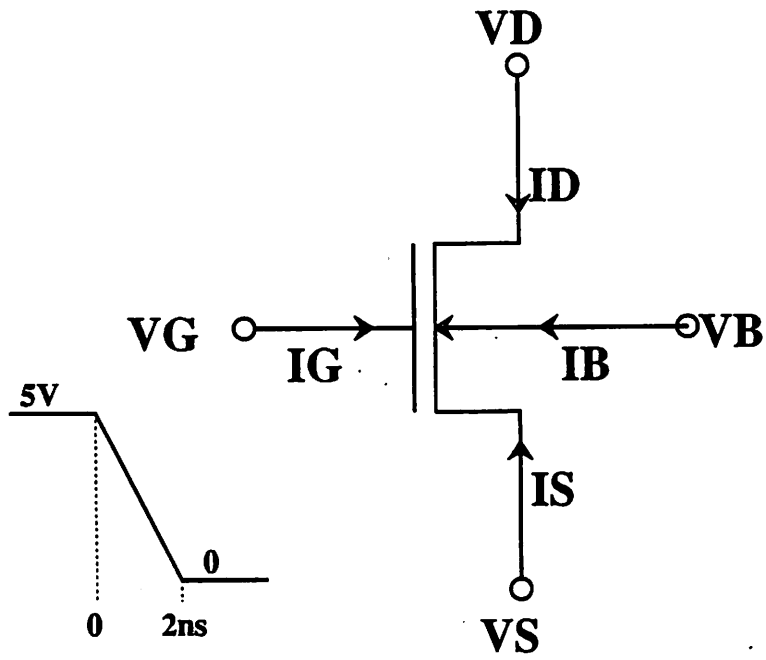
The SPICE3 input file for this turn-on transient is shown in Fig.5.4.(j).

5.3.5. Turn-off transient of a NMOSFET

Fig.5.5.(a) shows the circuit schematic of turn-off transient of an NMOSFET. The gate voltage V_G ramps from 5V at t=0 to 0V at t=2ns and remains at 0V after t=2ns. Except the gate voltage waveform, other parameters are exactly the same as those in Fig.5.4.(a). Since V_D=2V, V_S=0 and V_B=0, the MOSFET goes through linear region, saturation region and then cut off region, as time goes on.

Fig.5.5.(b) shows the drain current waveform. This work, PISCES and 1-D numerical simulation show that the current changes continuously throughout all the time except at t=2ns when there is a discontinuity in the slope of gate voltage and hence the current through the gate-drain overlap capacitance is discontinuous. But all the QS models show discontinuities in drain current at t=0 and t=1.65 ns (turn-off time). Although 1-D numerical simulation shows good agreements with this work, PISCES shows a slight discrepancy from this work. This discrepancy is considered to be due to the effect of drain junction geometry which is not considered in this work or 1-D numerical simulation.

Fig.5.5.(c) shows the waveform of intrinsic drain charge normalized by WLC_{OX}. Comparison of the drain charge waveform for turn-on(Fig.5.4.(c)) and turn-off(Fig.5.5.(c)) during the time interval between t=0 and t=2ns show that charge waveform of QS



W/L=10um/10um

VD=2V VB=VS=0

Fig.5.5.(a). The circuit schematic for the turn-off transient of a NMOSFET.

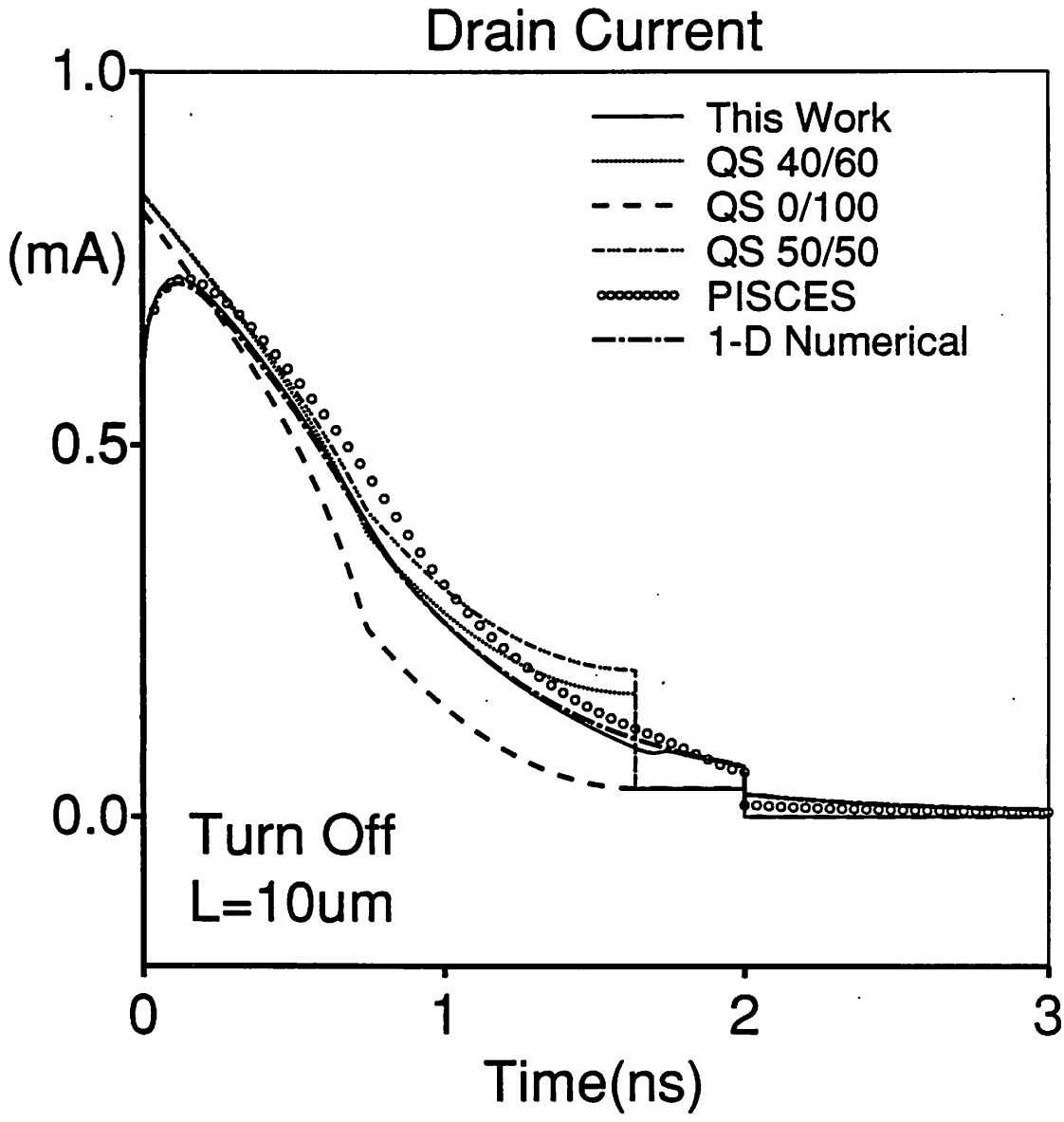


Fig.5.5.(b). The drain current waveform for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

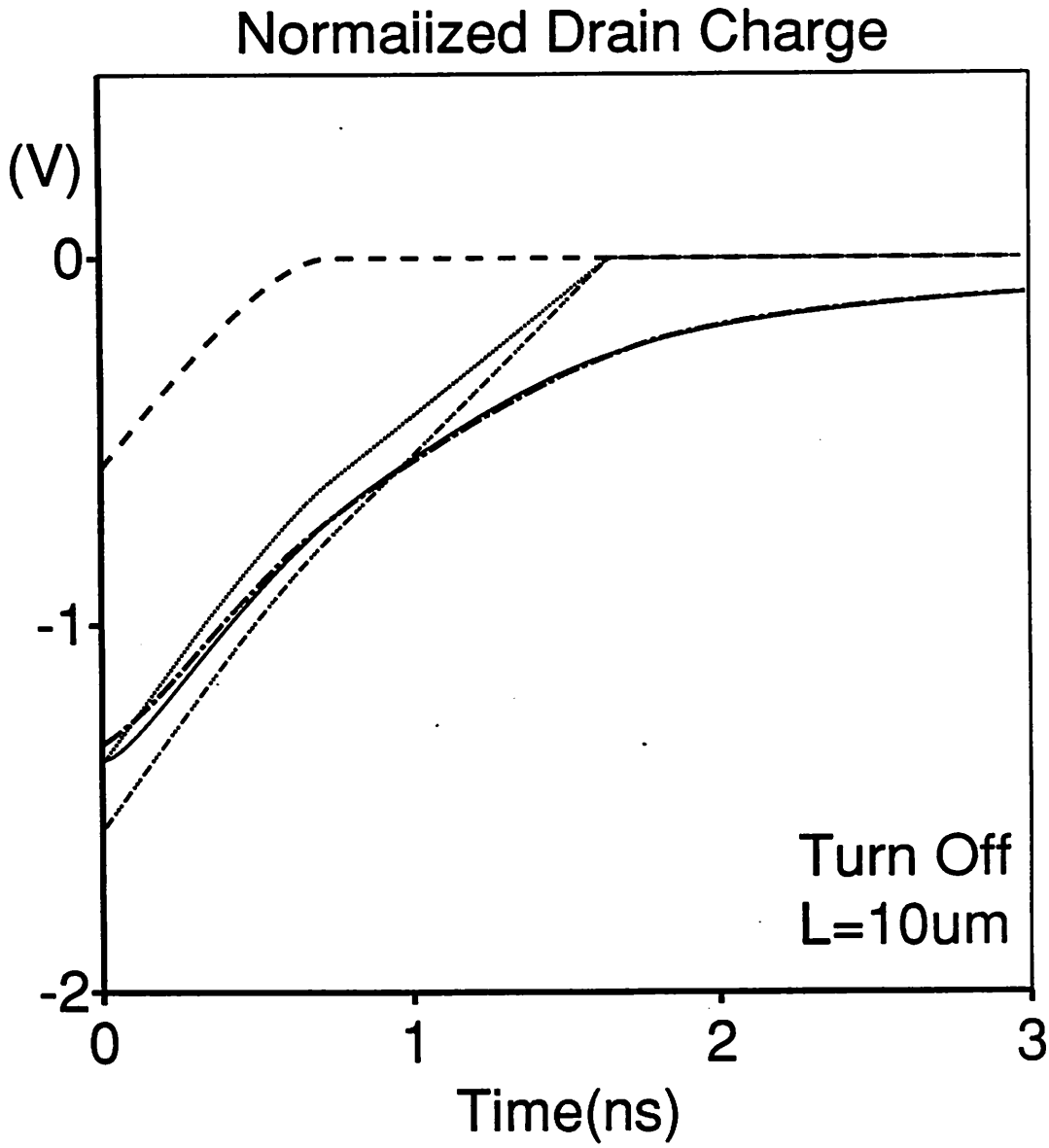


Fig.5.5.(c). The waveform of the intrinsic drain charge normalized by WLCox, for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

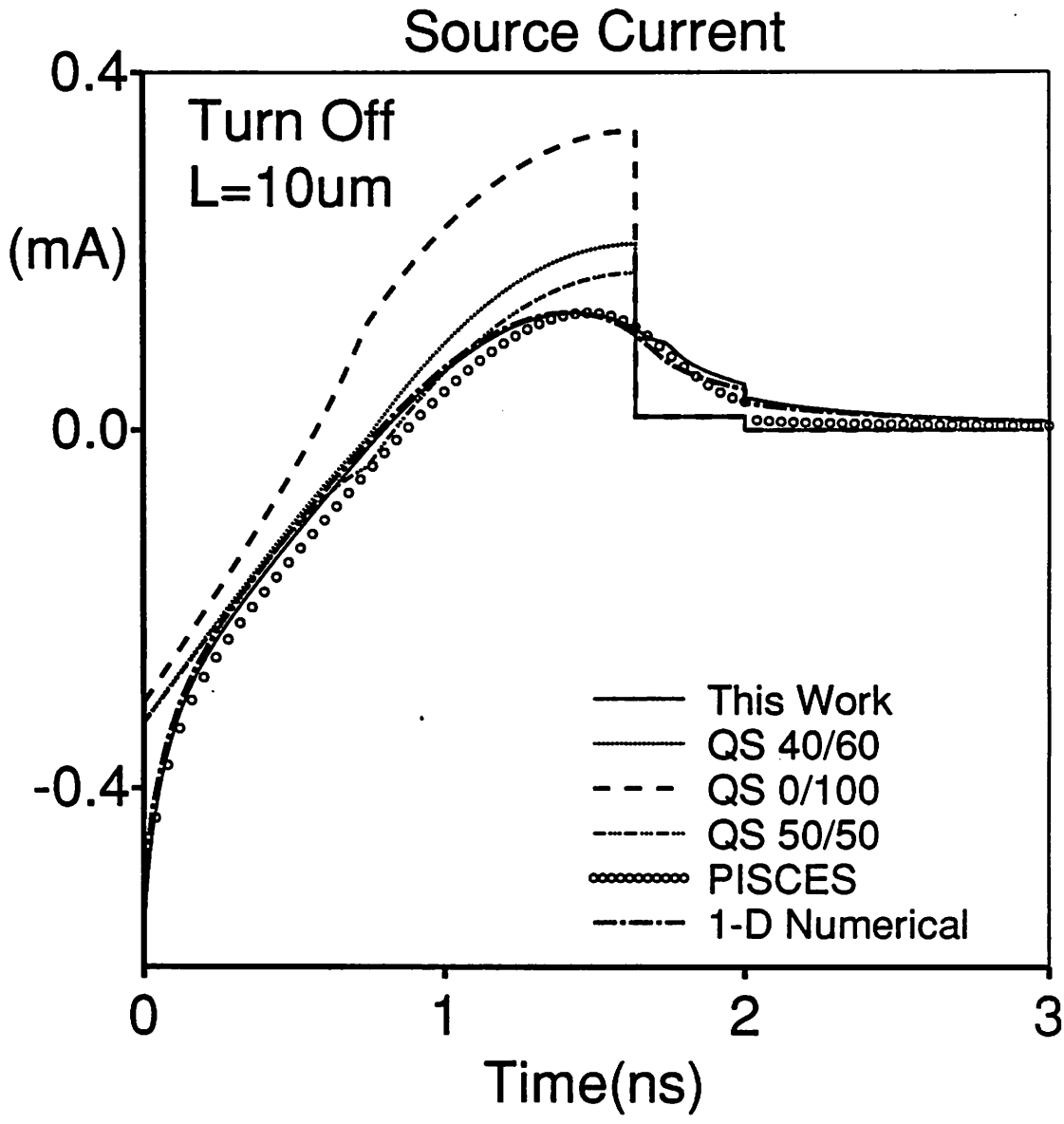


Fig.5.5.(d). The source current waveform for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

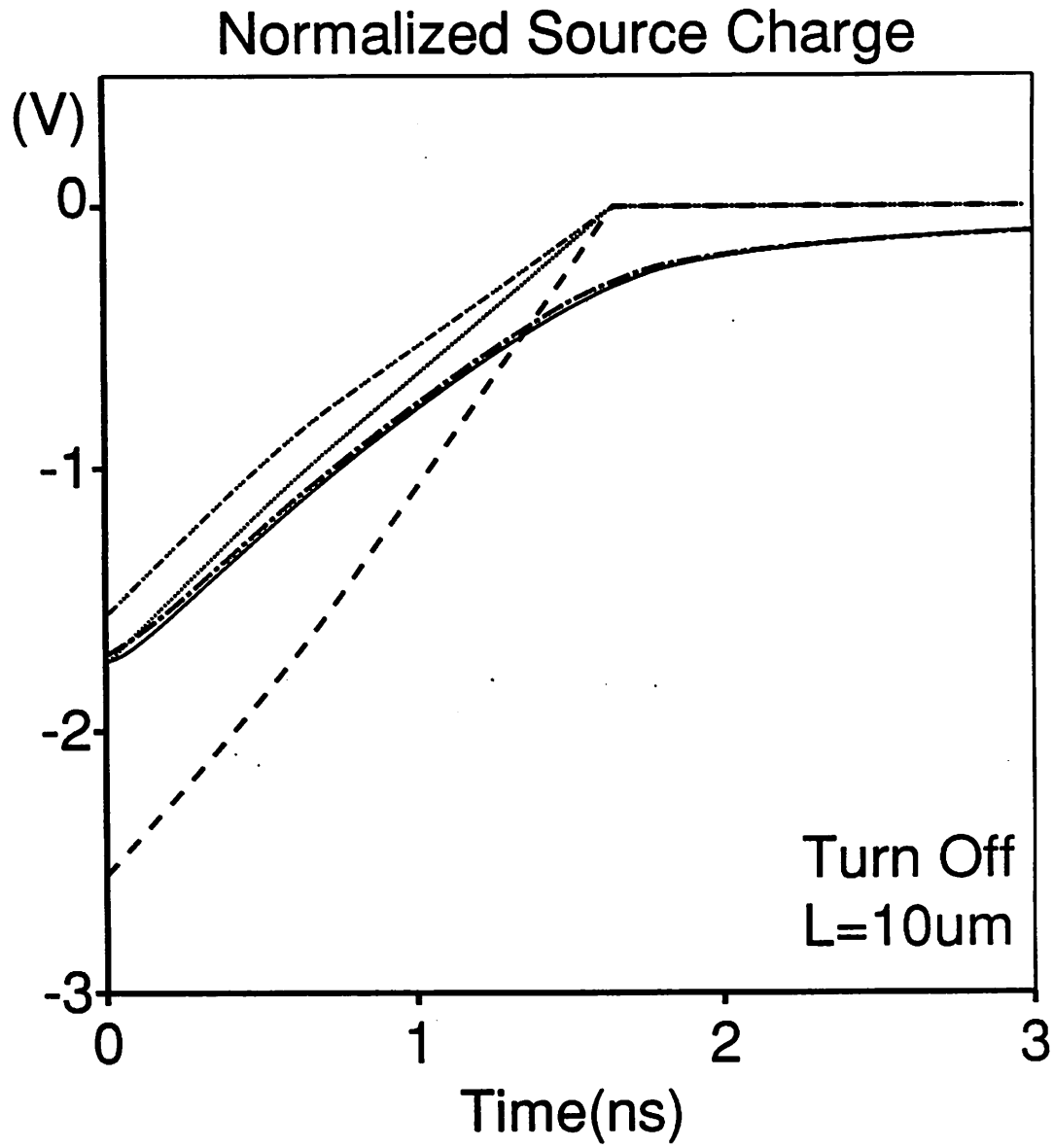


Fig.5.5.(e). The waveform of the intrinsic source charge normalized by WLCox, for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

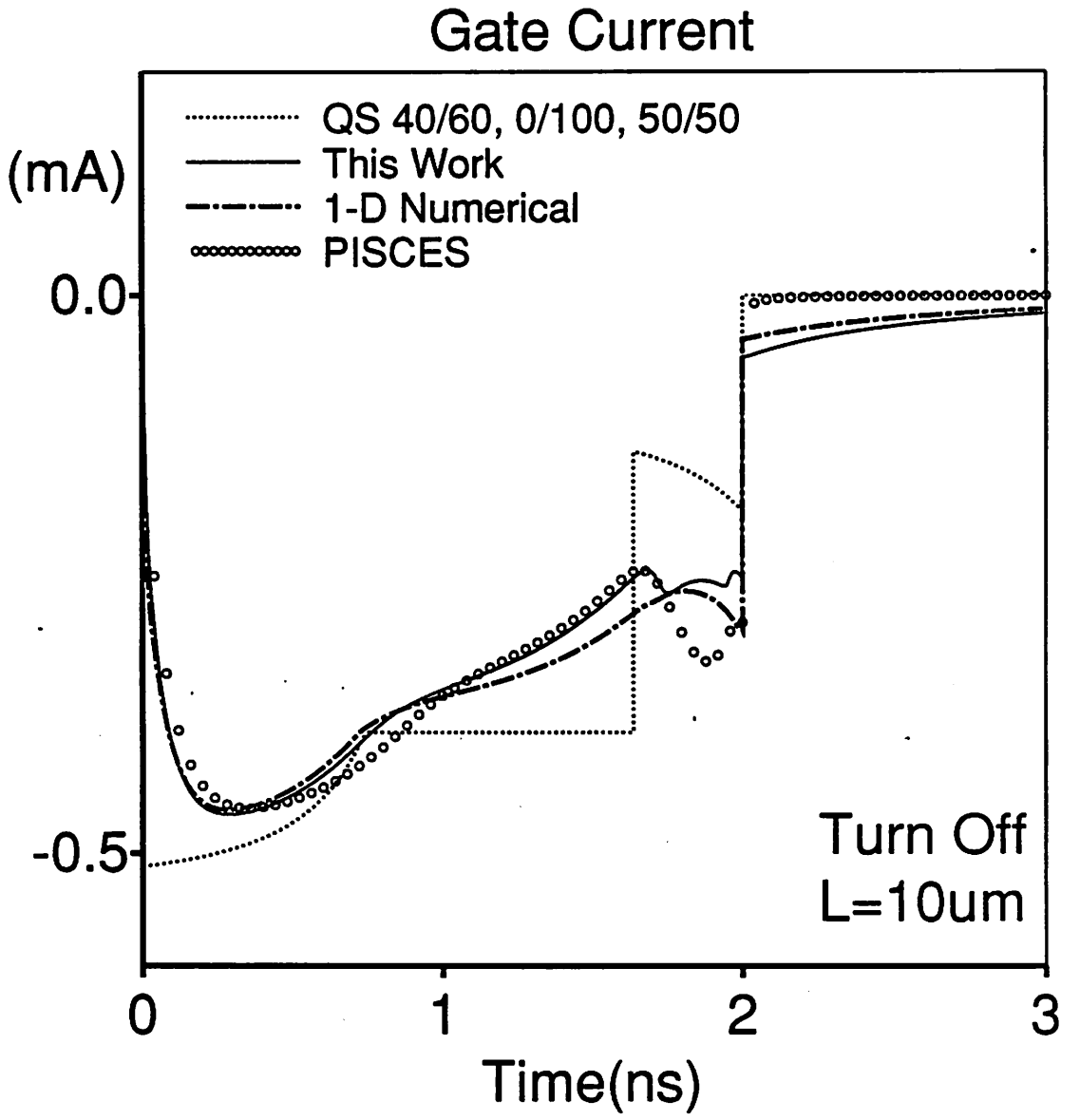


Fig.5.5.(f). The gate current waveform for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

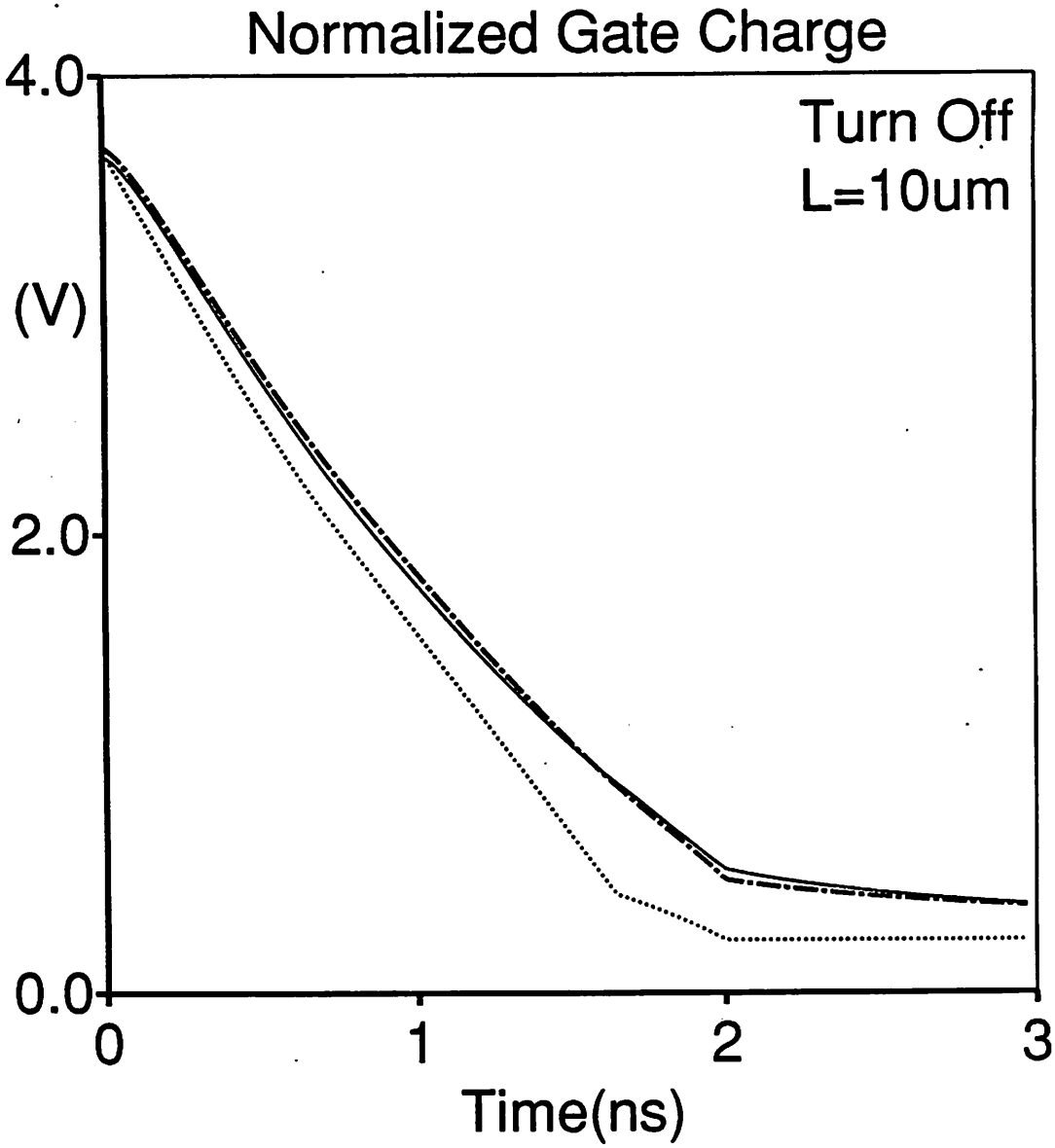


Fig.5.5.(g). The waveform of the intrinsic gate charge normalized by WLCox, for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

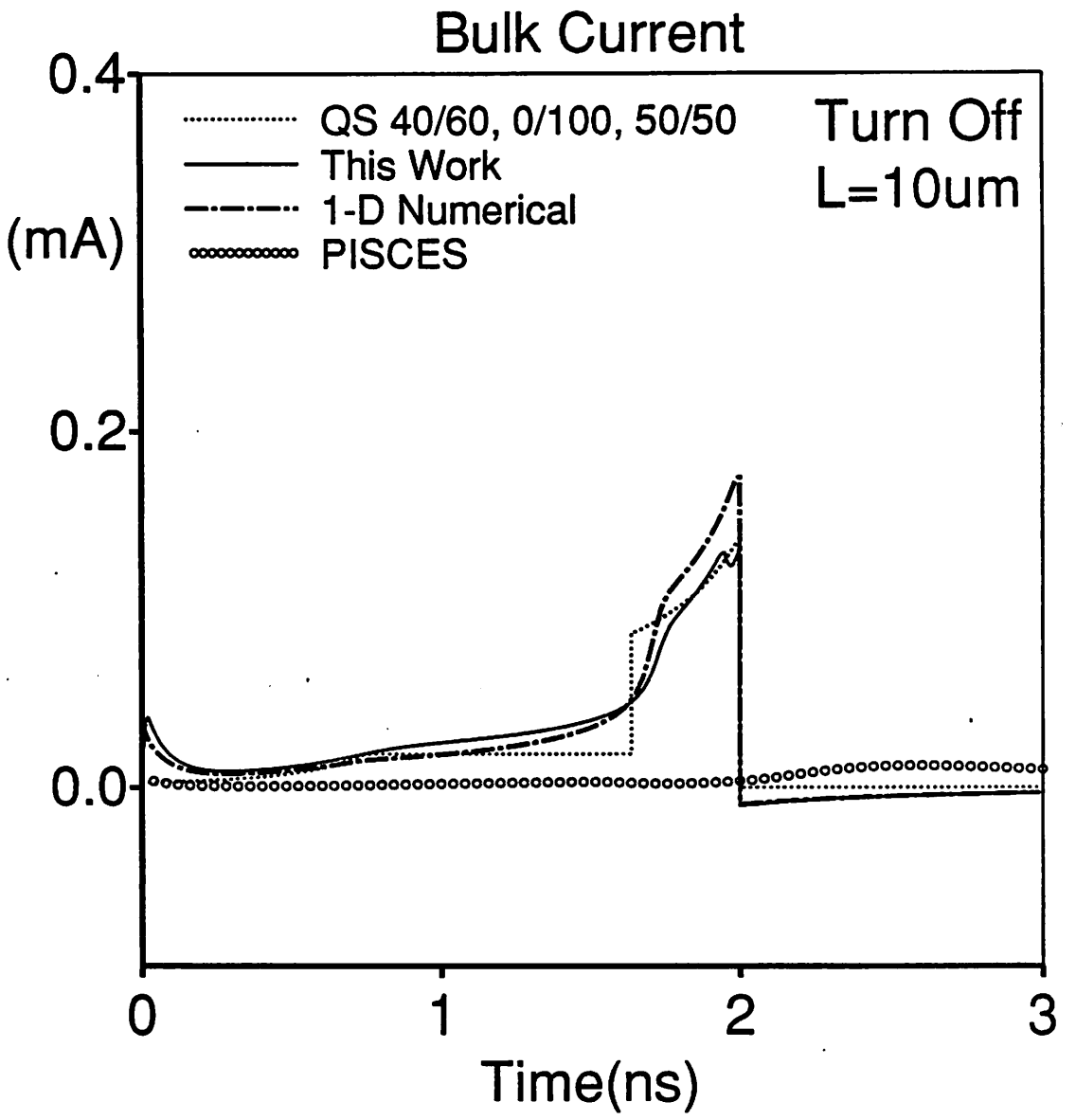


Fig.5.5.(h). The bulk current waveform for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

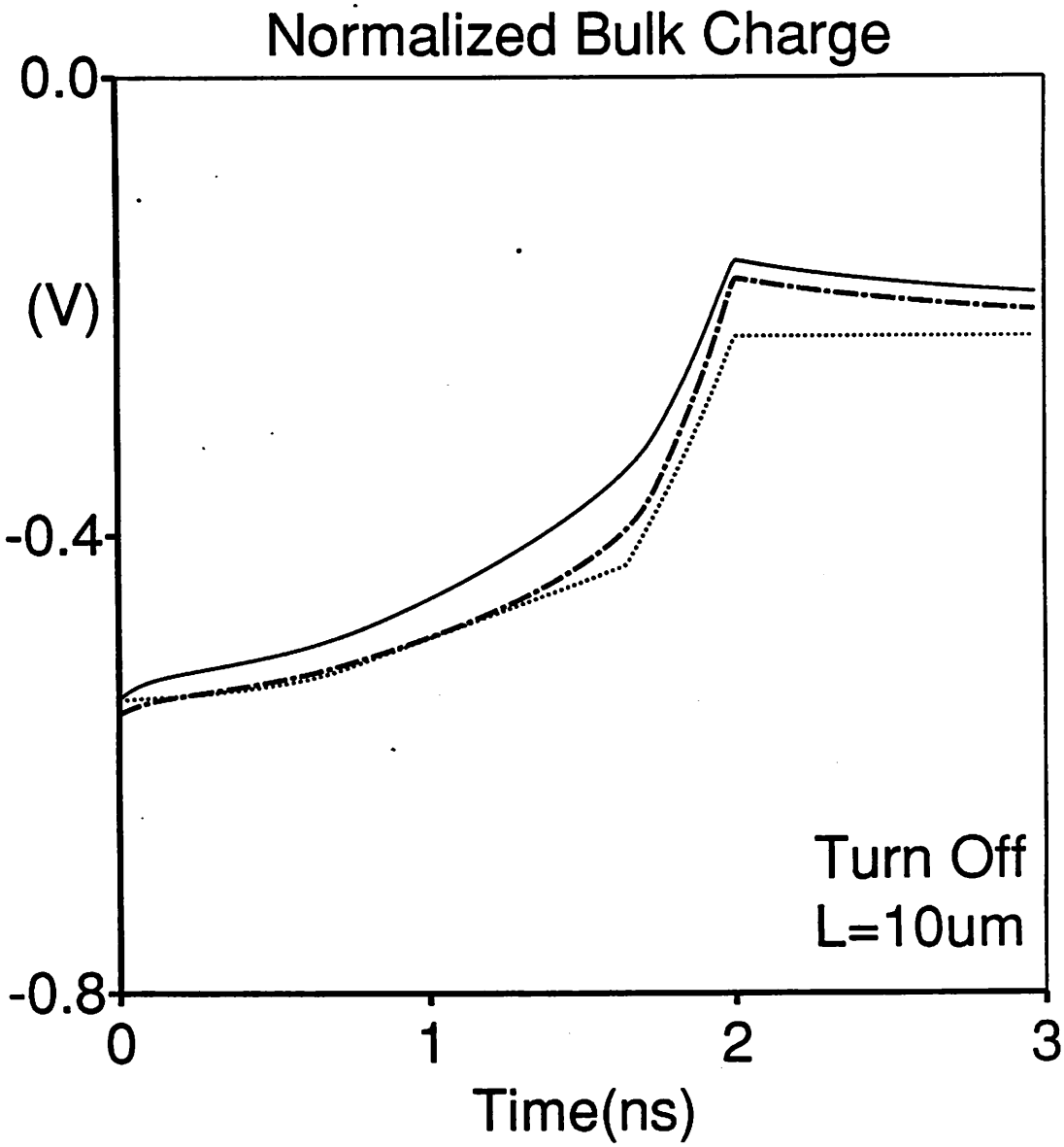


Fig.5.5.(i). The waveform of the intrinsic bulk charge normalized by WLCox, for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

SPICE3 Input Listings of
Turn-Off Transient of an NMOSFET

```
turnoff transient of a NMOSFET
m1 1 2 3 4 enh w=10u l=10u
VG 2 0  dc 5 PWL 0 5 2N 0 1 0
VD 1 0 2
VS 3 0 0
VB 4 0 0
.TRAN .01N 3N 0 0.01N
.PRINT TRAN I(VG) I(VD) I(VS) I(VB)
*
* model parameters for NQS model
.model enh nmos level=6 vfb=-0.5
+       tox=0.018 nsub=2.1e16 u0=500
+       lambda=0 js=1e-10 cgdo=1500p
+       cgso=600p mintol=0.1 qtrtol=1
*
* model parameters for BSIM
*    xpart=0 ----> QS 40/60
*    xpart=1 ----> QS 0/100
*    xpart=2 ----> QS 50/50
*.model enh nmos level=4 xpart=0
*+ vfb=-0.5 phi=0.96037 k1=0.435
*+ muz=500 mus=500 tox=0.018 vdd=5
*+ cgdo=1500p cgso=600p
* convergence paramters for BSIM SPICE3
*.OPT ACCT VNTOL=1N RELTOL=0.000002
*+ ABSTOL=1E-15 CHGTOL=1E-18
.end
```

Fig.5.5.(j). The SPICE3 input file for the turn-off transient of the NMOSFET shown in Fig.5.5.(a).

models for turn-on is essentially the same as that for turn-off and they are just flipped due to the opposite polarity of gate voltage ramp, but the charge waveform of this work for turn-on is completely different from that for turn-off. From this observation, we can see that the node charge of QS models is a function of applied node voltages only but the node charge of this work(NQS model) is a function of history as well as applied node voltages.

Fig.5.5.(d) shows the source current waveform. All the QS models show discontinuities in intrinsic source current at t=0 and t=1.65 ns (turn-off time). But this work, PISCES and 1-D simulation show that the intrinsic source current is continuous throughout all the time. By 'intrinsic source current', we mean the source current excluding the current components due to stray capacitances such as overlap and junction capacitances. At t=2ns, all the models show the same discontinuity in source current since the current through overlap capacitance is discontinuous due to the discontinuity of gate voltage slope.

Fig.5.5.(e) shows the waveform of intrinsic source charge normalized by WLC_{OX}. Good agreements can be observed between this work and 1-D numerical simulation.

Fig.5.5.(f) shows the gate current waveform. This work, PISCES and 1-D numerical simulation show that the intrinsic gate current is continuous throughout all the time. But the QS model show big discontinuities in intrinsic source currents at t=0 and t=1.65 ns(turn-off time). Since the gate voltage is a ramp voltage and all the other node voltages are DC voltages, the gate current is directly proportional to the gate capacitance. Hence, from Fig.5.5.(f), we can see that the effective loading by the gate capacitance predicted by this work, PISCES and 1-D simulation is much smaller than that predicted by the QS model, especially in linear and saturation region for the fast turn-off.

Fig.5.5.(g) shows the waveform of the intrinsic gate charge normalized by WLC_{OX}. Good agreements can be observed between this work and 1-D numerical simulation.

Fig.5.5.(h) shows the bulk current waveform. This work shows good agreements with 1-D numerical simulation. PISCES gives relatively very small bulk current. This is

considered to be due to the bulk resistance which is not considered in this work, 1-D simulation or the QS model.

Fig.5.5.(i) shows the waveform of the intrinsic bulk charge normalized by WLC_{OX}.

The SPICE3 input file for this turn-off transient is shown in Fig.5.5.(j).

5.3.6. Simulation of CMOS inverter chain and CMOS ring oscillator

Delay times of a CMOS inverter chain have been simulated using this work and other QS models and the results have been compared. Model parameters have been adjusted to guarantee the same DC characteristics among all the models considered. Fig.5.6.(a) and Fig.5.6.(b) show the DC transfer curve and the DC supply current of a single CMOS inverter simulated using this work, BSIM(QS 40/60, 0/100, 50/50) and the Meyer model. The Meyer model refers to the SPICE level-2 DC model with Meyer capacitance model. Good agreements among models have been obtained in DC characteristics.

Fig.5.6.(c) and Fig.5.6.(d) show the output voltage waveforms after 10 stages of CMOS inverter chain for the rising pulse input and the falling pulse input respectively as shown in the insert. The unit delay time per inverter stage and percent differences from this work are tabulated in Table.5.3. The QS 50/50 model is not shown in Table 5.3 due to convergence problem. This work gives the smallest delay time and the QS 40/60 model gives the largest delay time among models compared in Table.5.3. The reason why this work gives the smallest delay time is considered to be due to the fact that the effective loading by the gate capacitance is much smaller in this work than in QS(quasi-static) models, for the fast transient as shown in Fig.5.4.(f) and Fig.5.5.(f).

Delay times of each model for the rising input and the falling input are almost equal to each other, as can be seen in Table.5.3.

The SPICE3 input file for the CMOS inverter chain is shown in Fig.5.6.(e).

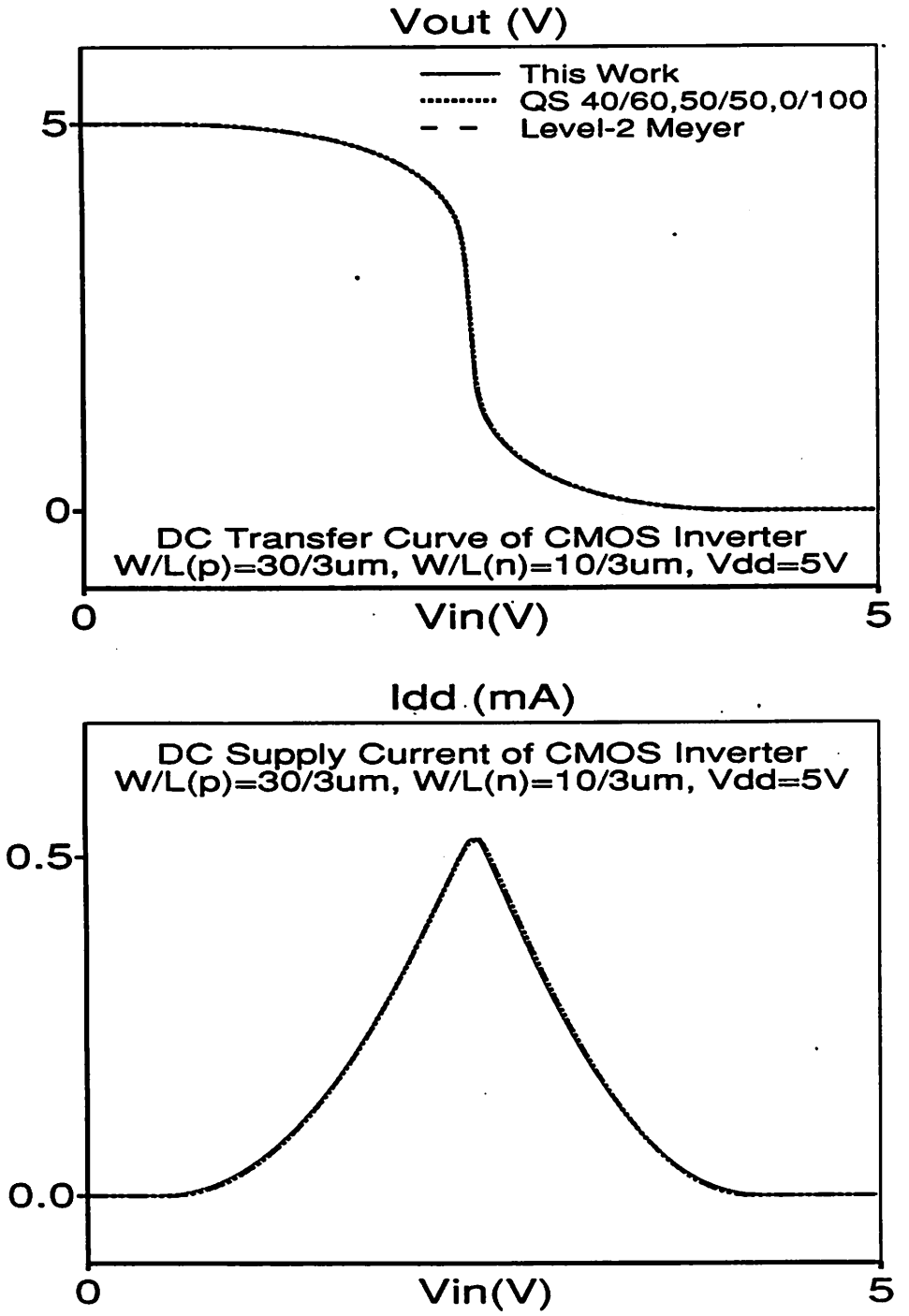


Fig.5.6.(a). Comparison of DC transfer curves of a CMOS inverter simulated using this work and other QS models.

Fig.5.6.(b). Comparison of DC supply current of a CMOS inverter simulated using this work and other QS models.

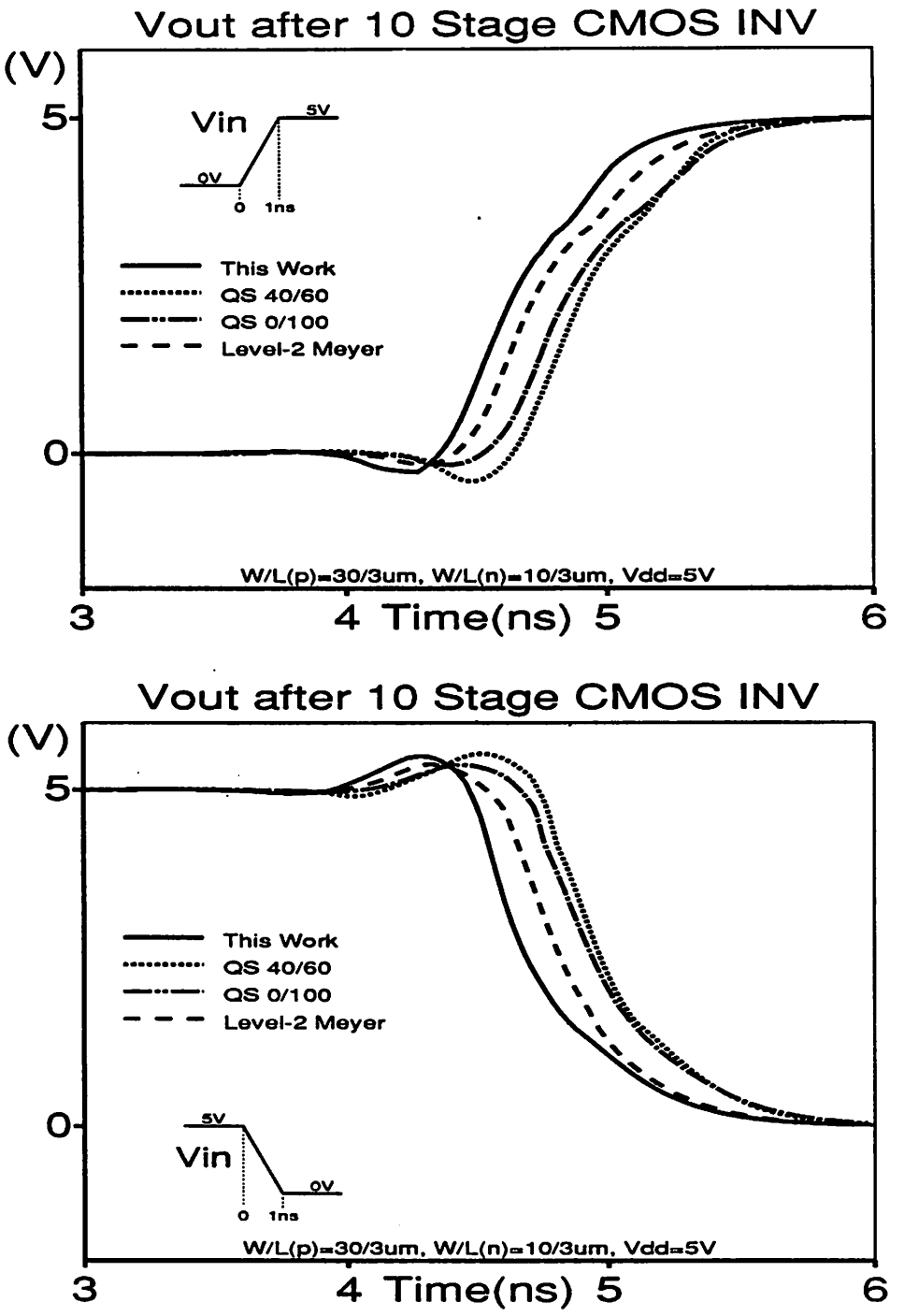


Fig.5.6.(c). Comparison of the output voltage waveform after 10 stages of CMOS inverter chain for the rising pulse input, simulated using this work and other QS models.

Fig.5.6.(d). The same comparison as in Fig.5.6.(c), for the falling pulse input.

<SPICE3 Input File for CMOS Inverter Chain>

```
transient analysis of 10 stage CMOS INV
m1 6 3 0 0 nch w=10u l=3u ad=60f as=60f
m2 6 3 5 5 pch w=30u l=3u ad=180f as=180f
m3 7 6 0 0 nch w=10u l=3u ad=60f as=60f
m4 7 6 5 5 pch w=30u l=3u ad=180f as=180f
m5 8 7 0 0 nch w=10u l=3u ad=60f as=60f
m6 8 7 5 5 pch w=30u l=3u ad=180f as=180f
m7 9 8 0 0 nch w=10u l=3u ad=60f as=60f
m8 9 8 5 5 pch w=30u l=3u ad=180f as=180f
m9 10 9 0 0 nch w=10u l=3u ad=60f as=60f
m10 10 9 5 5 pch w=30u l=3u ad=180f as=180f
m11 11 10 0 0 nch w=10u l=3u ad=60f as=60f
m12 11 10 5 5 pch w=30u l=3u ad=180f as=180f
m13 12 11 0 0 nch w=10u l=3u ad=60f as=60f
m14 12 11 5 5 pch w=30u l=3u ad=180f as=180f
m15 13 12 0 0 nch w=10u l=3u ad=60f as=60f
m16 13 12 5 5 pch w=30u l=3u ad=180f as=180f
m17 14 13 0 0 nch w=10u l=3u ad=60f as=60f
m18 14 13 5 5 pch w=30u l=3u ad=180f as=180f
m19 15 14 0 0 nch w=10u l=3u ad=60f as=60f
m20 15 14 5 5 pch w=30u l=3u ad=180f as=180f
m21 16 15 0 0 nch w=10u l=3u ad=60f as=60f
m22 16 15 5 5 pch w=30u l=3u ad=180f as=180f
vdd 5 0 dc 5
vpulse 3 0 dc 0 pwl 0 0 1n 5  1 5
.ic v(3)=0 v(6)=5 v(7)=0 v(8)=5 v(9)=0 v(10)=0 v(11)=0
+   v(12)=5 v(13)=0 v(14)=5 v(15)=0 v(16)=5
.TRAN .01N 10n 0 0.01n
.print tran v(15)
*******************************************************
* Model Parameters for This Work
*******************************************************
.model nch nmos level=6 vfb=-0.77 tox=0.018 nsub=2e16
+       u0=500 lambda=0.03 pb=0.7 pbsw=0.7 js=1e-4
+       cgdo=150p cgso=150p cj=3e-4 cjsw=8e-10
+       mintol=1 qtrtol=80
```

* (continued in the next page)

```
.model pch pmos level=6 vfb=0.22 tox=0.018 nsub=6e15
+      u0=180 lambda=0.05 pb=0.7 pbsw=0.7 js=1e-4
+      cgdo=150p cgso=150p cj=2e-4 cjsw=5e-10
+      mintol=1 qtrtol=80
*
*****************************************************
* Model Parameters for BSIM(QS 40/60) in SPICE 3
*        xpart=0 ----> BSIM 40/60
*        xpart=1 ----> BSIM 0/100
*        xpart=2 ----> BSIM 50/50
*****************************************************
*.model nch nmos level=4 xpart=0 vfb=-0.73
*+       phi=0.882 k1=0.425 muz=500 mus=520 tox=0.018
*+       vdd=5 n0=1 cgdo=150p cgso=150p cj=3e-4
*+       cjsw=8e-10 js=1e-12 pb=0.7 pbsw=0.7
*.model pch pmos level=4 xpart=0 vfb=-0.22
*+       phi=0.807 k1=0.233 muz=180 mus=201 tox=0.018
*+       vdd=5 n0=1 cgdo=150p cgso=150p cj=2e-4
*+       cjsw=5e-10 js=1e-12 pb=0.7 pbsw=0.7
**
*****************************************************
* Model Parameters for SPICE Level-2 DC model
*      with  Meyer capacitance model
*****************************************************
*.model nch nmos level=2 vto=0.502 phi=0.956
*+       gamma=0.425 uo=480 lambda=0.03 tox=0.018u
*+       nfs=5e10 cgdo=150p cgso=150p cj=3e-4
*+       cjsw=8e-10 js=1e-12 pb=0.7
*.model pch pmos level=2 vto=-0.812 phi=0.874
*+       gamma=0.233 uo=170 lambda=0.05 tox=0.018u
*+       nfs=5e10 cgdo=150p cgso=150p cj=2e-4
*+       cjsw=5e-10 js=1e-12 pb=0.7
.end
```

Fig.5.6.(e). The SPICE3 input file for the delay time simulation of the CMOS inverter chain whose output voltage waveforms are shown in Fig.5.6.(c) and Fig.5.6.(d) respectively.

(continued from the preceding page)

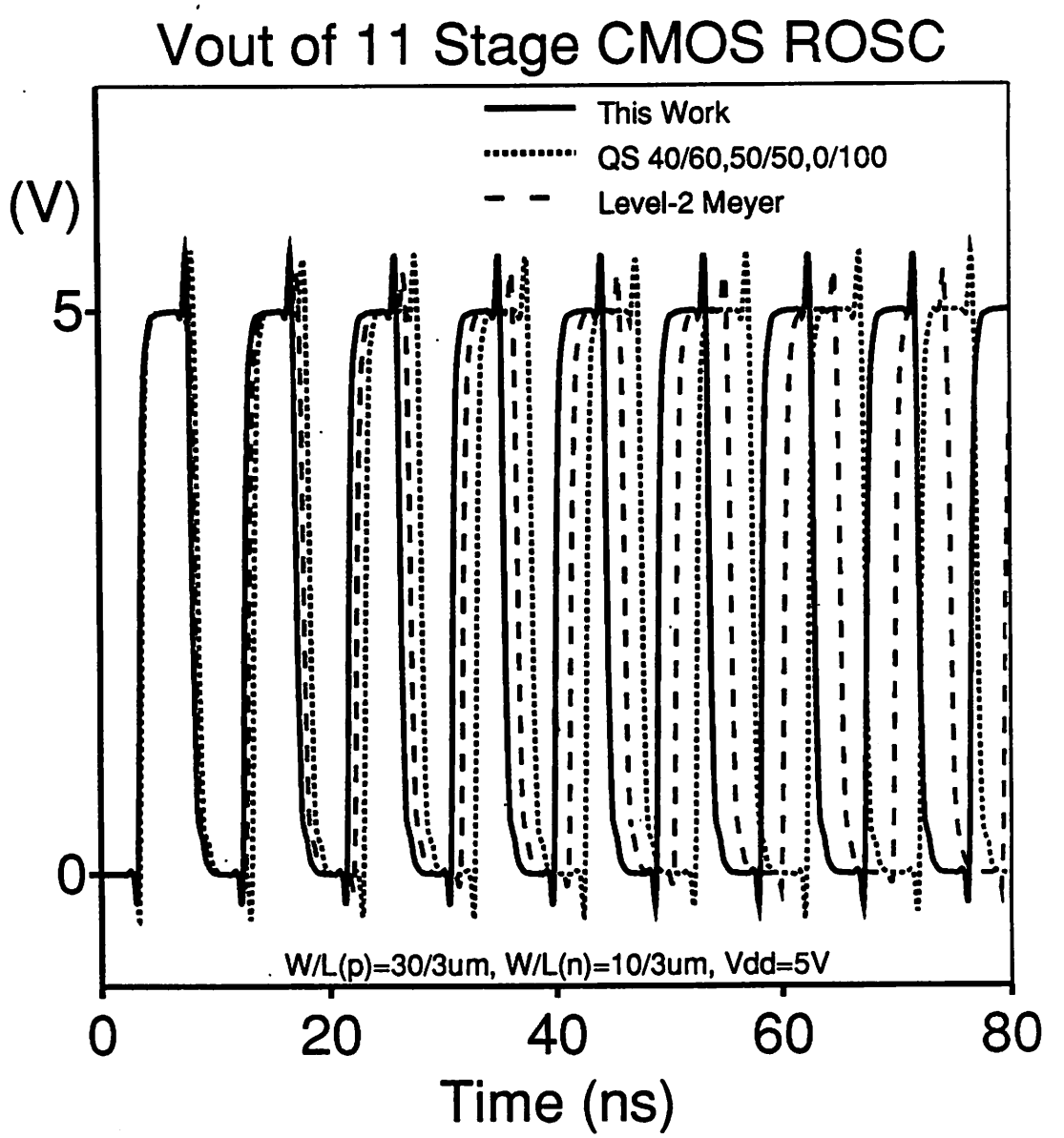


Fig.5.7.(a). The output voltage waveform of an 11-stage CMOS ring oscillator.

<SPICE3 Input File for CMOS Ring Oscillator>

```
11 stage cmos ring oscillator
m1 1 2 0 0 nch w=10u l=3u ad=60f as=60f
m2 1 2 5 5 pch w=30u l=3u ad=60f as=60f
m3 7 1 0 0 nch w=10u l=3u ad=60f as=60f
m4 7 1 5 5 pch w=30u l=3u ad=60f as=60f
m5 8 7 0 0 nch w=10u l=3u ad=60f as=60f
m6 8 7 5 5 pch w=30u l=3u ad=60f as=60f
m7 9 8 0 0 nch w=10u l=3u ad=60f as=60f
m8 9 8 5 5 pch w=30u l=3u ad=60f as=60f
m9 10 9 0 0 nch w=10u l=3u ad=60f as=60f
m10 10 9 5 5 pch w=30u l=3u ad=60f as=60f
m11 11 10 0 0 nch w=10u l=3u ad=60f as=60f
m12 11 10 5 5 pch w=30u l=3u ad=60f as=60f
m13 12 11 0 0 nch w=10u l=3u ad=60f as=60f
m14 12 11 5 5 pch w=30u l=3u ad=60f as=60f
m15 13 12 0 0 nch w=10u l=3u ad=60f as=60f
m16 13 12 5 5 pch w=30u l=3u ad=60f as=60f
m17 14 13 0 0 nch w=10u l=3u ad=60f as=60f
m18 14 13 5 5 pch w=30u l=3u ad=60f as=60f
m19 15 14 0 0 nch w=10u l=3u ad=60f as=60f
m20 15 14 5 5 pch w=30u l=3u ad=60f as=60f
m21 2 15 0 0 nch w=10u l=3u ad=60f as=60f
m22 2 15 55 55 pch w=30u l=3u ad=60f as=60f
vdd 5 0 dc 5
vpulse 55 0 dc 0 pwl 0 0 1n 5 1 5
.ic v(2)=0 v(1)=5 v(7)=0 v(8)=5 v(9)=0 v(10)=5
+ v(11)=0 v(12)=5 v(13)=0 v(14)=5  v(15)=0
.TRAN .01N 80N
.print tran v(11)
.OPTION ACCT
*****************************************************
* Model Parameters for This Work                    *
*****************************************************
.model nch nmos level=6 vfb=-0.77 tox=0.018 nsub=2e16
+       u0=500 lambda=0.03 pb=0.7 pbsw=0.7 js=1e-4
+       cgdo=150p cgso=150p cj=3e-4 cjsw=8e-10
+       mintol=1 qtrtol=80
```

* (continued in the next page)

```
.model pch pmos level=6 vfb=0.22 tox=0.018 nsub=6e15
+      u0=180 lambda=0.05 pb=0.7 pbsw=0.7 js=1e-4
+      cgdo=150p cgso=150p cj=2e-4 cjsw=5e-10
+      mintol=1 qtrtol=80
*
*****************************************************
* Model Parameters for BSIM(QS 40/60) in SPICE 3    *
*          xpart=0 ----> BSIM 40/60                 *
*          xpart=1 ----> BSIM 0/100                 *
*          xpart=2 ----> BSIM 50/50                 *
*****************************************************
*.model nch nmos level=4 xpart=0 vfb=-0.73
*+       phi=0.882 k1=0.425 muz=500 mus=520 tox=0.018
*+       vdd=5 n0=1 cgdo=150p cgso=150p cj=3e-4
*+       cjsw=8e-10 js=1e-12 pb=0.7 pbsw=0.7
*.model pch pmos level=4 xpart=0 vfb=-0.22
*+       phi=0.807 k1=0.233 muz=180 mus=201 tox=0.018
*+       vdd=5 n0=1 cgdo=150p cgso=150p cj=2e-4
*+       cjsw=5e-10 js=1e-12 pb=0.7 pbsw=0.7
*
*****************************************************
* Model Parameters for SPICE Level-2 DC model       *
*      with  Meyer capacitance model                *
*****************************************************
*.model nch nmos level=2 vto=0.502 phi=0.956
*+       gamma=0.425 uo=480 lambda=0.03 tox=0.018u
*+       nfs=5e10 cgdo=150p cgso=150p cj=3e-4
*+       cjsw=8e-10 js=1e-12 pb=0.7
*.model pch pmos level=2 vto=-0.812 phi=0.874
*+       gamma=0.233 uo=170 lambda=0.05 tox=0.018u
*+       nfs=5e10 cgdo=150p cgso=150p cj=2e-4
*+       cjsw=5e-10 js=1e-12 pb=0.7
.end
```

Fig.5.7.(b). The SPICE3 input file for the simulation of the 11-stage CMOS ring oscillator whose output voltage waveforms are shown in Fig.5.7.(a).

(continued from the preceding page)

Model	Unit Delay Time	% Difference
This Work	415 ps	0 %
Meyer	432 ps	+4.1 %
QS 40/60	445 ps	+7.2%
QS 0/100	444 ps	+7.0%
QS 50/50	444 ps	+7.0%

Table.5.4. The unit delay time per inverter stage computed from the CMOS ring oscillator simulation, and the percent differences of each model from this work

	This Work	Meyer	QS 40/60	QS 0/100	QS 50/50
Run Time(sec)	1235	148	139	127	174
# of Iterations	17959	6007	5262	4716	6045
Total Time Pts.	5561	1147	1266	1332	1246
Reject Time Pts	410	371	395	415	400
Load Time(sec)	1096	131	123	111	156

Table.5.5. Comparison of run statistics of CMOS ring oscillator simulation

Model	Rising Input(Fig.5.6.(c))	Falling Input(Fig.5.6.(d))
This Work	416 ps (0%)	419 ps (0%)
Meyer	425 ps (+2.2%)	431 ps (+2.9%)
QS 40/60	443 ps (+6.5%)	447 ps (+6.7%)
QS 0/100	437 ps (+5.0%)	444 ps (+6.0%)

Table.5.3. The unit delay time per stage of CMOS inverter chain and percent differences from this work

Fig.5.7.(a) shows the output voltage waveform of an 11-stage CMOS ring oscillator consisting of the same CMOS inverters as in the CMOS inverter chain example shown in Fig.5.6. QS 40/60, 0/100, and 40/60 models showed almost the same voltage waveform. The unit delay time per inverter stage and percent differences from this work are computed from Fig.5.7.(a) and are shown in Table.5.4.

The unit delay time in Table.5.4 agrees fairly well with the unit delay time in Table.5.3, which is computed from the inverter chain simulation. The run statistics of the CMOS ring oscillator simulation is shown in Table.5.5. The SPICE3 input file for this ring oscillator simulation is shown in Fig.5.7.(b). *QTRTOL*=80 is used in this work. *QTRTOL* is the model parameter for the time step control which is discussed in Appendix 11. Smaller values of *QTRTOL* (30 and 50) was also used in the simulation and it didn't show any appreciable differences in the output voltage waveform, and the maximum error in delay time among those different *QTRTOL*'s was less than 0.1%.

5.3.7. Turn-off transient of a feedback pass transistor

Fig.5.8.(a) shows the circuit schematic for the turn-off transient of a feedback pass transistor which is commonly encountered in switched analog circuits. An ideal OP

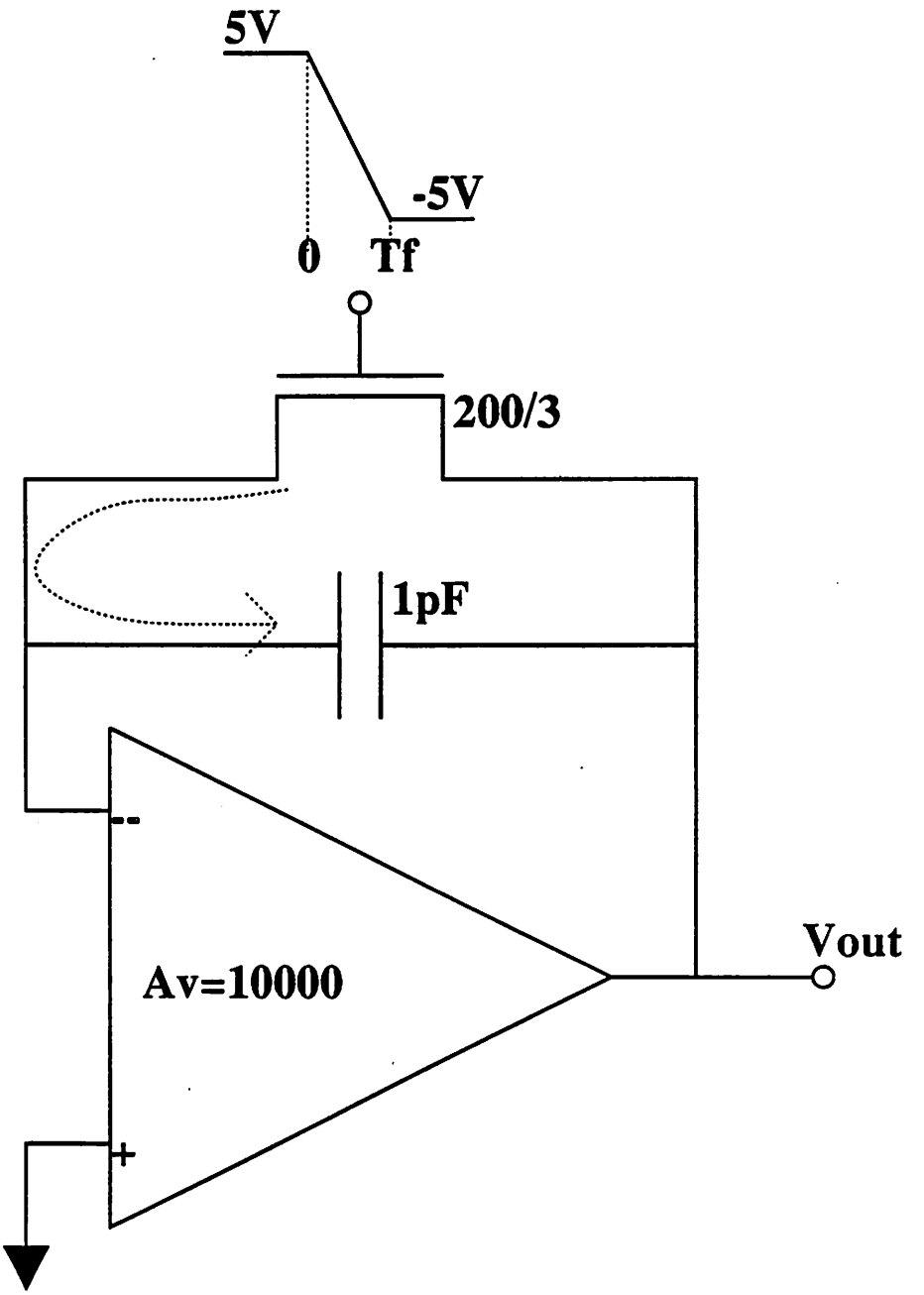


Fig.5.8.(a). The circuit schematic for the turn-off transient of a feedback pass-transistor.

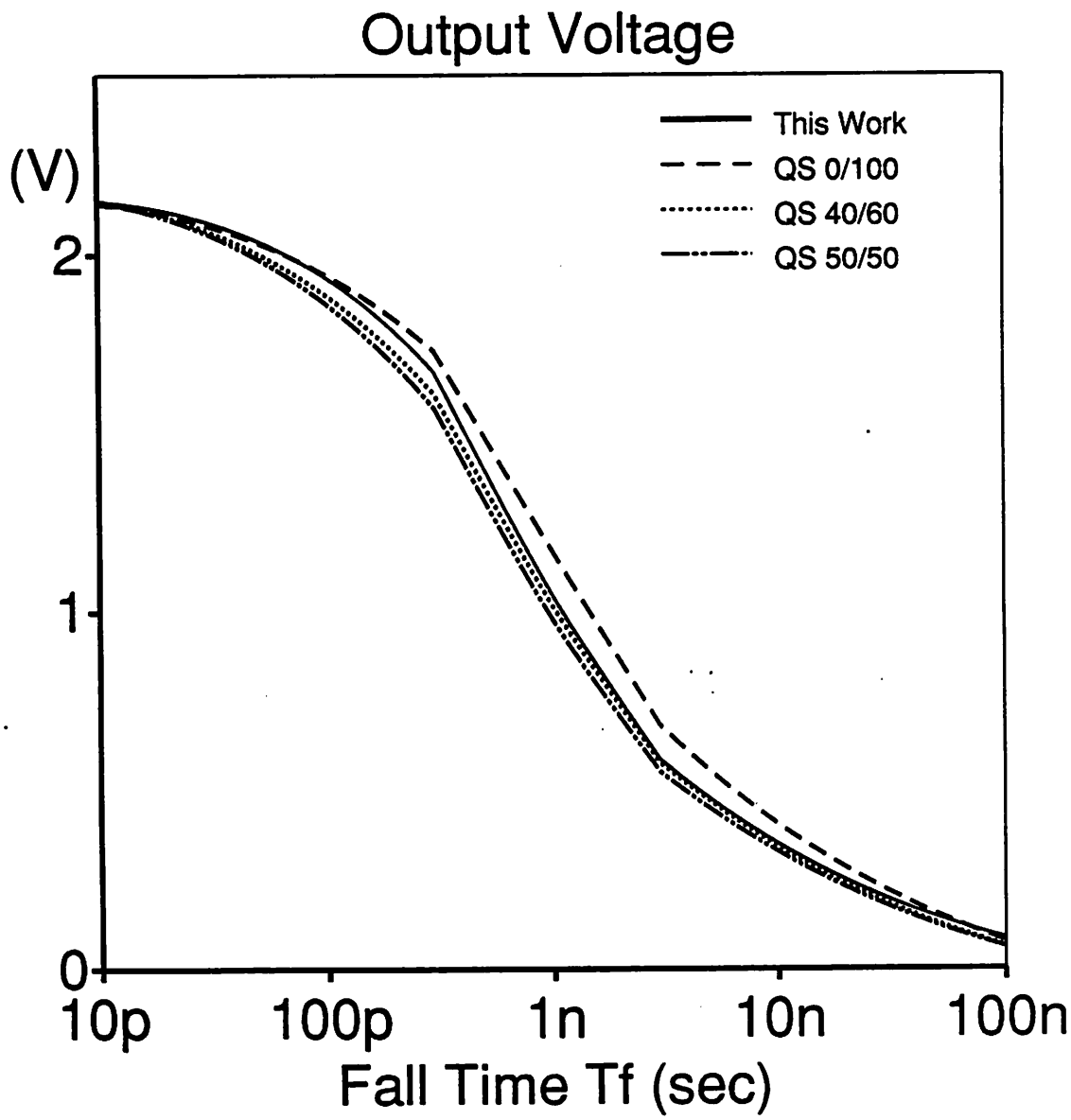


Fig.5.8.(b). The output voltage versus the fall time **Tf**, for the turn-off transient of a feedback pass-tr. shown in Fig.5.8.(a).

```
< SPICE3 Input File for the Turn-Off Transient of Feedback Pass Tr.>
turn-off transient of a feedback pass tr.
m1 2 4 3  5 nch w=200u l=3u
c1 2 3 1pF
E1 3 0 2 0 -10000
vb 5 0 dc -5
Vg2 4 0 dc 5 pwl 0 5 1n -5 1 -5
.tran 0.01n 3n
.ic v(2)=0 v(3)=0
.print tran v(3)
.opt gmin=1e-11
*
*****************************************
*   Model Parameters for This Work
*****************************************
.model nch nmos level=6 vfb=-0.77
+      tox=0.018 nsub=2e16 u0=500
+      lambda=0.03 pb=0.7 pbsw=0.7
+      js=1e-15 cgdo=0 cgso=0 cj=0 cjsw=0
+      mintol=0.5 qtrtol=5
*
*****************************************
*   Model Parameters for QS 40/60 Model
*****************************************
*.model nch nmos level=4
*+ xpart=0 vfb=-0.77 phi=0.958 k1=0.425 muz=500 mus=520
*+ tox=0.018 vdd=5 n0=1 cgdo=0 cgso=0
*+ cj=0 cjsw=0 js=1e-15 pb=0.7 pbsw=0.7
*.opt reltol=1e-4
.end
```

Fig.5.8.(c). The SPICE3 input file for the turn-off transient of a feedback pass-tr. shown in Fig.5.8.(a).

Amp is used to concentrate on the operation of the pass transistor. The voltage gain of the OP Amp is 80 dB throughout all the frequency range, and the output impedance is zero and the input impedance is infinity.

When the MOSFET is turned off, part of the channel charge is dumped on the capacitor as indicated by the dotted line shown in Fig.5.8.(a). The other part of the channel charge goes to the ground through the output node of OP Amp. Due to the charge injection of the pass tr., as indicated by the dotted line, the voltage of the OP Amp (-) input becomes negative and V_{OUT} becomes positive due to the OP Amp action. Hence, during the turn-off transient, the OP Amp (-) input becomes source node and the OP Amp output node becomes the drain node of the pass tr..

The source node is connected to the ground through the capacitor whose value is multiplied by the voltage gain due to Miller effect. And the drain node is shorted to the ground since the output impedance of the OP Amp is zero. For the slow signals, since the impedance of the Miller capacitor is high, almost all the channel charge goes to the ground through the output node of OP Amp. Hence, the output voltage V_{OUT} is small for slow signals (large Tf). RC time constant of the pass tr. with V_G=5V and the Miller capacitor is about 500ns. As the fall time Tf decreases, more charge is injected to the source node since the impedance of the Miller capacitor becomes smaller. Hence the output voltage increases as Tf is decreased, as shown in Fig.5.8.(b). When Tf is very small, half of the channel charge goes to the capacitor. As a reference, the RC time constant of the pass tr. with V_G=5V and the capacitor(1pF) is about 50 ps.

For the medium values of Tf (that is, 50 ps < Tf < 500 ns), appreciable differences among models can be observed in Fig.5.8.(b). The QS 0/100 model shows the largest output voltage since the QS 0/100 model assigns most of the channel charge to the source node, that is, the OP Amp (-) input node in this example. The QS 50/50 model gives the smallest output voltage. This work gives the output voltage which is close to that of QS 40/60 model for slow signals (Tf > 3ns) and is close to that of QS 0/100 model for fast signals (Tf < 300ps).

Fig.5.8.(c) shows the SPICE3 input file for this example.

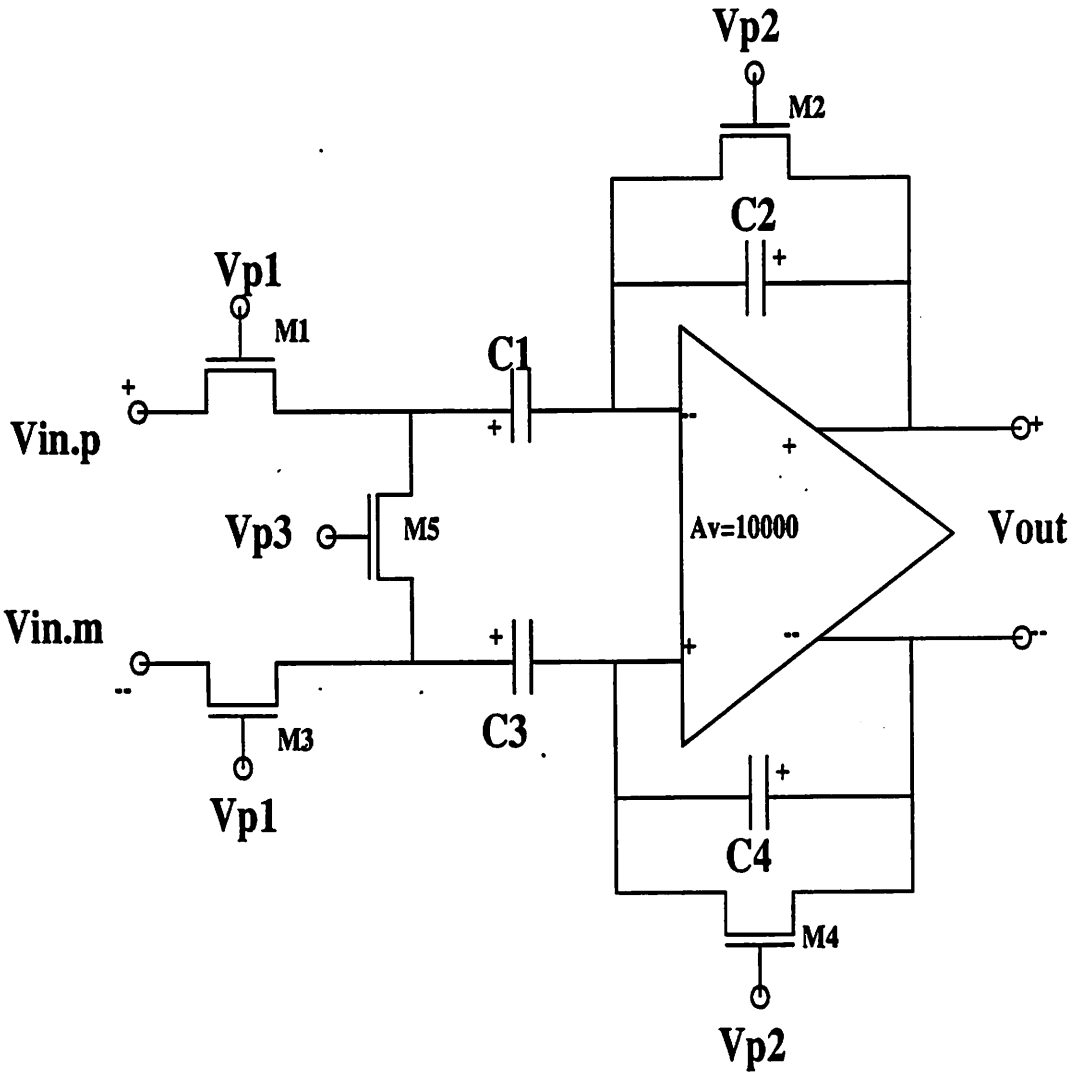
5.3.8. Simulation of a differential sample-hold circuit

Fig.5.9.(a) shows the circuit diagram of a differential sample-hold circuit which is used in MOS A/D converter circuits [5.5]. Input voltage waveforms are shown in Fig.5.9.(b) and Fig.5.9.(c).

When V_{P1} and V_{P2} are 5V and V_{P3} is -5V, input voltages *Vin.p* and *Vin.m* are stored(sampled) into capacitors $C1$ and $C3$ respectively and the output voltage V_{OUT} is zero. We call this time period the sampling time. When V_{P1} and V_{P2} become -5V and V_{P3} becomes 5V, the charges stored in the capacitors $C1$ and $C3$ are dumped into capacitors $C2$ and $C4$ respectively, and V_{OUT} becomes equal to the sampled input voltage (*Vin.p*-*Vin.m*) in an ideal manner. During this time period, the output voltage doesn't change with time even when the input voltage is changing. We call this time period the hold time. But due to the charge injection of MOSFET's, there is some error between the output voltage at hold time and the sampled input voltage. Fig.5.9.(c) shows three kinds of differential input voltage waveforms, rising, falling and DC inputs. Rising and falling inputs are a half-period of sine wave with frequency of 25MHz and amplitude of 2V. V_O is the superimposed DC voltage. In all cases, *Vin.p* and *Vin.m* have the same magnitude with opposite polarity. Ideal OP Amp is used to concentrate on the charge injection problem. The voltage gain of the OP Amp is 80dB throughout all the frequency range and the input impedance is infinity and the output impedance is zero.

As shown in Fig.5.9.(b), M2 and M4 are turned off before M1 and M3 are turned off to eliminate the effect of charge injection of M1 and M3 on the output voltage.

Three time phases (1), (2), and (3) are shown in Fig.5.9.(b). The time phase (1) is the time period between t=10ns and t=11ns, during which V_{P1}=5V and V_{P3}=-5V and V_{P2} changes from 5V to -5V. Time phase (2) is the time period between t=12ns and t=13ns during which V_{P1} changes from 5V to -5V and V_{P2}=-5V and V_{P3}=-5V. The time phase (3) is the time period between t=14ns and t=15ns, during which V_{P1}=V_{P2}=-5V and V_{P3}



W/L (M1.M2,M3,M4)=200/3um

W/L (M5) = 20/3um

C1 = C2 = C3 = C4 = 1pF

Fig.5.9.(a). The circuit diagram for a differential sample-hold circuit which is a part of MOS A/D converter [5.5].

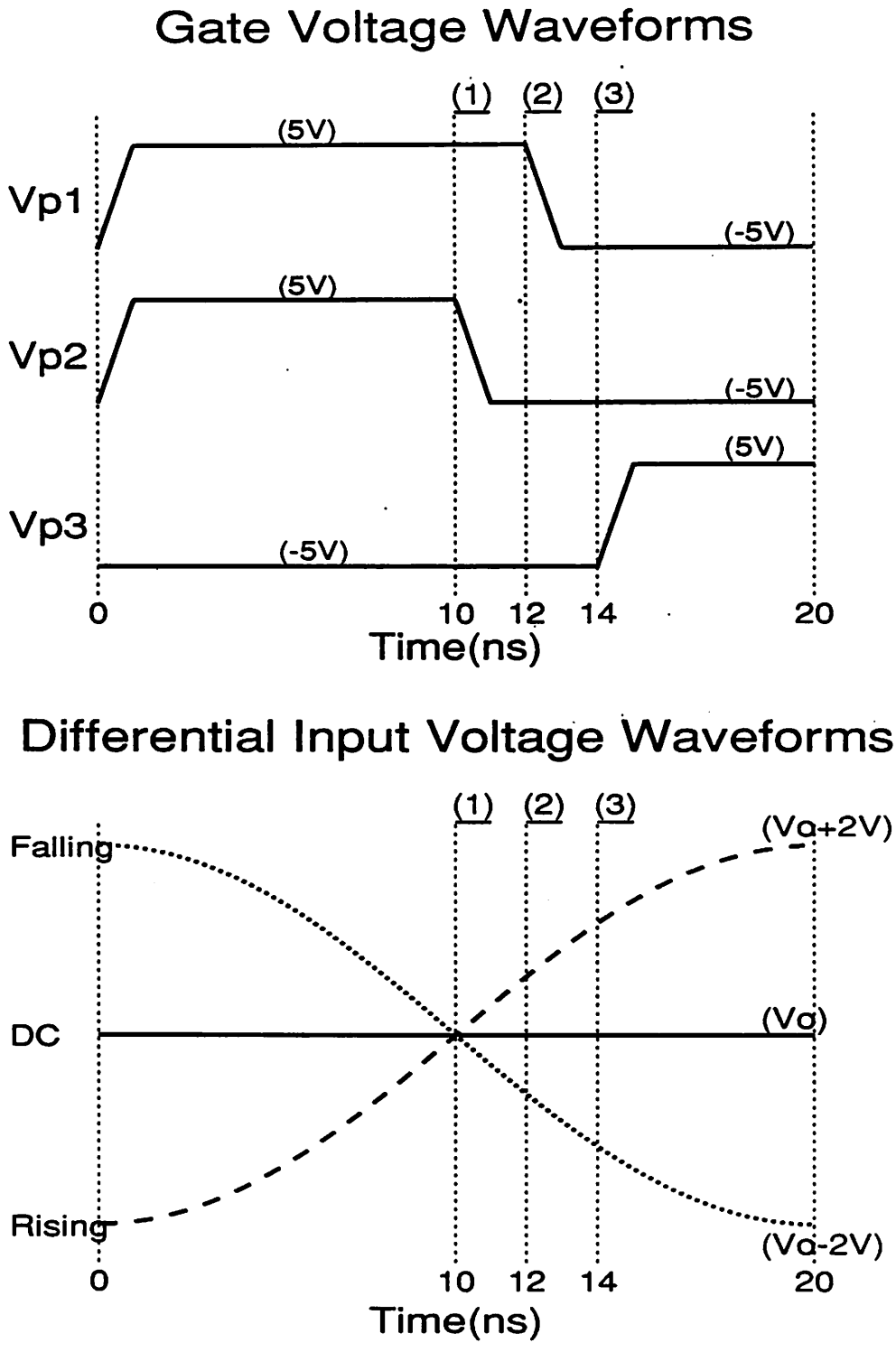


Fig.5.9.(b). The switching gate pulses Vp1, Vp2, and Vp3 for the simulation of the circuit shown in Fig.5.9.(a).

Fig.5.9.(c). Three kinds of differential input voltage waveforms.

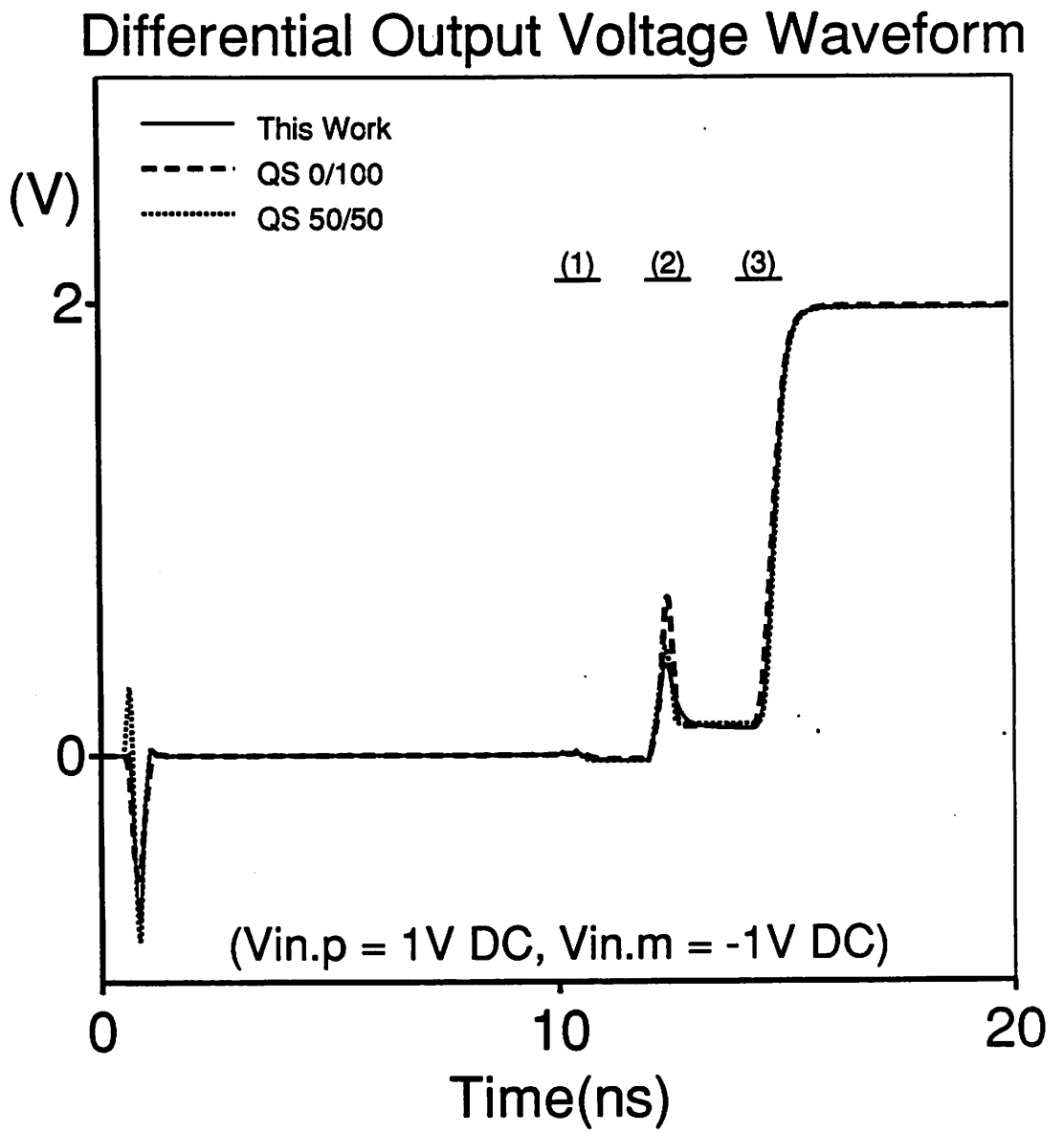


Fig.5.9.(d). The differential output voltage waveforms for the circuit shown in Fig.5.9.(a) with the switching gate pulses in Fig.5.9.(b) and Vin.p = 1V (DC) and Vin.m = -1V(DC).

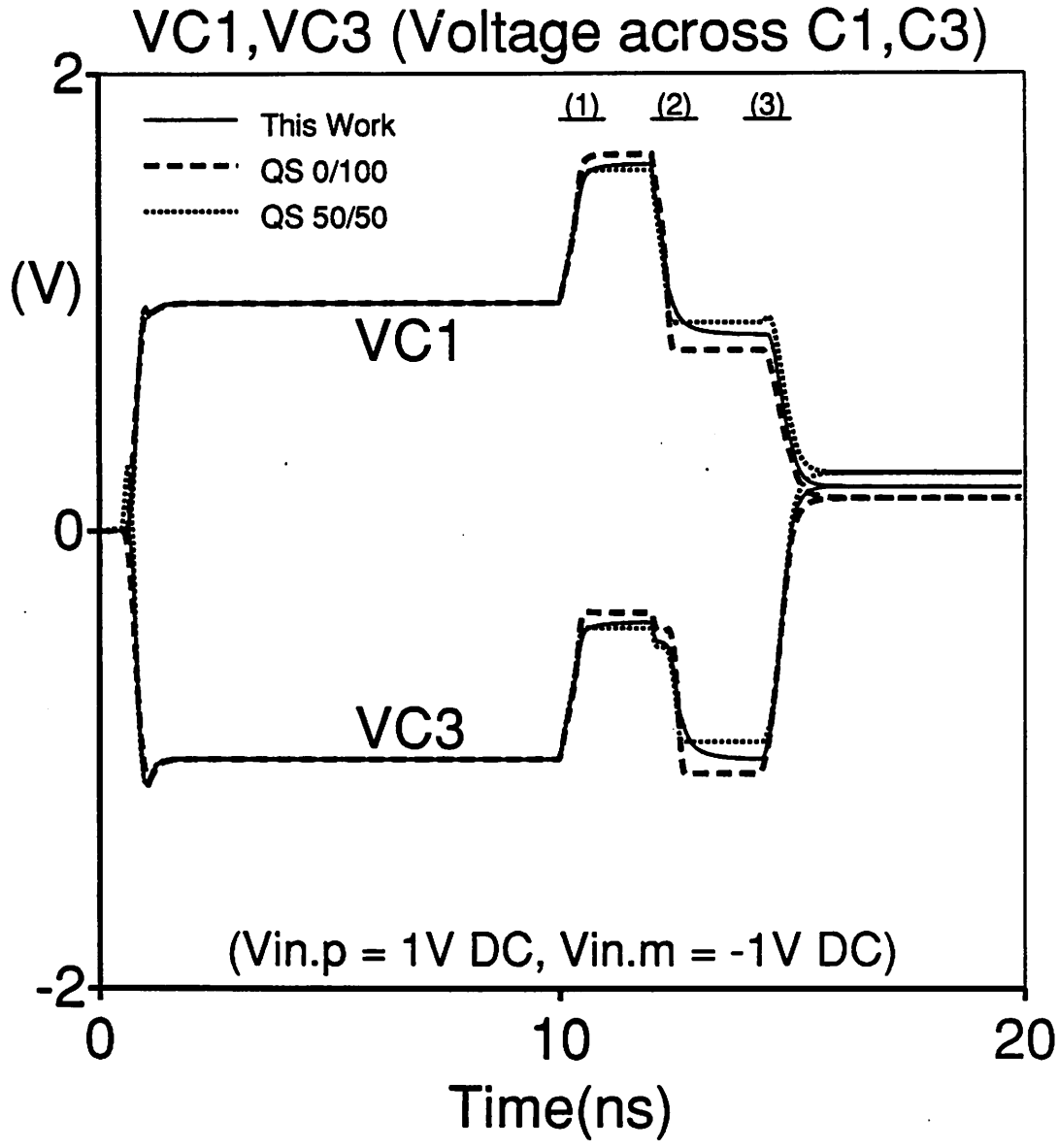


Fig.5.9.(e). The voltage waveforms across capacitors C1 and C3 with Vin.p = 1V(DC) and Vin.m = -1V(DC).

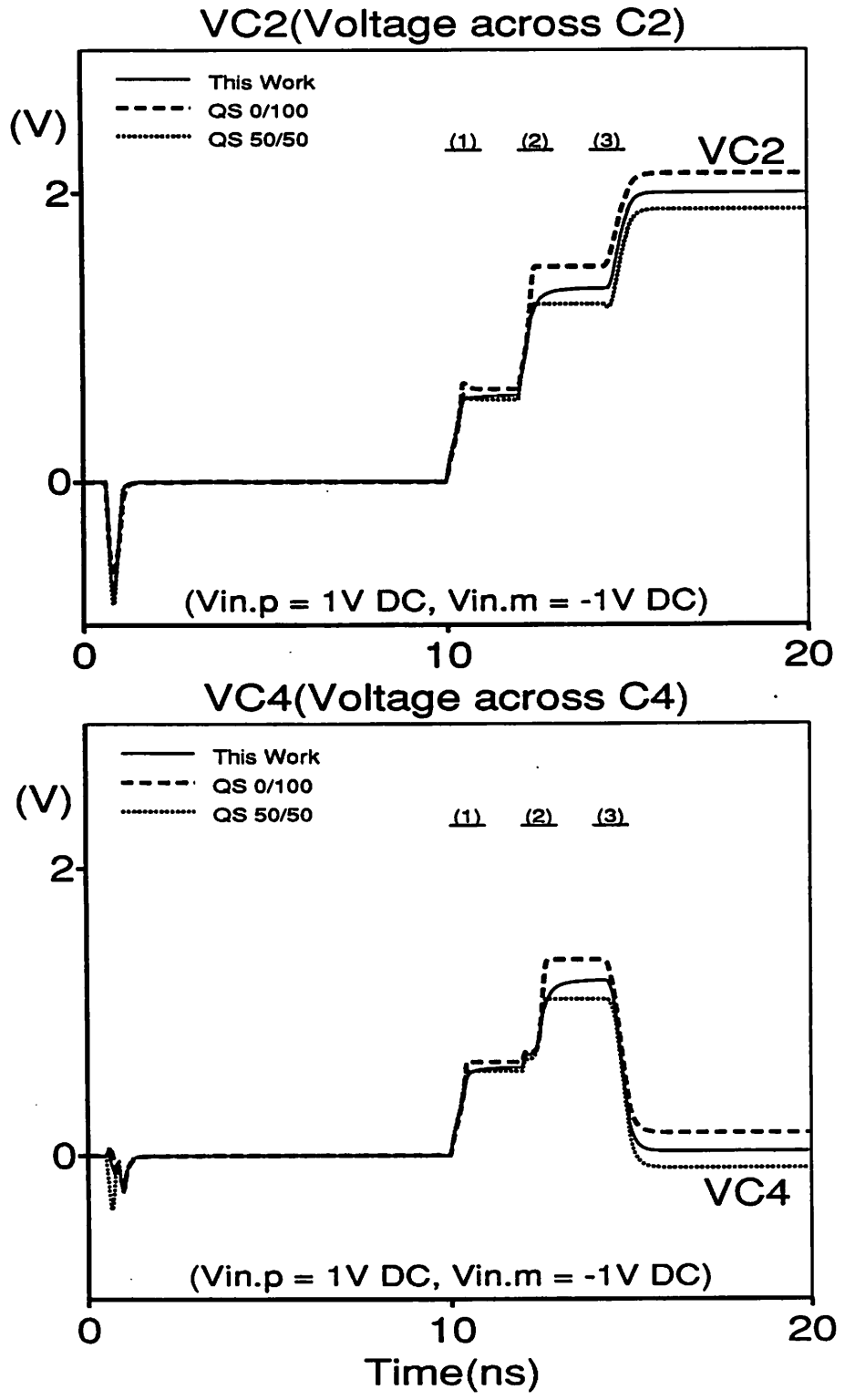


Fig.5.9.(f). The voltage waveforms across the capacitor C2.

Fig.5.9.(g). The voltage waveforms across the capacitor C4.

changes from -5V to 5V.

Fig.5.9.(d) shows the differential output voltage waveform predicted by this work, QS 0/100 and QS 50/50 models, for *Vin.p*=1V (DC) and *Vin.m*=-1V (DC). The QS 40/60 model is not shown for clarity. The output voltage is reset to 0 during sampling time (t < 10ns) and goes toward 2V during hold time (t > 15 ns).

Fig.5.9.(e), (f), and (g) show the waveforms of voltage across capacitors, $C1$, $C2$, $C3$ and $C4$, for *Vin.p*=1*V*(DC) and *Vin.m*=–1*V*(DC). The polarities of capacitor voltages are indicated by the (+) sign in Fig.5.9.(a).

During the sampling time (t < 10 ns), MOSFETs M1, M2, M3 and M4 are on and M5 is off, V_{C1} becomes equal to *Vin.p* and V_{C3} becomes equal to *Vin.m* as shown in Fig.5.9.(e), and V_{C2} and V_{C4} become equal to 0 as shown in Fig.5.9.(f) and (g). The differential input voltage *Vin.p–Vin.m* during this time period is called the sampled input voltage in the following discussions.

During the time phase (1), MOSFETs M2 and M4 are turned off. During this time period, the situation is essentially the same as the turn-off transient of a feedback pass-tr. discussed in the preceding section. Channel charge of M2 is injected into $C1$ and $C2$ and channel charge of M4 is injected into $C3$ and $C4$. Hence V_{C1}, V_{C2}, V_{C3} and V_{C4} increase with time during the time phase (1) and quite different results can be observed among different models.

During the time phase (2), MOSFETs M1 and M3 are turned off and channel charge is injected from M1 and M3 into $C1$ and $C3$ respectively. Hence V_{C1} and V_{C3} decreases and V_{C2} and V_{C4} increase with time.

During the time phase (3), M5 is turned on and charges stored in $C1$ and $C3$ are dumped into $C2$ and $C4$ respectively. The output voltage V_{OUT} which is equal to (V_{C2}-V_{C4}) changes approximately to the sampled input voltage. The voltages V_{C1} and V_{C3} become the same value which is close to 0 but is slightly positive to neutralize the negative channel charge of M5. Hence, the value of V_{C1} and V_{C3} become equal to (*channel charge of M5*) / ($C1$+$C3$).

During the time phase (3), the change of V_{C1} is equal to the negative of the change of V_{C2}, since $C1$ and $C2$ are connected in series and there is no other current path at the connection point. Similarly, the change of V_{C3} is equal to the change of V_{C4}, during this time interval.

The QS 40/60 model is not shown in Fig.5.9.(e), (f), and (g) for clarity. The QS 40/60 model gives voltage waveforms which are always located between the QS 0/100 and the QS 50/50 models and is closer to the QS 50/50 model.

All the changes of capacitor voltages during the time interval after the turn-off of M2 and M4 and before the turn-on of M5, become common mode and do not show up after M5 is turned on. Hence only the channel charge injections which occurred during the time phase (1) affects the final output error voltage. This advantage has been obtained by turning off M2 and M4 prior to M1 and M3. If M1 and M3 are turned off at the same time as M2 and M4, the output error voltage increases by about three times that of the scheme shown here.

Fig.5.9.(h) shows the simplified circuit diagram of the differential sample-hold circuit in Fig.5.9.(a) during the time phase (1) (10ns < t < 11ns). At t=10ns, MOSFETs M2 and M4 are on and V_{C2} and V_{C4} are 0 and V_{C1} is 1V and V_{C3} is -1V. During the initial part of the time phase (1) (10ns < t < 10.45ns), M2 and M4 are turned off and V_{C1} and V_{C3} increase with time due to injected charges $Q1$ and $Q3$ in Fig.5.9.(h). V_{C2} and V_{C4} increase due to injected charges $Q2$ and $Q4$ and also due to the OP Amp action as explained in the preceding section on the turn-off transient of a feedback pass-tr. During this time interval, $(V_{C1}-1V)$ and $(V_{C3}+1V)$ are voltage components due to injected charges, on capacitors $C1$ and $C3$ respectively.

During this time interval, the QS 0/100 model shows the largest charge injection and the QS 50/50 model shows the smallest charge injection among three models shown in the enlarged figures, Fig.5.9.(i) and (j). In Fig.5.9.(i), we can see that more charge is injected into $C3$ than into $C1$. This is due to the difference of polarity between V_{C1} and V_{C3}.

During the latter part of the time phase (1) (10.45ns < t < 11ns), M2 and M4 are off and charges on the capacitors are re-distributed as the steady state is approached. At the steady state when the capacitive current is settled down to 0, the voltage across source and drain nodes of M1 and M3 are 0. Also the voltage across OP Amp input nodes are essentially 0 due to the virtual ground. Hence we can find that (V_{C1} - 1V) = (V_{C3} + 1V) using the Kirchhoff voltage law.

Fig.5.9.(i) shows the charge re-distribution during the latter part of the time phase (1) (10.45ns < t < 11ns). In other words, different quantities of charges are injected on $C1$ and $C3$ during the initial part of the time phase (1) (10ns < t < 10.45ns). And the injected charges are re-distributed during the latter part of the time phase (1), so that the injected charges on $C1$ and $C3$ during the time phase (1) become equal to each other. Since more electron charge is injected into $C3$ than into $C1$ during the initial part of time phase (1), more negative charge is injected into $C4$ than into $C2$ during the re-distribution step. Hence, V_{C4} becomes larger than V_{C2} during the latter part of time phase (1) as shown in Fig.5.9.(j), due to the above mentioned mechanism and also due to OP Amp action.

Since any changes after the turn-off of M2 and M4, do not show up on the differential output error voltage after the turn-on of M5, the differential output voltage at t=12ns ((V_{C2}-V_{C4}) shown in Fig.5.9.(j)) must be equal to the differential output error voltage at t=20ns (shown in Fig.5.9.(k)).

To see this effect more clearly, the above mentioned voltage values are tabulated in Table.5.6.

Fig.5.9.(l) shows the differential output error voltage with respect to DC differential input voltage *Vin* where $Vin = Vin.p - Vin.m$. An exact symmetry with respect to center point can be observed for every model due to the differential symmetric circuit configuration. No overlap or junction capacitance is included in the simulation. Inclusion of those stray capacitances ruins the exact symmetry. For $Vin > 0$, the output error voltage is negative as explained in the preceding discussions in this section. For $Vin <$

0, the output error voltage is positive due to the symmetry. All the models show that the output error voltage is a non-linear function of DC input voltage *Vin*. The QS 0/100 model gives the smallest output error voltage and the QS 50/50 model gives the largest output error voltage. This work gives the result which is close to that of the QS 40/60 model.

Fig.5.9.(m) shows the output error voltage for three kinds of input voltage waveforms shown in Fig.5.9.(c). V_O is the superimposed differential DC voltage. Again an exact symmetry with respect to center point can be observed. The QS 50/50 model is not shown because the convergence couldn't be obtained for *Vin*:rising and *Vin*:falling.

For *Vin*:rising, all the models predict that the output error voltage is shifted in the positive direction. This is partially due to the fact that the input voltage increases with time during the sampling period (10ns < t < 10.45ns) and the real sampled voltage is slightly larger than V_O. But the difference in output error voltages between *Vin*:DC and *Vin*:rising is not constant for different V_O's.

This work gives results which are close to those of the QS 40/60 model. But the QS 0/100 model gives results which are much different from the other two models.

SPICE3 input file for the differential sample-hold circuit is shown in Fig.5.9.(m). The run statistics of the differential sample-hold circuit are shown in Table.5.7. This work takes about 3 times longer than other QS models.

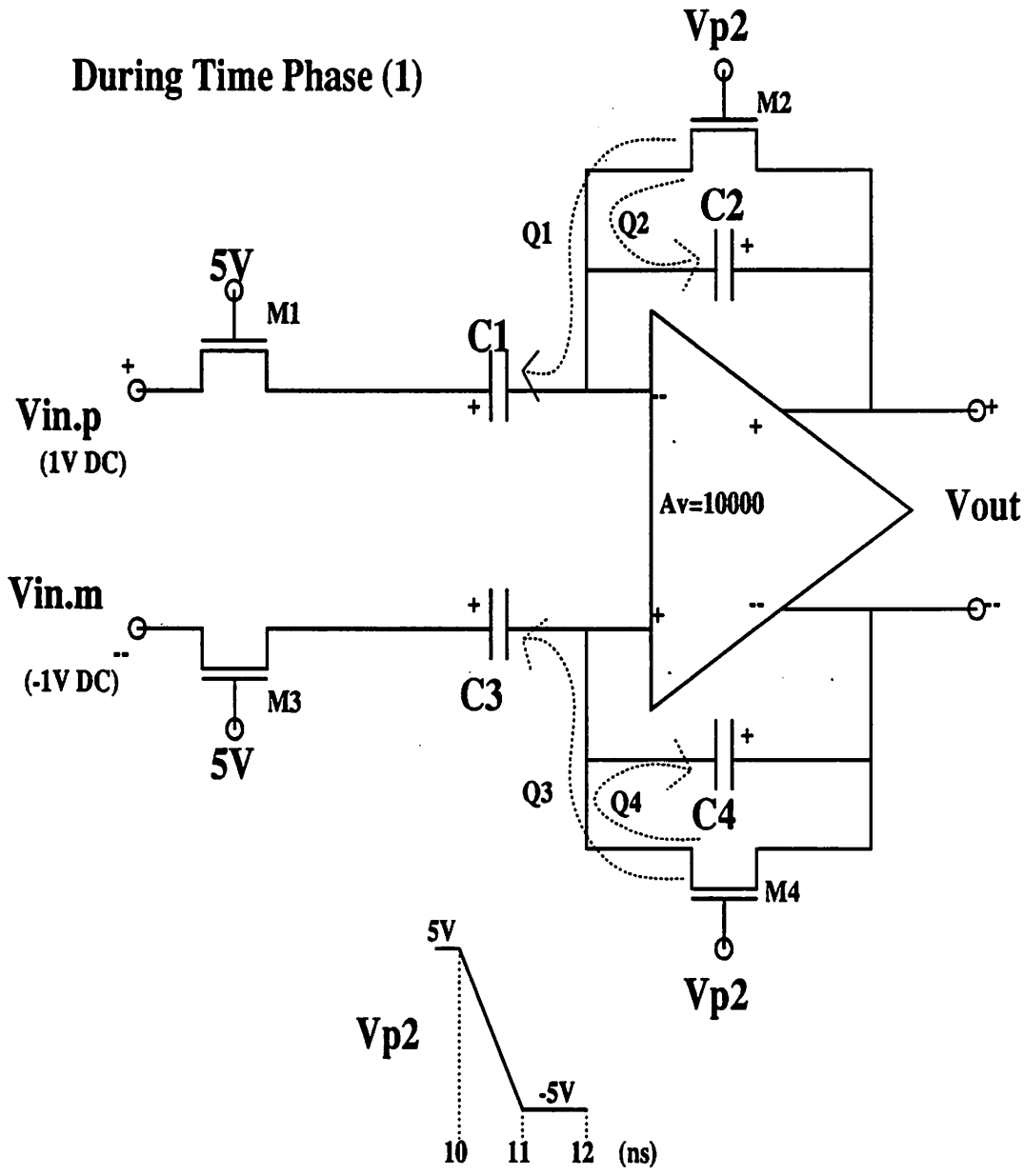


Fig.5.9.(h). The simplified circuit schematic of the differential sample-hold circuit in Fig.5.9.(a) during the time phase (1) (10ns ≤ t ≤ 11ns).

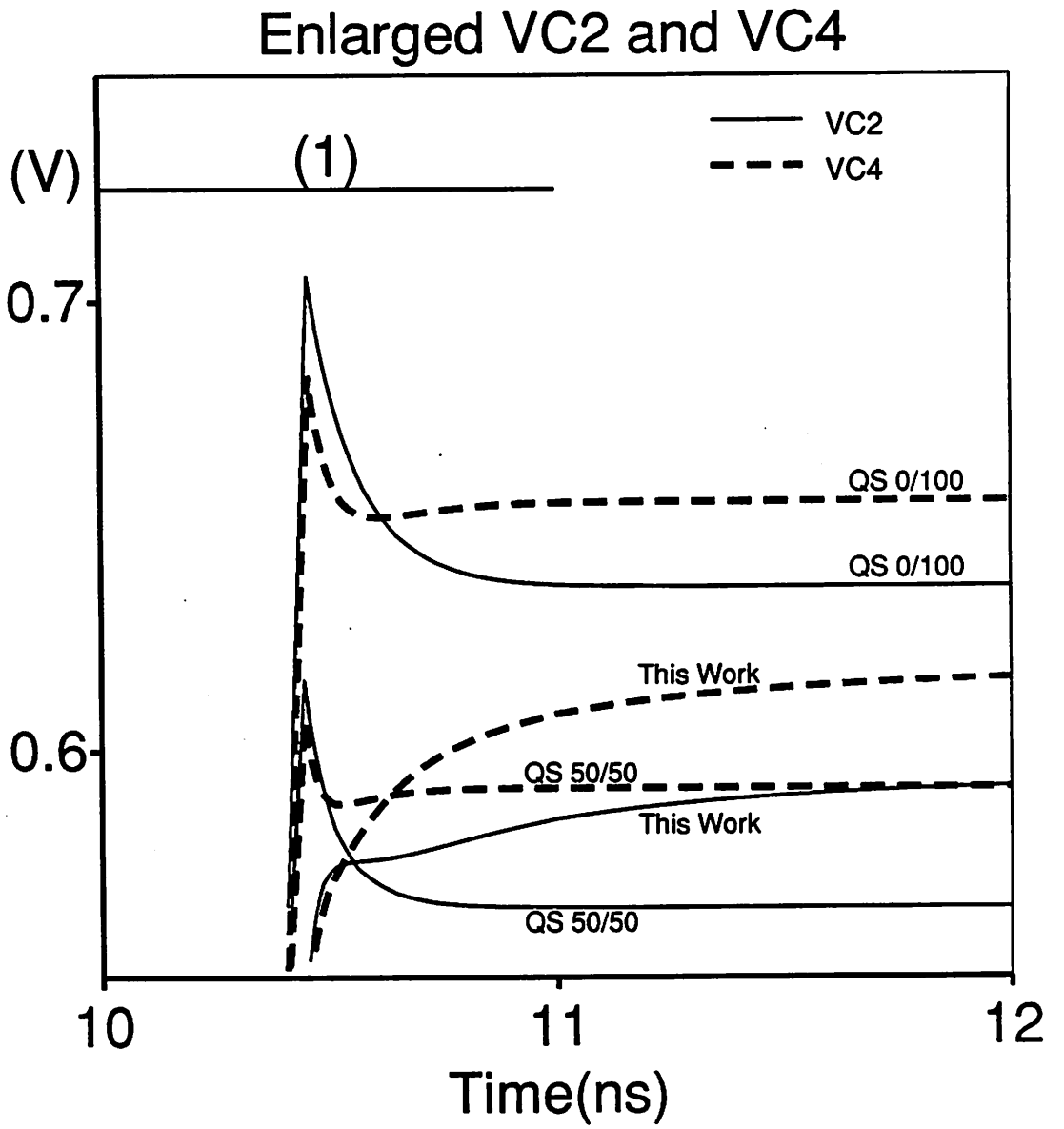


Fig.5.9.(i). Enlarged voltage waveforms, (VC1-1V) and (VC3+1V) during the time phase (1). VC1 and VC3 are voltages across capacitors C1 and C3 respectively. Vin.p = 1V(DC) and Vin.m = 1V(DC).

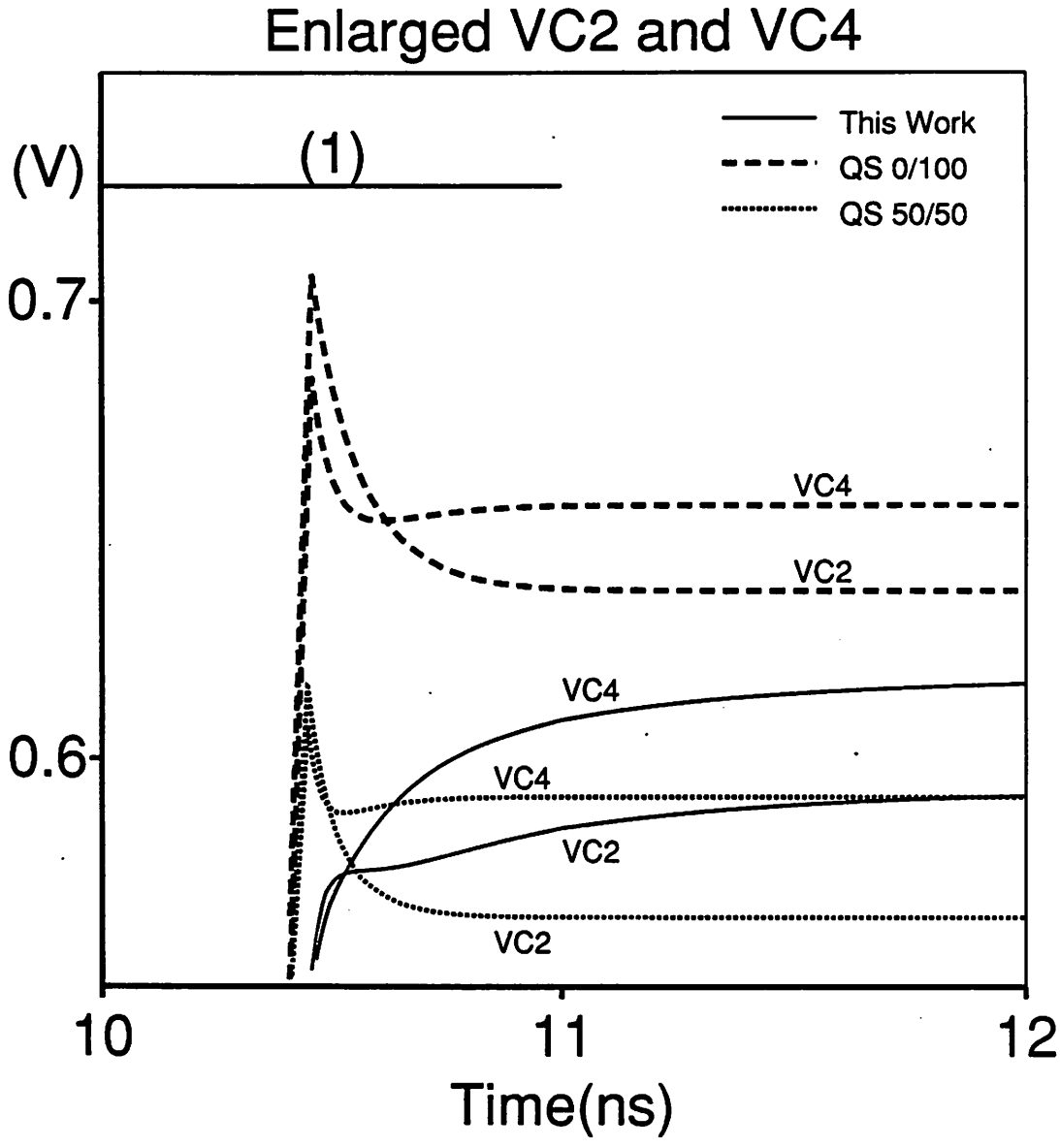


Fig.5.9.(j). Enlarged voltage waveforms, VC2 and VC4, across capacitors C2 and C4 respectively, during the time phase (1), with Vin.p = 1V(DC) and Vin.m = -1V(DC).

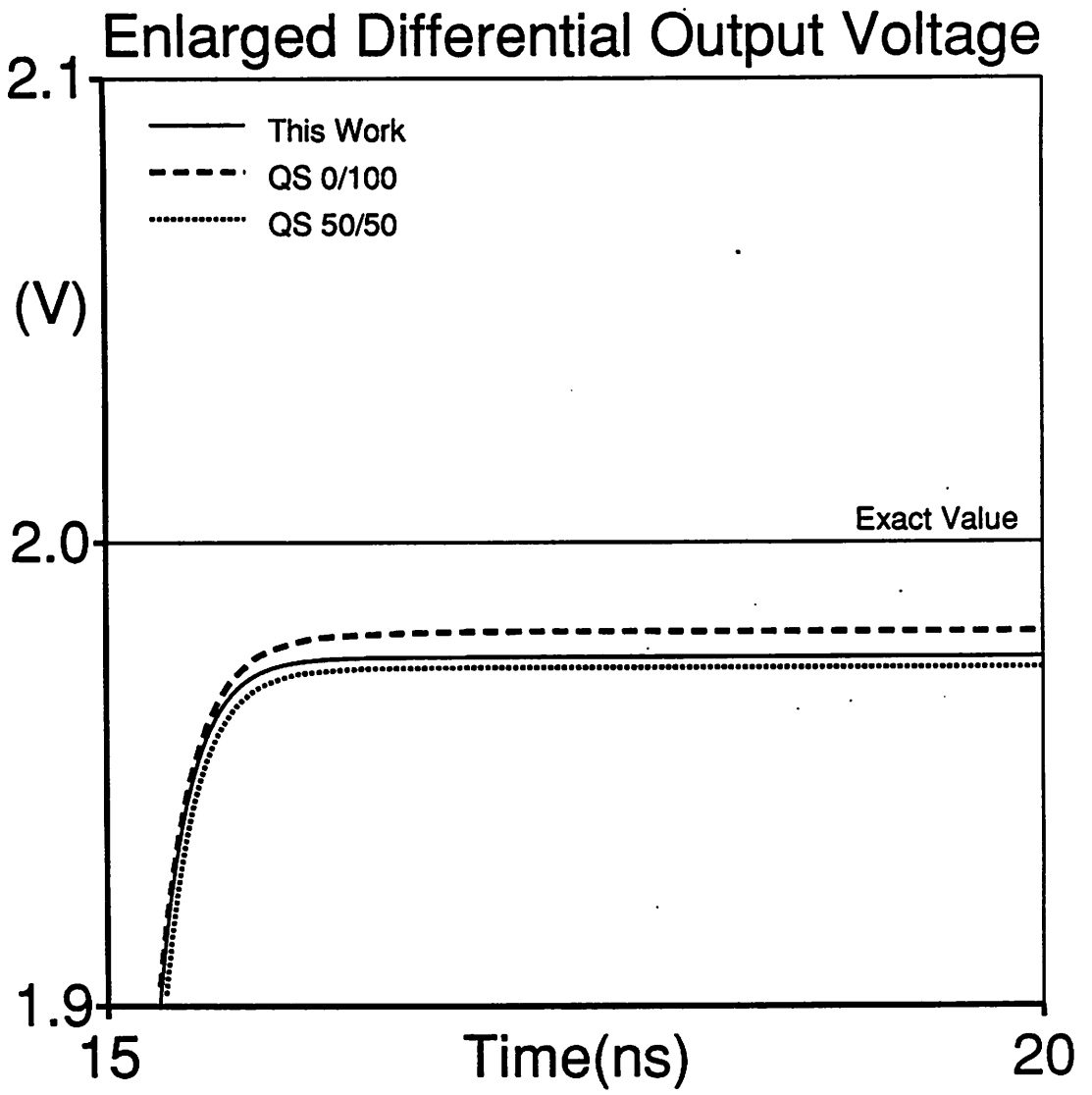


Fig.5.9.(k). Enlarged final differential output voltage for the circuit in Fig.5.9.(a), with Vin.p = 1V(DC) and Vin.m = -1V(DC). This waveform is enlarged from that in Fig.5.9.(d).

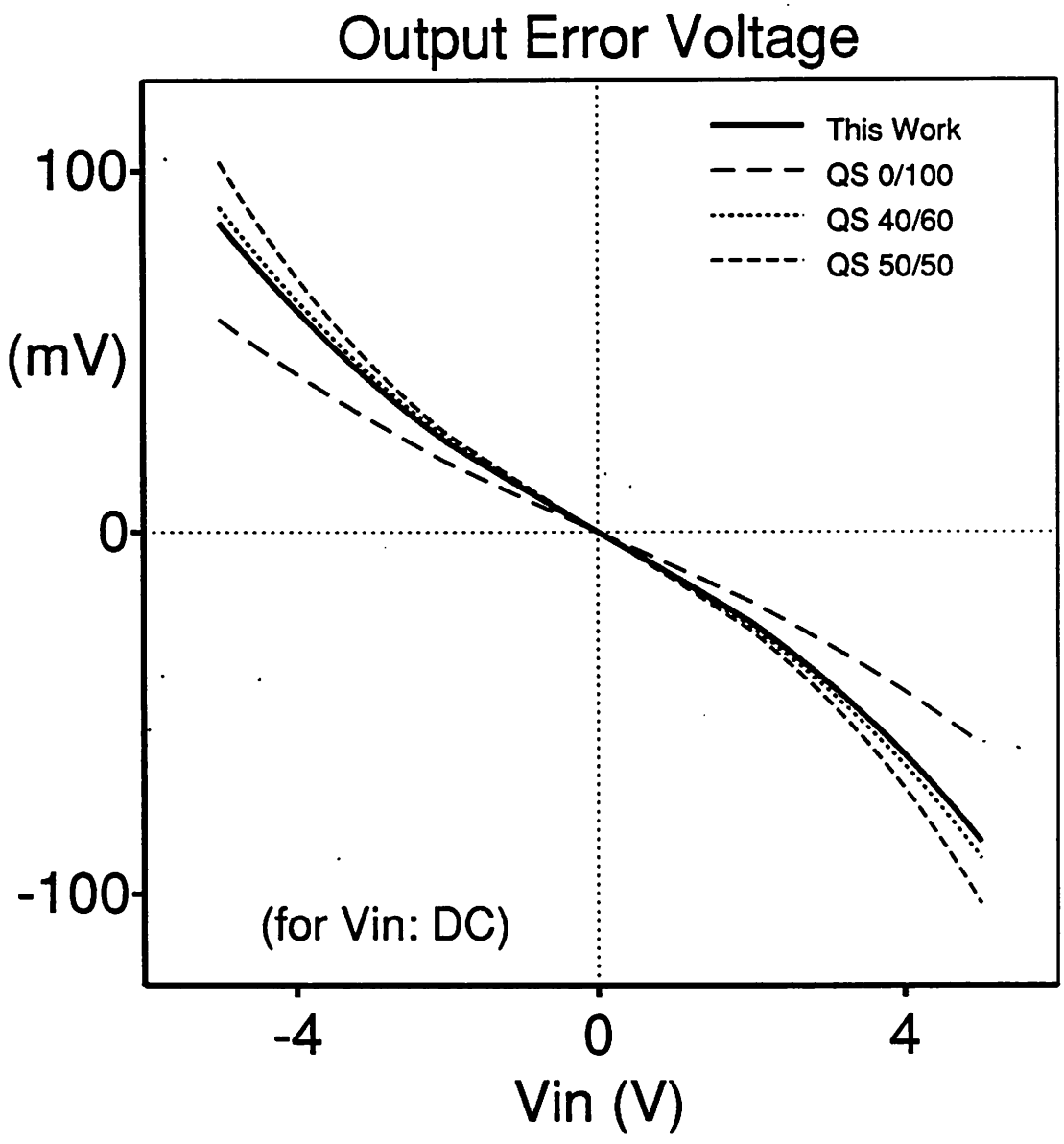


Fig.5.9.(l). The output error voltage versus the differential input voltage Vin, for the differential sample-hold circuit in Fig.5.9.(a).

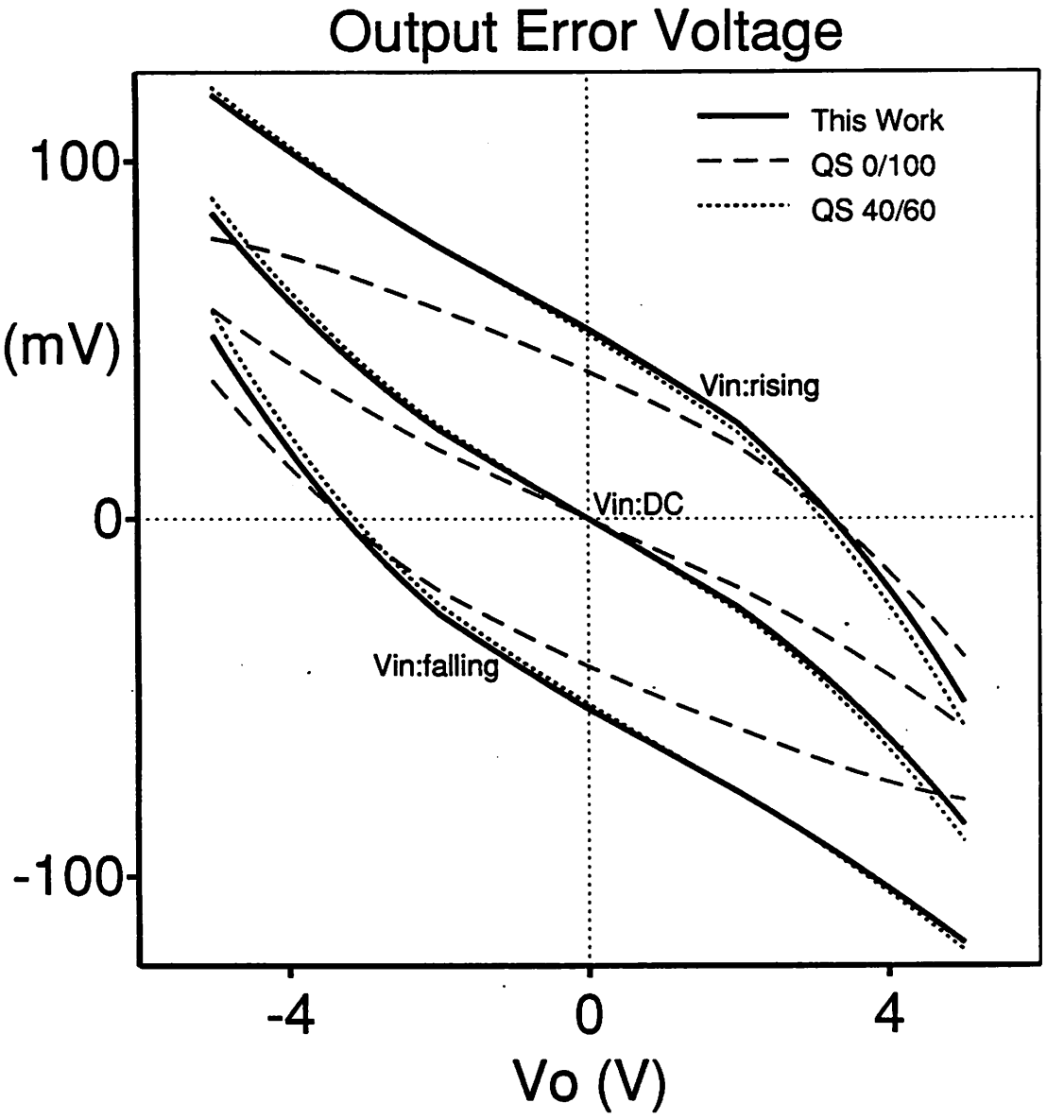


Fig.5.9.(m). The output error voltage versus the superimposed DC voltage Vo for three kinds of input waveforms shown in Fig.5.9.(c).

<SPICE3 Input File for Diff. S/H Circuit>

```
Sample and Hold Circuit for MOS A/D Converter
M1 1  20 2 50 nch w=200u l=3u ad=1200p as=1200p
M2 3  30 4 50 nch w=200u l=3u ad=1200p as=1200p
M3 5  20 6 50 nch w=200u l=3u ad=1200p as=1200p
M4 7  30 8 50 nch w=200u l=3u ad=1200p as=1200p
M5 2  40 6 50 nch w=20u l=3u ad=1200p as=1200p
C1 2 3  1.0pF
C2 3 4  1.0pF
C3 6 7  1.0pF
C4 7 8  1.0pF
*
*Ideal OP amp with differential I/O
* pos. inp.  (7)     neg. inp. (3)
* pos. out.  (4)     neg. out. (8)
* Voltage gain of 80dB
Eop1 4 0 7 3  5000
Eop2 8 0 7 3 -5000
*
* more realistic OP Amp for future simulation
* DC gain 10000(80dB)
* Unity gain freq.  3.2GHz
*Eop1 9 0 7 3 5000
*Rop1 9 4 100k
*Cop1 4 0 5pF
*Eop2 10 0  7 3 -5000
*Rop2 10 8 100k
*Cop2 8  0  5pF
*
* Voltage Sources
Vp1 20 0 pulse -5 5  0  1n 1n 11n 20n
Vp2 30 0 pulse -5 5  0  1n 1n 9n 20n
Vp3 40 0 pulse -5 5 14n 1n 0n 5n 20n
Vbb 50 0 dc -5
```

* (continued in the next page)

```
* sampled at Vin=2V DC
VinP 1 0 dc   1.0
VinM 5 0 dc  -1.0
*  sampled at Vin=1V falling edge
*VinP 1 0 dc 0.0 sin ( 0.5 -1.0 25MEG -30n 0)
*VinM 5 0 dc 0.0 sin (-0.5  1.0 25MEG -30n 0)
*  sampled at Vin=1V rising edge
*VinP 1 0 dc 0.0 sin ( 0.5  1.0 25MEG -30n 0)
*VinM 5 0 dc 0.0 sin (-0.5 -1.0 25MEG -30n 0)
.tran 0.1n 20n
.ic v(2)=0 v(3)=0 v(4)=0 v(6)=0 v(7)=0 v(8)=0
.print tran v(2,3) v(6,7) v(4,3) v(8,7) v(4,8)
.opt acct
*
**********************************************
*   Model Parameters for This Work
**********************************************
.model nch nmos level=6 vfb=-0.77 tox=0.018
+      nsub=2e16 u0=500 lambda=0.03 pb=0.7
+      pbsw=0.7 js=1e-15 cgdo=0 cgso=0 cj=0
+      cjsw=0 mintol=1 qtrtol=30
*
**********************************************
*   Model Parameters for BSIM SPICE3
*       xpart = 0 ----> BSIM 40/60
*       xpart = 1 ----> BSIM 0/100
*       xpart = 2 ----> BSIM 50/50
**********************************************
*.model nch nmos level=4 xpart=0 vfb=-0.77
*+     phi=0.958 k1=0.425 muz=500 mus=520
*+     tox=0.018 vdd=5 n0=1 cgdo=0 cgso=0
*+     cj=0 cjsw=0 js=1e-15 pb=0.7 pbsw=0.7
*
.end
```

Fig.5.9.(n). The SPICE3 input file for the simulation of the circuit in Fig.5.9.(a).

Model	V_{C2}(12n)	V_{C4}(12n)	V_{C2}(12n)-V_{C4}(12n)	$V_{OUT.ERROR}$(20n)
This Work	592.0	616.3	-24.3	-24.8
QS 40/60	578.3	655.5	-18.9	-19.3
QS 0/100	636.6	591.7	-26.4	-27.0
QS 50/50	565.3	603.9	-25.6	-26.0

[unit: mV]

Table.5.6. Table to show that the differential output error voltage at t=20ns, $V_{OUT.ERROR}$(20n), is equal to the differential output voltage, V_{C2}(12n)-V_{C4}(12n), at the end of the time phase (1).

Model	This Work	QS 40/60	QS 0/100	QS 50/50
Total Run Time(sec)	26.8	8.8	8.4	11.1
# of Iterations	1390	854	855	1021
Total Time Points	553	339	346	364
Rejected Time Pts	38	70	74	89
Load Time(sec)	20.4	5.5	5.7	7.0

Table.5.7. Comparison of run statistics for the simulation of the differential sample-hold circuit shown in Fig.5.9.(a). The SPICE3 input file for this comparison is shown in Fig.5.9.(m).

5.4. Implementation into SPICE3

This charge conserving NQS(non-quasistatic) MOSFET model, using the simple DC current model shown in eq. (5.40), has been implemented in SPICE3B.1 [5.15]. It is based on the charge sheet formulation which is described in Chapter 2. The details of this implementation is shown in Appendix.11.

< Charge conserving NQS charge model combined with Level-2 DC model >

Also this NQS charge model has been combined with the SPICE Level-2 DC model [5.13], and has been implemented into SPICE3B.1 as a level=5 model. Essentially any DC model can be combined with this NQS charge model, but the SPICE Level-2 DC model is chosen because it includes all the secondary effects and also because the level-2 model parameters are widely available in the industry and it makes the new model easy to use.

In this implementation, the DC drain current shown in (5.40) and also in (A11.7), is replaced by the DC drain current computed using the SPICE Level-2 DC model [5.13]. And the drain saturation voltage V_{DSSAT} in (A11.1) is replaced by V_{DSSAT} computed from the SPICE Level-2 DC model. This new V_{DSSAT} includes all the short channel and narrow channel effects.

In this way, all the short channel effects, such as the velocity saturation effect, the V_{GS} dependence of mobility and all other aspects considered by the SPICE Level-2 model are also included in the implementation. It is true that, while the DC model includes all the short channel effects, the charge model is still based on the long channel theory although short channel effects are partially included in the charge model through V_{DSSAT}. This compromise should be judged in the light of the fact that all the conventional charge (capacitance) models in SPICE (Meyer, Ward-Dutton, BSIM) don't include any short channel effects either.

The details of this implementation are shown in Appendix.12.

5.5. Conclusion

An analytic charge conserving NQS(non-quasistatic) model for long channel MOSFETs has been derived and implemented in SPICE3. It is based on an approximate solution to the current continuity equation. Analytic equations are derived for node charges to guarantee charge conservation.

Comparison of this work with PISCES(2-D device simulation) and 1-D numerical solution of current continuity equation shows good agreements, but QS(quasistatic) models show quite different results from this work.

Channel charge partitioning ratios are shown to be a complex function of history, signal rise and fall times and node voltages. But QS models give partitioning ratios which are functions of node voltages only and hence incorrect partitioning ratios for fast transients.

Also this NQS charge model has been combined with the SPICE Level-2 DC model and has been implemented in SPICE3. In this way, all the short channel effects considered by the SPICE Level-2 DC model have been included in the implementation.

Circuit simulations using this model in SPICE3 are about three to four times slower than those using conventional MOSFET models.

References

[5.1]. D.A.Hodges, P.R.Gray, R.W.Brodersen, "Potential of MOS Technologies for Analog Integrated Circuits", IEEE Journal of Solid State Circuits, SC-13, 1978, pp.285-294

[5.2]. D.E.Ward, R.W.Dutton, "A Charge-Oriented Model for MOS Transistor Capacitances", IEEE Journal of Solid State Circuits, SC-13, No.5, pp.703-707, Oct. 1978

[5.3]. P.Yang, B.D.Epler, P.K.Chatterjee, "An Investigation of the Charge Conservation Problem for MOSFET Circuit Simulation", IEEE Journal of Solid State Circuits, SC-18, No.1, pp.128-138, Feb. 1983

[5.4]. K.L.Lee, "Low Distortion Switched Capacitor Filters", ERL Memo M86/12, University of California, Berkeley, Feb. 1986

[5.5]. P.Li, M.Chin, P.R.Gray, R.Castello, "Ratio-Independent Algorithmic Analog-to-Digital Conversion Techniques", IEEE Journal of Solid State Circuits, SC-19, pp.828-836, 1984

[5.6] C.Turchetti, P.Mancini, G.Masetti, "A CAD-Oriented Non-Quasistatic Approach for the Transient Analysis of MOS IC's", IEEE Journal of Solid State Circuits, SC-21, pp.827-836, Dec. 1986

[5.7]. H.J.Park, P.K.Ko, C.Hu, "A Non-Quasistatic MOSFET Model for SPICE", IEEE IEDM Technical Digest, pp.652-655, Dec. 1987

[5.8]. H.J.Park, P.K.Ko, C.Hu, "A Non-Quasistatic MOSFET Model for SPICE - Transient Analysis", will be published in IEEE Trans. Electron Devices, March 1989

[5.9]. H.J.Park, P.K.Ko, C.Hu, "A Non-Quasistatic MOSFET Model for SPICE - AC Analysis", submitted to IEEE Trans. Electron Devices

[5.10]. J.E.Meyers, "MOS Models and Circuit Simulation", RCA Review, Vol.32, pp.42-63, Mar. 1971

[5.11]. B.J.Sheu, D.L.Scharfetter, P.K.Ko, "SPICE2 Implementation of BSIM", University of California, Berkeley, ERL Memo. M85/42, May 1985

[5.12]. M.R.Pinto, C.S.Rafferty, R.W.Dutton, PISCES-II User's Manual, Stanford Electronics Lab., Stanford University, Palo Alto, CA, 1984

[5.13]. A.Vladmirescu, S.Liu, "The Simulation of MOS Integrated Circuits using SPICE2G.6", University of California, Berkeley, ERL Memo. M80/7, Oct. 1980

[5.14]. subroutine "MOSQ2", SPICE2G.6 source code, Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley, 1983

[5.15]. T.L.Quarles, A.R.Newton, D.O.Pederson, A.Sangiovanni-Vincentelli, "SPICE3B1 User's Guide", Electronics Research Lab., University of California, Berkeley, Nov. 1987

[5.16]. L.W.Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits", ERL Memo M520, University of California, Berkeley, May 1975

[5.17]. E.Cohen, "Program Reference for SPICE2", ERL Memo. ERL-M592, 1976, University of California, Berkeley

[5.18]. T.L.Quarles, Ph.D. Dissertation, will be published, ERL Memo., 1989

Chapter 6

SUMMARY AND CONCLUSION

Analytic equations have been derived for charge sheet and non-quasistatic MOSFET models, and have been implemented in SPICE3.

6.1. Charge sheet MOSFET capacitance model

An analytic charge sheet capacitance model for short channel MOSFETs has been derived and implemented in SPICE3. It is based on the surface potential formulation in which an approximation is made to find the surface potential without iterations and without losing accuracy. An automatic gate capacitance measurement system with r.m.s. resolution of 14aF has been set up, for the comparison of this model with measured data and extraction of model parameters. Excellent agreements have been observed between the gate capacitances predicted by this model and the measured data for long and short channel MOSFETs respectively. The CPU time for model computation is about twice longer in this model than in conventional SPICE models.

6.2. NQS(Non-quasistatic) MOSFET model

Analytic non-quasistatic MOSFET models have been derived based on the approximate solution of the current continuity equation. They have been implemented in SPICE3 and several circuits are simulated using these NQS models, PISCES, 1-D numerical simulation of current continuity equation, and other QS models. Good agreements have been observed among these NQS models, PISCES, and 1-D numerical solution, but QS models give quite different results from NQS models.

Two NQS transient models are shown in this thesis. In one transient model described in Chapter 3, the current continuity equation is converted to diffusion equation and analytic equations have been derived for node currents from the solution of this diffusion equation. But it turns out that it doesn't conserve charges. In the other transient model described in Chapter 5, analytic equations are derived for node charges. Hence it is a charge-based model and it guarantees the charge conservation.

The channel charge partitioning ratio is shown to be not constant in saturation region, but is a complex function of history, signal transition rate, and node voltages. But all the QS models show that the channel charge partitioning ratio is a function of node voltages only.

These NQS transient models implemented in SPICE3 have been applied to several example circuits including the channel charge injection problem of switched analog circuits, and the results are compared with those from PISCES, 1-D numerical solution, and other QS models. Excellent agreements have been observed among this work, PISCES, and 1-D numerical solution of current continuity equation.

An analytic NQS AC MOSFET model has been derived from the solution of the above mentioned diffusion equation. This work and 1-D numerical solution show that the AC channel charge partitioning ratio in saturation region is 40/60 at low frequencies and becomes 0/100 at very high frequencies compared to the inverse of channel transit time. 40/60 and 0/100 refer to the ratio between AC drain and source charges.

Appendix 1

Program to compute surface potentials for long channel MOSFETs

This program is the implementation of the algorithm in Fig.2.4 for finding the surface potential as a function of $(V_{GB}-V_{FB})$ and V_{CB} for long channel MOSFETs.

```
/* program to compute the surface potential for long channel MOSFETs
*/
#define N_ROW        201
#define LINEAR       1
#define PARABOLIC    2
#define CUBIC        3
#include <math.h>
#include <stdio.h>

/* Global Variables */
double vgb_vfb[N_ROW],psis[N_ROW];
double Phi;
double beta, Gamma;
double Vso,Vgo,VsoRef,VgoRef;
double a[N_ROW],b[N_ROW],c[N_ROW],d[N_ROW];
int    Nmax;

/* main routine */
main()    {
double Vgb_Vfb,Phic_Vbs,PSIS,dPSISdVgs;
double dPSISdVds,dPSISdVbs;
FILE   *fp,*fpg,*fpd,*fpb;

  fp = fopen("VS","w");
  fpg = fopen("DG","w");
  fpd = fopen("DD","w");
  fpb = fopen("DB","w");
  PSISSetup();

  for( Phic_Vbs = 0.0; Phic_Vbs <= 5.0 ; Phic_Vbs += 1.) {
    for(Vgb_Vfb = -2.; Vgb_Vfb <= 8. ; Vgb_Vfb += 0.05)    {
      SurfacePotential(Vgb_Vfb,Phic_Vbs,&PSIS,
          &dPSISdVgs,&dPSISdVds,&dPSISdVbs);
      fprintf(fp,"%f  %f 0,Vgb_Vfb,PSIS);
```

```
        fprintf(fpg,"%f  %f 0,Vgb_Vfb,dPSISdVgs);
        fprintf(fpd,"%f  %f 0,Vgb_Vfb,dPSISdVds);
        fprintf(fpb,"%f  %f 0,Vgb_Vfb,dPSISdVbs);
    }
  fprintf(fp,"%c Phic_Vbs = %f %c 0,'"',Phic_Vbs,'"');
  fprintf(fpg,"%c Phic_Vbs = %f %c 0,'"',Phic_Vbs,'"');
  fprintf(fpd,"%c Phic_Vbs = %f %c 0,'"',Phic_Vbs,'"');
  fprintf(fpb,"%c Phic_Vbs = %f %c 0,'"',Phic_Vbs,'"');
  }
}

/* function to compute the surface potential PSISP, and their
/* derivatives, dPSISdVgsP, dPSISdVdsP, dPSISdVbsP
*/
SurfacePotential(Vgb_Vfb,Phic_Vbs,PSISP,dPSISdVgsP,dPSISdVdsP,dPSISdVbsP)
double  Vgb_Vfb,Phic_Vbs;
double  *PSISP,*dPSISdVgsP,*dPSISdVdsP,*dPSISdVbsP;
{
    double Vs1,Vg1,VGoPrime,VSoPrime,Qn,dVG,dVG_dVG;
    double Vs1Ref,Vg1Ref,dVg1RdVDB,dVGoPdVDB,dVSoPdVGB,dVSoPdVDB;
    double dPSISdVGB,dPSISdVDB,dQndVGB,dQndVDB;
    double SqrtVGB,SqrtVs1,SqrtVs1Ref,SqrtVGB_Qn,temp,SqrtVSoP;
    double Gamma_Gamma;
    int i;
  Vs1 = Vso + Phic_Vbs;
  SqrtVs1 = sqrt(Vs1 - 1./beta);
  Vg1 = Vs1 + Gamma * SqrtVs1;
  Gamma_Gamma = Gamma * Gamma ;

  /* Accumulation Region */
  if(Vgb_Vfb < Vgo )      {
      i = Search_Vgb_Vfb(Vgb_Vfb);
      dVG = Vgb_Vfb - vgb_vfb[i];
      dVG_dVG = dVG * dVG;
      *PSISP = a[i]*dVG*dVG_dVG+b[i]*dVG_dVG +c[i]*dVG+d[i];
      dPSISdVGB = 3.*a[i]*dVG_dVG + 2.*b[i]*dVG + c[i];
      dPSISdVDB = 0.0;

  /* Inversion Region */
   } else if(Vgb_Vfb >= Vg1 )  {
      Vs1Ref = VsoRef + Phic_Vbs;
      SqrtVs1Ref = sqrt(Vs1Ref - 1./beta);
      Vg1Ref = Vs1Ref + Gamma * SqrtVs1Ref;
      VGoPrime = VgoRef + Vgb_Vfb - Vg1Ref;
      dVg1RdVDB = (1. + 0.5 * Gamma / SqrtVs1Ref);
      dVGoPdVDB = - dVg1RdVDB;
      i = Search_Vgb_Vfb(VGoPrime);
      dVG = VGoPrime - vgb_vfb[i];
      dVG_dVG = dVG * dVG;
      VSoPrime = a[i]*dVG*dVG_dVG+b[i]*dVG_dVG +c[i]*dVG+d[i];
      dVSoPdVGB = 3. * a[i] * dVG_dVG + 2. * b[i] * dVG + c[i];
      dVSoPdVDB = dVSoPdVGB * dVGoPdVDB;
      SqrtVSoP = sqrt(VSoPrime - 1./beta);
```

```
    /* Qn : inversion charge  / Cox (Volt) */
    Qn = VGoPrime - VSoPrime - Gamma * SqrtVSoP;
    if(Qn < 0.0 )     {
          Qn = 0.0;
          dQndVGB = 0.0;
          dQndVDB = 0.0;
        } else {
          temp = 1. + 0.5 * Gamma / SqrtVSoP;
          dQndVGB = 1.0 - temp * dVSoPdVGB;
          dQndVDB = dVGoPdVDB - temp * dVSoPdVDB;
        }
    SqrtVGB_Qn = sqrt(Vgb_Vfb-Qn+Gamma_Gamma*0.25-1./beta);
    *PSISP = Vgb_Vfb - Qn + 0.5 * Gamma_Gamma - Gamma * SqrtVGB_Qn;
    dPSISdVGB = 1. - dQndVGB - 0.5*Gamma/SqrtVGB_Qn*(1.-dQndVGB);
    dPSISdVDB = - dQndVDB + 0.5*Gamma/SqrtVGB_Qn*dQndVDB;

  /* Depletion Region */
     } else   {
    SqrtVGB = sqrt( Vgb_Vfb + 0.25 * Gamma_Gamma - 1. / beta );
    *PSISP = Vgb_Vfb + 0.5 * Gamma_Gamma
                - Gamma * SqrtVGB;
    dPSISdVGB = 1. - 0.5 * Gamma / SqrtVGB;
    dPSISdVDB = 0.0;
     }
  if(dPSISdVDB > 1.0) dPSISdVDB = 1.0;
  *dPSISdVgsP = dPSISdVGB;
  *dPSISdVdsP = dPSISdVDB;
  *dPSISdVbsP = -dPSISdVGB - dPSISdVDB;
}

/* function to set up the cubic spline function for the surface potential at Vcb=0
*/
PSISSetup()   {
  double Cox,PSIS,Nsub,Tox,Ni;
  double Esi, q, Eox, Term1, Term2, Term3, Term4, Term5;
  int    N,Search_Vgb_Vfb();

  Nsub = 2.5e16;
  Tox = 200.0;
  Ni = 1.45e10;
  Esi = 8.854e-14 * 11.7;
  Eox = 8.854e-14 * 3.9;
  q = 1.602e-19;
  beta = 1.0 / 0.026;
  Phi = 2. * .026 * log( Nsub / Ni);
  Tox = Tox * 1.e-8;
  Cox = Eox / Tox;
  Gamma = sqrt(2. * q * Nsub * Esi )/Cox;

  Vso = 0.5 * Phi;
  Vgo = Vso + Gamma * sqrt(Vso-1./beta);
  VsoRef = Phi;
  VgoRef = VsoRef + Gamma * sqrt(VsoRef - 1./beta);
```

```
  N= 0;
  for ( PSIS = -0.4; PSIS < 1.1; PSIS += 0.01 )  {
          Term1 = exp(-beta*PSIS) - 1.0;
          Term2 = exp(beta*(PSIS-Phi));
          Term3 = exp(beta*(-Phi));
          Term5 = PSIS + 1./beta*(Term1+Term2-Term3);
          if(Term5 > 1.e-10) Term4 = Gamma * sqrt(Term5);
          else               Term4 = 0.0;
          if(PSIS < 0.0) Term4 = - Term4;
          vgb_vfb[N] =   PSIS + Term4;
          psis[N] = PSIS;
          N ++;
         }
        Nmax = N - 1;

  cubic_spline(LINEAR);

}             /* End of PSIS setup */


int Search_Vgb_Vfb(Vgb_Vfb)
    double Vgb_Vfb;
{
    int Half,Min,Max,Nmax_1,Prev_Half;
    Nmax_1 = Nmax - 1 ;
    if(Vgb_Vfb < vgb_vfb[2]) return(1);
    else if(Vgb_Vfb > vgb_vfb[Nmax_1]) return(Nmax_1);
    else {            /* Binary Search */
    Half = 1;
    Min  = 1;
    Max  = Nmax;
    Prev_Half = 32000;
    Iter:
      Half = (Min + Max)/2;
      if(Vgb_Vfb > vgb_vfb[Half]) Min = Half;
      else                        Max = Half;
      if(Half == Prev_Half) return(Half);
      else  {
        Prev_Half = Half;
        goto Iter;
      }
    }
}

/* function to compute cubic spline coefficients A[i,j] (j=1,2,3,4)
/* from a given set of (x,y) data, vgb_vfb[i] and psis[i]
*/
cubic_spline(Method)
int Method;
{
     double dX1,dX2,dY1,dY2,dXN_2,dXN_1,dXi;
     double A[N_ROW][5];
```

```
double S[N_ROW];
int    i,j,N,N_1,N_2,ip1,ip2,im1;

/*    Method = LINEAR      S(1) = S(Nmax) = 0
/*    Method = PARABOLIC  S(1) = S(2)
/*                        S(Nmax) = S(Nmax-1)
/*    Method = CUBIC       S(1) S(Nmax)
/*                              extrapolated
*/
N = Nmax;
N_2 = N - 2;
N_1 = N - 1;
dX1 = vgb_vfb[2] - vgb_vfb[1];
dY1 = (psis[2] - psis[1]) / dX1 * 6.0;
for(i=1; i <= N_2 ; i++)     {
 ip2 = i + 2;
 ip1 = i + 1;
 dX2 = vgb_vfb[ip2] - vgb_vfb[ip1];
 dY2 = (psis[ip2] - psis[ip1]) / dX2 * 6.0;
 /* matrix coefficient */
      A[i][1] = dX1;
 A[i][2] = 2. * (dX1 + dX2);
 A[i][3] = dX2;
 /* R.H.S. vector */
 A[i][4] = dY2 - dY1;
 dX1 = dX2;
 dY1 = dY2;
 }

 /* Adjust some coefficients */
 switch(Method)       {
 case LINEAR:
     break;
      case PARABOLIC:
     A[1][2] = A[2][2] + vgb_vfb[2]
                    - vgb_vfb[1];
     A[N_2][2] = A[N_2][2] + vgb_vfb[N]
                        - vgb_vfb[N_1];
     break;
      case CUBIC:
     dX1 = vgb_vfb[2] - vgb_vfb[1];
     dX2 = vgb_vfb[3] - vgb_vfb[2];
     A[1][2] = (dX1 + dX2) * (dX1 + 2.*dX2)/dX2;
     A[1][3] = (dX2 * dX2 - dX1 * dX1) / dX2;
     dXN_2 = vgb_vfb[N_1] - vgb_vfb[N_2];
     dXN_1 = vgb_vfb[N] - vgb_vfb[N_1];
     A[N_2][1] = (dXN_2 * dXN_2 - dXN_1 * dXN_1)
              / dXN_2;
     break;
 }

 /* solve the tridiagonal matrix equation */
 for(i=2; i <= N_2 ; i++)        {
```

```
im1 = i - 1;
A[i][2] = A[i][2] - A[i][1] / A[im1][2] * A[im1][3];
A[i][4] = A[i][4] - A[i][1] / A[im1][2] * A[im1][4];
}

/* Back substitution */
A[N_2][4] = A[N_2][4] / A[N_2][2];
for( i=2; i <= N_2 ; i++)    {
j = N_1 - i;
A[j][4] = (A[j][4] - A[j][3] * A[j+1][4]) / A[j][2];
}

/* Put the values to S vector */
for(i=1; i <= N_2 ; i++)
S[i+1] = A[i][4];

/* Get S[1] S[N] according to Method */
switch(Method)    {
case LINEAR:
    S[1] = 0.0;
    S[N] = 0.0;
    break;
     case PARABOLIC:
    S[1] = S[2];
    S[N] = S[N_1];
    break;
     case CUBIC:
    S[1] = ((dX1 + dX2) * S[2] + dX1 * S[3]) / dX2;
    S[N] = ((dXN_2 + dXN_1)*S[N_1] - dXN_1 * S[N_2]) / dXN_2;
         break;
}

/* Compute the cubic spline coefficients
/* for each interval
*/
for( i=1 ; i <= N; i++)    {
  ip1 = i + 1;
  dXi = vgb_vfb[ip1] - vgb_vfb[i];
  a[i] = (S[ip1] - S[i]) / 6. / dXi;
  b[i] = 0.5 * S[i];
  c[i] = (psis[ip1] - psis[i]) / dXi - (2. * dXi * S[i] + dXi * S[ip1]) / 6.;
   d[i] = psis[i];
}

}
```

Appendix 2

Implementation of the CSM model into SPICE3

This charge sheet MOSFET (CSM) model has been implemented into SPICE3B1. SPICE3 is very modular and each device routine is quite independent of other device model routines. The details of implementation are almost the same as those in Appendix 12 which shows the implementation details of the NQS (Non-Quasistatic) model into SPICE3.

Both SPICE2 and SPICE3 use the modified nodal analysis to solve the circuit equation (A2.1).

$$\mathbf{A}\cdot\mathbf{X} = \mathbf{B} \tag{A2.1}$$

where $\mathbf{X}$ is a column matrix for node voltages and currents of independent voltage sources. Since there are no voltage sources in the equivalent circuit of MOSFET, (A2.1) is simplified to (A2.2) for the part of circuits consisting of MOSFETs only.

$$\mathbf{Y}\cdot\mathbf{V} = \mathbf{I}_{EQ} \tag{A2.2}$$

where $\mathbf{V}$ is a column matrix for node voltages, $\mathbf{Y}$ is a square admittance matrix and $\mathbf{I}_{EQ}$ is a column matrix for equivalent currents incident to each node.

(1) DC analysis

The equivalent circuit of MOSFET for DC analysis is shown in Fig.A2.1. In Fig.A2.1, *gs* and *gd* are conductances of source, drain series resistances respectively. S and D are external source and drain nodes as specified in SPICE input. S' and D' are internal source node and internal drain node respectively. G and B are gate and bulk nodes respectively. *cdrain* is DC drain current I_D which is shown in (2.25). *cbs* and *cbd* are DC source, drain junction currents respectively. *cdrain*, *cbs* and *cbd* are nonlinear functions of node voltages. To solve the nonlinear circuit equation by the nonlinear

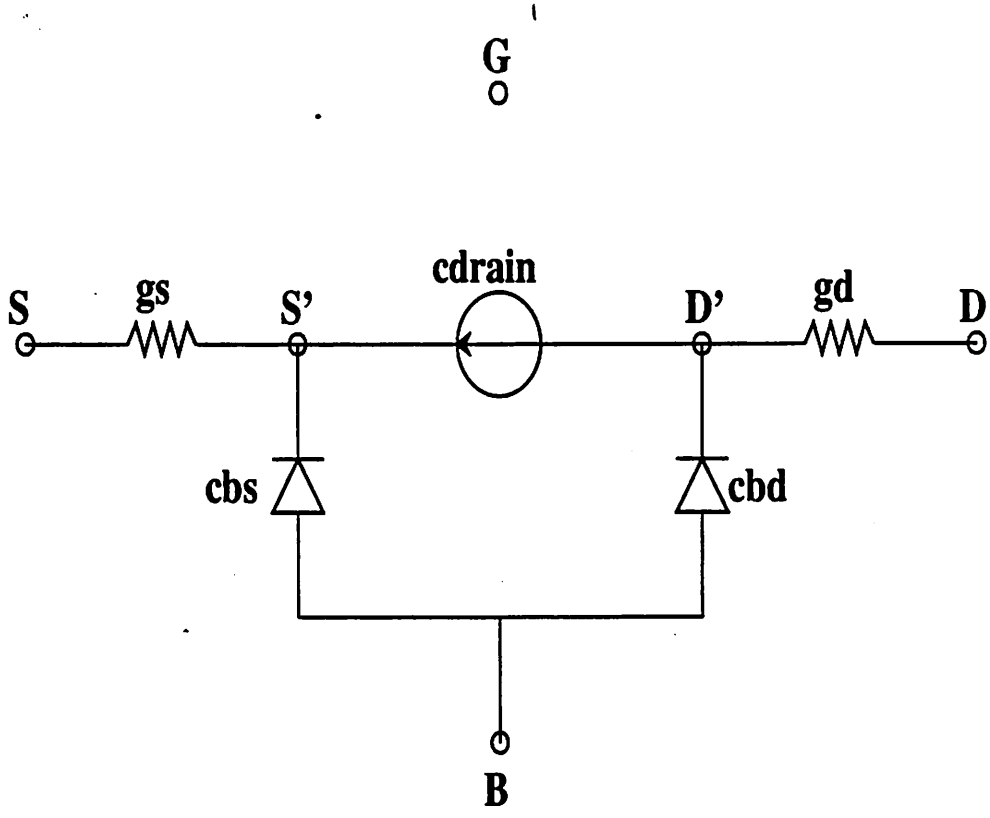


Fig.A2.1. The equivalent circuit of a MOSFET for DC analysis.

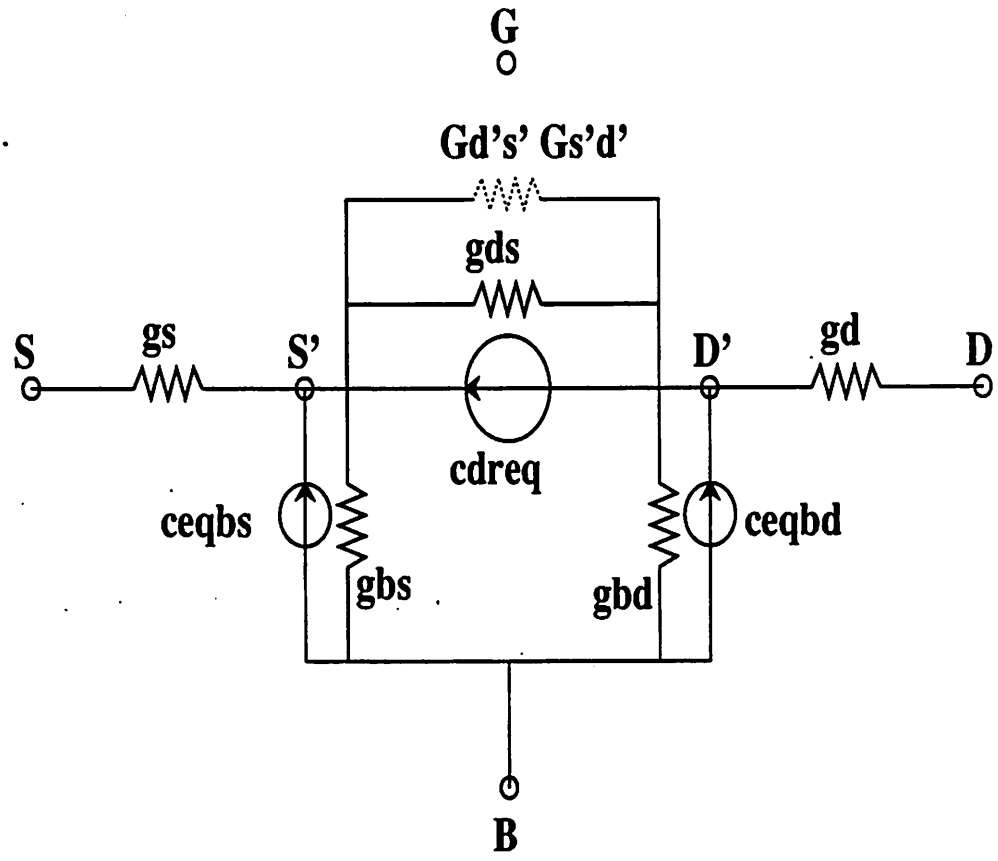


Fig.A2.2. The companion model of the MOSFET DC equivalent circuit in Fig.A2.1.

Newton-Raphson iteration method, the equivalent circuit in Fig.A2.1 is conveterted to its companion model as shown in Fig.A2.2. where *gs* and *gd* are source, drain series resistances respectively and *gbs* and *gbd* are the conductance of source, drain junctions respectively. $G_{D'S'}$ and $G_{S'D'}$ are non-reciprocal conductances.

$$G_{D'S'} = xnrm^{*}(gm+gmbs) \tag{A2.3}$$

$$G_{S'D'} = xrev^{*}(gm+gmbs) \tag{A2.4}$$

where *gm*, *gmbs* and *gds* are $\partial I_D/\partial V_{GS}$, $\partial I_D/\partial V_{BS}$ and $\partial I_D/\partial V_{DS}$ respectively. The equation for I_D is shown in (2.25) and V_{GS}, V_{BS} and V_{DS} are gate to source, bulk to source and drain to source branch voltages respectively. *xnrm* and *xrev* are indicators to whether the drain, source nodes as specified in SPICE input are normal or inverted.

$$xnrm = \begin{cases} 1 & when\ \ V_{D'S'} \geq 0 \\ 0 & when\ \ V_{D'S'} < 0 \end{cases} \tag{A2.5}$$

$$xrev = \begin{cases} 0 & when\ \ V_{D'S'} \geq 0 \\ 1 & when\ \ V_{D'S'} < 0 \end{cases} \tag{A2.6}$$

The coefficients of the admittance matrix **Y** and the right hand side vector $\mathbf{I_{EQ}}$ in (A2.2) are shown in Table.A2.1.

The symbols used in Table A2.1 are as follows. where

$$G_{D'G} = (xnrm-xrev)^{*}gm \tag{A2.7}$$

$$G_{D'B} = (xnrm-xrev)^{*}gmbs - gbd \tag{A2.8}$$

$$G_{D'D'} = xrev^{*}(gm+gmbs) + gds + gbd \tag{A2.9}$$

$$G_{D'S'} = -xnrm^{*}(gm+gmbs) - gds \tag{A2.10}$$

$$G_{S'G} = -(xnrm-xrev)^{*}gm \tag{A2.11}$$

$$G_{S'B} = -(xnrm-xrev)^{*}gmbs - gbs \tag{A2.12}$$

$$G_{S'D'} = -xrev^{*}(gm+gmbs) - gds \tag{A2.13}$$

$$G_{S'S'} = xnrm^{*}(gm+gmbs) + gds + gbs \tag{A2.14}$$

$$cdreq = (xnrm-xrev)^{*}(cdrain-gm^{*}V_{GS}-gmbs^{*}V_{BS}) - gds^{*}V_{DS} \tag{A2.15}$$

$$ceqbd = cbd - gbd^{*}V_{BD} \tag{A2.16}$$

$$ceqbs = cbs - gbs^{*}V_{BS} \tag{A2.17}$$

cdreq, *ceqbd* and *ceqbs* in (A2.15), (A2.16) and (A2.17) are multiplied by (-1) for PMOSFET. The components of **Y** and $\mathbf{I_{EQ}}$ matrices in Table.A2.1 are added to the components of matrices computed for the other part of the circuit.

	D	*S*	*G*	*B*	*D'*	*S'*	*RHS*
D	*gd*				*-gd*		0
S		*gs*				*-gs*	0
G			0	0	0	0	0
B			0	*gbd+gbs*	*-gbd*	*-gbs*	*-ceqbs-ceqbd*
D'	*-gd*		$G_{D'G}$	$G_{D'B}$	$G_{D'D'}$	$G_{D'S'}$	*-cdreq+ceqbd*
S'		*-gs*	$G_{S'G}$	$G_{S'B}$	$G_{S'D'}$	$G_{S'S'}$	*-cdreq+ceqbd*

Table.A2.1. **Y** matrix and *RHS* vector components for the DC part of MOSFET companion model in Fig.A2.2.

(2) transient analysis

For transient analysis, we need capacitance components as well as the previously discussed DC components. Fig.A2.3 shows the equivalent circuit for capacitance components of MOSFET. *capbs* and *capbd* are source, drain junction capacitances and *qbs*

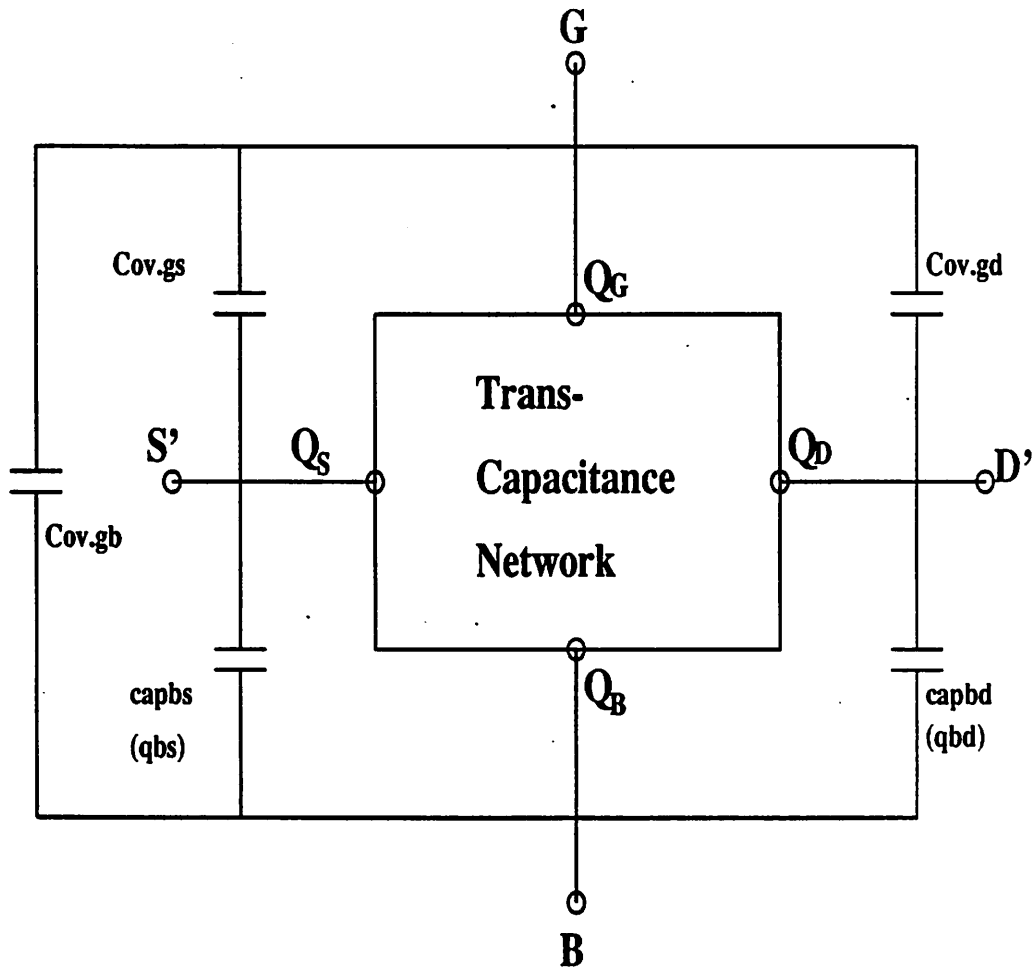


Fig.A2.3. The equivalent circuit of capacitance components for charge based MOSFET models.

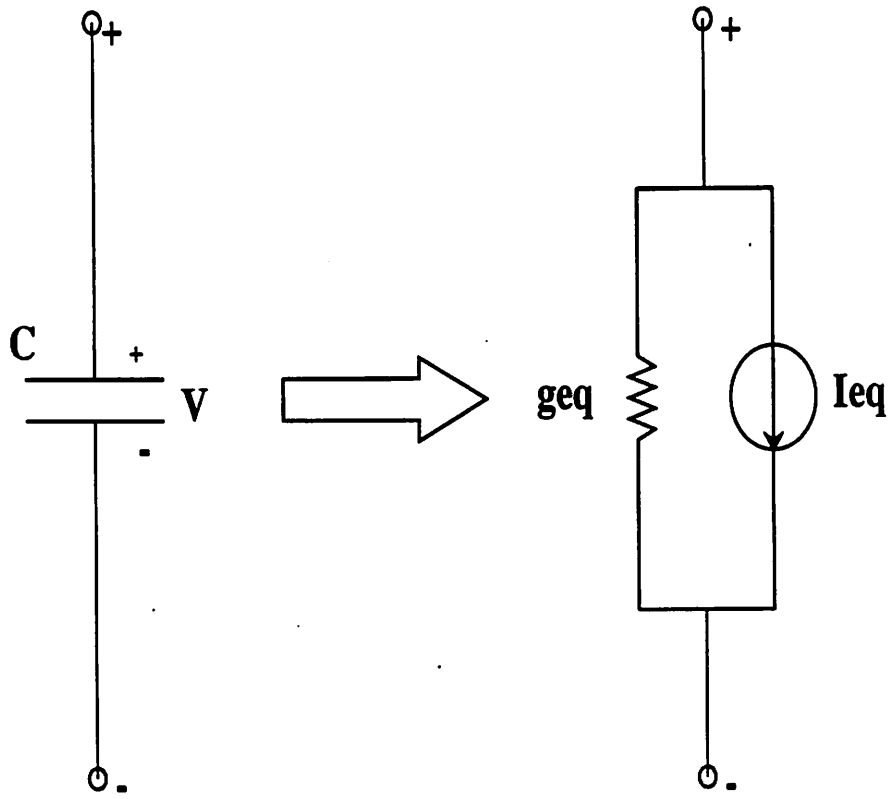


Fig.A2.4. Conversion of a linear capacitor into its companion model for the transient analysis of circuit simulations.

and *qbd* are source, drain junction charges respectively. $C_{ov.gs}$, $C_{ov.gd}$ and $C_{ov.gb}$ are overlap capacitances. Q_G and Q_B are gate and bulk charges whose equations can be found in (2.42) and (2.45) respectively. Q_D and Q_S are the drain and the source charges whose equations can be found in (2.46) and (2.37) respectively. When $V_{D'S'} < 0$, Q_D and Q_S are interchanged, so that Q_D and Q_S remain to be charges associated to the drain and source nodes as specified in the SPICE input.

Before discussing the companion model of the equivalent circuit in Fig.A2.3, we discuss the companion model of a linear capacitor as shown in Fig.A2.4. The current through a linear capacitor can be written as

$$I = \frac{dQ(V)}{dt} \tag{A2.18}$$

Q is a charge on the capacitor and V is an applied voltage across the capacitor. To find the current from (A2.18), integration schemes such as Trapezoidal or Gear integrations [2.6] are used. For simplicity, the Trapezoidal integration scheme is used throughout this work. Hence,

$$I_n = -I_{n-1} + \frac{2}{h}\cdot(Q_n - Q_{n-1}) \tag{A2.19}$$

where the subscrpt (n) refers to the present time point and the subscript $(n-1)$ refers to the previous time point. h is the time step which is the time interval between the previous time point and the present time point. Then the conductance *geq* and the equivalent current source *Ieq* can be written as

$$geq = \frac{\partial I_n}{\partial V_n} = \frac{2}{h}\cdot\frac{\partial Q_n}{\partial V_n} = \frac{2C}{h} \tag{A2.20}$$

$$Ieq = I_n - geq\cdot V_n = I_n - \frac{2C}{h}\cdot V_n \tag{A2.21}$$

Using this approach the equivalent circuit in Fig.A2.3 can be converted to its companion model in Fig.A2.5.

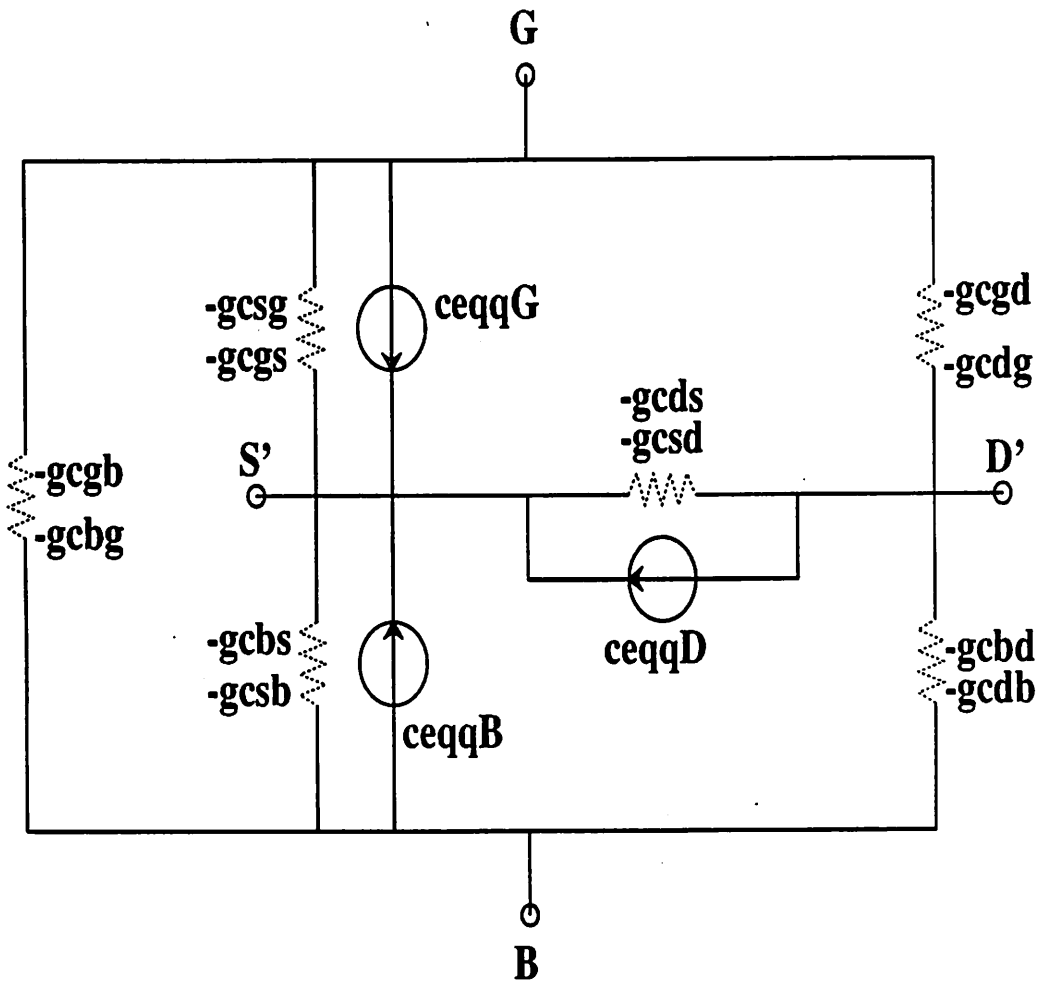


Fig.A2.5. The companion model for the MOSFET capacitance components for the transient analysis. Dotted zig-zag lines represent non-reciprocal conductances.

By including the charges on the overlap capacitor and the source, drain junctions, the apparent node charges are defined to be

$$q_G = Q_G + C_{ov.gd}*V_{GD} + C_{ov.gs}*V_{GS} + C_{ov.gb}*V_{GB} \quad \text{(A2.22)}$$

$$q_B = Q_B - C_{ov.gb}*V_{GB} + qbd + qbs \quad \text{(A2.23)}$$

$$q_D = Q_D - C_{ov.gd}*V_{GD} - qbd \quad \text{(A2.24)}$$

Then the capacitive node currents *cqgate*, *cqbulk* and *cqdrn* can be computed from

$$(cqgate)_n = -(cqgate)_{n-1} + \frac{2}{h}\cdot(q_{G.n} - q_{G.n-1}) \quad \text{(A2.25)}$$

$$(cqbulk)_n = -(cqbulk)_{n-1} + \frac{2}{h}\cdot(q_{B.n} - q_{B.n-1}) \quad \text{(A2.26)}$$

$$(cqdrn)_n = -(cqdrn)_{n-1} + \frac{2}{h}\cdot(q_{D.n} - q_{D.n-1}) \quad \text{(A2.27)}$$

The trapezoidal integration method is used in (A2.25), (A2.26) and (A2.27). Then the **Y** matrix and the *RHS* vector can be computed as shown in Table.A2.2.

	D	*S*	*G*	*B*	*D'*	*S'*	*RHS*
D	0				0		0
S		0				0	0
G			*gcgg*	*gcgb*	*gcgd*	*gcgs*	*−ceqqg*
B			*gcbg*	*gcbb*	*gcbd*	*gcbs*	*−ceqqb*
D'	0		*gcdg*	*gcdb*	*gcdd*	*gcds*	*−ceqqd*
S'		0	*gcsg*	*gcsb*	*gcsd*	*gcss*	*−ceqqs*

Table.A2.2. Components of **Y** matrix and **RHS** vector for the capacitance componets of MOSFET for transient analysis

where $gcgg=(C_{GG}+C_{ov.gd}+C_{ov.gs}+C_{ov.gb})*(2/h)$ (A2.28)

$gcgd = (C_{GD}-C_{ov.gd})*(2/h)$ (A2.29)

$gcgs = (C_{GS}-C_{ov.gs})*(2/h)$ (A2.30)

$gcgb = (C_{GB}-C_{ov.gb})*(2/h)$ (A2.31)

$gcbg = (C_{BG}-C_{ov.gb})*(2/h)$ (A2.32)

$gcbd = (C_{BD}-capbd)*(2/h)$ (A2.33)

$gcbs = (C_{BS}-capbs)*(2/h)$ (A2.34)

$gcbb = (C_{BB}+capbd+capbs+C_{ov.gb})*(2/h)$ (A2.35)

$gcdg = (C_{DG}-C_{ov.gd})*(2/h)$ (A2.36)

$gcdd = (C_{DD}+C_{ov.gd}+capbd)*(2/h)$ (A2.37)

$gcds = C_{DS}*(2/h)$ (A2.38)

$gcdb = (C_{DB}-capbd)*(2/h)$ (A2.39)

$gcsg = (C_{SG}-C_{ov.gs})*(2/h)$ (A2.40)

$gcsd = C_{SD}*(2/h)$ (A2.41)

$gcss = (C_{SS}+C_{ov.gs}+capbs)*(2/h)$ (A2.42)

$gcsb = (C_{SB}-capbs)*(2/h)$ (A2.43)

C_{XG}, C_{XD}, C_{XS} and C_{XB} are $\partial Q_X/\partial V_G$, $\partial Q_X/\partial V_D$, $\partial Q_X/\partial V_S$ and $\partial Q_X/\partial V_B$ respectivcly, where X ε (G, D, S, B).

The equivalent current sources can be computed from

$$ceqqg = (cqgate)_n - gcgg * V_{GB.n} - gcgd * V_{DB.n} - gcgs * V_{SB.n} \quad \text{(A2.44)}$$

$$ceqqb = (cqbulk)_n - gcbg * V_{GB.n} - gcbd * V_{DB.n} - gcbs * V_{SB.n} \quad \text{(A2.45)}$$

$$ceqqd = (cqdrn)_n - gcdg * V_{GB.n} - gcdd * V_{DB.n} - gcds * V_{SB.n} \quad \text{(A2.46)}$$

$$ceqqs = -(ceqqg + ceqqd + ceqqb) \quad \text{(A2.47)}$$

where $V_{GB.n}$, $V_{DB.n}$ and $V_{SB.n}$ are branch voltages computed at the previous iterations (and at the present time point). For PMOSFET, *ceqqg*, *ceqqb*, *ceqqd* and *ceqqs* are multiplied by (-1). For transient analysis, the matrix components in Table.A2.2 are added to the components for DC part in Table.A2.1 and also to the components from the other parts of the circuit.

(3) AC analysis

The components of **Y** matrix and *RHS* vector for AC analysis are shown in Table.A2.3.

	D	*S*	*G*	*B*	*D'*	*S'*	*RHS*
D	0				0		0
S		0				0	0
G			*j·xcgg*	*j·xcgb*	*j·xcgd*	*j·xcgs*	0
B			*j·xcbg*	*j·xcbb*	*j·xcbd*	*j·xcbs*	0
D'	0		*j·xcdg*	*j·xcdb*	*j·xcdd*	*j·xcds*	0
S'		0	*j·xcsg*	*j·xcsb*	*j·xcsd*	*j·xcss*	0

Table.A2.3. Components of **Y** matrix and **RHS** vector for the capacitance componets of MOSFET for AC analysis, where *j* is the imaginary number notation.

$$xcgg = \omega \cdot (C_{GG} + C_{ov.gd} + C_{ov.gs} + C_{ov.gb}) \quad \text{(A2.48)}$$

$$xcgd = \omega\cdot(C_{GD}-C_{ov.gd}) \quad \text{(A2.49)}$$

$$xcgs = \omega\cdot(C_{GS}-C_{ov.gs}) \quad \text{(A2.50)}$$

$$xcgb = \omega\cdot(C_{GB}-C_{ov.gb}) \quad \text{(A2.51)}$$

$$xcbg = \omega\cdot(C_{BG}-C_{ov.gb}) \quad \text{(A2.52)}$$

$$xcbd = \omega\cdot(C_{BD}-capbd) \quad \text{(A2.53)}$$

$$xcbs = \omega\cdot(C_{BS}-capbs) \quad \text{(A2.54)}$$

$$xcbb = \omega\cdot(C_{BB}+C_{ov.gb}+capbd+capbs) \quad \text{(A2.55)}$$

$$xcdg = \omega\cdot(C_{DG}-C_{ov.gd}) \quad \text{(A2.56)}$$

$$xcdd = \omega\cdot(C_{DD}+capbd+C_{ov.gd}) \quad \text{(A2.57)}$$

$$xcds = \omega\cdot C_{DS} \quad \text{(A2.58)}$$

$$xcdb = \omega\cdot(C_{DB}-capbd) \quad \text{(A2.59)}$$

$$xcsg = \omega\cdot(C_{SG}-C_{ov.gs}) \quad \text{(A2.60)}$$

$$xcsd = \omega\cdot C_{SD} \quad \text{(A2.61)}$$

$$xcss = \omega\cdot(C_{SS}+capbs+C_{ov.gs}) \quad \text{(A2.62)}$$

$$xcsb = \omega\cdot(C_{SB}+capbs+C_{ov.gs}) \quad \text{(A2.63)}$$

where ω is the signal angular frequency. For AC analysis, the components in Table.A2.3 are added to the components for DC part in Table.A2.1 and also to the components from other parts of the circuit.

Appendix 3

Program to drive the automatic gate capacitance measurement system

The following is the program for the gate capacitance measurement system which is described in Section 2.4 of Chapter 2.

```
 5   !_____________________________________________________________
10   ! The Main Program to measure CGT,CGD,CGS,CGB vs. VGS with VDS as parameters
12   !            by  Hong J. Park     Dec. 1985
15   !_____________________________________________________________
20    COM Lockin,Hp4145,Scanner
25    ! Set up GPIB(HPIB) addresses
30       Lockin=709   ! GPIB ADDRESS OF 5206 LOCK IN AMP.
40       Hp4145=707   ! GPIB ADDRESS OF HP4145A
50       Scanner=718  ! GPIB ADDRESS OF 706 SCANNER
60       Scalefactor=1.0E-11/2.590!  F/V ( CALIBRATED FROM 10pF STANDARD Cap)
70    INPUT "Nmeasure ?",Nmeasure
80 Start:        !
90    INPUT "CGT,CGB,CGD,CGS OR QUIT ?",Name$
100   IF Name$="QUIT" THEN Endpgm
105          !
110 Setup:    ! output Command of Sensitivity Adjustment to 5206 lock in amp.
120   CALL Setup5206
130   CALL Setuphp4145(Name$)
140   CALL Connectswitch(Name$)
150   CALL Getns(Ns,Rmax,Name$)
160      Scale=Scalefactor*FNCapscale(Ns)/2000
170      Cmax=Rmax*Scale
180      Cmax=Cmax*1.2
210      !
220      !
230   ! MEASUREMENT
250   INPUT "VDSMIN, VDSMAX, VDSSTEP  ? ",Vdsmin,Vdsmax,Vdsstep
260   INPUT "VGSMIN, VGSMAX, VGSSTEP  ? ",Vgsmin,Vgsmax,Vgsstep
270   INPUT "VBS ?",Vbs
370   !
380      Nd=(Vdsmax-Vdsmin)/Vdsstep
390      Ng=(Vgsmax-Vgsmin)/Vgsstep
400      ALLOCATE Ans(Nd,Ng)
405   !
```

```
410  ! SETUP GRAPHICS
420     Cmin=0
430     Cstep=(Cmax-Cmin)/4
440     CALL Grsetup(Vgsmin,Vgsmax,Cmin,Cmax,Cstep,Name$,Vbs)
450     !
460     !
470  ! Apply Source and Bulk Voltages
480     CALL Apply(Name$,"VS",0.)
500     CALL Apply(Name$,"VB",Vbs)
505  ! Apply Drain and Gate Voltages
510  FOR I=0 TO Nd
520        Vds=Vdsmin+Vdsstep*I
540        CALL Apply(Name$,"VD",Vds)
550        WAIT .6     !  WAIT DURING 20 TIME CONSTANTS
560        FOR J=-3 TO Ng
570           Vgs=Vgsmin+J*Vgsstep
590           CALL Apply(Name$,"VG",Vgs)
600           WAIT .15      ! WAIT DURING 5 TIME CONSTANTS
610           Response=0
620           FOR K=1 TO Nmeasure
630           CALL Outlockin("Q1")
640           Response=FNGetlockin+Response
650           NEXT K
660           Response=Response/Nmeasure
670           IF ABS(Response)>2000 THEN PRINT " OUT OF RANGE ";Response
680           IF J<0 THEN 720
690           Ans(I,J)=Response*Scale
710           PLOT Vgs,Ans(I,J)
720        NEXT J
730        PENUP
740  NEXT I
745  ! Reset the bias to eliminate the unnecessary stress to the Device
750        CALL Apply(Name$,"VG",0.)
760        CALL Apply(Name$,"VD",0.)
770        CALL Apply(Name$,"VB",0.)
780     BEEP
790     PRINT "MEASUREMENT DONE"
800     GOTO 1080
810     PRINT "DUMP DATA INTO DISKETTE"
820     INPUT "DATA FILE NAME ?",Datafile$
830     INPUT "W,L (um) ? ",W,L
850     Length=100+(Ng+1)*(Nd+1)*4
860     FOR N=0 TO Nd
870        FOR N1=0 TO Ng
880           Length=Length+LEN(VAL$(Ans(N,N1)))
890        NEXT N1
900     NEXT N
910     Rec=Length/256*1.5
920     CREATE ASCII Datafile$,INT(Rec)+(Rec-INT(Rec)>1.E-6)
930     ASSIGN @File TO Datafile$;FORMAT ON
940     OUTPUT @File;Name$
950     OUTPUT @File;W,L
960     OUTPUT @File;Vbs
```

```
970   OUTPUT @File;Vdsmin,Vdsstep,Nd
980   OUTPUT @File;Vgsmin,Vgsstep,Ng
990   FOR I=0 TO Nd
1000      Vds=Vdsmin+Vdsstep*I
1010      FOR J=0 TO Ng
1020         Vgs=Vgsmin+Vgsstep*J
1030         OUTPUT @File;Ans(I,J)
1040      NEXT J
1050   NEXT I
1060   !
1070   GOTO Endpgm
1080   INPUT "  WANT TO PLOT AGAIN ( Y OR N ) ",Q$
1090   IF Q$="N" THEN GOTO 1200
1100   INPUT "CMIN,CMAX,CSTEP ?",Cmin,Cmax,Cstep
1110   CALL Grsetup(Vgsmin,Vgsmax,Cmin,Cmax,Cstep,Name$,Vbs)
1120   FOR I=0 TO Nd
1130      Vds=Vdsmin+Vdsstep*I
1140      FOR J=0 TO Ng
1150         Vgs=Vgsmin+Vgsstep*J
1160         PLOT Vgs,Ans(I,J)
1170      NEXT J
1180      PENUP
1190   NEXT I
1200   GOTO 810
1210 Endpgm:    !
1220     CALL Outlockin("S 0")
1230   END
1250   !
1260   !
1270   !_____________________________________________________________
1280   ! The routine to send the character string C$ to the Lock-in Amp
1285   !_____________________________________________________________
1290    SUB Outlockin(C$)
1300    COM Lockin,Hp4145,Scanner
1310   ! SEND COMMAND TO HP4145A
1320 Spoll1:S9=SPOLL(Lockin)
1330   !IF PREVIOUS COMMAND WAS COMPLETED, SEND NEW COMMAND
1340      IF BIT(S9,0)=0 THEN Spoll1
1350      OUTPUT Lockin USING "K";C$
1360    SUBEND
1370   !
1380   !_____________________________________________________________
1384   ! The routine to read in data from Lock-in Amp
1386   !_____________________________________________________________
1390    DEF FNGetlockin
1400       COM Lockin,Hp4145,Scanner
1410   ! GET RESPONSE FROM 5206 LOCK IN AMP.
1420       I9=0
1430 Spoll2:L9=SPOLL(Lockin)
1440   ! IF OUTPUT READY READ IN THE DATA
1450       IF BIT(L9,7)=0 THEN Loop4
1460       I9=I9+1
1470       WAIT .01
```

```
1480      ENTER Lockin;A
1490      WAIT .01
1500 Loop4:IF BIT(L9,0)=0 THEN Spoll2
1510      IF I9>1 THEN PRINT "MORE THAN ONE RESPONSES FROM LOCKIN AMP  ";I9;" RESPOS
1520      RETURN A
1530   FNEND
1540   !
1550   !
1560   !_____________________________________________________________
1564   ! The routine to send the character string D$ to HP4145
1566   !_____________________________________________________________
1570   SUB Outhp4145(D$)
1580      COM Lockin,Hp4145,Scanner
1590   ! SEND COMMAND TO HP4145A
1600      OUTPUT Hp4145 USING "K";D$
1610   SUBEND
1620   !
1630   !_____________________________________________________________
1634   ! The routine to set up constants of the Lock-in Amp
1636   !_____________________________________________________________
1640   SUB Setup5206
1650    ! SETUP 5206 LOCK IN AMP.
1660      Outlockin("C 10")                    ! LINEFEED IS DELIMITER
1670      Outlockin("D 0")                     ! DISPLAY CH 1
1680      Outlockin("M 1")                     ! ANALOG DISPLAY(R THETA)
1690      Outlockin("X 0")                     ! EXPAND OFF
1700      Outlockin("O 0,0")                   ! TURN OFFSET OFF
1710      Outlockin("F 0")                     ! FREQUENCY BROAD BAND
1720      Outlockin("J 650,1")                 ! SET FREQ. AT 6.5 KHz
1730      Outlockin("P 0,200")                 ! SET PHASE 0 DEGREE
1740      Outlockin("T 7,0")                   ! TIME CONST 30MS 12DB/OCT
1750   SUBEND
1760    !
1770    !
1780    !_____________________________________________________________
1784    ! The routine to find the full range of the measurement
1786    !_____________________________________________________________
1790   SUB Getns(Ns,Rmax,Name$)
1800   !
1810   Ns=0
1820   CALL Apply(Name$,"VS",0.)
1830   CALL Apply(Name$,"VB",0.)
1840   SELECT Name$
1850         CASE "CGT"
1860            CALL Apply("CGT","VG",-5.0)
1870            CALL Apply("CGT","VD",0.)
1880         CASE "CGD"
1890            CALL Apply("CGD","VG",5.0)
1900            CALL Apply("CGD","VD",0.)
1910         CASE "CGS"
1920            CALL Apply("CGS","VG",5.0)
1930            CALL Apply("CGS","VD",5.0)
1940         CASE "CGB"
```

```
1950            CALL Apply("CGB","VG",-5.0)
1960            CALL Apply("CGB","VD",0.)
1970        END SELECT
1980        WAIT 1
1990 Begin:   !
2000        CALL Outlockin("S "&VAL$(Ns))
2010        WAIT 1
2020        CALL Outlockin("Q1")
2030        Response=FNGetlockin
2040        PRINT "NS=";Ns;"      Response= ";Response
2050        IF Response<2000 THEN Check2
2060        Ns=Ns-1
2070        IF Ns<0 THEN PRINT "ERROR: CAPACITANCE OVERFLOW"
2080        GOTO Begin
2090 Check2:  IF Response>800 THEN GOTO Finish
2100         Ns=Ns+1
2110         GOTO Begin
2120 Finish:  Rmax=Response
2130     SUBEND
2140  !
2150  !_____________________________________________________________
2154  ! The routine to convert Ns into an appropriate constant
2156  !_____________________________________________________________
2160  DEF FNCapscale(Ns)
2170  IF (Ns<0) OR (Ns>20) THEN PRINT "ERROR: NS= ";Ns
2180  IF Ns=0 THEN RETURN 5
2190  IF Ns=1 THEN RETURN 2
2200  IF Ns=2 THEN RETURN 1
2210  IF Ns=3 THEN RETURN 5.E-1
2220  IF Ns=4 THEN RETURN 2.E-1
2230  IF Ns=5 THEN RETURN 1.E-1
2240  IF Ns=6 THEN RETURN 5.E-2
2250  IF Ns=7 THEN RETURN 2.E-2
2260  IF Ns=8 THEN RETURN 1.E-2
2270  IF Ns=9 THEN RETURN 5.E-3
2280  IF Ns=10 THEN RETURN 2.E-3
2290  IF Ns=11 THEN RETURN 1.E-3
2300  IF Ns=12 THEN RETURN 5.E-4
2310  IF Ns=13 THEN RETURN 2.E-4
2320  IF Ns=14 THEN RETURN 1.E-4
2330  IF Ns=15 THEN RETURN 5.E-5
2340  IF Ns=16 THEN RETURN 2.E-5
2350  IF Ns=17 THEN RETURN 1.E-5
2360  IF Ns=18 THEN RETURN 5.E-6
2370  IF Ns=19 THEN RETURN 2.E-6
2380  IF Ns=20 THEN RETURN 1.E-6
2390  FNEND
2400  !
2410  !
2414  !_____________________________________________________________
2415  ! The setup routine for the graphic display of the measured data
2416  !_____________________________________________________________
2420  SUB Grsetup(Vgsmin,Vgsmax,Cmin,Cmax,Cstep,Name$,Vbs)
```

```
2430  !COM Vgsmin,Vgsmax,Cmin,Cmax,Cstep
2440  GINIT
2450  GRAPHICS ON
2460  CSIZE 5
2470  MOVE 70,3
2480  LABEL "VGS[V]"
2490  MOVE 0,85
2500  LABEL Name$&" [F]"
2510  CSIZE 3
2520  MOVE 0,80
2530  LABEL "   (VBS:"&VAL$(Vbs)&" V)"
2540  CSIZE 5
2550  VIEWPORT 30,110,15,95
2560  WINDOW Vgsmin,Vgsmax,Cmin,Cmax
2570  AXES .2,Cstep/5,Vgsmin,Cmin,5,5
2580  AXES .2,Cstep/5,Vgsmax,Cmax,5,5
2590  CLIP OFF
2600  LORG 6
2610  LDIR 0
2620  FOR Vgs=Vgsmin TO Vgsmax STEP 2
2630  MOVE Vgs,0
2640  LABEL Vgs
2650  NEXT Vgs
2660  LORG 8
2670  FOR C=Cmin TO Cmax STEP Cstep
2680  MOVE Vgsmin+0.,C
2690  LABEL USING "D.DDE";C
2700  NEXT C
2710  !LORG 5
2720  SUBEND
2730  !
2740  !
2750  !
2754  !_____________________________________________________________
2755  ! The routine to apply bias to each MOSFET node
2756  !_____________________________________________________________
2760  SUB Apply(Name$,V$,Value)
2770  COM Lockin,Hp4145,Scanner
2780  SELECT Name$
2790  CASE "CGT"
2800      SELECT V$
2810      CASE "VG"
2820         ! APPLY VGS VIA SMU 1 AUTO RANGE, COMPLIANCE 0.1 A
2830         CALL Outhp4145("US DV 1,0,"&VAL$(Value)&",0.1")
2840      CASE "VB"
2850         ! APPLY VBS VIA SMU 2 AUTO RANGE, COMPLIANCE 0.1 A
2860         CALL Outhp4145("US DV 2,0,"&VAL$(Value)&",0.1")
2870      CASE "VD"
2880         ! APPLY VDS VIA SMU 3 AUTO RANGE, COMPLIANCE 0.1 A
2890         CALL Outhp4145("US DV 3,0,"&VAL$(Value)&",0.1")
2900      CASE "VS"
2910         ! APPLY SOURCE VTG. WITH VALUE 0.0 VOLT
2920         ! CALL Outhp4145("US DS 1,0")
```

```
2930        END SELECT
2940    CASE "CGB"
2950        SELECT V$
2960        CASE "VG"
2970            CALL Outhp4145("US DV 4,0,"&VAL$(Value)&",0.1")
2980        CASE "VB"
2990            CALL Outhp4145("US DV 1,0,"&VAL$(Value)&",0.1")
3000        CASE "VD"
3010            CALL Outhp4145("US DV 3,0,"&VAL$(Value)&",0.1")
3020        CASE "VS"
3030            ! CALL Outhp4145("US DV 2,0,0.0,0.1")
3040        END SELECT
3050    CASE "CGD"
3060        SELECT V$
3070        CASE "VG"
3080            CALL Outhp4145("US DV 4,0,"&VAL$(Value)&",0.1")
3090        CASE "VB"
3100            CALL Outhp4145("US DV 2,0,"&VAL$(Value)&",0.1")
3110        CASE "VD"
3120            CALL Outhp4145("US DV 1,0,"&VAL$(Value)&",0.1")
3130        CASE "VS"
3140            ! CALL Outhp4145("US DV 3,0,"&VAL$(Value)&",0.1")
3150        END SELECT
3160    CASE "CGS"
3170        SELECT V$
3180        CASE "VG"
3190            CALL Outhp4145("US DV 4,0,"&VAL$(Value)&",0.1")
3200        CASE "VB"
3210            CALL Outhp4145("US DV 2,0,"&VAL$(Value)&",0.1")
3220        CASE "VD"
3230            CALL Outhp4145("US DV 3,0,"&VAL$(Value)&",0.1")
3240        CASE "VS"
3250            CALL Outhp4145("US DV 1,0,0.0,0.1")
3260        END SELECT
3270    END SELECT
3280    SUBEND
3290    !
3300    !
3304 !_____________________________________________________________
3305 ! The routine to set up constants in HP4145
3306 !_____________________________________________________________
3310 SUB Setuphp4145(Name$)
3320 COM Lockin,Hp4145,Scanner
3330 SELECT Name$
3340 CASE "CGT"
3350        ! ASSIGN SMU 1 TO 'VG' voltage src. (var. 1)
3360        OUTPUT Hp4145;"DE CH 1,'VG','IG',1,1"
3370        ! ASSIGN SMU 2 TO 'VB' COMMON GND (CONSTANT)
3380        OUTPUT Hp4145;"DE CH 2,'VB','IB',1,3"
3390        ! ASSIGN SMU 3 TO 'VD' VOLTAGE SRC (VAR 2)
3400        OUTPUT Hp4145;"DE CH 3,'VD','ID',1,2"
3410        OUTPUT Hp4145;"DE CH 4"
3420        ! ASSIGN VS 1 TO 'VS'
```

```
3430      ! OUTPUT Hp4145;"DE VS 1,'VS',3"
3440  CASE "CGB"
3450      ! ASSIGN SMU 4 TO 'VG' VOLTAGE SRC (VAR 1)
3460      ! ASSIGN SMU 1 TO 'VB' VOLTAGE SRC
3470      OUTPUT Hp4145;"DE CH 1,'VB','IB',1,3"
3480      OUTPUT Hp4145;"DE CH 2"
3490      ! ASSIGN SMU 3 TO 'VD' VOLTAGE SRC (VAR 2)
3500      OUTPUT Hp4145;"DE CH 3,'VD','ID',1,2"
3510      ! ASSIGN SMU 2 TO 'VS' GND (CONSTANT)
3520      ! OUTPUT Hp4145;"DE CH 2,'VS','IS',3,3"
3530      OUTPUT Hp4145;"DE CH 4,'VG','IG',1,1"
3540  CASE "CGD"
3550      ! ASSIGN SMU 1 TO 'VD' VOLTAGE SRC (VAR 2)
3560      OUTPUT Hp4145;"DE CH 1,'VD','ID',1,2"
3570      ! ASSIGN SMU 2 TO 'VB' CONSTANT VOLTAGE SRC
3580      OUTPUT Hp4145;"DE CH 2,'VB','IB',1,3"
3590      OUTPUT Hp4145;"DE CH 3"
3600      ! ASSIGN SMU 3 TO 'VS' COMMON GND
3610      ! OUTPUT Hp4145;"DE CH 3,'VS','IS',3,3"
3620      ! ASSIGN SMU 4 TO 'VG' VOLTAGE SRC (VAR 1)
3630      OUTPUT Hp4145;"DE CH 4,'VG','IG',1,1"
3640  CASE "CGS"
3650      ! ASSIGN SMU 1 TO 'VS' CONSTANT COM GND
3660      !OUTPUT Hp4145;"DE CH 1,'VS','IS',3,3"
3670      OUTPUT Hp4145;"DE CH 1,'VS','IS',1,3"
3680      ! ASSIGN SMU 2 TO 'VB' CONSTANT VOLTAGE SRC
3690      OUTPUT Hp4145;"DE CH 2,'VB','IB',1,3"
3700      ! ASSIGN SMU 3 TO 'VD' VOLTAGE SRC (VAR 2)
3710      OUTPUT Hp4145;"DE CH 3,'VD','ID',1,2"
3720      ! ASSIGN SMU 4 TO 'VG' VOLTAGE SRC (VAR 1)
3730      OUTPUT Hp4145;"DE CH 4,'VG','IG',1,1"
3740  END SELECT
3750  !
3760  WAIT 1
3770  ! THE REMAINING CHANNELS ARE NOT USED.
3780  OUTPUT Hp4145;"DE VS 1"
3790  OUTPUT Hp4145;"VS 2"
3800  OUTPUT Hp4145;"VM 1"
3810  OUTPUT Hp4145;"VM 2"
3820  OUTPUT Hp4145;"SS HT 3.0"     ! HOLD TIME 3.0 SEC
3830  OUTPUT Hp4145;"DT 1.0"     ! DELAY TIME 1.0 SEC
3840  OUTPUT Hp4145;"IT 3"       ! SELECT LONG INTEGRATION TIME
3850  !
3860  SUBEND
3870  !
3874  !_____________________________________________________________________
3875  ! The routine to connect switches in switching matrix scanner
3876  !            corresponding to the measurement mode Name$
3877  !_____________________________________________________________________
3880  SUB Connectswitch(Name$)
3890  COM Lockin,Hp4145,Scanner
3900  !
3910  !   C1: AC + SMU1
```

```
3920 !   C2: SMU2
3930 !   C3: SMU3
3940 !   C4: SMU4
3950 !   C5: LOW
3960 !   R1: GND
3970 !   R2: BULK
3980 !   R3: SOURCE
3990 !   R4: DRAIN
4000 !
4010 !
4020   SELECT Name$
4030   CASE "CGT"
4040         ! CONNECT (C1,R1) (C2,R2) (C3,R4) (C5,R3)
4050         REMOTE Scanner
4060         CLEAR Scanner      ! OPEN ALL THE CHANNELS
4070         OUTPUT Scanner;"A0X" ! SELECT MATRIX MODE
4080         OUTPUT Scanner;"RX" ! RESET, OPEN ALL THE CHANNELS
4090         OUTPUT 718;"C 001 1 C 002 2 C 003 4 C 005 3 X"
4100   CASE "CGB"
4110         ! CONNECT (C1,R2) (C5,R3) (C3,R4) (C4,R1)
4120         REMOTE Scanner
4130         CLEAR Scanner      ! OPEN ALL THE CHANNELS
4140         OUTPUT Scanner;"A0X" ! SELECT MATRUX MODE
4150         OUTPUT Scanner;"RX"  ! RESET, OPEN ALL THE CHANNELS
4160         OUTPUT Scanner;"C 001 2 C 005 3 C 003 4 C 004 1 X"
4170   CASE "CGD"
4180         ! CONNECT (C1,R4) (C2,R2) (C5,R3) (C4,R1)
4190         REMOTE Scanner
4200         CLEAR Scanner
4210         OUTPUT Scanner;"A0X"
4220         OUTPUT Scanner;"RX"
4230         OUTPUT Scanner;"C 001 4 C 002 2 C 005 3 C 004 1 X"
4240   CASE "CGS"
4250         ! CONNECT (C1,R3) (C2,R2) (C3,R4) (C4,R1)
4260         REMOTE Scanner
4270         CLEAR Scanner
4280         OUTPUT Scanner;"A0X"
4290         OUTPUT Scanner;"RX"
4300         OUTPUT Scanner;"C 001 3 C 002 2 C 003 4 C 004 1 X"
4310   END SELECT
4320   !
4330   SUBEND
```

Appendix 4

Algorithm to find surface potentials for long channel MOSFETs with non-uniform substrate doping concentration

For conventional MOSFETs, the channel implantation is performed to adjust the threshold voltage and this results in the non-uniform substrate doping concentrations [2.19]. The substrate doping concentration profile is approximated into a step profile as shown in Fig.A4.1 to derive analytic model equations. When the substrate is depleted beyond $x1$, that is, the depth of the depleted substrate x_D is larger than $x1$, the electric field at the surface of silicon, E_S, can be written as

$$E_S = \frac{qN2}{\varepsilon_S}\cdot(x_D - x1) + \frac{qN1}{\varepsilon_S}\cdot x1 \qquad \text{(A4.1)}$$

The surface potential Ψ_S can be computed by integrating the electric field from the neutral bulk to the surface of silicon. Hence,

$$\Psi_S = \frac{qN2}{2\varepsilon_S}\cdot(x_D{}^2 - x1^2) + (\gamma_1\cdot A)^2 \qquad \text{(A4.2)}$$

When $x_D = x1$, we have $\Psi_S = (\gamma_1\cdot A)^2$ where $\gamma_1 = \sqrt{2\varepsilon_S q\cdot N1}\,/\,C_{OX}$ and $A = (\varepsilon_{OX}\cdot x1)^2/(2\varepsilon_S T_{OX})^2$. A is a dimensionless quantity. In order for x_D to be larger than $x1$, Ψ_S must be larger than $(\gamma_1\cdot A)^2$. Referring to Fig.2.4 for the uniform substrate doping concentration, we need the following conditions for $x_D > x1$.

$$V_{CB} > ((\gamma_1\cdot A)^2 - \Phi_F) \quad \text{and} \quad (V_{GB} - V_{FB}) > \gamma_1{}^2\cdot A\cdot(A+1) \qquad \text{(A4.3)}$$

$$\textit{where} \quad \Phi_F = 0.5\cdot V_t\cdot\ln((N1\cdot N2)/n_i{}^2) \qquad \text{(A4.4)}$$

For V_{CB}=0, we assume that x_D is always less than $x1$, which is usually true for modern

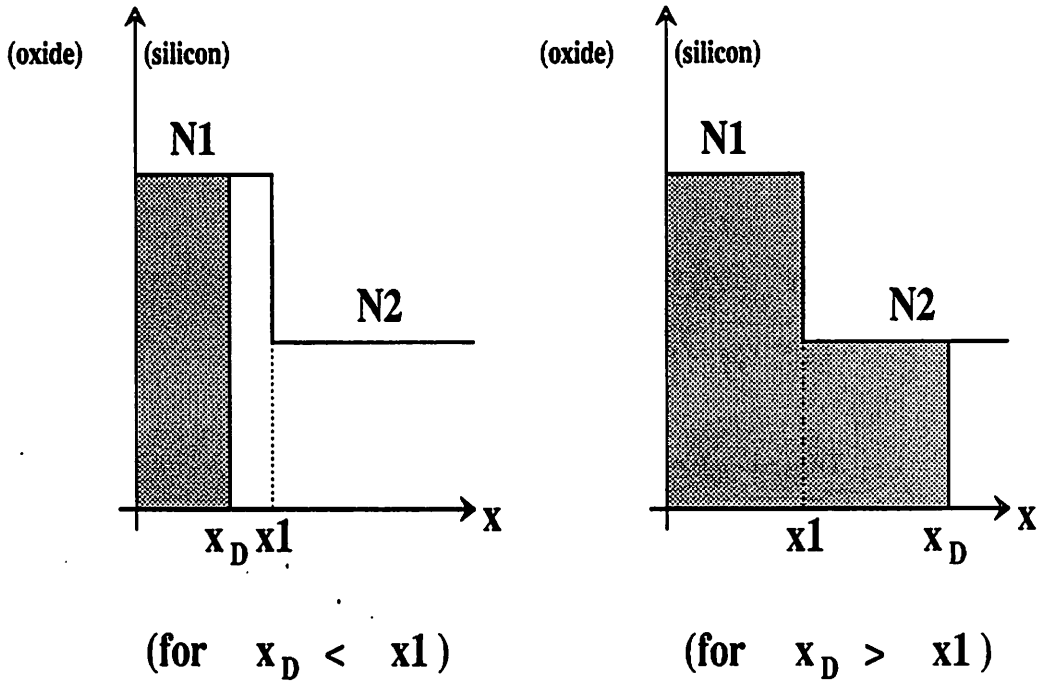


Fig.A4.1. The substrate doping concentration profile of a MOSFET, where x_D is the depletion depth and $x1$ is the boundary where the substrate doping concentration is discontinuous. A step profile is assumed for the doping profile. The shaded areas represent the depleted regions.

MOSFETs. Otherwise it can be analyzed in the same way as the case of uniform substrate doping concentration by adjusting V_{FB}. Hence the surface potential for long channel MOSFETs with non-uniform substrate doping concentration can be computed by using the following schemes.

if $V_{CB} < ((\gamma_1 A)^2 - \Phi_F)$ *or* $(V_{GB}-V_{FB}) < \gamma_1^2 A(A+1)$

if $(V_{GB}-V_{FB}) < (\Phi_F+\gamma_1\sqrt{\Phi_F-V_t})$

$\Psi_S = CSF(V_{GB}-V_{FB})$

else if $(V_{GB}-V_{FB}) < (\Phi_F+V_{CB}+\gamma_1\sqrt{\Phi_F+V_{CB}-V_t})$

$\Psi_S = V_{GB}-V_{FB}+0.5\gamma_1^2-\gamma_1\sqrt{V_{GB}-V_{FB}+0.25\gamma_1^2-V_t}$

else

$(V_{GB}-V_{FB})' = V_{GB}-V_{FB}-V_{CB}-\gamma_1\sqrt{2\Phi_F+V_{CB}-V_t}+\gamma_1\sqrt{2\Phi_F-V_t}$

$\Psi_S' = CSF((V_{GB}-V_{FB})')$

$Q_n' = (V_{GB}-V_{FB})'-\Psi_S'-\gamma_1\sqrt{\Psi_S'-V_t}$

$\Psi_S = V_{GB}-V_{FB}-Q_n'+0.5\gamma_1^2-\gamma_1\sqrt{V_{GB}-V_{FB}-Q_n'+0.25\gamma_1'-V_t}$

else

if $(V_{GB}-V_{FB}) < (\Phi_F+\gamma_1\sqrt{\Phi_F-V_t})$

$\Psi_S = CSF(V_{GB}-V_{FB})$

else if $(V_{GB}-V_{FB}) < (\Phi_F+V_{CB}+A(\gamma_1^2-\gamma_2^2)$

$+\gamma_2\sqrt{\Phi_F+V_{CB}-(\gamma_1^2-\gamma_2^2)A^2-V_t})$

$\Psi_S = V_{GB}-V_{FB}+0.5\gamma_2^2-A(\gamma_1^2-\gamma_2^2)$

$-\gamma_2\sqrt{V_{GB}-V_{FB}+0.25\gamma_2^2-A(A+1)(\gamma_1^2-\gamma_2^2)-V_t}$

else

$(V_{GB}-V_{FB})' = V_{GB}-V_{FB}-V_{CB}-A(\gamma_1^2-\gamma_2^2)$

$-\gamma_2\sqrt{2\Phi_F+V_{CB}-A^2(\gamma_1^2-\gamma_2^2)-V_t}+\gamma_1\sqrt{2\Phi_F-V_t}$

$\Psi_S' = CSF((V_{GB}-V_{FB})')$

$Q_n' = (V_{GB}-V_{FB})'-\Psi_S'-\gamma_1\sqrt{\Psi_S'-V_t}$

$\Psi_S = V_{GB}-V_{FB}-Q_n'+0.5\gamma_2^2-A(\gamma_1^2-\gamma_2^2)$

$$-\gamma_2\sqrt{V_{GB}-V_{FB}-Q_n{}'+0.25\gamma_2{}^2-A\,(A+1)(\gamma_1{}^2-\gamma_2{}^2)-(\gamma_1/\gamma_2)^2V_t}$$

Comparison with measured data

The surface potentials computed using the scheme described above, are compared with the measured data in Fig.A4.2. The measured data of surface potential are extracted in the same way as described in Chapter 2, Section 2.5.1. Fig.A4.2.(a) shows the surface potential Ψ_S versus V_{GB}, for a long NMOSFET with non-uniform substrate doping concentration. The surface potentials calculated using the scheme for the uniform doping concentration (Fig.2.4) are also shown for comparison. The derivatives of surface potential with respect to bias, $\partial\Psi_S/\partial V_G$, $\partial\Psi_S/\partial V_C$, and $\partial\Psi_S/\partial V_B$, are shown in Fig.A4.2.(b), Fig.A4.2.(c), and Fig.A4.2.(d) respectively.

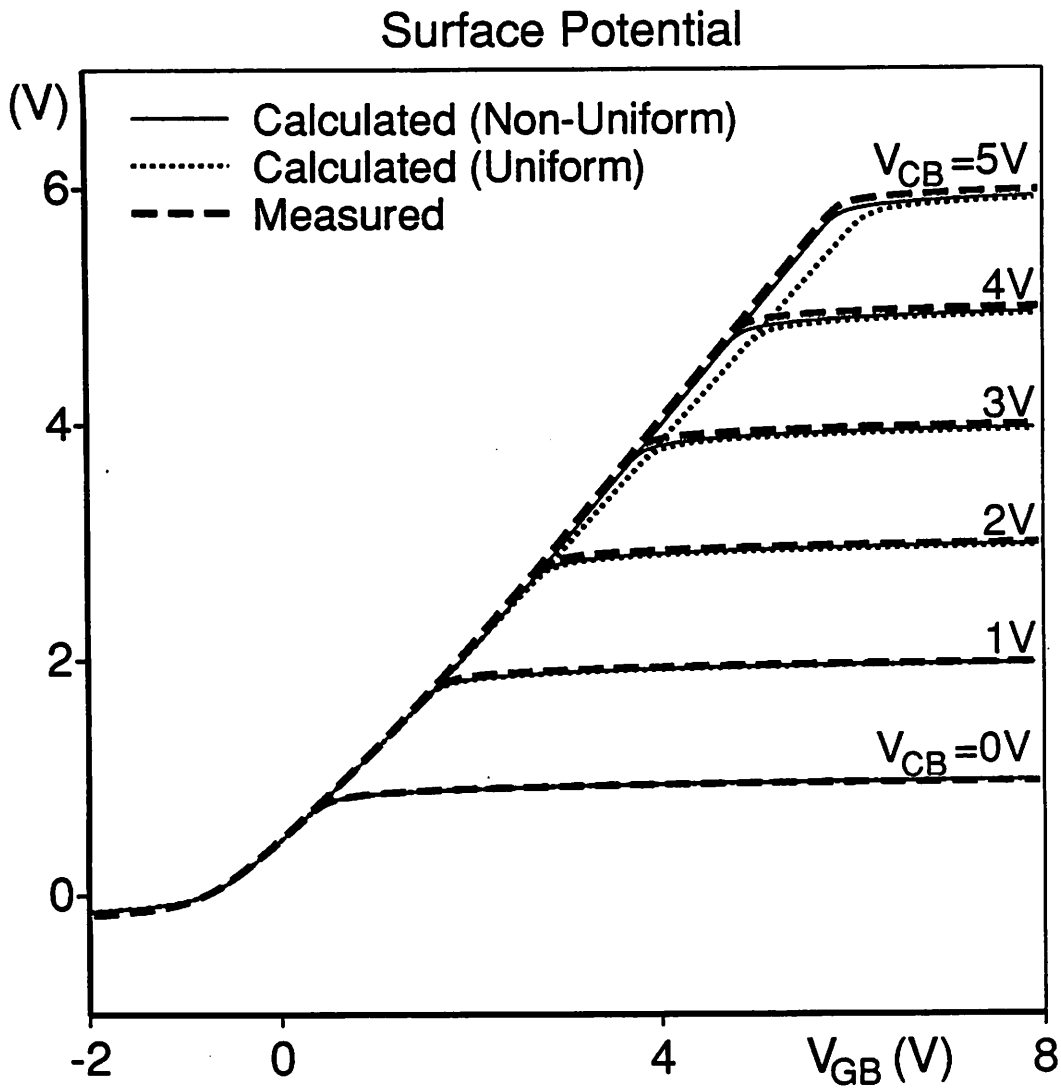


Fig.A4.2.(a). The surface potential Ψ_S versus V_{GB} with V_{CB} as parameters, for a NMOSFET with a non-uniform substrate doping concentration profile. Measured data are taken for a NMOSFET whose drain and source nodes are tied together. V_C represents the voltage of the channel node which is the tied drain source node. The surface potential curve computed from the uniform doping theory is also shown for comparison.

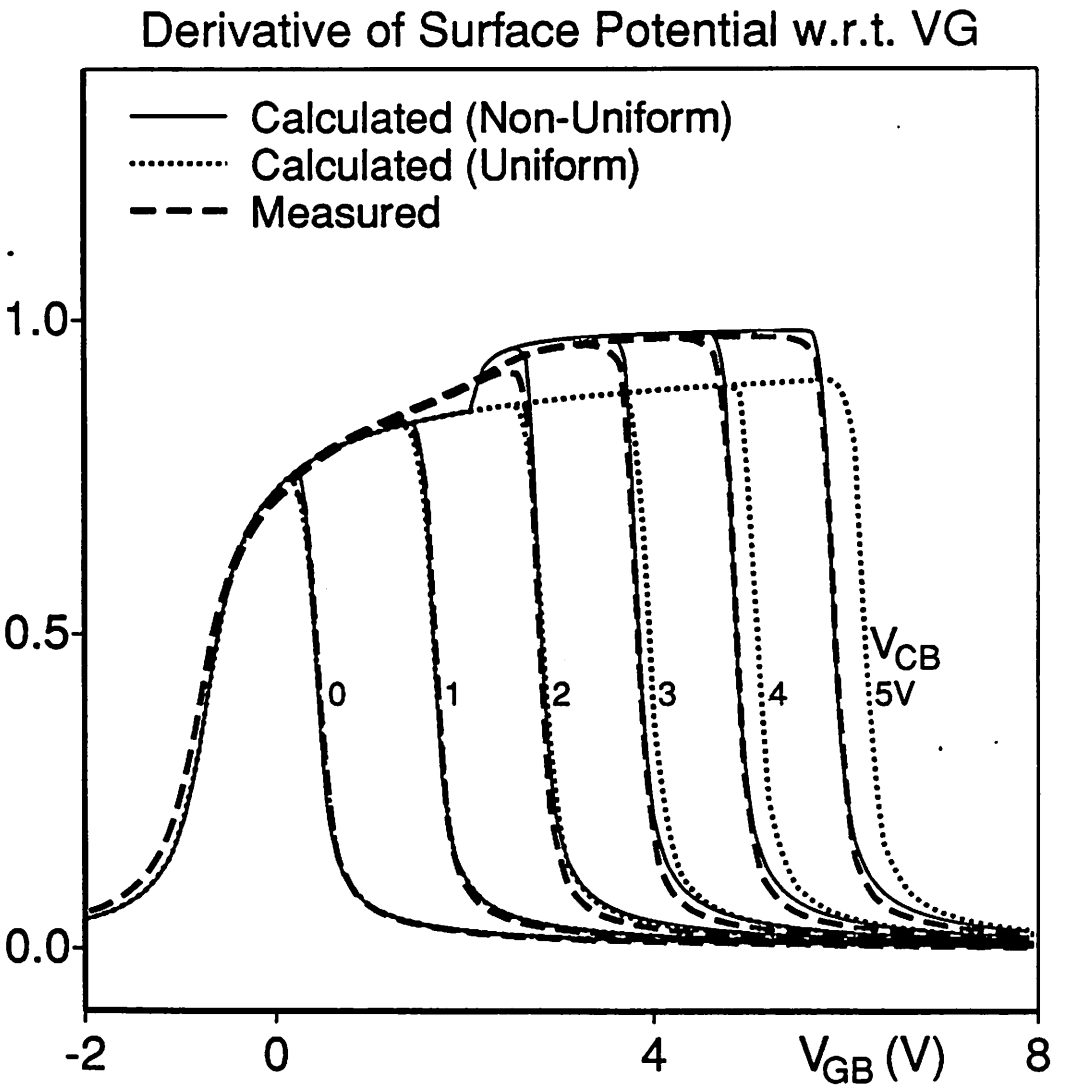


Fig.A4.2.(b). $\partial\Psi_S/\partial V_G$ versus V_{GB} for the NMOSFET with a non-uniform substrate doping concentration profile, whose surface potential curve is shown in Fig.A4.2.(a).

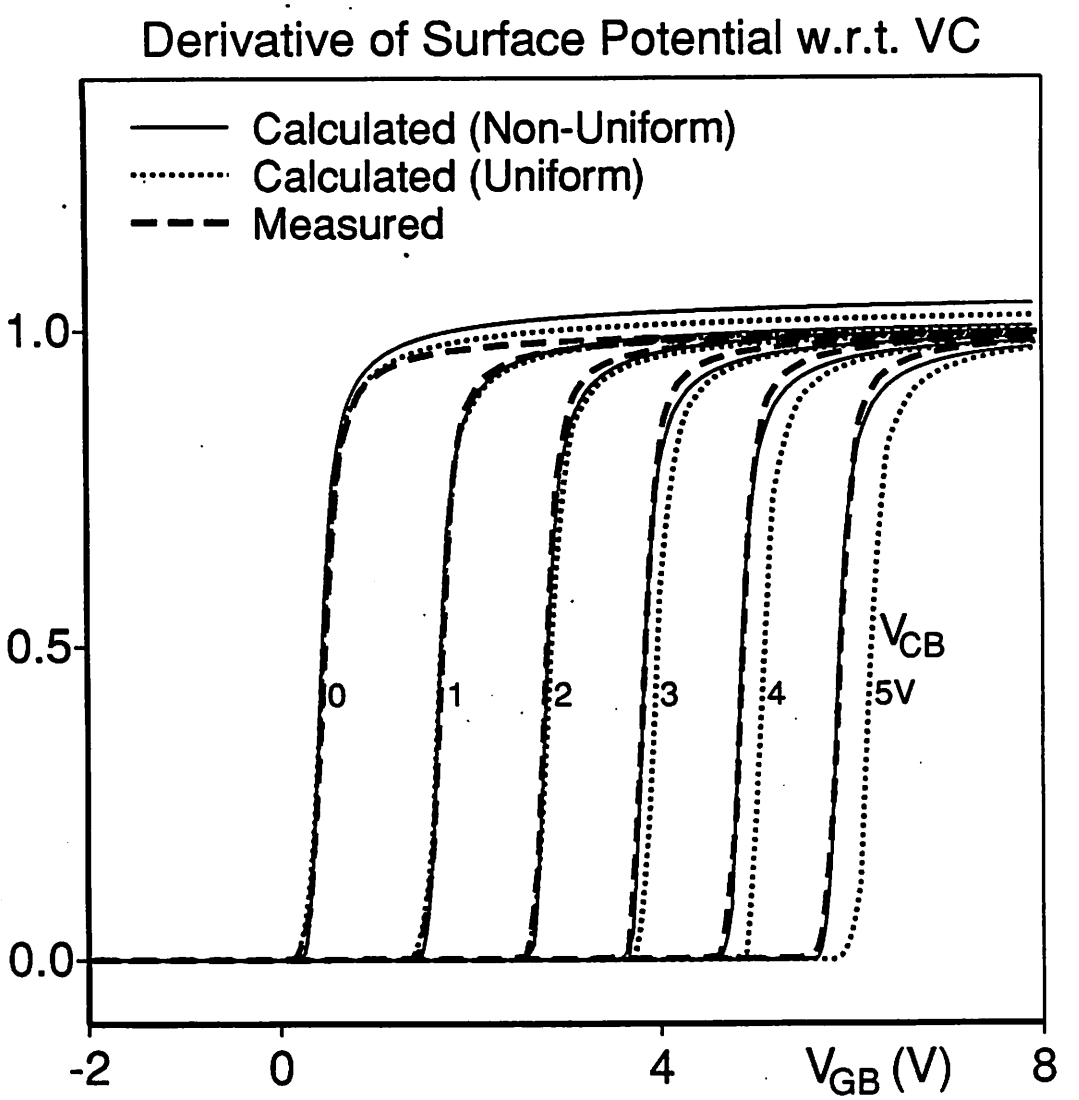


Fig.A4.2.(c). $\partial\Psi_S/\partial V_C$ versus V_{GB} for the NMOSFET with a non-uniform substrate doping concentration profile, whose surface potential curve is shown in Fig.A4.2.(a).

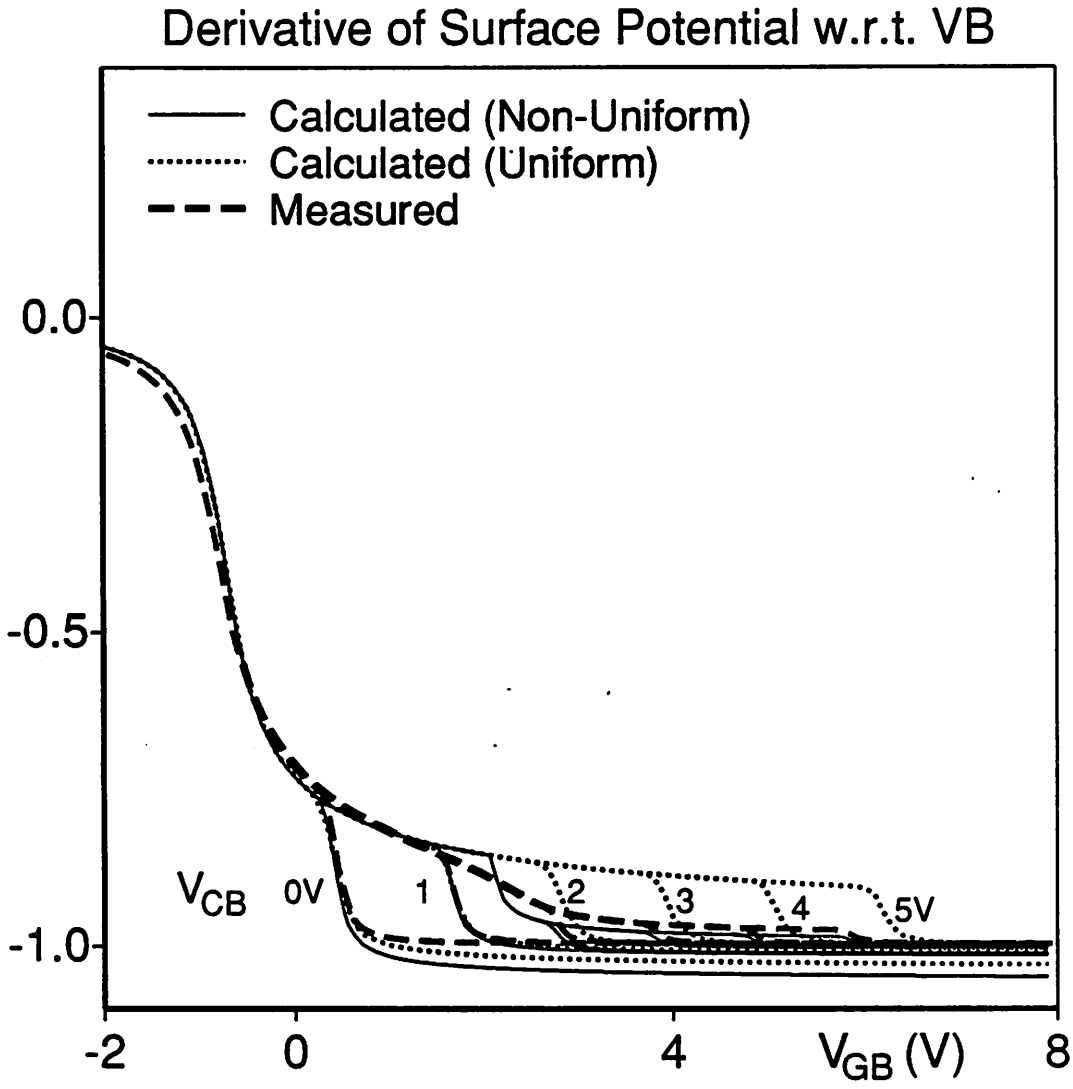


Fig.A4.2.(d). $\partial\Psi_S/\partial V_B$ versus V_{GB} for the NMOSFET with a non-uniform substrate doping concentration profile, whose surface potential curve is shown in Fig.A4.2.(a).

Appendix 5

Limiting routines for the CSM model implemented in SPICE3

In SPICE2 and SPICE3, the non-linear Newton Raphson iterations are used to solve the circuit equation. In the Newton Raphson iteration, the initial condition must be close enough to the real solution in order to get the convergence. For this purpose, the node voltages of MOSFET are limited between iterations both in SPICE2 and SPICE3, so that they don't digress far away from the real solution. In this work, a new current limiting routine is developed and both the voltage limiting routine and the current limiting routine are incorporated in the implementation of the CSM model into SPICE3 and the simulation reults are compared with each model.

(1) voltage limiting routines in SPICE

The gate to source voltage V_{GS} is limited by the subroutine *FETLIM* in SPICE2 and by the function *DEVfetlim* in SPICE3. Fig.A5.1 illustrates the function of the V_{GS} limiting routine in SPICE. *Vold* is the value of V_{GS} computed from the previous iteration and *Vnew* is the value of V_{GS} computed from the present iteration. When *Vnew* falls in the forbidden region, it is clipped to the closest boundary of the allowed region as indicated by the arrow in Fig.A5.1.

The drain to source voltage V_{DS} is limited by the subroutine *LIMVDS* in SPICE2 and by the function *DEVlimvds* in SPICE3. Fig.A5.2 illustrates the function of the V_{DS} limiting routine in SPICE. Fig.A5.2 is basically the same as Fig.A5.1 except the fact that there is no limitings when *Vold* < 0.

The bulk to source voltage V_{BS} is limited by the subroutine *PNJLIM* in SPICE2 and by the function *DEVpnjlim* in SPICE3. This V_{BS} limiting routine is derived from the diode

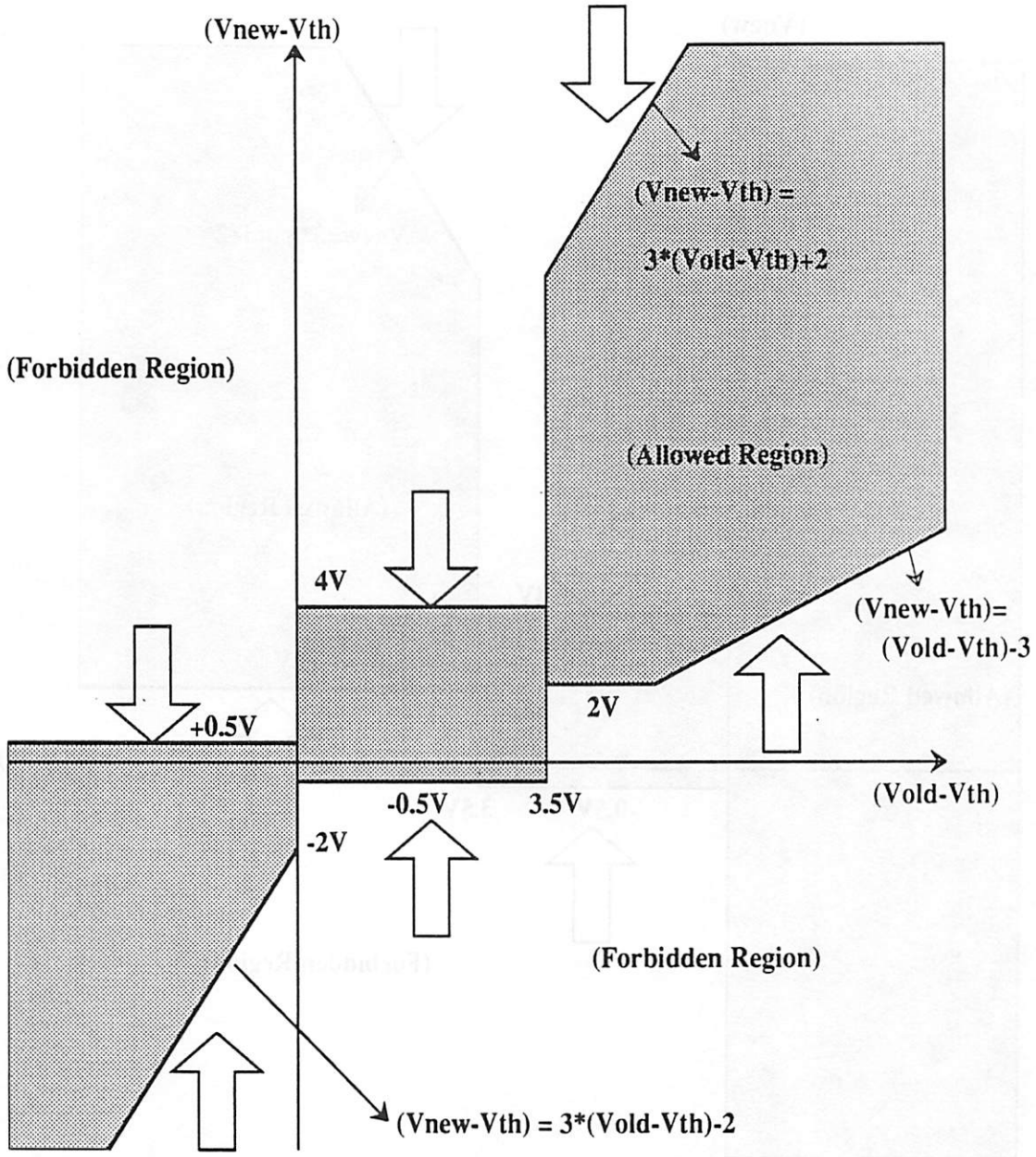


Fig.A5.1. The pictorial representation of the limiting routine for V_{GS} in SPICE [2.20]. *Vth* is the threshold voltage and all the constant numbers are in units of [V].

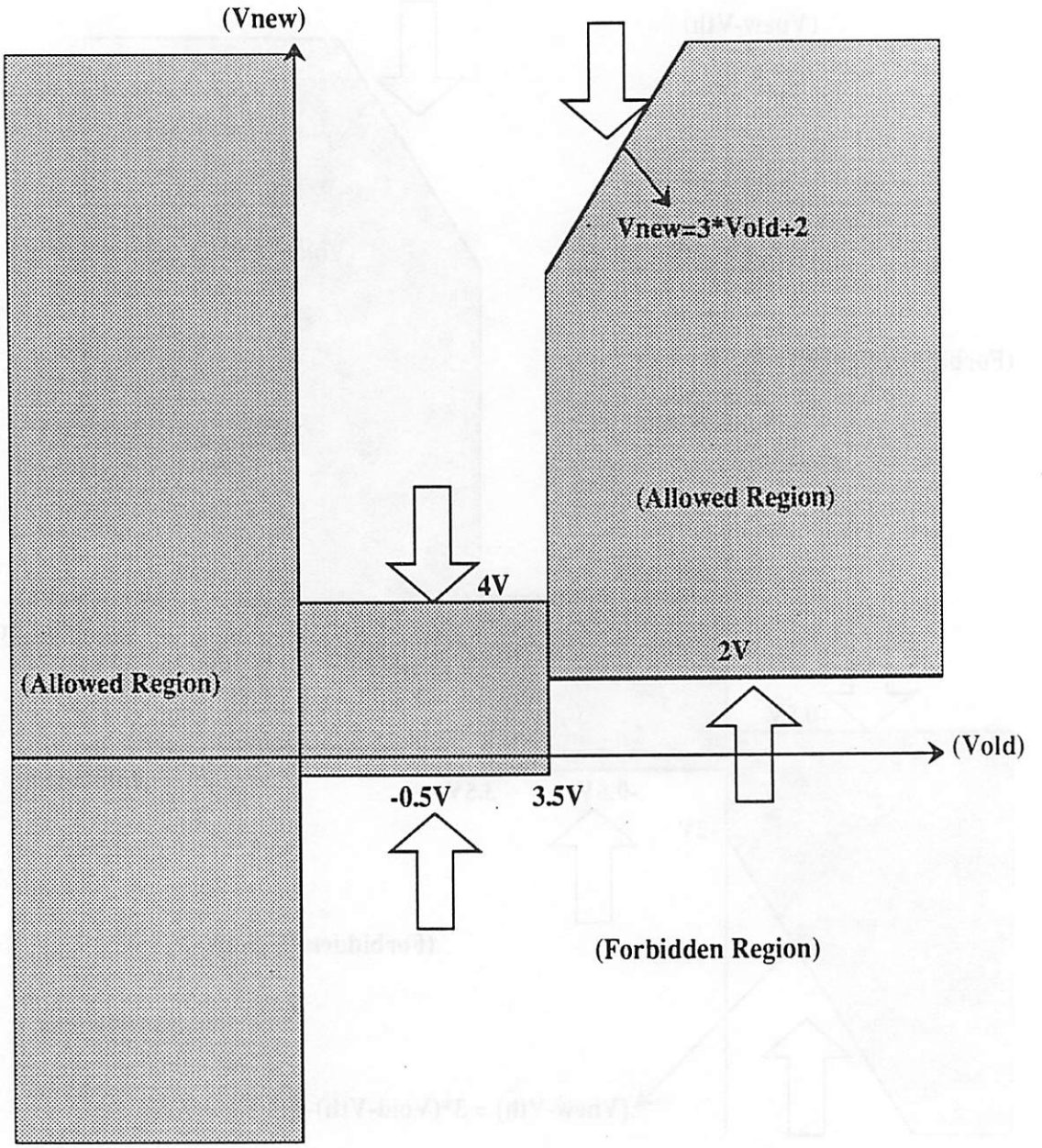


Fig.A5.2. The pictorial representation of the limiting routine for V_{DS} in SPICE.

limiting routine which is based on the exponential I-V characteristics and more explanations on them can be found in [2.20].

(2) current limiting routine

The conventional voltage limiting routines in SPICE are based on heuristic approaches as shown in Fig.A5.1 and Fig.A5.2. A new current limiting routine is developed in this work and it is demonstrated that the current limiting routine can be practically used in simulations.

The current limiting routine limits the change of the DC drain current I_D per iteration, such that,

$$|\Delta I_{D.MAX}| = \max(ABSCUR * \mu_0 C_{OX} * \frac{W}{L} , RELCUR * I_{D.OLD}) \qquad \text{(A5.1)}$$

where *ABSCUR* and *RELCUR* are model parameters which can be specified in the SPICE input model card. The default values for *ABSCUR* and *RELCUR* are 5 and 1 respectively. $I_{D.OLD}$ is the DC drain current computed from the previous iteration. The maximum change allowed for V_{GS} per iteration is

$$|\Delta V_{GS.MAX}| = \min\left[\frac{\Delta I_{D.MAX}}{g_{M.OLD}} , 10^{30}V\right] \qquad \text{(A5.2)}$$

where $g_{M.OLD}$ is the DC transconductance $\partial I_D/\partial V_{GS}$ computed from the previous iteration. $10^{30}V$ is a heuristic factor to prevent a blow up when $g_{M.OLD}$ is very small.

The maximum change allowed for V_{DS} per iteration is

$$|\Delta V_{DS.MAX}| = \min\left[\frac{\Delta I_{D.MAX}}{g_{DS.OLD}} , 10^{30}V\right] \qquad \text{(A5.3)}$$

where $g_{DS.OLD}$ is the DC drain conductance $\partial I_D/\partial V_{DS}$ computed from the previous iteration.

For limiting V_{BS}, the conventional diode limiting routine *DEVpnjlim* is used.

Appendix 6

SPICE3 input listings

for the examples in Chapter 2

<The SPICE3 input listing for Table.2.2>

```
Transient analysis of Depletion Load NMOS ROSC
* subckt for an E-D NMOS inverter
.subckt dinv 1 2 3
m1 2 1 0 0 nch w=10u l=3u ad=36p as=36p pd=12u ps=18u
m2 3 2 2 0 dep w=3u  l=3u ad=36p pd=18u
.ends dinv
*
x1 1 2 vtrig dinv
x2 2  3 vdd dinv
x3 3  4 vdd dinv
x4 4  5 vdd dinv
x5 5  6 vdd dinv
x6 6  7 vdd dinv
x7 7  8 vdd dinv
x8 8  9 vdd dinv
x9 9 10 vddp dinv
x10 10 11 vdd dinv
x11 11  1 vdd dinv
*
vdd vdd 0 dc 5
vddp vddp 0 dc 5
vtrig vtrig 0 dc 0 pwl 0 0 5n 5 1 5
*
.tran 0.1n 100n
.print tran v(10)
.option acct
*
* model parameters for the Charge Sheet MOSFET model
*_____________________________________________________
.model nch nmos level=5 vfb=-1.12 tox=0.050 nsub=2e16 u0=480
+          lambda=0.00 gamma=1.18 phi=0.75
.model dep nmos level=5 vfb=-5.24 tox=0.050 nsub=2e16 u0=480
+          lambda=0.00 gamma=1.18 phi=0.75
*
* model parameters for the SPICE level-2 with Meyer cap. model
*_____________________________________________________
```

```
*.model nch nmos level=2 vto=1.02 tox=0.050u nsub=2e16 u0=480
*+          lambda=0.00 gamma=1.18 phi=0.75 nfs=5e10
*.model dep nmos level=2 vto=-3.10 tox=0.050u nsub=2e16 u0=480
*+          lambda=0.00 gamma=1.18 phi=0.75 nfs=5e10
* model parameters for the BSIM 40/60 (xpart=0)
*
*.model nch nmos level=4   xpart = 0
*+ vfb = -0.752 phi = 0.75 k1 = 1.18 k2 = 0.0 n0=1
*+ muz = 480 mus = 480 tox = 0.050 temp = 27 vdd = 5
*.model dep nmos level=4
*+ vfb = -4.87 phi = 0.75 k1 = 1.18 k2 = 0.0 n0=1
*+ muz = 480 mus = 480 tox = 0.050 temp = 27 vdd = 5 xpart = 0
*
.end
```

<The SPICE3 input listing for Table.2.3>

```
MOSAMP2 - NMOS OP AMP - TRANSIENT ANALYSIS
* IN UNITY GAIN CONFIGURATION
*
.OPTIONS ACCT ABSTOL=10N  VNTOL=10N itl1=200
.TRAN 0.1US 10US
*
M1  15 15  1 32 M W=88.9U  L=25.4U
M2   1  1  2 32 M W=12.7U  L=266.7U
M3   2  2 30 32 M W=88.9U  L=25.4U
M4  15  5  4 32 M W=12.7U  L=106.7U
M5   4  4 30 32 M W=88.9U  L=12.7U
M6  15 15  5 32 M W=44.5U  L=25.4U
M7   5 20  8 32 M W=482.6U L=12.7U
M8   8  2 30 32 M W=88.9U  L=25.4U
M9  15 15  6 32 M W=44.5U  L=25.4U
M10  6 21  8 32 M W=482.6U L=12.7U
M11 15  6  7 32 M W=12.7U  L=106.7U
M12  7  4 30 32 M W=88.9U  L=12.7U
M13 15 10  9 32 M W=139.7U L=12.7U
M14  9 11 30 32 M W=139.7U L=12.7U
M15 15 15 12 32 M W=12.7U  L=207.8U
M16 12 12 11 32 M W=54.1U  L=12.7U
M17 11 11 30 32 M W=54.1U  L=12.7U
M18 15 15 10 32 M W=12.7U  L=45.2U
M19 10 12 13 32 M W=270.5U L=12.7U
M20 13  7 30 32 M W=270.5U L=12.7U
M21 15 10 14 32 M W=254U   L=12.7U
M22 14 11 30 32 M W=241.3U L=12.7U
M23 15 20 16 32 M W=19U    L=38.1U
M24 16 14 30 32 M W=406.4U L=12.7U
M25 15 15 20 32 M W=38.1U  L=42.7U
M26 20 16 30 32 M W=381U   L=25.4U
M27 20 15 66 32 M W=22.9U  L=7.6U
*
CC 7  9 40PF
CL 66 0 70PF
*
VIN 21 0 PULSE(0 5 1NS 1NS 1NS 5US 10US)
VCCP 15 0 DC +15
VDDN 30 0 DC -15
VB 32 0 DC -20
*
.PRINT TRAN V(20) V(66)
*
* initial conditions for node voltages
.IC
+ V( 1)=  9.6384  V( 2)=-11.7269  V( 4)=-12.0573  V( 5)=  9.4891
+ V( 6)=  9.6563  V( 7)=-11.9377  V( 8)= -2.3574  V( 9)=-12.5579
```

```
+ V(10)= -9.8382  V(11)=-12.0735  V(12)= -8.8779  V(13)=-12.2137
+ V(14)=-12.5253  V(15)= 15.0000  V(16)=-11.9220  V(20)=  0.0000
+ V(21)= 0.0000   V(30)=-15.0000  V(32)=-20.0000  V(66)=  0.0000
*
*_____________________________________________________________
* Charge Sheet MOSFET model parameters
.model M nmos level=5 vfb=-1 tox=0.11 nsub=2.2e15 u0=575
+         vsat=1e7 ecrit=4.9e5 dl=5.00 dw=0.0
+         cgso=1.5N cgdo=1.5N relcur=1.0 abscur=10
*
*_____________________________________________________________
* SPICE level 2 Meyer capacitance model
*.model M nmos level=2 vto=0.573 tox=0.11u nsub=2.2e15 u0=575
*+         cgso=1.5N cgdo=1.5N ld=2.5u nfs=5e10
*
*_____________________________________________________________
* BSIM 40/60 model (xpart=0)
*.model m nmos level=4 xpart = 0
*+         vfb = -0.72 phi = 0.62 k1 = 0.85 k2 = 0.0 n0=1
*+         muz = 575 mus = 575 tox = 0.11 temp = 27 vdd = 5
*+         dl=5.00 cgso=1.5N cgdo=1.5N
*
.END
```

<The SPICE3 input listing for Table.2.5>

```
MOSAMP1 - MOS AMPLIFIER - AC ANALYSIS
.OPT ACCT
.OPT ABSTOL=10N VNTOL=10N gmin=1e-9
.AC DEC 10 100 1000MEG
M1  15 15  1 32 M W=88.9U  L=25.4U
M2   1  1  2 32 M W=12.7U  L=266.7U
M3   2  2 30 32 M W=88.9U  L=25.4U
M4  15  5  4 32 M W=12.7U  L=106.7U
M5   4  4 30 32 M W=88.9U  L=12.7U
M6  15 15  5 32 M W=44.5U  L=25.4U
M7   5  0  8 32 M W=482.6U L=12.7U
M8   8  2 30 32 M W=88.9U  L=25.4U
M9  15 15  6 32 M W=44.5U  L=25.4U
M10  6 21  8 32 M W=482.6U L=12.7U
M11 15  6  7 32 M W=12.7U  L=106.7U
M12  7  4 30 32 M W=88.9U  L=12.7U
M13 15 10  9 32 M W=139.7U L=12.7U
M14  9 11 30 32 M W=139.7U L=12.7U
M15 15 15 12 32 M W=12.7U  L=207.8U
M16 12 12 11 32 M W=54.1U  L=12.7U
M17 11 11 30 32 M W=54.1U  L=12.7U
M18 15 15 10 32 M W=12.7U  L=45.2U
M19 10 12 13 32 M W=270.5U L=12.7U
M20 13  7 30 32 M W=270.5U L=12.7U
M21 15 10 14 32 M W=254U   L=12.7U
M22 14 11 30 32 M W=241.3U L=12.7U
M23 15 20 16 32 M W=19U    L=38.1U
M24 16 14 30 32 M W=406.4U L=12.7U
M25 15 15 20 32 M W=38.1U  L=42.7U
M26 20 16 30 32 M W=381U   L=25.4U
M27 20 15 66 32 M W=22.9U  L=7.6U
CC 7 9 40PF
CL 66 0 70PF
VIN 21 0 DC 0 AC 1
VCCP 15 0 DC +15
VCCN 30 0 DC -15
VB 32 0 DC -20
.print  ac VDB(66)
.print  ac VP(66)
*
*_____________________________________________________
* model parameters for Charge Sheet MOSFET Model
.model M nmos level=5 vfb=-1 tox=0.11 nsub=2.2e15 u0=575
+         vsat=1e7 ecrit=4.9e5 dl=5.00 dw=0.0 satfactor=15
+         cgso=1.5N cgdo=1.5N relcur=1.0 abscur=10
*
*_____________________________________________________
* model parameters for SPICE level 2 Meyer capacitance model
```

```
*.model M nmos level=2 vto=0.573 tox=0.11u nsub=2.2e15 u0=575
*+          ld=2.50u cgso=1.5N cgdo=1.5N nfs=5e10
*
*_____________________________________________________________
* model parameters for BSIM 40/60 (xpart=0) model
*.model m nmos level=4  xpart = 0
*+ vfb = -1. phi = 0.62 k1 = 0.85 k2 = 0.0 n0=1
*+ muz = 575 mus = 575 tox = 0.11 temp = 27 vdd = 15
*+ cgdo=1.5n cgso=1.5n dl=5.0
*
.END
```

Appendix 7

Programs for Computing MOSFET Capacitances using Different Models

Programs for computing MOSFET capacitances are shown for the Meyer model and the Ward-Dutton model, as implemented in SPICE2.

A7.1. Program for the Meyer capacitance model

```
/* program to compute the MOSFET capacitances using the Meyer model
 *              Copyright Hong J. Park, 1987
 */
#include <stdio.h>
#include <math.h>

/* main program */
main() {
double Vgs,Vds,Vbs, Vth, Vdsat, Phi, Cox;
double Cgd,Cgs,Cgg,Cgb, Gamma, Vbin;
FILE *fpg,*fpd,*fps,*fpb;

    fpg = fopen("Cgg","w");
    fpd = fopen("Cgd","w");
    fps = fopen("Cgs","w");
    fpb = fopen("Cgb","w");

    Vth = 0.776;
    Phi = 0.967;
    Gamma = 1.087;
    Cox = 2.34e-12;
    Vbs = 0;
    Vbin = Vth - Gamma * sqrt(Phi);
    for(Vds = 0; Vds <= 5 ; Vds += 2 ) {
     for(Vgs = -2 ; Vgs <= 8 ; Vgs += .025) {
        /*Vdsat = Vgs - Vth;*/
        Vdsat = Vgs - Vbin + 0.5 * Gamma*Gamma
            - sqrt(.5*Gamma*Gamma*(2.*(Vgs-Vbs-Vbin+Phi)+.5*Gamma*Gamma));
        CMEYER(Vgs,Vgs-Vds,Vgs-Vbs,Vth,Vdsat,Cox,Phi,&Cgd,&Cgs,&Cgb);
        Cgg = Cgd + Cgs + Cgb;
        fprintf(fpg,"%e %e0,Vgs,Cgg);
```

```
        fprintf(fpd,"%e %e0,Vgs,Cgd);
        fprintf(fps,"%e %e0,Vgs,Cgs);
        fprintf(fpb,"%e %e0,Vgs,Cgb);
        }
        fprintf(fpg,"%c %c0,'"','"');
        fprintf(fpd,"%c %c0,'"','"');
        fprintf(fps,"%c %c0,'"','"');
        fprintf(fpb,"%c %c0,'"','"');
    }
}

/* subroutine to compute capacitance values for applied biases
 * This routine is directly converted from subroutine CMEYER of
 * SPICE2G.6 into a C language program
 */
    CMEYER (VGS,VGD,VGB,VON,VDSAT,COX,PHI,CGDP,CGSP,CGBP)
    double VGS,VGB,VGD,VON,VDSAT,COX,PHI;
    double *CGDP,*CGBP,*CGSP;
    {
    double VONS,VBS,VDBSAT,VDB,VDS,VGBT;
    double VDDIF, VDDIF1, VDDIF2;
    double COVLGB,COVLGD,COVLGS;
    double CGD,CGS,CGB;
    COVLGB = 0.0;
    COVLGD = 0.0;
    COVLGS = 0.0;
    VONS=VON;
    VBS=VGS-VGB;
    VDBSAT=VDSAT-VBS;
    VDB=VGB-VGD;
    VDS=VGS-VGD;
    VGBT=VGS-VONS;
    if (VGBT > (-PHI)) goto 1100;
    CGB=COX+COVLGB;
    CGD=COVLGD;
    CGS=COVLGS;
    goto 1430;
1100: if (VGBT > -PHI/2.0) goto 1200;
    CGB= -VGBT*COX/PHI+COVLGB;
    CGD=COVLGD;
    CGS=COVLGS;
    goto 1430;
1200: if (VGBT > 0.0) goto 1300;
    CGB= -VGBT*COX/PHI+COVLGB;
    CGD=COVLGD;
    CGS=COX/(7.5e-1*PHI)*VGBT+COX/1.5e0+COVLGS;
    goto 1430;
1300: if (VDBSAT > VDB) goto 1400;
    CGB=COVLGB;
    CGD=COVLGD;
    CGS=COX/1.5e0+COVLGS;
    goto 1430;
1400: VDDIF=2.0e0*VDBSAT-VDB;
```

```
    VDDIF1=VDBSAT-VDB-1.0e-12;
    /*VDDIF2=VDDIF*VDDIF; Original */
    VDDIF2=VDDIF*VDDIF + 1.0e-12;  /* Modified */
    CGD=COX*(1.0e0-VDBSAT*VDBSAT/VDDIF2)/1.5e0+COVLGD;
    CGS=COX*(1.0e0-VDDIF1*VDDIF1/VDDIF2)/1.5e0+COVLGS;
    CGB=COVLGB;

 l430: *CGBP = CGB;
       *CGDP = CGD;
       *CGSP = CGS;
}
```

A7.2. Program for the Ward-Dutton capacitance model

```
/*
 * program to print out the MOSFET capacitances  using the
 * Ward-Dutton model implemented in SPICE2
 *          Copyright (c) 1986 Hong J. Park
 */

#define GRAPH 1
#include <stdio.h>
#include <math.h>
#define CONSTvt  0.026
#define MAX(a,b) ((a) > (b) ? (a) : (b))
#define MIN(a,b) ((a) < (b) ? (a) : (b))

/* global Variables */
double Vfb= -1.26;
double Phi=0.967;
double Gamma=1.087;
double WLCox=2.34e-12;
double XQC = 0.4;
double Vth;


#ifdef GRAPH
/* File Pointers for graph files are declared as global variables */
FILE *fileID;
FILE *fileGM;
FILE *fileGDS;
FILE *fileGMBS;
FILE *fileQGATE;
FILE *fileQBULK;
FILE *fileQSRC;
FILE *fileQDRN;
FILE *fileCGGB;
FILE *fileCGDB;
FILE *fileCGSB;
FILE *fileCGBB;
FILE *fileCDGB;
```

```
FILE *fileCDDB;
FILE *fileCDSB;
FILE *fileCDBB;
FILE *fileCBGB;
FILE *fileCBDB;
FILE *fileCBSB;
FILE *fileCBBB;
FILE *fileCSGB;
FILE *fileCSDB;
FILE *fileCSSB;
FILE *fileCSBB;
#endif GRAPH


/* subroutine to compute charges and capacitances using
 * the Ward-Dutton model, this routine is directly converted
 * from subroutine MOSQ2 in SPICE2 into a C language program
 */
    MOSQ2(VDS,VBS,VGS,VTH,GAMASD,COX,PHI,GMptr,GDSptr,GMBSptr,
      QGptr,QBptr,QDptr,CGGBptr,CGDBptr,CGSBptr,CBGBptr,CBDBptr,
      CBSBptr,CDGBptr,CDDBptr,CDSBptr,CDRAINptr)

    double VDS,VBS,VGS,VTH,GAMASD,COX,PHI;
    double *QGptr,*CGGBptr,*CGDBptr,*CGSBptr;
    double *QBptr,*CBGBptr,*CBDBptr,*CBSBptr;
    double *QDptr,*CDGBptr,*CDDBptr,*CDSBptr;
    double *GMptr,*GDSptr,*GMBSptr,*CDRAINptr;

{
   double VBD,VGB,VD,VS,VG,VSP5,VSAT,VS2,VS3,VS5,VS1P5,VS2P5;
   double VE,DVEDVD,DVEDVG,VE2,VE3,VE5,VEP5,VE1P5,VE2P5;
   double TERM0,TERM1,TERM2,TERM3,TERM4,TERM5,TERM6,TERM7;
   double TERM10,TERM11,TERM12,TERM20,TERM21,TERM22;
   double ARGN,ARGD,ARGD2,DGNDVE,DDDVE,DQGDVE,DGNDVS,DDDVS;
   double GAMMA2,SQARG;
   double QG,QB,QD,QS;
   double CGGB,CGDB,CGSB;
   double CBGB,CBDB,CBSB;
   double CDGB,CDDB,CDSB;
   double CSGB,CSDB,CSSB;
   double CGBB,CDBB,CBBB,CSBB;
   double QC,QCSAT,dQCdVg,dQCdVd,dQCdVs;
   double dQCSATdVg,dQCSATdVd,dQCSATdVs;
   double Tmp,SqrtVS;
   double VDSAT,VBIN;

/*C
/*C    INITIALIZE CHARGES;
/*C    CHANGE REFERENCE VOLTAGES FOR CHARGE COMPUTATION
/*C
*/
    VBIN = Vfb + Phi;
    QG=0.0e0;
```

```
      QB=0.0e0;
      VBD=VBS-VDS;
      VGB=VGS-VBS;
      VD=MAX(PHI-VBD,1.0e-8);
      VS=MAX(PHI-VBS,1.0e-8);
      VG=VGB-VBIN+PHI;
      VSP5=sqrt(VS);
/*C;
/*C     DETERMINE OPERATING REGION;
/*C;
*/
      if (VGS <= VTH) goto L1100;
/*C;
/*C     COMPUTE CHARGES FOR "ON" REGION;
/*C;
*/
L1020: VDSAT = VG + 0.5 * GAMASD*GAMASD
           - sqrt(.5*GAMASD*GAMASD*(2.*VG+.5*GAMASD*GAMASD)) - VS;
      VSAT=VDSAT+VS;
      VS2=VS*VS;
      VS3=VS2*VS;
      VS5=VS3*VS2;
      VS1P5=VS*VSP5;
      VS2P5=VS1P5*VS;
 L1025: if (VD >= VSAT) goto L1035;
      VE=VD;
 L1030: DVEDVD=1.0e0;
      DVEDVG=0:0e0;
      goto L1040;
 L1035: VE=VSAT;
      DVEDVD=0.0e0;
      DVEDVG=0.0e0;
 L1040: VE2=VE*VE;
      VE3=VE2*VE;
      VE5=VE2*VE3;
      VEP5=sqrt(VE);
      VE1P5=VE*VEP5;
      VE2P5=VE1P5*VE;
      TERM0=VE+VS;
      TERM1=VEP5+VSP5;
      TERM2=VEP5*VSP5;
      TERM3=VE2+VS2;
      TERM4=VE*VS;
      TERM5=TERM0*TERM1;
      TERM6=(TERM3+TERM4)+TERM2*TERM0;
      TERM7=(TERM3+TERM4)*TERM1;
      TERM10=VEP5+0.5e0*VSP5;
      TERM11=1.5e0*VE+VSP5*TERM10;
      TERM12=2.0e0*VE1P5+VSP5*TERM11;
      TERM20=0.5e0*VEP5+VSP5;
      TERM21=1.5e0*VS+VEP5*TERM20;
      TERM22=2.0e0*VS1P5+VEP5*TERM21;
      ARGN=0.5e0*VG*TERM5-0.4e0*GAMASD*TERM6-TERM7/3.0e0;
```

```
      ARGD=VG*TERM1-GAMASD*(TERM0+TERM2)/1.5e0-0.5e0*TERM1*TERM0;
      ARGD2=ARGD*ARGD;
      QG=COX*(VG-ARGN/ARGD);
      DGNDVE=0.5e0*VG*TERM11-0.4e0*GAMASD*TERM12-
        (2.5e0*VE2+VSP5*TERM12)/3.0e0;
      DDDVE=0.5e0*VG-GAMASD*TERM10/1.5e0-0.5e0*TERM11;
      DQGDVE= -COX/ARGD*(DGNDVE-(VG-QG/COX)*DDDVE);
      DGNDVS=0.5e0*VG*TERM21-0.4e0*GAMASD*TERM22-
        (2.5e0*VS2+VEP5*TERM22)/3.0e0;
      DDDVS=0.5e0*VG-GAMASD*TERM20/1.5e0-0.5e0*TERM21;
      CGDB= -COX/(ARGD*VEP5)*(DGNDVE-(VG-QG/COX)*DDDVE)*DVEDVD;
      CGSB= -COX/(ARGD*VSP5)*(DGNDVS-(VG-QG/COX)*DDDVS);
      CGGB=COX*(1.0e0-TERM1/ARGD*(0.5e0*TERM0-VG+QG/COX));
      ARGN=VG*(TERM0+TERM2)/1.5e0-0.5e0*GAMASD*TERM5-0.4e0*TERM6;
      DGNDVE=VG*TERM10/1.5e0-0.5e0*GAMASD*TERM11-0.4e0*TERM12;
      DGNDVS=VG*TERM20/1.5e0-0.5e0*GAMASD*TERM21-0.4e0*TERM22;
      QB= -GAMASD*COX*ARGN/ARGD;
      CBDB= -COX/(VEP5*ARGD)*(QB/COX*DDDVE+GAMASD*DGNDVE)*DVEDVD;
      CBSB= -COX/(VSP5*ARGD)*(QB/COX*DDDVS+GAMASD*DGNDVS);
      CBGB= -COX/ARGD*(GAMASD*(TERM0+TERM2)/1.5e0+QB/COX*TERM1);
      QC = -(QG+QB);
      dQCdVg = -(CGGB + CBGB);
      dQCdVd = -(CGDB + CBDB);
      dQCdVs = -(CGSB + CBSB);
      Tmp = -2./3. * COX;
      SqrtVS = sqrt(VS);
      QCSAT = Tmp * (VG-VS-GAMASD*SqrtVS);
      dQCSATdVg = Tmp;
      dQCSATdVd = 0.0;
      dQCSATdVs = Tmp * (-1. - GAMASD * 0.5 / SqrtVS);
      QD = XQC * QCSAT + (1.5 - 2.*XQC)*(QC-QCSAT);
      CDGB = XQC * dQCSATdVg + (1.5-2*XQC)*(dQCdVg-dQCSATdVg);
      CDDB = XQC * dQCSATdVd + (1.5-2*XQC)*(dQCdVd-dQCSATdVd);
      CDSB = XQC * dQCSATdVs + (1.5-2*XQC)*(dQCdVs-dQCSATdVs);
      QS = (1.-XQC)*QCSAT + (2.*XQC-0.5)*(QC-QCSAT);
      CSGB = (1.-XQC) * dQCSATdVg + (2*XQC-0.5)*(dQCdVg-dQCSATdVg);
      CSDB = (1.-XQC) * dQCSATdVd + (2*XQC-0.5)*(dQCdVd-dQCSATdVd);
      CSSB = (1.-XQC) * dQCSATdVs + (2*XQC-0.5)*(dQCdVs-dQCSATdVs);
      goto L2000;
/*C;
/*C  FINISH SPECIAL CASES;
/*C;
*/
 L1100: if (VG > 0.0e0) goto L1110;

      /* Accumulation Region */
      QG=COX*VG;
      CGGB=COX;
      goto L1120;

 L1110:  /* Depletion Region */
      GAMMA2=GAMASD*0.5e0;
      SQARG=sqrt(GAMMA2*GAMMA2+VG);
```

```
    QG=GAMASD*COX*(SQARG-GAMMA2);
    CGGB=0.5e0*COX*GAMASD/SQARG;

L1120:  /* Depletion, Accumulation Region */
    QB= -QG;
    CBGB= -CGGB;
    CGDB=0.0e0;
    CGSB=0.0e0;
    CBDB=0.0e0;
    CBSB=0.0e0;
    QD = 0.0;
    QS = 0.0;
    CDGB = CDDB = CDSB = CDBB = 0.0;
    CSGB = CSDB = CSSB = CSBB = 0.0;
/*C;
/*C  FINISHED;
/*C;
*/
 L2000:   *QGptr = QG;
     *QDptr = QD;
     *QBptr = QB;

     *CGGBptr = CGGB;
     *CGDBptr = CGDB;
     *CGSBptr = CGSB;

     *CDGBptr = CDGB;
     *CDDBptr = CDDB;
     *CDSBptr = CDSB;

     *CBGBptr = CBGB;
     *CBDBptr = CBDB;
     *CBSBptr = CBSB;

 }


/* main routine and house keeping stuffs*/

main() {

double vgs_from,vgs_to,vgs_step;
double vds_from,vds_to,vds_step;
double vbs_from,vbs_to,vbs_step;
double vds,vbs,vgs;
double gm,gds,gmbs;
double qg,qb,qd,qsrc;
double cggb,cgdb,cgsb,cgbb;
double cbgb,cbdb,cbsb,cbbb;
double cdgb,cddb,cdsb,cdbb;
double csgb,csdb,cssb,csbb;
double cdrain,v1;
char Quote='"';
```

```
FILE *fp;

printf("VGS: from to step0);
scanf("%f %f %f",&vgs_from,&vgs_to,&vgs_step);
printf("VDS: from to step0);
scanf("%f %f %f",&vds_from,&vds_to,&vds_step);
printf("VBS: from to step0);
scanf("%f %f %f",&vbs_from,&vbs_to,&vbs_step);
if(vbs_to > 0 ) vbs_to = - vbs_to;
if(vbs_step > 0 ) vbs_step = - vbs_step;

#ifdef GRAPH
   fileID = fopen("ID","w");
   fileGM = fopen("GM","w");
   fileGDS = fopen("GDS","w");
   fileGMBS = fopen("GMBS","w");
   fileQGATE = fopen("QGATE","w");
   fileQBULK = fopen("QBULK","w");
   fileQSRC = fopen("QSRC","w");
   fileQDRN = fopen("QDRN","w");
   fileCGGB = fopen("CGGB","w");
   fileCGDB = fopen("CGDB","w");
   fileCGSB = fopen("CGSB","w");
   fileCGBB = fopen("CGBB","w");
   fileCDGB = fopen("CDGB","w");
   fileCDDB = fopen("CDDB","w");
   fileCDSB = fopen("CDSB","w");
   fileCDBB = fopen("CDBB","w");
   fileCBGB = fopen("CBGB","w");
   fileCBDB = fopen("CBDB","w");
   fileCBSB = fopen("CBSB","w");
   fileCBBB = fopen("CBBB","w");
   fileCSGB = fopen("CSGB","w");
   fileCSDB = fopen("CSDB","w");
   fileCSSB = fopen("CSSB","w");
   fileCSBB = fopen("CSBB","w");
#endif GRAPH


for(vbs=vbs_from;vbs>=vbs_to;vbs += vbs_step)    {
    Vth = Vfb + Phi + Gamma * sqrt(Phi - vbs);
    for(vds=vds_from;vds<=vds_to;vds += vds_step)
    {

        for(vgs=vgs_from;vgs<=vgs_to;vgs += vgs_step)
        {
        MOSQ2(vds,vbs,vgs,Vth,Gamma,WLCox,Phi,&gm,&gds,&gmbs,
        &qg,&qb,&qd,&cggb,&cgdb,&cgsb ,&cbgb,&cbdb,&cbsb,&cdgb,&cddb
        ,&cdsb,&cdrain);
        cdrain = gm = gds = gmbs = 0.0;

        v1 = vgs;      /* the First Variable */
        qsrc = - (qg+qd+qb);
```

```
            cgbb = - (cggb+cgdb+cgsb);
            cdbb = - (cdgb+cddb+cdsb);
            cbbb = - (cbgb+cbdb+cbsb);
            csgb = - (cggb+cdgb+cbgb);
            csdb = - (cgdb+cddb+cbdb);
            cssb = - (cgsb+cdsb+cbsb);
            csbb = - (cgbb+cdbb+cbbb);

#ifdef GRAPH

    fprintf( fileID,"%g %g 0,v1,cdrain);
    fprintf( fileGM,"%g %g 0,v1,gm);
    fprintf( fileGDS,"%g %g 0,v1,gds);
    fprintf( fileGMBS,"%g %g 0,v1,gmbs);
    fprintf( fileQGATE,"%g %g 0,v1,qg);
    fprintf( fileQBULK,"%g %g 0,v1,qb);
    fprintf( fileQSRC,"%g %g 0,v1,qsrc);
    fprintf( fileQDRN,"%g %g 0,v1,qd);
    fprintf( fileCGGB,"%g %g 0,v1,cggb);
    fprintf( fileCGDB,"%g %g 0,v1,cgdb);
    fprintf( fileCGSB,"%g %g 0,v1,cgsb);
    fprintf( fileCGBB,"%g %g 0,v1,cgbb);
    fprintf( fileCDGB,"%g %g 0,v1,cdgb);
    fprintf( fileCDDB,"%g %g 0,v1,cddb);
    fprintf( fileCDSB,"%g %g 0,v1,cdsb);
    fprintf( fileCDBB,"%g %g 0,v1,cdbb);
    fprintf( fileCBGB,"%g %g 0,v1,cbgb);
    fprintf( fileCBDB,"%g %g 0,v1,cbdb);
    fprintf( fileCBSB,"%g %g 0,v1,cbsb);
    fprintf( fileCBBB,"%g %g 0,v1,cbbb);
    fprintf( fileCSGB,"%g %g 0,v1,csgb);
    fprintf( fileCSDB,"%g %g 0,v1,csdb);
    fprintf( fileCSSB,"%g %g 0,v1,cssb);
    fprintf( fileCSBB,"%g %g 0,v1,csbb);
#endif GRAPH

        }

#ifdef GRAPH

    fprintf( fileID,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileGM,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileGDS,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileGMBS,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileQGATE,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileQBULK,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileQSRC,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileQDRN,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileCGGB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileCGDB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileCGSB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileCGBB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    fprintf( fileCDGB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
```

```
        fprintf( fileCDDB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCDSB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCDBB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCBGB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCBDB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCBSB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCBBB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCSGB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCSDB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCSSB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
        fprintf( fileCSBB,"%c %s %3.1f%c 0,Quote,"VDS=",vds,Quote);
    #endif GRAPH
        }

      }
    }
```

Appendix 8

Derivation and solution of equation (3.22)

Substituting (3.17) and (3.19) into the diffusion equation (3.10), we have

$$\frac{dP_S}{dt} - \left(\frac{dP_S}{dt} - \frac{dP_D}{dt}\right)\cdot\frac{y}{L} + \sum_{n=1}^{N}\frac{dA_n}{dt}\cdot\sin\left(n\pi\frac{y}{L}\right)$$

$$= -\left\{D_0 + D_1\cdot\cos\left(\pi\frac{y}{L}\right)\right\}\sum_{n=1}^{N}A_n\,\frac{(n\pi)^2}{L^2}\cdot\sin\left(n\pi\frac{y}{L}\right) \qquad \text{(A8.1)}$$

The argument for the time dependence (t) has been dropped from P_S, P_D, A_n, D_0 and D_1 for simplicity. Expanding the y dependence of the first two terms in the left hand side of (A1), that is, 1 and y/L, in terms of Fourier sine series and matching the coefficients of each sine series term in both sides of the equation (A8.1), we have

$$\frac{dA_n}{dt} = -\frac{D_0}{L^2}\cdot n^2\cdot\pi^2\cdot A_n - \frac{D_1}{2L^2}\cdot(n-1)^2\cdot\pi^2\cdot A_{(n-1)}$$

$$-\frac{D_1}{2L^2}\cdot(n+1)^2\cdot\pi^2\cdot A_{(n+1)} - \frac{2}{n\pi}\cdot\frac{dP_S}{dt} + (-1)^n\cdot\frac{2}{n\pi}\cdot\frac{dP_D}{dt} \qquad \text{(A8.2)}$$

where n = 1, 2, 3, 4, ···, N. $A_{(0)}$ and $A_{(N+1)}$ are taken to be 0 in (A8.2).

The interactions between coefficients $\{A_n\}$ with different values of n are due to the cosine term in (3.19). Rewriting (A8.2) in a matrix equation, we have the state equation (A8.3), which is the same as (3.22).

$$\frac{d\mathbf{A}}{dt} = \mathbf{D}\cdot\mathbf{A} + \frac{dP_D}{dt}\,\mathbf{C_D} + \frac{dP_S}{dt}\,\mathbf{C_S} \qquad \text{(A8.3)}$$

The coefficients of the matrices in (A8.3) can be derived from (A8.2) as.

$$\mathbf{D}_{(n,n)} = -\frac{D_0}{L^2}\cdot n^2\cdot\pi^2 \qquad \text{(A8.4)}$$

$$\mathbf{D}_{(n,n+1)} = -\frac{D_1}{2L^2}\cdot(n+1)^2\cdot\pi^2 \qquad \text{(A8.5)}$$

$$\mathbf{D}_{(n,n-1)} = -\frac{D_1}{2L^2}\cdot(n-1)^2\cdot\pi^2 \tag{A8.6}$$

$$C_{S(n)} = -\frac{2}{n\pi} \tag{A8.7}$$

$$C_{D(n)} = (-1)^n\cdot\frac{2}{n\pi} \tag{A8.8}$$

where n = 1, 2, 3, 4, ···.

Using the trapezoidal integration method, eq. (A8.3) can be rewritten as.

$$\mathbf{A}^{(h)} - \mathbf{A}^{(h-1)} = \frac{k}{2}\cdot(\mathbf{D}^{(h-1)}.\mathbf{A}^{(h-1)} + \mathbf{D}^{(h)}.\mathbf{A}^{(h)}) + \Delta P_D\cdot\mathbf{C}_D + \Delta P_S\cdot\mathbf{C}_S \tag{A8.9}$$

where $\Delta P_D = P_D{}^{(h)} - P_D{}^{(h-1)}$ (A8.10)

$$\Delta P_S = P_S{}^{(h)} - P_S{}^{(h-1)} \tag{A8.11}$$

where k is the time step. The superscript (h) represents the present(new) time point and (h-1) represents the previous(old) time point. Eq. (A8.9) can be rewritten as

$$(\mathbf{I} - \frac{k}{2}\cdot\mathbf{D}^{(h)})\cdot\mathbf{A}^{(h)} = (\mathbf{I} + \frac{k}{2}\cdot\mathbf{D}^{(h-1)})\cdot\mathbf{A}^{(h-1)} + \Delta P_D\cdot\mathbf{C}_D + \Delta P_S\cdot\mathbf{C}_S \tag{A8.12}$$

Since the coefficients of the matrix $\mathbf{D}^{(h)}$ are complex functions of the coefficients of $\mathbf{A}^{(h)}$ and are unknown at this solution step, the same matrix $\mathbf{D}$ is used for $\mathbf{D}^{(h)}$ and $\mathbf{D}^{(h-1)}$. The coefficients of $\mathbf{D}$ are computed from (3.20), (3.21), (A8.4), (A8.5) and (A8.6) using $P_D{}^{(h)}$, $P_S{}^{(h)}$ and $\mathbf{A}^{(h-1)}$. Then (A8.9) can be rewritten as

$$(\mathbf{I} - \frac{k}{2}\cdot\mathbf{D})\cdot(\mathbf{A}^{(h)} - \mathbf{A}^{(h-1)}) = k\ \mathbf{D}\cdot\mathbf{A}^{(h-1)} + \Delta P_D\ \mathbf{C}_D + \Delta P_S\ \mathbf{C}_S \tag{A8.13}$$

Eq. (A8.13) can be decomposed into three equations.

$$\frac{\partial\mathbf{A}}{\partial t} = (\mathbf{I} - \frac{k}{2}\mathbf{D})^{-1}\cdot\mathbf{D}\cdot\mathbf{A}^{(h-1)} \tag{A8.14}$$

$$\frac{\partial\mathbf{A}}{\partial P_D} = (\mathbf{I} - \frac{k}{2}\mathbf{D})^{-1}\cdot\mathbf{C}_D \tag{A8.15}$$

$$\frac{\partial\mathbf{A}}{\partial P_S} = (\mathbf{I} - \frac{k}{2}\mathbf{D})^{-1}\cdot\mathbf{C}_S \tag{A8.16}$$

Since the same matrix $(\mathbf{I} - 0.5\mathrm{k}\ \mathbf{D})$ is used in (A8.14), (A8.15) and (A8.16), these three matrix equations can be solved in one matrix inversion step. $\mathbf{A}^{(h)}$ can be found from

$$\mathbf{A}^{(h)} = \mathbf{A}^{(h-1)} + \frac{\partial \mathbf{A}}{\partial t}\cdot k + \frac{\partial \mathbf{A}}{\partial P_D}\cdot \Delta P_D + \frac{\partial \mathbf{A}}{\partial P_S}\cdot \Delta P_S \qquad \text{(A8.17)}$$

Since $\partial \mathbf{A}/\partial P_D$ and $\partial \mathbf{A}/\partial P_S$ are available and the bias dependences of P_D and P_S can be computed using (3.13) and (3.14), we can get the bias dependencies of $\mathbf{A}$ which will be used to form the Jacobian matrix for the Newton-Raphson iterations in solving the circuit equations.

Appendix 9

Time step control for the simple NQS model

Based on the local truncation error, a new time step control scheme has been derived. For the time step control, the equation (3.22) has been simplified to

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{D} \cdot \mathbf{A} \tag{A9.1}$$

Fig.A9.1 shows the time variation of P_S and other coefficients $\{A_n\}$ for the turn-off transient of the MOSFET shown in Fig.3.5. Since A_1 is dominant over other coefficients by at least an order of magnitude, only the truncation error of A_1 is considered, for simplicity.

$$\frac{\partial A_1}{\partial t} = D_{11} \cdot A_1 + D_{12} \cdot A_2 \tag{A9.2}$$

The local truncation error(L.T.E.) of the trapezoidal integration method to solve (A9.2) is

$$\text{L.T.E.} = \frac{1}{12} \cdot k^3 \cdot (D_{11} \cdot A_1 + D_{12} \cdot A_2)^{(2)} \tag{A9.3}$$

where k is the time step and the superscript (2) in the right hand side of (A9.3) represents the second derivative with respect to time. This second derivative is computed from the divided difference method [3.28]. So the time step k is chosen such that the local truncation error lies within some tolerance as shown in (A9.4).

$$\text{L.T.E.} < \text{QTRTOL} * \text{RELTOL} * (|A_1| + \text{MINTOL}) \tag{A9.4}$$

RELTOL is specified in the option card of SPICE input and QTRTOL and MINTOL are new model parameters. For the overlap capacitance, junction charges and bulk charges, the original time step control scheme in SPICE [3.28] is used.

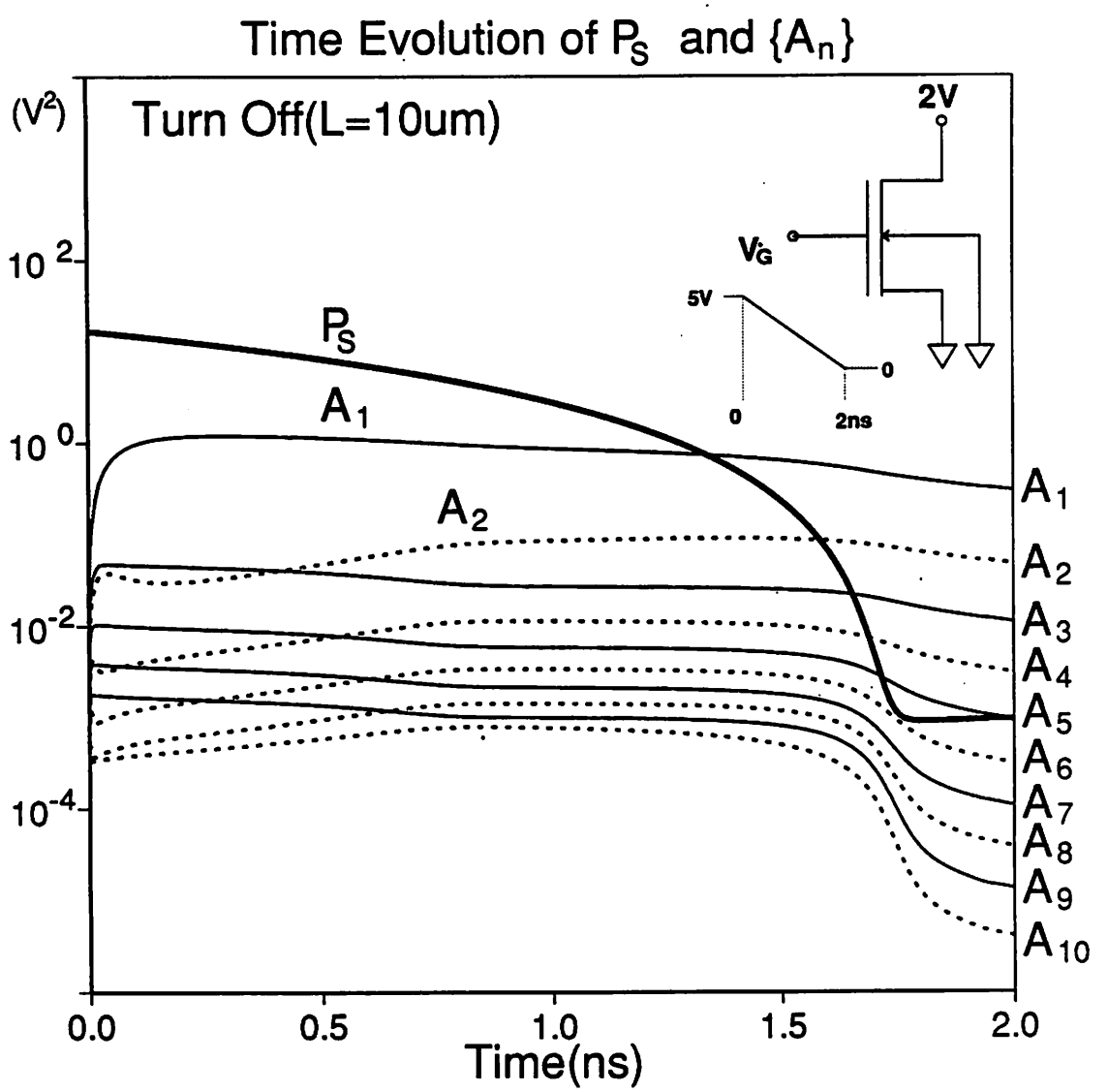


Fig.A9.1. Time evolution of P_S and $\{A_n\}$ for the turn-off transient of the MOSFET shown in Fig.2.5. Solid lines represent P_S and $\{A_n\}$ with odd n and dotted lines represent $\{A_n\}$ with even n.

Appendix 10

Derivation of drain, source currents and drain, source charges from the current continuity equation

Although the derivation of I_D, I_S, Q_D and Q_S from the current continuity equation had been shown in [3.2] and was repeated in [3.10], the detailed derivation steps are shown in this appendix.

Equations for I_D, I_S, Q_D and Q_S can be derived from the current continuity equation (A10.1) by applying an integration with y twice and an integration by parts once in both sides of (A10.1).

$$\frac{\partial I_y(y,t)}{\partial y} = WC_{OX}\cdot\frac{\partial Q_n'(y,t)}{\partial t} \tag{A10.1}$$

where y is the lateral dimension along the channel from source toward drain as shown in Fig.3.1. $Q_n'(y,t)$ is the normalized inversion charge density, $I_y(y,t)$ is the channel current at y as shown in Fig.3.1.

Applying an integration with y to both sides of (A10.1) from 0 to y, we have

$$I_y(y,t) - I_y(0,t) = WC_{OX}\cdot\frac{\partial}{\partial t}\int_0^y Q_n'(y',t)\,dy' \tag{A10.2}$$

Applying another integration with y from 0 to L in both sides of (A10.2), we have

$$L\cdot I_y(0,t) = \int_0^L I_y(y,t)\,dy - WC_{OX}\cdot\int_0^L\int_0^y Q_n'(y',t)dy'\,dy \tag{A10.3}$$

Applying an integration by parts to the double integration shown in the right hand side of (A10.3), we have

$$-\int_0^L\int_0^y Q_n'(y',t)\,dy'\,dy = \int_0^L \frac{\partial}{\partial y}(L-y)\cdot\int_0^y Q_n'(y',t)\,dy'\,dy$$

$$= \int_0^L (L-y)\cdot Q_n'(y,t)\, dy \qquad \text{(A10.4)}$$

Substituting (3.6) and (A10.4) into (A10.3), we have

$$I_y(0,t) = -I_{DC}(t) + \frac{dQ_S(t)}{dt} \qquad \text{(A10.5)}$$

where $I_{DC}(t) = -\frac{1}{L}\cdot\int_0^L I_y(y,t)\, dy$

$$= \frac{W}{L}\cdot C_{OX}\cdot\frac{\mu_n}{2F_B}\cdot(P_S(t) - P_D(t)) \qquad \text{(A10.6)}$$

$$Q_S(y,t) = -WC_{OX}\cdot\int_0^L (1-\frac{y}{L})\cdot Q_n'(y,t)\, dy \qquad \text{(A10.7)}$$

To find $I_y(L,t)$, we apply $\int_y^L dy'$ in both sides of (A10.1). Then,

$$I_y(L,t) = I_y(y,t) + WC_{OX}\cdot\frac{\partial}{\partial t}\int_y^L Q_n'(y',t)\, dy' \qquad \text{(A10.8)}$$

Applying $\int_0^L dy$ in both sides of (A10.8), we have

$$L\cdot I_y(L,t) = \int_0^L I_y(y,t)\, dy + WC_{OX}\cdot\frac{\partial}{\partial t}\int_0^L\int_y^L Q_n'(y',t)\, dy'\, dy \qquad \text{(A10.9)}$$

Applying an integration by parts to the double integration in the right hand side of (A10.9), we have

$$\int_0^L\int_y^L Q_n'(y',t)dy'\, dy = \int_0^L \frac{\partial}{\partial y}(y)\cdot\int_y^L Q_n'(y',t)\, dy'\, dy$$

$$= \int_0^L y\cdot Q_n'(y,t)\, dy \qquad \text{(A10.10)}$$

Substituting (A10.6) and (A10.10) into (A10.9), we have

$$I_y(L,t) = -I_{DC}(t) - \frac{dQ_D(t)}{dt} \qquad \text{(A10.11)}$$

where $Q_D(t) = -WC_{OX}\cdot\int_0^L \frac{y}{L}\cdot Q_n'(y,t)\, dy \qquad \text{(A10.12)}$

Since the incident currents are taken to be positive as shown in Fig.3.1, the drain and source currents $I_D(t)$ and $I_S(t)$ can be represented as

$$I_D(t) = -I_y(L,t)$$

$$= I_{DC}(t) + \frac{dQ_D(t)}{dt} \tag{A10.13}$$

$$I_S(t) = I_y(0,t)$$

$$= -I_{DC}(t) + \frac{dQ_S(t)}{dt} \tag{A10.14}$$

Equations for $I_{DC}(t)$, $Q_D(t)$ and $Q_S(t)$ are shown in (A10.6), (A10.12) and (A10.7) respectively.

In some literature [5.2], $I_{DC}(t)$ in (A10.6) is called 'transport current'.

Appendix 11

Implementation of the charge-conserving NQS MOSFET model into SPICE3

This charge-conserving NQS(non-quasistatic) MOSFET model, which is described in Chapter 5, has been implemented into SPICE3 based on the charge sheet formulation which is described in Chapter 2.

Equations for node currents

To compute the node currents, we need the surface potentials at source and drain ends. The surface potentials at source and drain ends of the channel, Ψ_{SO} and Ψ_{SL}, are computed, following the charge-sheet formulation which is described in Chapter 2. Ψ_{SL} is computed from the applied drain-to-source voltage V_{DS} and the drain saturation voltage V_{DSSAT}, by using the scheme shown in Chapter 2, Section 2.2.2. In this work shown in Chapter 5 and in this appendix, the long channel approximation is used to find V_{DSSAT}. Hence,

$$V_{DSSAT} = \frac{V_{GST}}{F_B} + V_t \tag{A11.1}$$

where

$$V_{GST} = V_{GB} - V_{FB} - \Psi_{SO} - \gamma\sqrt{\Psi_{SO} - V_t} \tag{A11.2}$$

$$= 0 \qquad (for\ \Psi_{SO} \le \Phi_F)$$

$$F_B = 1 + \frac{0.5\gamma}{\sqrt{\Psi_{SO} - V_t}} \cdot \left[1 - \frac{1}{1.744 + 0.8364(\Psi_{SO} - V_t)}\right] \tag{A11.3}$$

$$= 1 + \frac{0.5\gamma}{\sqrt{\Phi_F - V_t}} \cdot \left[1 - \frac{1}{1.744 + 0.8364(\Phi_F - V_t)}\right] \qquad (for\ \Psi_{SO} \le \Phi_F)$$

$$\Phi_F = V_t \cdot \ln(\frac{N_{SUB}}{n_i}) \qquad \text{(A11.4)}$$

where V_{GST} is conceptually equivalent to $(V_{GS}-V_{TH})$ in the conventional notation and is non-negative, F_B is the correction factor due to the bulk effect and F_B is slightly larger than 1.0, V_t is the thermal voltage kT/q, N_{SUB} is the substrate doping concentration, and n_i is the intrinsic carrier concentration.

From Ψ_{SO} and Ψ_{SL}, two quantities P_S and P_D are computed from

$$P_S = (V_{GST} + F_B \cdot V_t)^2 \qquad \text{(A11.5)}$$

$$P_D = (V_{GST} + F_B \cdot V_t + F_B \cdot (\Psi_{SL}-\Psi_{SO}))^2 \qquad \text{(A11.6)}$$

The DC drain current I_{DC} is computed from

$$I_{DC} = \frac{W}{L} \cdot C_{OX} \cdot \frac{\mu_n}{2F_B} \cdot (P_S - P_D) \cdot (1+\lambda \cdot V_{DS}) \qquad \text{(A11.7)}$$

where λ is the empirical channel length modulation factor. The instantaneous node currents can be computed from

$$I_D = I_{DC} + \frac{dQ_D}{dt} \qquad \text{(A11.8)}$$

$$I_S = -I_{DC} + \frac{dQ_S}{dt} \qquad \text{(A11.9)}$$

$$I_G = \frac{dQ_G}{dt} \qquad \text{(A11.10)}$$

$$I_B = -(I_G + I_D + I_S) \qquad \text{(A11.11)}$$

Solution of the state equation

To compute the charges, we need to solve the state equation.

$$\frac{d\mathbf{B}}{dt} = \mathbf{D} \cdot \mathbf{B} + \mathbf{G_S} \cdot \frac{dP_S}{dt} + \mathbf{G_D} \cdot \frac{dP_D}{dt} \qquad \text{(A11.12)}$$

where **B** is a column matrix for the coefficients $\{B_n\}$, **D** is a tri-diagonal square matrix,

and G_S and G_D are column matrices. The column matrix **B** is derived from the column matrix **A** in the state equation (5.25), using (A11.13). Hence, coefficients $\{B_n\}$ are used only inside the model computation routine and coefficients $\{A_n\}$ are used outside of the model computation routine.

$$B_n = A_n \qquad (for\ n=1,3,5,7,9) \tag{A11.13}$$

$$= (xnrm - xrev) \cdot A_n \qquad (for\ n=2,4,6,8,10)$$

where *xnrm* and *xrev* are the same as those in eq. (A2.5) and (A2.6) of Appendix 2 and are repeated here for clarity.

$$xnrm = 1 \qquad (for\ V_{DS} \geq 0) \tag{A11.14}$$

$$= 0 \qquad (for\ V_{DS} < 0)$$

$$xrev = 0 \qquad (for\ V_{DS} \geq 0) \tag{A11.15}$$

$$= 1 \qquad (for\ V_{DS} < 0)$$

where V_{DS} is the applied drain-to-source voltage and D and S refer to drain and source as specified in SPICE input, that is, D and S in

Mx D G S B model_name w=x l=x ad=x as=x pd=x ps=x

Coefficients of the tri-diagonal square matrix **D** are computed from

$$D_{i,(i-1)} = -\frac{\mu_n}{2F_B} \cdot \frac{\pi^2}{L^2} \cdot i^2 \cdot D1 \tag{A11.16}$$

$$D_{i,i} = -\frac{\mu_n}{2F_B} \cdot \frac{\pi^2}{L^2} \cdot i^2 \cdot D0 \tag{A11.17}$$

$$D_{i,(i+1)} = -\frac{\mu_n}{2F_B} \cdot \frac{\pi^2}{L^2} \cdot i^2 \cdot D1 \tag{A11.18}$$

(for i=1, 2, 3, ..., 10, but $D_{1,0} = D_{10,11} = 0$)

$$D0 = \frac{4}{3} \cdot \sqrt{P_S} \cdot \frac{1-(1-a)^{1.5}}{a} + \frac{2}{\pi} \cdot \sum_{m=1}^{5} \frac{B^{(old)}_{2m-1}}{2m-1} \tag{A11.19}$$

$$= 2\sqrt{P_S}\cdot(1-\frac{a}{4}) + \frac{2}{\pi}\cdot\sum_{m=1}^{5}\frac{B^{(old)}_{2m-1}}{2m-1} \qquad (for\ a < 10^{-5})$$

$$where \quad a = 1 - \frac{P_D}{P_S} \tag{A11.20}$$

$\{B^{(old)}_{2m-1}\}$ are $\{B_{2m-1}\}$ at the previous time point. $D1$ is set to 0 when a=0, that is, when cutoff or when V_{DS}=0.

$$D1 = (\sqrt{P_S} - \frac{D0}{2}) \qquad (\text{for } a \geq 0.1) \tag{A11.21}$$

$$= (\sqrt{P_S} - \frac{D0}{2})\cdot(1-(1-\frac{a}{0.1})^2) \qquad (\text{for } 0 < a < 0.1)$$

$$= 0.0 \qquad (for\ a \leq 0.0)$$

The coefficients of the column matrices $\mathbf{G_S}$ and $\mathbf{G_D}$ are computed from

$$G_{S.2m} = -\frac{(1-\frac{a}{2})\cdot C1 + (b + 2(1-a)\cdot\frac{db}{da})\cdot C2}{2m\pi\sqrt{P_S}} \tag{A11.22}$$

$$G_{S.2m-1} = -\frac{(1+\frac{a}{2})\cdot C1 - (b+2(1-a)\cdot\frac{db}{da})\cdot(1-\frac{4}{\pi^2(2m-1)^2})\cdot C2}{(2m-1)\pi\cdot\sqrt{P_S}} \tag{A11.23}$$

$$G_{D.2m} = \frac{C1\cdot(1+2\frac{db}{da})}{2m\pi\cdot\sqrt{P_S}} \tag{A11.24}$$

$$G_{D.2m-1} = -\frac{C1 + 2\cdot\frac{db}{da}\cdot(1-\frac{4}{\pi^2(2m-1)^2})\cdot C2}{(2m-1)\pi\cdot\sqrt{P_S}} \tag{A11.25}$$

(for m = 1, 2, 3, 4, 5)

where

$$b = 3 - \frac{3}{4}a - \frac{2}{a}\cdot(1-(1-a)^{1.5}) \tag{A11.26}$$

$$= \frac{a^2}{8} \qquad (\text{for } a < 10^{-4})$$

db/da is the derivative of b with respect to a.

$C1$ and $C2$ are constants to compensate for the truncation error, that is, the error due to the truncation of the Fourier sine series in (5.2) into 10 terms, where C1 = 1.0420956770489359 and C2 = 1.0644098681085510.

By appling the trapezoidal integration scheme to (A11.12), we can compute the partial derivatives of **B** as follows.

$$(\mathbf{I} - \frac{k}{2}\mathbf{D}) \cdot \frac{\partial \mathbf{B}}{\partial t} = \mathbf{D} \cdot \mathbf{B}^{(old)} \tag{A11.27}$$

$$(\mathbf{I} - \frac{k}{2}\mathbf{D}) \cdot \frac{\partial \mathbf{B}}{\partial P_S} = \mathbf{G_S} \tag{A11.28}$$

$$(\mathbf{I} - \frac{k}{2}\mathbf{D}) \cdot \frac{\partial \mathbf{B}}{\partial P_D} = \mathbf{G_D} \tag{A11.29}$$

where k is the time step, and $\mathbf{B}^{(old)}$ is the column matrix of coefficients $\{B_n\}$ computed at the previous time point. The new **B** at the present time point can be computed from the partial derivatives of **B** in (A11.27), (A11.28), and (A11.29).

$$\mathbf{B} = \mathbf{B}^{(old)} + \frac{\partial \mathbf{B}}{\partial t} \cdot k + \frac{\partial \mathbf{B}}{\partial P_S} \cdot \Delta P_S + \frac{\partial \mathbf{B}}{\partial P_D} \cdot \Delta P_D \tag{A11.30}$$

ΔP_S and ΔP_D are changes of P_S and P_D respectively during the present time step, that is, the time interval $[t-k, t]$, where t is the present time point.

The derivatives of **B** which are used in computing the derivatives of node currents with respect to applied biases can be computed from

$$\frac{\partial \mathbf{B}}{\partial V_X} = \frac{\partial \mathbf{B}}{\partial P_S} \cdot \frac{\partial P_S}{\partial V_X} + \frac{\partial \mathbf{B}}{\partial P_D} \cdot \frac{\partial P_D}{\partial V_X} \tag{A11.31}$$

(where $V_X = V_G, V_D, V_S, V_B$)

Equations for node charges

If $a > 10^{-3}$, drain and source charges Q_D and Q_S can be computed from

$$Q_D = -WLC_{OX} \cdot$$

$$\left[\sqrt{P_S}\cdot\left(\frac{2}{3}\cdot\frac{1-(1-a)^{1.5}}{a^2}-\frac{2}{5}\cdot\frac{1-(1-a)^{2.5}}{a^2}\right)-\frac{F_B V_t}{2}+\sum_{n=1}^{10}(-1)^{(n-1)}\cdot\frac{B_n}{n\pi}\right] \quad \text{(A11.32)}$$

$$Q_S = -WLC_{OX}\cdot$$

$$\left[\sqrt{P_S}\cdot\left(-\frac{2}{3}\cdot(1-a)\cdot\frac{1-(1-a)^{1.5}}{a^2}+\frac{2}{5}\cdot\frac{1-(1-a)^{2.5}}{a^2}\right)-\frac{F_B V_t}{2}+\sum_{n=1}^{10}\frac{B_n}{n\pi}\right] \quad \text{(A11.33)}$$

If $a \le 10^{-3}$, Q_D and Q_S are computed from

$$Q_D = -WLC_{OX}\cdot\left[0.5\sqrt{P_S}\cdot\left(1-a\cdot\left(\frac{1}{3}+\frac{a}{16}\right)\right)+\sum_{n=1}^{10}(-1)^{(n-1)}\cdot\frac{B_n}{n\pi}\right] \quad \text{(A11.34)}$$

$$Q_S = -WLC_{OX}\cdot\left[0.5\sqrt{P_S}\cdot\left(1-a\cdot\left(\frac{1}{6}+\frac{a}{48}\right)\right)+\sum_{n=1}^{10}\frac{B_n}{n\pi}\right] \quad \text{(A11.35)}$$

The gate and bulk charges Q_G and Q_B are computed from

$$Q_G = WLC_{OX}\cdot\left(V_{GB}-V_{FB}-\Psi_{SO}-\frac{V_{GST}}{F_B}-\frac{Q_D+Q_S}{F_B}\right) \quad \text{(A11.36)}$$

$$Q_B = -(Q_G+Q_D+Q_S) \quad \text{(A11.37)}$$

Moving boundary condition

The flag for the moving boundary condition is set to 1, if the following three conditions are met.

$$If \quad \frac{dV_{GST}}{dt} > \frac{\mu_n\cdot(V_{GST})^2}{F_B L^2} \quad \text{(A11.38)}$$

$$And\ If \quad V_{GST} > 10^{-3}\cdot F_B\cdot V_t \quad \text{(A11.39)}$$

$$And\ If \quad P_D-(xnrm\cdot P_D^{(old)}+xrev\cdot P_S^{(old)}) < (F_B\cdot V_t)^2 \quad \text{(A11.40)}$$

If the flag for the moving boundary condition is set, the drain charge Q_D computed from (A11.32) or (A11.34) is changed to Q'_D, so that the drain current I_D doesn't change during the moving boundary period. The state equation $I=dQ/dt$ can be solved using different integration schemes. Hence,

If (ckt->CKTorder == 1) (Backward Euler)

$$Q'_D = Q_D^{(old)} + k\cdot(-I_D + I_D^{(old)}) \tag{A11.41}$$

$$C_{DX} = -k\cdot\frac{\partial I_D}{\partial V_X} \tag{A11.42}$$

Else If (ckt->CKTorder == 2) (Trapezoidal)

$$Q'_D = Q_D^{(old)} + k\cdot(-I_D + I_D^{(old)}) \tag{A11.43}$$

$$C_{DX} = -k\cdot\frac{\partial I_D}{\partial V_X} \tag{A11.44}$$

(where X = G, D, S, B)

C_{DX} represents the drain capacitance. The superscript (*old*) represents the previous time point and k is the time step. (ckt->CKTorder) in SPICE3 indicates what integration scheme is used to solve the state equation $I=dQ/dt$. (ckt->CKTorder) changes during one simulation. It is set to 1 at the initial time point and at the first time point after the break point where slopes of input voltage waveform are discontinuous. Under other circumstances, it is usually set to 2.

Coefficients $\{B_n\}$ are adjusted to guarantee the continuity of the drain charge between the moving-boundary and the non-moving-boundary period. This adjustment has been done so that it doesn't change the values of other node charge components. Hence,

$$B'_{2m-1} = B_{2m-1} - \frac{0.5\pi(Q'_D - Q_D)}{WLC_{OX}\cdot C3}\cdot\frac{1}{2m-1} \tag{A11.45}$$

$$B'_{2m} = B_{2m} + \frac{0.5\pi(Q'_D - Q_D)}{WLC_{OX}\cdot C3}\cdot\frac{1}{2m} \tag{A11.46}$$

(for m = 1, 2, 3, 4, 5)

where

$$C3 = \sum_{m=1}^{5}\frac{1}{2m-1^2} = 1.1838649533887629 \tag{A11.47}$$

$$C4 = \sum_{m=1}^{5}\frac{1}{(2m)^2} = 0.3659027777777778 \tag{A11.48}$$

$\{B_{2m}\}$ and $\{B_{2m-1}\}$ are coefficients computed from (A11.30) without considering the effect of the moving boundary condition. $\{B'_{2m}\}$ and $\{B'_{2m-1}\}$ are new coefficients which have been corrected to include the moving boundary effect.

Similarly, Q_D is the drain charge computed without considering the moving boundary condition, and Q'_D is the new drain charge which includes the moving boundary effect.

Time step control

To get the accuracy and the reasonable CPU time, we need a time step control scheme.

Two kinds of state equations are solved in this work.

One type of the state equation is (A11.49) which is shown in (A11.8), (A11.9) and (A11.10).

$$I = \frac{dQ}{dt} \tag{A11.49}$$

Conventional time step control scheme in SPICE [5.16] is used for this type of state equation.

The other type of state equation is shown in eq. (5.26) in Chapter 5. It is repeated here for clarity.

$$\frac{d\mathbf{A}(t)}{dt} = \mathbf{D}(t)\cdot\mathbf{A}(t) + \mathbf{G_S}(t)\cdot\frac{d\mathbf{P_S}(t)}{dt} + \mathbf{G_D}(t)\cdot\frac{d\mathbf{P_D}(t)}{dt} \tag{A11.50}$$

To solve (A11.50), the values at the previous time point (t_0) are used for the matrices $\mathbf{D}$, $\mathbf{G_S}$, and $\mathbf{G_D}$. Using the trapezoidal integration, (A11.50) can be approximated as

$$\begin{aligned}\Delta\mathbf{A} &= \mathbf{A}(t) - \mathbf{A}(t_0) \\ &\approx \frac{k}{2}\,\mathbf{D}(t_0)\cdot(\mathbf{A}(t)+\mathbf{A}(t_0)) + \mathbf{G_S}(t_0)\cdot(\mathbf{P_S}(t)-\mathbf{P_S}(t_0)) + \mathbf{G_D}(t_0)\cdot(\mathbf{P_D}(t)-\mathbf{P_D}(t_0))\end{aligned} \tag{A11.51}$$

where t is the present time point and t_0 is the previous time point, and k is the present time step, that is, $k = t - t_0$. Using the Taylor series expansion, the approximated value $\Delta A\,|_{(approx)}$ can be derived from (A11.51) as.

$$\Delta A\,|_{(approx)} = -k \cdot (D(t_0)\cdot A(t_0) + G_S(t_0)\cdot P'_S(t_0) + G_D(t_0)\cdot P'_D(t_0))$$

$$- \frac{k^2}{2} \cdot (D(t_0)\cdot A''(t_0) + G_S(t_0)\cdot P''_S(t_0) + G_D(t_0)\cdot P''_D(t_0))$$

$$+ O(k^3) \qquad \text{(A11.52)}$$

The superscripts (′) and (″) denote the first and the second time derivatives respectively. The exact value $\Delta A\,|_{(exact)}$ can be derived from (A11.50) using the Taylor series expansion.

$$\Delta A\,|_{(exact)} = - k \cdot (D(t_0)\cdot A(t_0) + G_S(t_0)\cdot P'_S(t_0) + G_D(t_0)\cdot P'_D(t_0))$$

$$- \frac{k^2}{2} \cdot (D(t_0)\cdot A''(t_0) + D'(t_0)\cdot A'(t_0) + G_S(t_0)\cdot P''_S(t_0) + G_S'(t_0)\cdot P'_S(t_0) + G_D(t_0)\cdot P''_D(t_0)$$

$$+ G_D'(t_0)\cdot P'_D(t_0)) + O(k^3) \qquad \text{(A11.53)}$$

Hence the local truncation error LTE can be written as

$$LTE = norm(\Delta A\,|_{(exact)} - \Delta A\,|_{(approx)})$$

$$= \frac{k^2}{2} \cdot norm\left[D'(t_0)\cdot A(t_0) + G_S'(t_0)\cdot P_S'(t_0) + G_D'(t_0)\cdot P_D'(t_0)\right] \qquad \text{(A11.54)}$$

Since A_1 is dominant over other coefficients of $\{A_n\}$ for n = 2, 3, ..., 10, under most of operating conditions, only A_1 is considered in the time step control for simplicity. Hence,

$$LTE = \frac{k^2}{2} \cdot \left|\frac{dD_{11}}{dt}\cdot A_1 + \frac{dD_{12}}{dt}\cdot A_2 + \frac{dG_{S.1}}{dt}\cdot\frac{dP_S}{dt} + \frac{dG_{D.1}}{dt}\cdot\frac{dP_D}{dt}\right| \qquad \text{(A11.55)}$$

The time derivatives in (A11.50) are computed using the divided difference method [5.16]. The LTE in (A11.55) must be within the tolerance limit.

$$LTE \le QTRTOL \cdot (RELTOL \cdot (MINTOL + \max(A_1(t), A_1(t_0)))) \qquad \text{(A11.56)}$$

QTRTOL and *MINTOL* can be specified in the model card, and the default values are 10 and 0.1 (V^2) respectively.

Once the DC drain current, the node charges, and their derivatives with respect to node voltages are given, the subsequent implementation steps are the same as those in Appendix 2.

Appendix 12

Implementation of the NQS charge model combined with SPICE Level-2 DC model

The charge-conserving NQS(non-quasistatic) MOSFET charge model which is described in Chapter 5, has been combined with the SPICE Level-2 DC model [5.13] and has been implemented into SPICE3 as a level=5 model. Essentially any DC model can be combined with this NQS charge model, but the SPICE Level-2 DC model is chosen because it includes all the secondary effects and also because the level-2 model parameters are widely available in the industry and it makes the new model easy to use.

In this implementation, the DC drain current shown in (5.40) and also in (A11.7), is replaced by the DC drain current computed using the SPICE Level-2 DC model [5.13]. And the drain saturation voltage V_{DSSAT} in (A11.1) is replaced by V_{DSSAT} computed from the SPICE Level-2 DC model. This new V_{DSSAT} includes all the short channel and narrow channel effects.

In this way, all the short channel effects such as the velocity saturation effect and the junction geometry effect, the V_{GS} dependence of mobility and all other aspects considered by the SPICE Level-2 model are also included in the implementation. It is true that, while the DC model includes all the short channel effects, the charge model is still based on the long channel theory although the short channel effects are partially included in the charge model through V_{DSSAT}. This compromise should be judged in the light of the fact that all the conventional charge (capacitance) models in SPICE (Meyer, Ward-Dutton, BSIM) don't include any short channel effects either.

Model parameters

All the SPICE Level-2 model parameters except XQC (channel charge partitioning ratio) are recognized in this implementation. The DC current model is exactly the same as the SPICE Level-2 DC model and the charge model is the same as that shown in Chapter 5 except V_{DSSAT}. Long channel V_{DSSAT} is used in Chapter 5 but V_{DSSAT} of the SPICE Level-2 DC model is used in this implementation. Model parameters of this implementation (level=5 model) are as follows. Almost all the parameters are described in [5.15] but they are repeated here for clarity.

Name	Meaning	Default	Unit
VTO (VT0)	threshold voltage at V_{BS}=0	0.0	V
KP	$W/L \cdot C_{OX} \cdot \mu$ at $V_{GS}=V_{TH}$	$2 \cdot 10^{-5}$	A/V^2
GAMMA	$\sqrt{2\varepsilon_{si} qN_{SUB}}/C_{OX}$	0.0	$\sqrt{V}$
PHI	$2\Phi_F = 2V_t \cdot \ln(N_{SUB}/n_i)$	0.6	V
LAMBDA	channel length modulation factor	0.0	V^{-1}
RD	drain series resistance	0.0	Ω
RS	source series resistance	0.0	Ω
CBD	B-D junction cap. at V_{BD}=0	0.0	F
CBS	B-S junction cap. at V_{BS}=0	0.0	F
IS	S,D junction saturation current	10^{-14}	A
PB	S,D junction potential	0.8	V
CGSO	G-S overlap capacitance / W	0.0	F/m
CGDO	G-D overlap capacitance / W	0.0	F/m
CGBO	G-B overlap capacitance / L	0.0	F/m
RSH	S,D junction sheet resistance	0.0	$\Omega/square$
CJ	unit area S,D junction cap.	0.0	F/m^2
MJ	S,D jct. bottom grading coeff.	0.5	-
CJSW	unit length jct. sidewall cap.	0.0	F/m
MJSW	junction sidewall grading coeff.	0.33	-

JS	unit area jct. sat. current	0.0	A/m^2
TOX	gate oxide thickness	10^{-7}	m
NSUB	bulk doping concentration	0.0	cm^{-3}
NSS	slow surface state density for computing V_{TH}	0.0	cm^{-2}
NFS	fast surface state density for subthreshold conduction	0.0	cm^{-2}
TPG	type of gate material for computing V_{TH} +1: opposite to bulk -1: same as bulk 0: Al	1.0	-
XJ	S,D junction depth, flag for short channel jct. geometry effect	0.0	m
LD	S,D lateral diffusion	0.0	m
UO (U0)	surface mobility at $V_{GS}=V_{TH}$	600	$cm^2/V-s$
UCRIT	critical E-field for V_{GS} dependence of mobility	10^4	V/cm
UEXP	exponent for V_{GS} dependence of mobility	0.0	-
UTRA	V_{DS} effect on V_{GS} dependence of mobility, not used inside SPICE	0.0	-
VMAX	maximum drift velocity for computing velocity saturation effect	0.0	m/s
NEFF	multiplication factor for NSUB for computing channel length modulation with velocity saturation effect	1.0	-
FC	forward bias S,D jct. cap. coeff.	0.5	-
DELTA	narrow width effect on V_{TH}	0.0	-

QTRTOL	tolerance parameter for LTE, used in time step control		
MINTOL	tolerance parameter for LTE, used in time step control		
SATF	saturation factor for charge sheet formulation, factor A in eq.(2.15)	80	-

QTRTOL and MINTOL are tolerance parameters for the time step control shown in Appendix 11, eq.(A11.56). SATF is the saturation factor A of the saturation function which is shown in Chapter 2, eq.(2.15). The saturation function is used to find the surface potential Ψ_{SL}, the surface potential at the drain end of the channel. All the parameters except QTRTOL, MINTOL, and SATF are the same as the SPICE Level-2 model parameters [5.15].

Implementation into SPICE3B1

This NQS charge model combined with the Level-2 DC model has been implemented into SPICE3B.1 [5.15]. Compared to SPICE2 [5.17], SPICE3 is written so modularly that each model routine can be quite independent of other model routines with minimal changes in files outside of the model routine. Hence, it is much easier to add a new model into SPICE3 than into SPICE2. The general organization of SPICE3 will be published in [5.18]. All the model routines for this work are placed in the directory

DEV/NQS/

The files outside of the model routine which need changes are

DEV/Makefile

CKT/SIMinit.c

INP/INPpas2.c

INP/INPdomodel.c

INP/INPfindLev.c

In "DEV/Makefile", 'NQS' is added in the list of device models.

The file "SIMinit.c" contains the routine which initializes the data structure for the circuit to be simulated before starting the simulation. Two lines are inserted in "SIMinit.c" to add a new model. One is the following line in the variable declaration.

```
extern SPICEdev  NQSinfo;
```

The other is the following line in the lists of the data structure "SPICEdev *DEVices [ ]".

```
&NQSinfo
```

The file "INPpas2.c" contains the routine which processes SPICE3 input file, line by line, and checks whether the input is correctly specified or not. When the first character of the input line is 'M', the program switches to the part of the routine which checks the syntax of the MOSFET element, that is,

```
Mx D G S B model_type w=x l=x
```

The routine also checks whether the specified "model_type" is available in the program. Hence we add the following line in this part of the routine to let the routine know that the new model is available.

```
&&(thismodel->INPmodType != INPtypelook("NQS"))
```

The file "INPdomodel.c" contains the routine which identifies the model type and generates a data structure for the model. In the standard SPICE3B.1 release [5.15], four kinds of MOSFET models are available, that is, level-1(MOS1), level-2(MOS2), level-3(MOS3), and level-4(BSIM). To add this work as level-5 MOSFET model, we add the following three lines twice, once for 'nmos' and once for 'pmos'.

```
case 5:
        type = CKTtypelook("NQS");
        break;
```

The file "INPfindLev.c" contains the routine which extracts the level of MOSFET models from the SPICE3 input line and returns the value of the level to the calling routine. To add this work as a level-5 model, we add the following three lines inside the switch block of the routine.

```
case '5':
        *level = 5;
        return((char *)NULL);
```

The *return* statement above is just a house keeping stuff.

The model routine is decomposed into several files. Names of the files, the subroutines in the files, and their functions are as follows.

File	Subroutine	Function
NQSdefs.h	-	define data structures for MOSFET element

		and MOSFET model
NQS.c	-	define data structures for MOSFET element parameter table, model parameter table, and subroutines
NQSeval.c	NQSevaluate	model computation routine, computes node currents, node charges and their derivatives w.r.t. node voltages
NQSload.c	NQSload	process MOSFETs for DC and transient analysis, fill in Jacobian matrix and RHS vector by calling NQSevaluate. This routine is called once per iteration.
NQSacLoad.c	NQSacLoad	process MOSFETs for AC analysis
NQSpzLoad.c	NQSpzLoad	process MOSFETs for pole-zero analysis, not implemented yet
NQSsetup.c	NQSsetup	allocates memory for the sparse matrix elements for the given circuit
	NQSvsSetup	allocates memory for the cubic spline function and compute the cubic spline coefficients
	NQScubicSp	compute the cubic spline coefficients from a set of (x,y) data
NQStemp.c	NQStemp	default value processing for model parameters including the temperature effect
NQStrunc.c	NQStrunc	time step (LTE) control
NQSmosCap.c	NQSmosCap	process overlap capacitors for transient analysis
NQSparam.c	NQSparam	process SPICE3 input lines for MOSFET elements

NQSmParam.c	NQSmParam	process SPICE3 input lines for MOSFET models
NQSgetic.c	NQSgetic	grab initial condition
NQSdelete.c	NQSdelete	free the memory allocated to a data structure of MOSFET element
NQSmDelete.c	NQSmDelete	free the memory allocated to a data structure of MOSFET model
NQSdestroy.c	NQSdestroy	free the memory allocated to data structures of all MOSFET elements

Appendix 13

1-D numerical solution of current continuity equation

The current continuity equation (3.8) for the transient analysis has been solved using the 1-D numerical analysis, for the comparison with the non-quasistatic models described in Chapter 3 and Chapter 5.

From the current relation (3.1) and the current continuity equation (3.7), one can derive the modified current continuity equation (3.8), which is repeated here for clarity.

$$\frac{\partial Q'_n(y,t)}{\partial t} = \frac{\mu_n}{2F_B}\cdot\frac{\partial^2}{\partial y^2}\,(Q'_n(y,t) + F_B V_t)^2 \qquad \text{(A13.1)}$$

To solve (A13.1), one need the initial condition $Q'_n(y,0)$ and the two boundary conditions $Q'_n(0,t)$ and $Q'_n(L,t)$. The initial profile $\{Q'_n(y_i,0)\}$ can be computed from (A13.1) by setting the time derivative to be zero. Hence,

$$Q'_n(y_i,0) = -F_B V_t + \sqrt{(Q'_n(0,0) + F_B V_t)^2 - ((Q'_n(0,0) + F_B V_t)^2 - (Q'_n(L,0) + F_B V_t)^2)\cdot\frac{y_i}{L}} \qquad \text{(A13.2)}$$

where the subscript i represents the grid point and $Q'_n(0,0)$ and $Q'_n(L,0)$ are computed from applied biases using (3.3) and the charge sheet formulation described in Chapter 2.

Also the boundary conditions $Q'_n(0,t)$ and $Q'_n(L,t)$ are computed in the same way as $Q'_n(0,0)$ and $Q'_n(L,0)$.

The finite difference method is used for $\partial^2/\partial y^2$ in (A13.1) and the explicit (Forward Euler) integration method is used for $\partial/\partial t$. Hence, (A13.1) is converted to

$$\frac{Q'_n(y_i,t_{j+1}) - Q'_n(y_i,t_j)}{t_{j+1} - t_j} = \frac{\mu_n}{2F_B} \cdot \frac{1}{h^2} \cdot$$

$$\left[(Q'_n(y_{i-1},t_j) + F_B V_t)^2 - 2 \cdot (Q'_n(y_i,t_j) + F_B V_t)^2 + (Q'_n(y_{i+1},t_j) + F_B V_t)^2 \right] \quad \text{(A13.3)}$$

where h is the grid spacing in y direction. Uniform grids are used for $\{y_i\}$.

For the time step control, we refer the diffusion equation (3.10). It is well known that the time step must be less than or equal to $0.5 \cdot h^2/(diffusion\ coefficient)$ to get the stable solution of the diffusion equation [2.12]. Hence

$$(t_{j+1} - t_j) \leq \frac{h^2 F_B}{2 \mu_n} \cdot \frac{1}{(Q'_n(y_i,t_j) + F_B V_t) \mid_{MAX(y_i)}} \quad \text{(A13.4)}$$

By solving (A13.3), we can find the inversion charge density profile $\{Q'_n(y_i,t_{j+1})\}$ at $t=t_{j+1}$. Substituting the computed $\{Q'_n(y_i,t_{j+1})\}$ into (3.3), we can find the surface potential profile $\{\Psi_S(y_i,t_{j+1})\}$. Using (3.6), we can compute the instantaneous drain and source currents, $-I_y(L,t_{j+1})$ and $I_y(0,t_{j+1})$. From the computed profile $\{\Psi_S(y_i,t_{j+1})\}$, we can compute the total bulk charge $Q_B(t_{j+1})$ using (3.29) and the instantaneous bulk current $I_B(t_{j+1})$ using (3.32).