Computation of Drain and Substrate Currents in Ultra-Short-Channel nMOSFET's Using the Hydrodynamic Model

Khalid Rahmat, *Student Member, IEEE*, Jacob White, *Member, IEEE*, and Dimitri A. Antoniadis, *Fellow, IEEE*

***Abstract*—The goal of this work was to develop a robust and efficient numerical solution of the hydrodynamic model, which solves the energy balance equation, and to compare predictions of this model, using one set of parameters, with experimental nMOSFET characteristics for a range of channel lengths down to ultrashort channels. The substrate current was calculated by direct integration of the energy distribution function, which uses the computed temperature to obtain the number of high energy electrons. The drain current calculated using this method is accurate for a range of channel lengths and biases, and correctly predicts the observed enhanced transconductance for ultrashort-channel devices. The substrate current matches the experimental data for a range of channel lengths and biases above threshold with one set of physically reasonable parameters.**

I. Introduction

In this paper we use the basic formulation of [1] and keep the full energy balance equation but neglect the nonlinear convective term $(\mathbf{v}\cdot\nabla\mathbf{v})$ in the momentum conservation equation. We then modify the discretization of the energy balance equation to take into account explicitly the variation of the thermal conductivity with carrier concentration. This scheme is shown to be numerically more stable than the schemes proposed in [2] and [3], which produce instabilities when used with coarse meshes. We have implemented the hydrodynamic model for one carrier in steady state in a two-dimensional device simulator, and the simulated device characteristics match well with the experimental data for MOSFET's with 0.16–0.90-μm channel lengths. This was achieved by using one set of model parameters for all channel lengths and with the simulator calibrated at 0.90 μm [4].

We also present the results of a simple method to calculate the substrate current based on computed electron temperatures. The method predicts quite well the experimentally observed substrate current for MOSFET's over a range of channel lengths, but only for transistor biases significantly above threshold. Because the substrate current is a sensitive measure of the hot carrier population, accurate prediction of the substrate current provides an independent check on the validity of the computed solution to the energy equation.

Manuscript received May 22, 1991; revised June 5, 1992. This work was supported by the Defense Advanced Research Projects Agency under contracts N00014-87-K-825 and MDA972-88-K-008, and grants from IBM. This paper was recommended by Associate Editor D. Scharfetter.

The authors are with the Research Laboratory of Electronics and the Microsystems Technology Laboratory, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9206843.

II. Physical Model

Under suitable assumptions, the conservation laws for electron charge, momentum, and energy, for a single carrier, along with Poisson's equation can be written in simplified form [1] as

$$\nabla \cdot \mathbf{J}_n = qU \tag{1}$$

$$\mathbf{J}_n - \frac{\tau_n}{q}(\mathbf{J}_n \cdot \nabla)\left(\frac{\mathbf{J}_n}{n}\right) = qD_n\nabla n - q\mu_n n\nabla\left(\psi - \frac{k_BT}{q}\right) \tag{2}$$

$$\nabla \cdot \left[-\kappa\nabla T - \frac{\mathbf{J}_n}{q}(k_BT + w)\right] = \mathbf{E}\cdot\mathbf{J}_n - n\left(\frac{w - w_0}{\tau_w}\right) - wU \tag{3}$$

$$\nabla \cdot \mathbf{E} = \frac{q}{\epsilon}\left(n_i e^{[q(\phi_p - \psi)/k_BT_0]} - n + N_D - N_A\right) \tag{4}$$

In the preceding equations, n is the electron concentration, $\mathbf{J}_n$ is the electron current, and w is the average electron energy which is given by

$$w = \tfrac{1}{2}m_n^* v_n^2 + \tfrac{3}{2}k_BT \tag{5}$$

where v_n is the electron velocity, m_n^* is the effective electron mass, T is the electron temperature, and k_B is Boltzmann's constant. In (2), τ_n, D_n, and μ_n are the electron momentum relaxation time, diffusion constant, and mobility. In (3), τ_w, κ, and U are the energy relaxation time, thermal conductivity, and the net recombination rate per unit volume for the electrons. w_0 is the thermal energy of the electrons in equilibrium and equals $\frac{3}{2}k_BT_0$, where T_0 is the lattice temperature. In Poisson's equation, (4), $\mathbf{E}$ is the electric field, q is the absolute value of the electronic charge, and ϵ the permittivity of silicon. ψ and ϕ_p are the electrostatic potential and hole quasi-Fermi level, respectively. The hole quasi-Fermi level is assumed to be con-

stant and is determined by the biases on the contacts. Finally, the electrostatic potential and electric field **E** are related by $\mathbf{E} = -\nabla\psi$.

A. Relaxation Time Models

Given mobility, thermal conductivity, and energy relaxation time models, the preceding set of equations can be solved to determine current and temperature distributions in MOS devices. We used the standard Weidemann–Franz law thermal conductivity model, and a mobility model that is a function of doping and vertical field near the device surface (see Appendix), and inversely proportional to temperature [5]. For the energy relaxation time we used a constant value of 0.1 ps. It should be noted that this combination of a constant energy relaxation time and a mobility inversely proportional to temperature leads to two inconsistencies in the homogeneous case. To see this, consider that simplifying (1)–(5) assuming spatial homogeneity leads to a relationship between electric field and temperature given by

$$E^2 = \frac{3}{2}\frac{k_B T_0}{q}\left[\left(\frac{T}{T_0}\right)^2 - \left(\frac{T}{T_0}\right)\right]\frac{1}{\tau_w \mu_{n0}} \tag{6}$$

where it is assumed that $\frac{1}{2}mv_n^2 \ll \frac{3}{2}k_B T$. It then follows that, for high fields, the saturation velocity, $\mu|E|$, is given by

$$\left(\frac{3}{2}\frac{k_B T_0}{q}\frac{\mu_{n0}}{\tau_w}\right)^{1/2} \tag{7}$$

where μ_{n0} is often referred to as the low field mobility, but is, in fact, the mobility ignoring temperature effects.

Both Monte Carlo and experimental data indicate that the saturation velocity is independent of doping and surface effects; therefore, (7) suggests that τ_w should be proportional to μ_{n0} and not constant. However, the data also indicate that the electric field versus temperature relation is independent of doping at high fields, and therefore (6) suggests τ_w is *inversely* proportional to μ_{n0} and not constant. In [6], a modified model for mobility dependence on temperature was used which resolves this contradiction.

As most of the current flow in an MOS device is near the surface under highly inhomogeneous conditions, it is not clear that homogeneous considerations are of primary importance. In the bulk, the low field mobility is strongly dependent on the doping, whereas at the surface the low field mobility is essentially independent of the doping and completely determined by surface effects. It then follows from (6) and (7) that near the surface the saturation velocity and the electric field versus temperature relationship are independent of doping, which is consistent with Monte Carlo results.

An example of the impact of surface effects can be seen in the choice of the relaxation time. We found that an energy relaxation time of 0.1 ps yielded the best match to measured data over a range of channel lengths, and this value is much lower than the 0.2–0.4 ps predicted from Monte Carlo simulations of bulk silicon. The justification for the lower energy relaxation time is that it models the smaller mean free path for electrons near the surface, which has been observed experimentally [13]. The situation is not changed significantly when we use a mobility model which resolves the above-mentioned inconsistencies. Both in our own experiments with such a mobility model, and those reported in [6], best results were achieved with an energy relaxation time much lower than that determined from bulk simulations.

III. Discretization Scheme

In the discretization scheme used in [3], which is an extension of the work in [2], the problem is cast in terms of an energy flow density, **S**, defined by

$$\mathbf{S} = -\kappa\nabla T - \left(\frac{5}{2}k_B T\right)\frac{\mathbf{J}_n}{q}. \tag{8}$$

Thus the energy equation can be written as

$$\nabla\cdot\mathbf{S} = \nabla\cdot\left(\frac{m_n^*|\mathbf{J}_n|^2}{2q^2n^2} - q\psi\right)\frac{\mathbf{J}_n}{q} + (q\psi - w)U - n\left(\frac{w - w_0}{\tau_w}\right). \tag{9}$$

One advantage of this formulation is that a Scharfetter–Gummel or exponentially fit discretization scheme can be applied. This can be seen by projecting **S** onto an edge between nodes i and j,

$$S_{ij} = -\kappa\frac{dT}{dx_{ij}} - \frac{5}{2}\frac{k_B T}{q}J_{ij}. \tag{10}$$

Treating S_{ij}, J_{ij}, and κ as constant along the edge, (10) can be integrated analytically to obtain.

$$S_{ij} = -\frac{\kappa_{ij}}{d_{ij}}\left[B(\omega_{ij})T_j - B(-\omega_{ij})T_i\right], \tag{11}$$

where

$$\omega_{ij} = -\frac{5}{2}k_B\frac{J_{ij}}{q}\frac{d_{ij}}{\kappa_{ij}} \tag{12}$$

and $B(x) = x/(e^x - 1)$ is the Bernoulli function, κ_{ij} is an average thermal conductivity between the two nodes, and d_{ij} is the distance between nodes i and j.

A. Temperature Instabilities

The preceding discretization technique was implemented in a two-dimensional finite-box-based device simulator and used to simulate a short-channel MOSFET. To solve the nonlinear algebraic problem generated by the discretization, Newton's method was used combined with sparse Gaussian elimination to solve for the Newton updates. We observed that the temperatures computed using a coarse rectangular mesh with the discretization method described earlier exhibited numerical instabilities in certain regions of the device. In particular, the computed

temperatures oscillated in space, occasionally dipping *below* the lattice temperature. An example of this anomalous behavior is shown in Fig. 1. Although such instabilities can be eliminated by refining the mesh, this may not always be practical as the instability may artificially make the discretized problem more nonlinear, which in turn worsens the convergence of any iterative nonlinear solver, such as Newton's method. Without a converged solution, it may not be obvious where to add additional mesh points.

The source of this numerical instability is that the discretization of the energy equation in (11) and (12) inappropriately assumes that the thermal conductivity is a constant. To see why such an approximation leads to coarse-grid instability, consider computing the divergence of (8) assuming J_n, but *not* κ, is constant. The result is

$$\nabla \cdot \mathbf{S} = -\kappa \nabla^2 T - \left(\nabla \kappa + \left(\frac{5}{2} k_B\right) \frac{\mathbf{J}_n}{q}\right) \nabla T. \quad (13)$$

To be stable for coarse grids, a method for discretizing (13) must *upwind* the ∇T term, i.e., discretize ∇T in the upwind direction given by the sign of

$$\left(\nabla \kappa + \left(\frac{5}{2} k_B\right) \frac{\mathbf{J}_n}{q}\right). \quad (14)$$

Equation (12) does not include the $\nabla \kappa$ term, therefore, the resulting Scharfetter–Gummel scheme will not upwind correctly unless the $\nabla \kappa$ term can be ignored. This is not the case, as can be seen if we write the thermal conductivity as

$$\kappa = \left(\frac{5}{2} + c\right) k_B D_{n0} n, \quad (15)$$

where c is a constant that depends on the dominant scattering mechanism [5], and substitute this relation in (14) to yield

$$\left(\frac{5}{2} k_B\right) \left(\left(1 + \frac{2}{5} c\right) D_{n0} \nabla n + \frac{\mathbf{J}_n}{q}\right) \quad (16)$$

Clearly, $(1 + \frac{2}{5}c)\, D_{n0} \nabla n$ and $\mathbf{J}_n/q$ will be comparable when diffusion contributes significantly to current flow. In particular, this implies that the ∇T term in (13) may not be discretized in the upwind direction when ∇n is large. Our numerical experiments verify this, as the temperatures computed with the preceding approach oscillate in device regions where the electron concentration gradients are large.

B. Modified Energy Discretization

In this section, we develop a better stabilized discretization scheme for the energy equation. The convective term in (2)

$$\frac{\tau_n}{q} (\mathbf{J}_n \cdot \nabla) \left(\frac{\mathbf{J}_n}{n}\right) \quad (17)$$

is neglected explicitly, which makes it possible to substitute the expressions for the thermal conductivity, κ, and

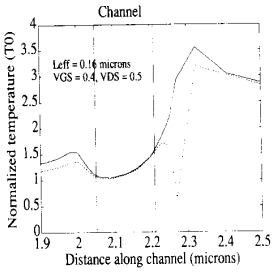


Fig. 1. Simulated electron temperature for a MOSFET with $L_{\text{eff}} = 0.16$ μm, along the device at a depth of 0.15 μm from the oxide interface using the discretization method described by (10) (dash) and (20) (solid).

electron current density, $\mathbf{J}_n$, into the energy equation as suggested in [7]. The justification for neglecting the convective term is that in MOSFET's where current flow is by majority carriers, this term is generally small compared with $\mathbf{J}_n$ in regions where the electron concentration is large.

Neglecting the convective term, (2) becomes

$$\mathbf{J}_n \simeq q \frac{k_B T_0}{q} \mu_n [r \nabla n + n \nabla (r - u)] \quad (18)$$

where $r = T/T_0$ is the electron temperature normalized by the lattice temperature and u is the normalized electrostatic potential given by $u = [q/(k_B T_0)]\psi$. Note (18) is identical to (6.4) in [2], once the convective term is neglected. The temperature dependence of the mobility model can be included explicitly into the current equation, as in

$$\mathbf{J}_n = q \frac{D_{n0}}{r} [r \nabla n + n \nabla (r - u)]. \quad (19)$$

Substituting both the expression for the current (19) and the thermal conductivity (15) into the energy flux equation yields

$$\mathbf{S} = -\tfrac{5}{2} k_B T_0 D_{n0} [r \nabla n + n ((2 + \tfrac{2}{5}c) \nabla r - \nabla u)]. \quad (20)$$

This expression for $\mathbf{S}$ has the same form as that for J_n, (19), but with a different coefficient in front of the ∇r term. Hence, the Scharfetter–Gummel method can be applied just as easily to this equation as to the current equation, with presumably equal success. Just as in the equation for $\mathbf{J}_n$, we have assumed that the electron temperature and electrostatic potential vary linearly between the two nodes. This assumption more naturally captures the physical variation of these variables, as it is the electron concentration that needs to be "exponentially fitted" rather than the electron temperature. It should be noted that the method is consistent with respect to temperature: in both the current and energy equations temperature is assumed to vary linearly; but it is inconsistent with respect to electron concentration: the assumed form of the spatial variation is different in the two equations.

The discretization of the right-hand side of (9) poses no special difficulties, and is handled in a conventional manner.

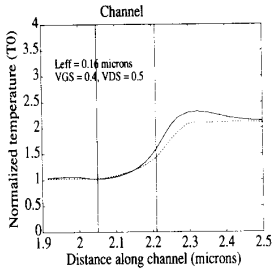


Fig. 2. Simulated electron temperature for a MOSFET with L_{eff} = 0.16 μm along the device at a depth of 0.15 μm from the oxide interface using the discretization method described by (10) (dash) and (20) (solid). Same as Fig. 1 but with a finer mesh.

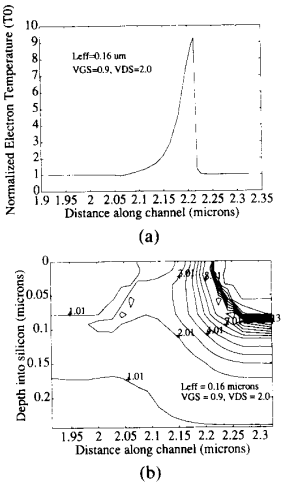


Fig. 3. Electron temperature normalized to the lattice temperature in the 0.16 μm MOSFET: (a) at the silicon-oxide interface (b) contour plot in the channel region.

Note that the above modified discretization has a disadvantage if the equation system is solved with an iterative scheme that decouples the current continuity from the energy equation. That is, given the electron concentration and the electric field, the original discretization scheme yields an energy equation that is linear in temperature, and, therefore, is easy to solve. However, the modified discretization scheme yields an energy equation which is exponentially nonlinear in temperature, and may not converge with an arbitrary electron concentration produced by a decoupled scheme.

Using the modified discretization scheme with the same mesh spacing and biases as used for the unstable case, the solution shown in Fig. 1 does not display any instability. Of course, this solution should not only be stable, but also accurate in the sense that it should be close to the "correct" solution. To check that this is the case, we computed solutions with much finer mesh (in both dimensions) using both the stable and unstable discretization techniques. The results obtained on this mesh are shown in Fig. 2. Note that the temperature profiles obtained from both methods on the finer mesh are similar, and that the stable discretization solution on the coarse mesh is at least qualitatively similar to the finer mesh solutions. The mesh size in the coarse case was 31 × 24 while it was 37 × 38 in the finer case, where most of the added mesh lines were placed in the region where the instability appeared.

Fig. 3 shows an example of the temperature calculated using the discretization method described here for the shortest channel device.

IV. Simulation Results

In this section we compare the results obtained from our two-dimensional simulator for devices with effective channel lengths from 0.16 μm to 0.90 μm with experimental data reported in [8]. Gate oxide thickness for these devices is 52 Å, the junction depth is about 0.09 μm, and the device width is 10 μm for all the simulated MOSFET's. One set of parameters was used for all devices, and a constant series resistance of 30 Ω was added to the source and drain of the simulated devices to account for extrinsic device resistance.

A. Drain Current Calculation

The drain current predicted by the simulator and that actually measured for three different channel lengths is shown in Fig. 4. The predicted current is within 10% of the measured value for all three channel lengths and all the biases.

A question of some technological interest is whether the hydrodynamic model is needed for the prediction of drain currents and, if so, at what channel length. It has been suggested that, as a result of velocity overshoot near the source, the current in an ultrashort channel may exceed the value predicted by the drift-diffusion model, which imposes velocity saturation. In the hydrodynamic model, of course, no such limitation is built-in and we should expect greater fidelity to experiment.

To answer this question, we performed simulations using the hydrodynamic model described earlier and the drift-diffusion model using an electric field dependent mobility given by (A4) in the Appendix. In Fig. 5 we plot the computed small-signal transconductance of MOSFET's with different channel lengths using the drift-diffusion and hydrodynamic models. The transconductance was calculated at a bias voltage of V_{DS} = 2.0 V and V_{GS} = 1.2 V. Clearly, for devices with channel lengths much longer than about 0.15 μm, the difference between the two simulations is not significant. For shorter channel lengths, the simulations based on the hydrodynamic model predict a more rapid increase in transconductance than those based on the drift-diffusion model. The results obtained by the hydrodynamic model are quite similar to the experimental data reported in [9] although exact comparison is difficult because of the different device structures.

B Substrate Current Calculation

Because the substrate current in MOS transistors at high drain biases is caused primarily by impact ionization, our approach for calculating the substrate current is to assume

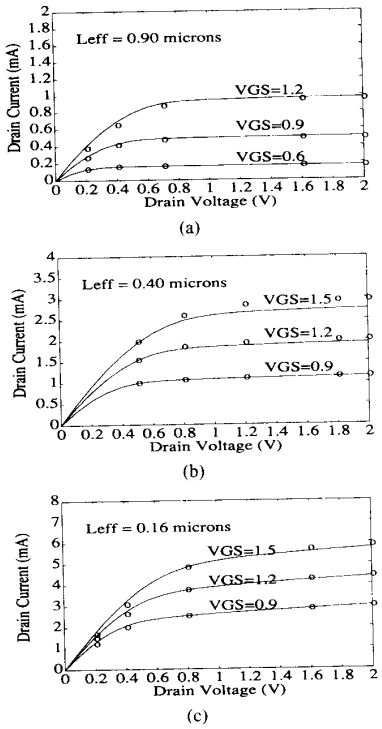


Fig. 4. Experimental (solid) and simulated (○) drain current for three MOSFET's with channel lengths of (a) 0.90 μm, (b) 0.40 μm, and (c) 0.16 μm.

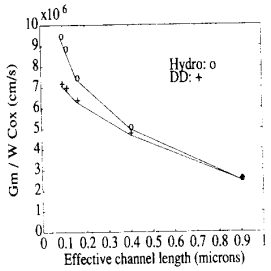


Fig. 5. Normalized transconductance at V_{GS} = 1.2 V, V_{DS} = 2.0 V using the hydrodynamic and drift-diffusion models for different channel lengths.

that it is proportional to the total number of electrons that have an energy above a threshold value. That is,

$$I_{sub} = C_{sub} W q \int_0^{L_x} \int_0^{L_y} dx\, dy\, n(x, y) \int_{\epsilon_{Thresh}}^{\infty} d\epsilon\, F(\epsilon, T(x, y)) \tag{21}$$

where ϵ and ϵ_{Thresh} are the electron and threshold energies, respectively; $F(\epsilon, T)$ is the product of the electron energy distribution as a function of temperature and the density of states; L_x, L_y, and W are the device length, height, and width, respectively; and C_{sub} is a proportionality constant. Note that w, which is the average electron energy for the ensemble, is distinct from ϵ, which is a random variable for the energy of each electron in the ensemble.

For the two-dimensional simulations discussed herein, uniformity along the width is assumed implicitly, and therefore integration with respect to the width is replaced by multiplication.

Under high field conditions, it is well known that the actual electron energy distribution is substantially different from Maxwellian [10]. Recent Monte Carlo studies [11], [12] suggest that the tail of the distribution function decays much faster than an exponential dependence. In [11], it is proposed that a more accurate model would be to use a cubic energy dependence in the exponent. This result was derived analytically using nonparabolic bands which, of course, also changes the density of states. Thus, in our notation, this leads to

$$F(\epsilon, T) = C_{dist}(T)\epsilon^{1.25} \exp\left(-\chi \frac{\epsilon^3}{T^{1.5}}\right), \tag{22}$$

where

$$C_{dist}(T) = 1 \Big/ \int_0^{\infty} \epsilon^{1.25} \exp\left(-\chi \frac{\epsilon^3}{T^{1.5}}\right). \tag{23}$$

The preceding substrate current model has only one free parameter, χ, in the distribution function, as we keep the threshold energy fixed at a value of 1.8 eV in (21), which is a reasonable value for the ionization energy. Changing χ is equivalent to scaling the energy axis; hence, χ can be thought of as a contraction or dilation factor for the distribution function. The value of the proportionality constant, C_{sub}, is determined by equating the measured current and the calculated current at one bias point for the longer channel device. This constant, which for our simulations was $8.10 \times 10^{11}\,s^{-1}$, is then used for all the other calculations.

To verify the accuracy of the preceding distribution function and choose an appropriate dilation factor, χ, we compared the experimentally measured impact ionization coefficient, α, in silicon with that obtained by using our approach. The generation rate per unit volume for electrons can be written as

$$G = \alpha n |v_n|, \tag{24}$$

where α is the impact ionization coefficient, n is the electron concentration, and v_n is the electron velocity usually taken to be the saturation velocity. In our approach, the generation rate per unit volume is

$$G = C_{sub}\, n \int_{\epsilon_{Thresh}}^{\infty} d\epsilon\, F(\epsilon, T). \tag{25}$$

Hence, the impact ionization coefficient is given by

$$\alpha = \frac{C_{sub}}{|v_n|} \int_{\epsilon_{Thresh}}^{\infty} d\epsilon\, F(\epsilon, T). \tag{26}$$

The impact ionization rate thus obtained would be a function of the electron temperature; the measured data and models proposed for the impact ionization rate are given as a function of electric field; therefore, it is nec-

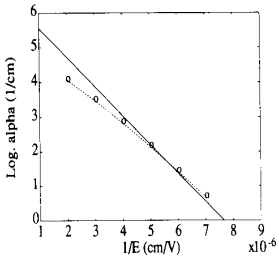


Fig. 6. Comparison of the impact ionization coefficient, α, as a function of the inverse electric field between experimental data (solid) and our model (26) (dash).

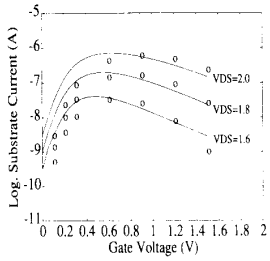


Fig. 8. Simulated (○) and measured (solid) substrate current for a device with $L_{\text{eff}} = 0.20$ μm as a function of the gate voltage for three different drain voltages.

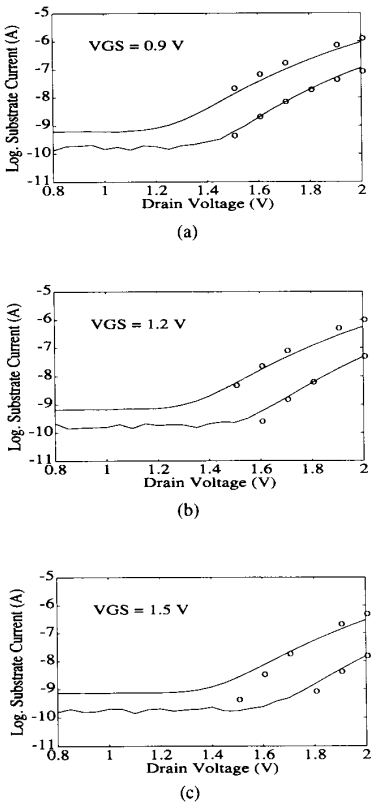


Fig. 7. Simulated (○) and measured (solid) substrate current for the 0.16-μm (upper curve) and 0.40-μm (lower curve) devices using the cubic distribution function. (a) $V_{GS} = 0.9$ V; (b) $V_{GS} = 1.2$ V; (c) $V_{GS} = 1.5$ V.

essary to convert the temperature dependence to a field dependence. To accomplish this, the relationship for the homogeneous case, given in (6), can be applied. Note that in our mobility model, the "low field mobility" is independent of the doping near the oxide interface, as surface effects dominate. Using the expression in (6) for electric field versus temperature with $\mu_{n0} = 450$ cm^2/(V · s), and the energy distribution function in (22) with a dilation factor of 1.0×10^5, where the energy is measured in electron volts and the temperature in Kelvin, (26) results in impact ionization coefficient versus inverse electric field curve as shown in Fig. 6. Experimental data from [13] are also shown in this figure. Although the two do not match exactly, this is to be expected due to the large uncertainty in the measured data and the approximations in the model.

Fig. 7 shows the simulated and measured substrate currents as a function of the drain voltage at different gate biases for both devices using the cubic distribution function with the parameters outlined earlier. Generation-recombination is ignored in our simulator so that only a comparison with the hot carrier part of the substrate current is appropriate.

Fig. 8 shows the substrate current as a function of the gate voltage with the drain voltage as a parameter. The threshold voltage for this device is about 0.23 V, and the accuracy of the calculated substrate current seriously degrades for biases near the threshold voltage, although the peak substrate current is predicted correctly. This inaccuracy cannot be attributed to errors in the threshold voltage because a careful comparison of the simulated and measured drain currents even at low gate biases shows very little error. An explanation for this effect can be found if we consider that, in the subthreshold regime with high drain voltage, the electric fields are extremely high (over 5×10^5 V/cm) in a very narrow region near the drain. Consequently, the carriers do not fully thermalize before entering the drain; therefore, the electron temperature in this region does not accurately reflect the tails of the distribution function.

V. Conclusion

The results in this paper demonstrate that the hydrodynamic model can be used successfully to simulate silicon MOSFET's with channel lengths as short as 0.16 μm. Equally important is the fact that one set of parameters was used to simulate devices with channel lengths varying from 0.90 μm down to 0.16 μm, and all the parameter values used were physically reasonable. We have also shown the importance of using a stable discretization method in the energy balance equation to avoid spurious numerical results. Finally, an approach to computing the substrate currents by the direct integration of the energy distribution function yielded a simple, but reasonably accurate, method.

APPENDIX

The expression for electron mobility which includes the effects of the scattering of the electrons by the lattice and ionized impurities is [15]:

$$\mu_{LI} = 88\left(\frac{T_0}{300}\right)^{-0.57} + \frac{1252\,(T_0/300)^{-2.33}}{1 + [N/(1.432 \times 10^{17})](T_0/300)^{-0.57}}, \quad \text{(A1)}$$

where N is the total doping concentration in cm^{-3}, and T_0 is the lattice temperature.

The effects of surface scattering are included by using an expression that depends on both the distance from the surface and the lateral electric field [14]

$$\mu_{LIS} = \frac{\mu_{LIS_{\min}} + (\mu_{LI} - \mu_{LIS_{\min}})(1 - f)}{1 + f(E_t/EREF)^{LIS_{\exp}}} \quad \text{(A2)}$$

where f is given by

$$f = \frac{2.0e^{-1 \times 10^{12}d^2}}{1.0 + e^{-2 \times 10^{12}d^2}} \quad \text{(A3)}$$

where d is the distance from the silicon–oxide interface in centimeters and E_t is the transverse electric field in volts per centimeter. The constants in the preceding expression are $\mu_{LIS_{\min}} = 638\ \text{cm}^2/(\text{V} \cdot \text{s})$; $EREF = 5.0 \times 10^5$ V/cm; $LIS_{\exp} = 1.30$.

The distance function, f, has a value of one at the surface, and its value approaches zero deeper in the bulk. Consequently, near the surface the mobility is dominated by surface effects whereas deeper into the bulk, the surface effects become negligible.

To include the high field effects in the drift-diffusion simulations we used the electric field dependence given by equation 4.1-56 in [15]:

$$\mu_{LISE} = \frac{2\,\mu_{LIS}}{1 + \sqrt{1 + (2\,\mu_{LIS}E/v_{\text{sat}})^2}}. \quad \text{(A4)}$$

ACKNOWLEDGMENT

The authors would like to thank J. Jacobs for his help on mobility models and calibration techniques, M. Reichelt for providing his drift-diffusion simulator, G. Shahidi for the experimental data, and F. Odeh for many discussions, we will miss his valuable guidance.

REFERENCES

[1] M. Rudan and F. Odeh, "Multi-dimensional discretization scheme for the hydrodynamic model of semiconductor devices," *COMPEL*, vol. 5, pp. 149–183, 1986.

[2] A. Forghieri, R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan, and G. Baccarani, "A new discretization strategy of the semiconductor equations comprising momentum and energy balance," *IEEE Trans. Computer-Aided Design*, vol. CAD-7, pp. 231–242, 1988.

[3] A. Gnudi and F. Odeh, "Engineering and numerics of the hydrodynamic model," invited talk presented at SIAM Meeting, Chicago, July 1990.

[4] K. Rahmat, J. White, and D. A. Antoniadis, "Computation of drain and substrate currents in ultra-short-channel NMOSFET's using the hydrodynamic model," *IEDM Tech. Digest*, pp. 115–118, 1991.

[5] G. Baccarani and M. R. Wordeman, "An investigation of steady-state velocity overshoot effects in Si and GaAs devices," *Solid State Electron.*, vol. 28, pp. 407–416, 1985.

[6] B. Meinerzhagen and W. L. Engl. "The influence of the thermal equilibrium approximation on the accuracy of classical two-dimensional numerical modeling of silicon submicrometer MOS transistors," *IEEE Trans. Electron Devices*, vol. ED-35, pp. 689–697, 1988.

[7] T. W. Tang, "Extension of the Scharfetter–Gummel algorithm to the energy balance equation," *IEEE Trans. Electron Devices*, vol. ED-31, pp. 1912–1914, 1984.

[8] G. G. Shahidi, D. A. Antoniadis, and H. I. Smith, "Reduction of channel hot-electron-generated substrate current in sub-150-nm channel length Si MOSFET's," *IEEE Electron Device Lett.*, vol. 9, pp. 497–499, 1988.

[9] G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, S. Rishton, and E. Ganin, "High transconductance and velocity overshoot in NMOS devices at the 0.1-μm gate-length level," *IEEE Electron Device Lett.*, vol. 9, pp. 464–466, 1988.

[10] J. M. Higman, K. Hess, C. G. Hwang, and R. W. Dutton, "Coupled Monte-Carlo drift diffusion analysis of hot-electron effects in MOSFETS," *IEEE Trans. Electron Devices*, vol. ED-36, pp. 930–937, 1989.

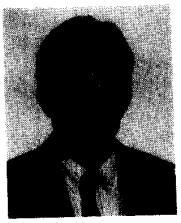
[11] D. Cassi and B. Riccò, "An analytical model of the energy distribution of hot electrons," *IEEE Trans. Electron Devices*, vol. ED-37, pp. 1514–1521, 1990.

[12] N. Goldsman, L. Hendrickson, and J. Frey, "Reconciliation of a hot-electron distribution function with the lucky electron-exponential model in silicon," *IEEE Electron Device Lett.*, vol. 11, pp. 472–474, 1990.

[13] J. W. Slotboom, G. Streutker, G. J. T. Davids, and P. B. Hartog, "Surface impact ionization in silicon devices," *IEDM Tech. Digest*, pp. 494–497, 1987.

[14] J. Jacobs, S. M., "Modeling the effects of Si/SiO_2 interface proximity and transverse field on carrier mobility in MOSFET's," thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1987.

[15] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Vienna: Springer-Verlag, 1984.



Khalid Rahmat (S'89) received the B.S. and M.S. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1984 and 1986, respectively. From 1986 to 1989 he was employed at Standard Microsystems Corporation, NY, in the Process Development group. He is currently working on his Ph.D. in the area of device physics and simulation at MIT.

Jacob White (S'80–M'83) received the B.S. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, and the S.M. and Ph.D. degrees in electrical engineering and computer science from the University of California, Berkeley.

He worked at the IBM T.J. Watson Research Center from 1985 to 1987, was the Analog Devices Career Development Assistant Professor at the Massachusetts Institute of Technology from 1987 to 1989, and is a 1988 Presidential Young Investigator. He is currently an Associate Professor at MIT and his research focuses on theoretical and practical aspects of serial and parallel numerical methods for problems in circuit design and fabrication.

Dimitri A. Antoniadis received the B.S. in physics from the National University of Athens in 1970, the M.S.E.E. degree in 1973 and Ph.D. in electrical engineering in 1976 from Stanford University. From 1969 to 1976 he conducted research in the area of measurements and modeling of the earth's ionosphere and thermosphere. Starting with the development of the SUPREM integrated circuit fabrication process simulator in 1976, his technical activity has been in the area of semiconductor devices and integrated circuit technology. He has worked on the physics of diffusion in silicon, silicon-on-insulator technology and devices, and device measurements. At present his research focuses on: a) extreme submicron and quantum-effect electronic devices, both in silicon and III-V materials; b) computer aids for the design and analysis of silicon devices and of the corresponding fabrication processes; c) novel concepts in semiconductor device fabrication. He is author and co-author of over 120 technical articles.

From 1970 to 1971 he was a Fellow of the National Research Institute, Athens. From 1976 to 1978 he was Research Associate and Instructor in the Department of Electrical Engineering at Stanford University. He joined the faculty at MIT in 1978, where he is Professor of Electrical Engineering. From 1984 to 1990 he was Director of the MIT Microsystems Technology Laboratories, which he helped design and establish.

Professor Antoniadis is a Fellow of IEEE, and a member of the Electrochemical Society and the Materials Research Society.