INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 15, 1221-1238 (1980)

# NUMERICAL ANALYSIS OF THE STEADY-STATE BEHAVIOUR OF A MOS FIELD EFFECT TRANSISTOR†

LEX A. AKERS
*Electrical Engineering Department, University of Nebraska, Lincoln, Nebraska, U.S.A.*

WILLIAM M. PORTNOY
*Electrical Engineering Department, Texas Tech University, Lubbock, Texas, U.S.A.*

## SUMMARY

A numerical scheme is developed to simulate the non-isothermal steady-state behaviour of a MOS field effect transistor. In a desire to develop a fast, stable numerical scheme, physical instabilities were eliminated by using a simplified device model. The numerical technique developed permits a computer solution of the majority carrier transport equation, the nonlinear heat conduction equation, in which the heat generation term is obtained from the solution of the transport equation, and a number of auxiliary differential equations.

The simplified model of the MOS transistor adopted will not, of course, produce any information on the actual operation of the short channel MOS transistor of practical interest today, but the numerical scheme can be extended to simulate short channel models that are of great practical interest.

## INTRODUCTION

The propensity to increase the number of circuit functions per chip has caused, since the development of the first planar transistor, an almost doubling every year of the number of devices per chip. The increase in chip area is partly responsible for this exponential growth, but the decrease in individual device size has been the major contributor.[1]

Hoeneisen and Mead[1] have reported the fundamental limitations which will eventually limit the miniaturization of MOS circuits. The maximum number of circuit functions per unit area is limited by either the power dissipation density or by the area occupied by transistors, interconnections and any passive devices. Therefore, to obtain the maximum number of circuit functions per unit area, both the transistor size and the power dissipation density must be reduced.

The area of a MOS transistor can be reduced by decreasing its channel length and width. This resulting small geometry MOS device has been found to be difficult to characterize because of its multi-dimensional nature.[2] Mathematical models of these small geometry MOS transistors must be at least two- and, more probably, three-dimensional. The three-dimensional model would allow the coupling of both the narrow and short channel effects. General analytical solutions probably do not exist and hence numerical methods are required to reveal the effects of small geometries on device design.

The power dissipation density can be included into such a model by coupling the heat conduction equation into the system of equations which describe the electrical behaviour of the

† This material is based upon work supported by the National Science Foundation under Grant ENG77-05369.

0029-5981/80/0815-1221$01.00

*Received 5 September 1979*
*Revised 14 November 1979*

MOS transistor. The temperature dependence of carrier mobility, density, thermal conductivity and other internal distributions must be included.

The modelling of constant temperature small geometry devices has in many cases proved to be numerically very difficult. Extensive computer time and numerical instabilities have been reported.[3,4] The coupling of another nonlinear equation into an already potentially unstable set of equations could generate extreme numerical instabilities. The problem is further complicated by the presence of physical instabilities resulting from device operation near punch through. It is critical that the numerical scheme be as stable and efficient as possible and any instabilities be contributed only to device operation.

In a desire to develop such a non-isothermal numerical scheme, it was decided to first eliminate the possibility of physical instabilities. This is accomplished by using a simplified MOS transistor model. The model adopted is similar to previously developed long channel models (e.g. Reference 5). This model, of course, will not produce any accurate information on the actual operation of small geometry MOS transistors of practical interest today, but will allow for the development of a stable and fast non-isothermal numerical scheme.

## SYSTEM DESCRIPTION, APPROXIMATIONS AND ASSUMPTIONS

The geometry of the model of the MOS field effect transistor (MOSFET) device used in this analysis is shown in Figure 1. The device consists basically of a $p$-type semiconductor substrate and two $N^+$ regions which have been formed by either diffusion or ion-implantation. One of the $N^+$ regions is labelled the source and the other the drain.

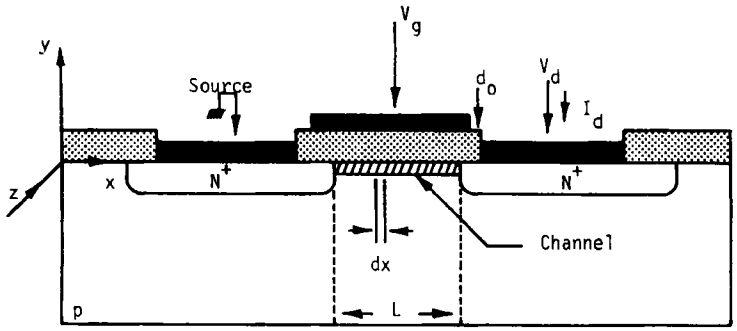


Figure 1. Cross-sectional cut

The source can be internally connected to the substrate or the substrate can be connected to a separate terminal and the MOSFET operated as a four-terminal device. Only the structure with the substrate shorted to the source will be considered. The middle electrode is a metal or a polysilicon electrode over an oxide layer. This electrode is commonly called the gate electrode.

The MOSFET will be assumed structurally perfect (deliberately introduced impurity densities will not be considered structural defects, although lattice strains and dislocations introduced during the diffusion of high concentrations of impurities, i.e. source and drain, fall into that category; these will be neglected).

The simple device structure shown in Figure 1 will be employed and the region of interest will be limited to the conducting channel between the source and drain. Several assumptions, which

are reasonable with respect to the normal operation of the MOSFET, are required to make the calculations manageable:

1. Ideal contacts and charge neutrality at the source and drain.
2. Ideal metal–insulator–semiconductor interface, i.e. no surface states, no trapped surface charge, no work function difference and the infinite resistivity of the insulator.
3. Impurities are fully ionized and their distribution remains constant in time.
4. The Boltzmann statistics are valid (non-degeneracy).

An $n$-channel device will be used as a model, in which hole current and carrier generation and recombination currents are neglected. The electron current will be assumed to be the sum of a drift and a diffusion component.

## CALCULATION OF AN AVERAGE CHANNEL DENSITY

To obtain a relationship between the carrier density at some point $x$ in the channel and the gate and drain voltages, it is necessary to know the surface energy as a function of gate voltage, $V_g$, when the drain voltage at that point is $V_d(x)$. If it is assumed that the transverse electric field, in the channel, $E(y)$, is greater than the longitudinal field, $E(x)$ (the so-called gradual channel approximation) the total charge induced in the inversion layer is:[6]

$$Q_n(x) = -[V_g - \psi_s]C_1 + (2\varepsilon_s q N_A \psi_s)^{1/2}, \tag{1}$$

where $C_1 \equiv \varepsilon_0/d_0$, the capacitance per unit area of the oxide, and $\psi_s(x)$ is the surface potential. $\varepsilon_0$ is the permittivity of the oxide, $\varepsilon_s$ is the permittivity of silicon, $N_A$ is the fixed acceptor density, and $q$ is the electronic charge.

The total charge per unit area within the inversion layer may be related to the channel carrier density by

$$Q_n = -q\int_0^{y_c} n_c(x, y)\,dy, \tag{2}$$

where $n_c(x, y)$ is the channel carrier density, and $y_c$ is the depth of the channel.

Defining the average channel density, $\bar{n}_c(x)$, as

$$\bar{n}_c(x) = \frac{1}{y_c}\int_0^{y_c} n_c(x, y)\,dy, \tag{3}$$

then, according to equation (2),

$$Q_n(x) = -q\bar{n}_c(x)y_c. \tag{4}$$

Therefore, the average channel density is

$$\bar{n}_c(x) = -\frac{Q_n}{qy_c}. \tag{5}$$

For a strongly inverted surface,

$$\psi_s \cong V_d(x) + 2\phi_B, \tag{6}$$

where $\phi_B$ is the bulk potential and[6]

$$y_c(x) = \left(\frac{2\varepsilon_s kT(x)}{n_i q^2}\right)^{1/2} \exp\left(\frac{-q\phi_B}{2kT(x)}\right). \tag{7}$$

## SYSTEM EQUATION

The basic one-dimensional equations which govern the steady-state transport of electrons in the channel are:

1. The continuity equation

$$\frac{\partial J_n(x)}{\partial x} = 0, \tag{8}$$

where $J_n(x)$ is the electron current density.

2. The transport equation

$$J_n(x) = q\mu_n(x)\bar{n}_c(x)E(x) + qD_n(x)\frac{\partial \bar{n}_c(x)}{\partial x}, \tag{9}$$

where $\mu_n(x)$ is the surface electron mobility, $D_n(x)$ is the diffusion constant, and $\bar{n}_c$ is defined by equation (5).

The transport equation transformed into a form explicitly dependent on the electron quasi-Fermi level, $\phi_n$, is

$$J_n(x) = -q\mu_n(x)\bar{n}_c(x)\left(\frac{\partial \phi_n(x)}{\partial x} - \frac{kT(x)}{q}\frac{1}{n_i(x)}\frac{\partial n_i}{\partial x} + \frac{k}{q}\ln\left(\frac{\bar{n}_c(x)}{n_i(x)}\right)\frac{\partial T(x)}{\partial x}\right). \tag{10}$$

The continuity and transport equations, together with the initial distribution and boundary conditions, determine the electrical steady-state behaviour in the channel. The current density, $J_n(x)$, and the electric field in the channel, $E(x)$, provide the forcing function for the heat equation. The differential equation describing steady-state heat conduction is

$$\frac{\partial}{\partial x}\left[\kappa(T)\frac{\partial T}{\partial x}\right] + H(x) = 0, \tag{11}$$

where $\kappa(T)$ is the thermal conductivity, and $H$ is the rate per unit volume at which heat is being generated, i.e.

$$H = J_n(x) \,.\, E(x) \tag{12}$$

(in these units, the mechanical equivalent of heat is unity).

These equations are not independent since the equations describing the electrical behaviour of the device contain temperature-dependent terms. The equations describing the electrical and temperature behaviour of the device form a system of simultaneous nonlinear partial differential equations.

In order to keep the total system tractable, one additional important approximation is required. It will be assumed that all the physical properties of the material, i.e. specific heat capacity, thermal conductivity, and so on, are isotropic and homogeneous, and at most a scalar function of temperature. The carrier mobility and diffusion constant depend on the position co-ordinate through the doping profile, the electric field and the temperature.

The boundary conditions chosen for the electrical response for the MOSFET are

$$\phi_n(0) = \phi_p(0) = \phi_B \tag{13}$$

$$\phi_n(L) = \phi_B(L) + V_d. \tag{14}$$

Because of the simplified one-dimensional model, the temperature boundary conditions will be unrealistic. A complete description would require dynamic (Neumann-type) boundary condi-

tions at the source and drain and a restrained (Dirichlet) boundary condition in a second dimension on the heat sink.

In the one-dimensional case, at least one boundary condition must be of the Dirichlet type. Therefore, a mixed set of boundary conditions were chosen. The source was forced to be thermally as well as electrically tied to the substrate, and the drain was to allow a dynamic boundary condition. Therefore

$$T(0) = T(\text{source}) \tag{15}$$

$$\left.\frac{\partial T}{\partial x}\right|_{x=L} = 0. \tag{16}$$

The nonlinear coefficients in equations (10) and (11) were estimated in the following manner

1. *Mobility.* To express current parallel to the MOSFET surface in the form of a product of the electric field, carrier concentration and mobility, requires the definition of an effective mobility which is a function of the normal surface potential, the parallel channel voltage, temperature, and the average surface mobility. The average $n$-type surface mobility has been measured to be[7] 348 $\text{cm}^2/\text{V sec}$ at 300 K. For an inverted channel, the effective mobility, as a function of gate voltage and temperature, has been found to be approximately[8]

$$\mu_n(T, V_g) = \frac{\mu_{ave}(T/300)^{\alpha}}{\left[1+\frac{kT}{q}\theta(V_g)\right]}, \tag{17}$$

where $\alpha = -1{\cdot}5$ and $\theta = 0{\cdot}074$ for silicon. The effect on mobility from the parallel electric field may be expressed by the factor[9]

$$\left[1+\left(\frac{E(x)}{E_c}\right)^2\right]^{-1/2}, \tag{18}$$

where $E_c$ is the critical electric field at the surface and is approximately $1\times10^7$ V/m.

The total effective mobility as a function of $V_g$, $E(x)$ and $T(x)$ is then

$$\mu_n(V_g, E(x), T(x)) = \frac{\mu_{ave}\left(\frac{T}{300}\right)^{-1{\cdot}5}}{\left[1+\frac{kT}{q}(0{\cdot}074)V_g\right]\left[1+\left(\frac{E(x)}{E_c}\right)^2\right]^{1/2}}. \tag{19}$$

2. *Diffusion constant.* The diffusion constant, $D_n(x)$, will be approximated by using the Einstein relationship

$$D_n(x) = \frac{kT(x)}{q}\mu_n(x). \tag{20}$$

3. *Thermal conductivity.* Graphical thermal conductivity data for silicon[10] were approximated by a continuous function by exponential fitting methods to obtain

$$\kappa(T) = 603\exp(-4{\cdot}46\times10^{-3}T)\left(\frac{W}{\text{m}\,K}\right). \tag{21}$$

4. *Intrinsic carrier concentration.* The expression for the intrinsic carrier concentration, $n_i$, is given by[10]

$$n_i = 3{\cdot}87\times10^{16}T^{3/2}\exp\left(\frac{-1{\cdot}21q}{2kT}\right)\text{cm}^{-3}. \tag{22}$$

## NUMERICAL ANALYSIS

To achieve a numerical solution of this system of nonlinear coupled partial differential equations requires the application of numerical methods. Both finite difference and elements methods are applicable, with the finite difference method being chosen for its coding convenience.

Some guide for existence and uniqueness is advisable before attempting a numerical solution. It is extremely difficult to demonstrate the existence of a solution and its uniqueness for the MOSFET system.[11] In this work, the existence of a unique solution for the system will be assumed. The results obtained seem to justify this assumption.

The system equations (8), (10) and (11) can be rearranged in an equivalent form which is more compatible with discretization by inserting the electron transport equation, equation (10), into the continuity equation, equation (8). It is frequently convenient to leave the heat conduction equation, equation (11), in a self-adjoint form for discretization rather than writing it in the expanded form. It is also important to note that the system equations are both of the parabolic type, which simplifies the numerical procedure as compared to mixed systems.

A second-order correct analogue of the first and second derivatives derived from the Taylor series was employed.[12] Since the boundary conditions are not completely of the Dirichlet type, the Neumann-type boundary conditions must be approximated. This derivative boundary condition was approximated by a second-order correct backwards difference scheme. A uniform mesh spacing, $Dx$, was used throughout the model.

This set of nonlinear partial differential equations together with the boundary conditions, equations (13)–(16), and the specified external voltage excitation, $V_g$, completely represent the MOSFET model. The electron quasi-Fermi potential, $\phi_n(x)$, and the temperature, $T(x)$, are the dependent variables for the system.

The system equations will be discretized to obtain values for the dependent variables at the mesh points.

The system equations written in finite difference form are:

$$
\begin{aligned}
&[[[\mu_n(i)\bar{n}_c(i)]/Dx^2]-[[[[[\mu_n(i+1)-\mu_n(i-1)]/2Dx]\bar{n}_c(i)]+\mu_n(i)[[\bar{n}_c(i+1)\\
&\quad-\bar{n}_c(i-1)]/2Dx]]/2Dx]]\phi_n(i-1)-2[\mu_n(i)^*\bar{n}_c(i)/Dx^2]\phi_n(i)+[[[\mu_n(i)\bar{n}_c(i)]/Dx^2]\\
&\quad+[[[[[\mu_n(i+1)-\mu_n(i-1)]/2Dx]\bar{n}_c(i)]+\mu_n(i)[[\bar{n}_c(i-1)]/2Dx]]/2Dx]]\phi_n(i+1)\\
&=[[[D_n(i)/n_i(i)][[n_i(i+1)-n_i(i-1)]/2Dx]-[k\mu_n(i)/q][\ln[\bar{n}_c(i)/n_i(i)]][[T(i+1)\\
&\quad-T(i-1)]/2Dx]-[k\mu_n(i)/q][[T(i+1)-T(i-1)]/2Dx]][[\bar{n}_c(i-1)]/2Dx]\\
&\quad+[[D_n(i)/n_i(i)][[n_i(i+1)-2n_i(i)+n_i(i-1)]/Dx^2]\\
&\quad-[D_n(i)/n_i(i)^2][[[n_i(i+1)-n_i(i-1)]/2Dx]^2]\\
&\quad+[k\mu_n(i)/q][1/n_i(i)][[n_i(i+1)-n_i(i-1)]/2Dx][[T(i+1)-T(i-1)]/2Dx]\\
&\quad+[kT(i)/q][1/n_i(i)][[\mu_n(i+1)-\mu_n(i-1)]/2Dx][[n_i(i+1)-n_i(i-1)]/2Dx]\\
&\quad-[k/q][[\mu_n(i+1)-\mu_n(i-1)]/2Dx]\ln[\bar{n}_c(i)/n_i(i)][[T(i+1)-T(i-1)]/2Dx]\\
&\quad-[k\mu_n(i)/q]\ln[\bar{n}_c(i)/n_i(i)][[T(i+1)-2T(i)+T(i-1)]/Dx^2]\\
&\quad+[k\mu_n(i)/q][1/n_i(i)][[n_i(i+1)-n_i(i-1)]/2Dx][[T(i+1)-T(i-1)]/2Dx]]\bar{n}_c(i)
\end{aligned}
\tag{23}
$$

and

$$\frac{\kappa(i-\frac{1}{2})}{Dx^2}T(i-1)-\left(\frac{\kappa(i+\frac{1}{2})}{Dx^2}+\frac{\kappa(i-\frac{1}{2})}{Dx^2}\right)T(i)$$

$$+\frac{\kappa(i+\frac{1}{2})}{Dx^2}T(i+1)=-J_n(i)E(i), \tag{24}$$

where $\kappa$ is the thermal conductivity evaluated at half space step:

$$\kappa(i+\tfrac{1}{2})=\kappa\left(\frac{T(i+1)+T(i)}{2}\right) \tag{25}$$

$$\kappa(i-\tfrac{1}{2})=\kappa\left(\frac{T(i)+T(i-1)}{2}\right). \tag{26}$$

It is advantageous to restrict the difference equations to those leading to a triagonal matrix. The triagonal matrix can be solved explicitly for the unknowns, eliminating any matrix operations. An algorithm developed by Thomas[12] was used to obtain rapid and accurate solutions to the triagonal system.

The nonlinear coefficients in equations (23) and (24) must be evaluated at the present iteration value of the dependent variables, but for the resulting finite difference equations to be linear, the expression must not contain coefficients of undetermined value at the iteration step.

Many methods[12] of evaluating the nonlinear coefficients are available, such as an evaluattion using the previous value, the forward projection of the coefficient, the backward Taylor series projection, and so on.

One of the simplest methods of obtaining the nonlinear coefficients while keeping the system linear is the method of iteration using the old value. If the nonlinear coefficients do not change very rapidly with the dependent variable, the coefficient evaluated at the previous iteration value is approximately equal to the true value of the coefficient. The equations are then iterated to within a predetermined tolerance. The method of evaluating the coefficient using the previous iteration value will be used throughout this work.

Since this system of equations must be solved by iteration, the value of the resubstituted dependent variable is of considerable importance. One popular method is direct resubstitution. In an effort to accelerate convergence, relaxation of the dependent variables was employed. Basically this method consists of calculating the value of each dependent variable and then modifying the value before resubstitution. The general form for this relaxation is:[13]

$$y_i^m = y_i^{m-1} + \lambda(y_i^{m^*} - y_i^{m-1}), \tag{27}$$

where $m$ is the current iteration and $m-1$ the preceding iteration. The quantity $y_i^{m^*}$ is the value of the dependent variable on the current iteration. $\lambda$ is called the relaxation factor and is a real number in the range of $0<\lambda<2$.

If $\lambda<1$, the current value of the dependent variable generated is a weighted average of the present and previous iteration values. This is termed underrelaxation and is important in making potentially non-convergence iterative processes converge.

The optimum value of $\lambda$ is rather difficult to calculate. In most circumstances, the choice for $\lambda$ can be found by trial and error.

The error introduced by replacing a continuous problem by a discrete problem for this nonlinear set is very difficult to calculate. Although the truncation errors, the error introduced by the first neglected term in the Taylor series, is of the order of $Dx^2$ per iteration,[12] a total error

analysis is complicated by the presence of several sources of errors of different natures making a realistic error estimation impossible.[14] We have determined that for different mesh lengths, the equations always converge towards the same limit.

A block diagram of the numerical sequence to obtain a solution is shown in Figure 2. An initial starting state description of the device, i.e. $\phi_n(x)$ and $T(x)$, is required to specify the device

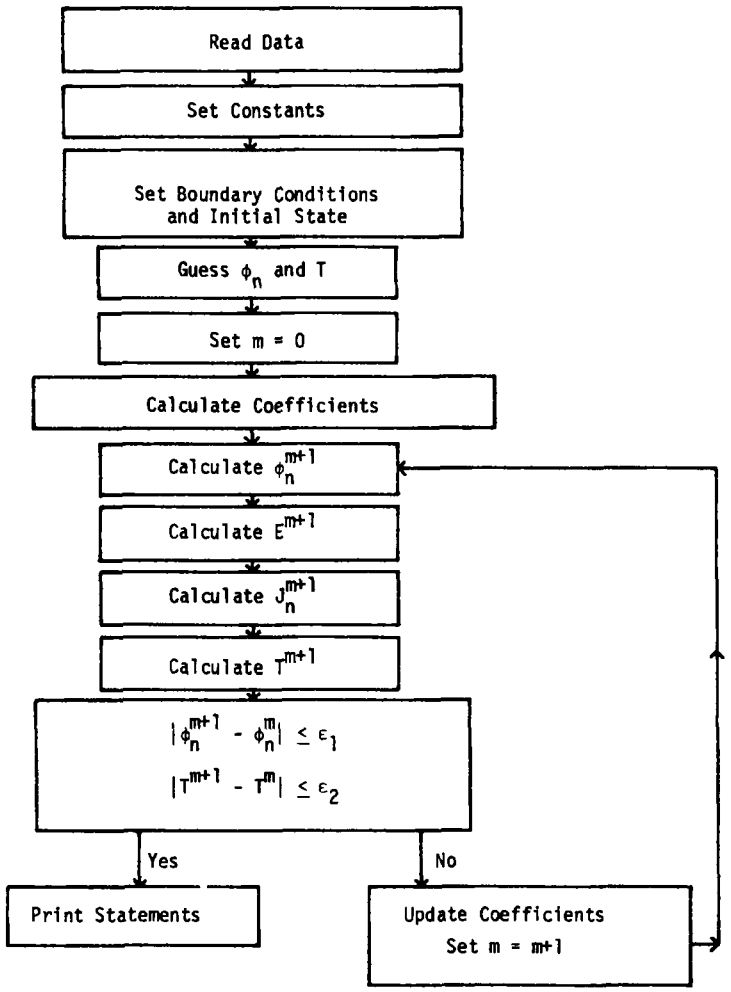


Figure 2. Numerical sequence

completely for the initial iteration. This initial state can be a zero bias state or a steady-state bias state. The values of the coefficients are obtained and the set of unknows is evaluated. An initial distribution of the variables is inserted into equation (23) and an improved $\phi_n(i)^{m+1}$ results. The improved electron quasi-Fermi level and the initial distribution for $T$ are inserted into equation (24) to provide an improved $T(i)^{m+1}$.

Upon obtaining an improved solution for the dependent variables, a check is made to see if a predetermined degree of consistency has been achieved in the iterative solution. If the desired degree of consistency has been achieved, the steady-state condition has been obtained.

## DISCUSSION AND RESULTS

The numerical analysis techniques described earlier were incorporated into a computer code. The results of this analysis provide the spatial properties of the dependent variables and coefficients for different gate and drain voltages and supplies the I–V plots.

In all the illustrations in this section, the same substrate doping level, source and drain doping levels, oxide depth and channel length were used; these were:

$$N_A = 1 \times 10^{15}\ \text{cm}^{-3} \tag{28}$$

$$d_o = 1 \times 10^{-7}\ \text{m} \tag{29}$$

$$L = 6 \times 10^{-6}\ \text{m}. \tag{30}$$

The optimum values of the relaxation factors were found by trial and error. Defining $\lambda\phi_n$ as the quasi-Fermi potential relaxation factor and $\lambda_T$ as the temperature relaxation factor, Figure 3

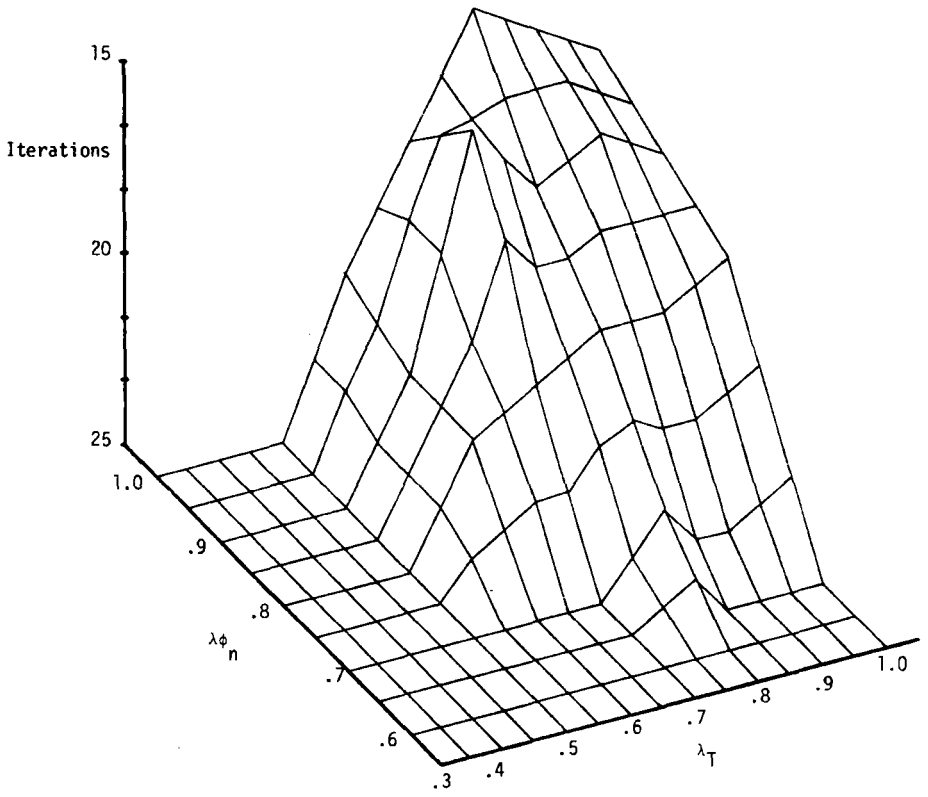


Figure 3. $\phi_n$ and $T$ relaxation vs. number of iterations

illustrates the effect of relaxation on the iteration rate. The proper choice of relaxation factors in many cases increased the convergence rate of the system of equations by more than a factor of 2. Relaxation factors of $\lambda\phi_n = 1{\cdot}0$ and $\lambda_T = 0{\cdot}8$ were used for this work.

Figure 4 is a plot of the electron quasi-Fermi potential, $\phi_n(x)$, vs. $x$ for a constant gate voltage and various drain voltages. It should be noted that, as $V_d$ increases, the gradient in $\phi_n$ in the drain region rapidly increases. This rapid change in $\phi_n$ is caused by the widening of the drain depletion region as $V_d$ increases.[6]

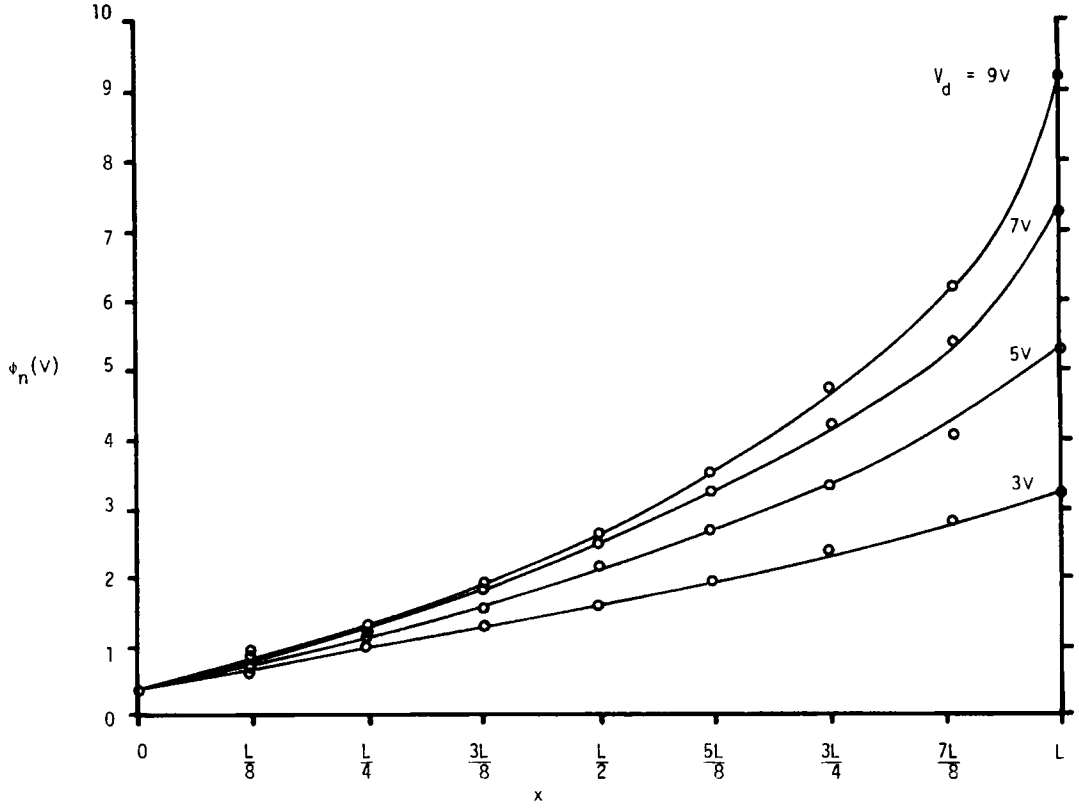


Figure 4. $\phi_n$ vs. $x$ for a constant gate voltage, $V_g = 13$ V, and various drain voltages

The power dissipated in the device, $J_n(x) \cdot E(x)$, will cause the temperature of the MOSFET to rise. Figure 5 illustrates $T(x)$ vs. $x$ for a fixed gate voltage and varying drain voltages. As $V_d$ increases, the temperature increases, since the current and the electric field are increasing. In Figure 6, $V_g$ is varied and $V_d$ is held constant. As $V_g$ increases, the number of carriers brought to the surface will increase, causing $J_n(x)$ to increase and the temperature to rise.

The spatial variations of $\bar{n}_c(x)$ for a constant $V_g$ and varying $V_d$ are shown in Figure 7. As expected, the channel density curve is always a maximum at the source and decreases approaching the drain.

Since the uniqueness or even the existence of a solution for the describing system of equations could not be obtained, verification of the computer generated quasi-Fermi potential and

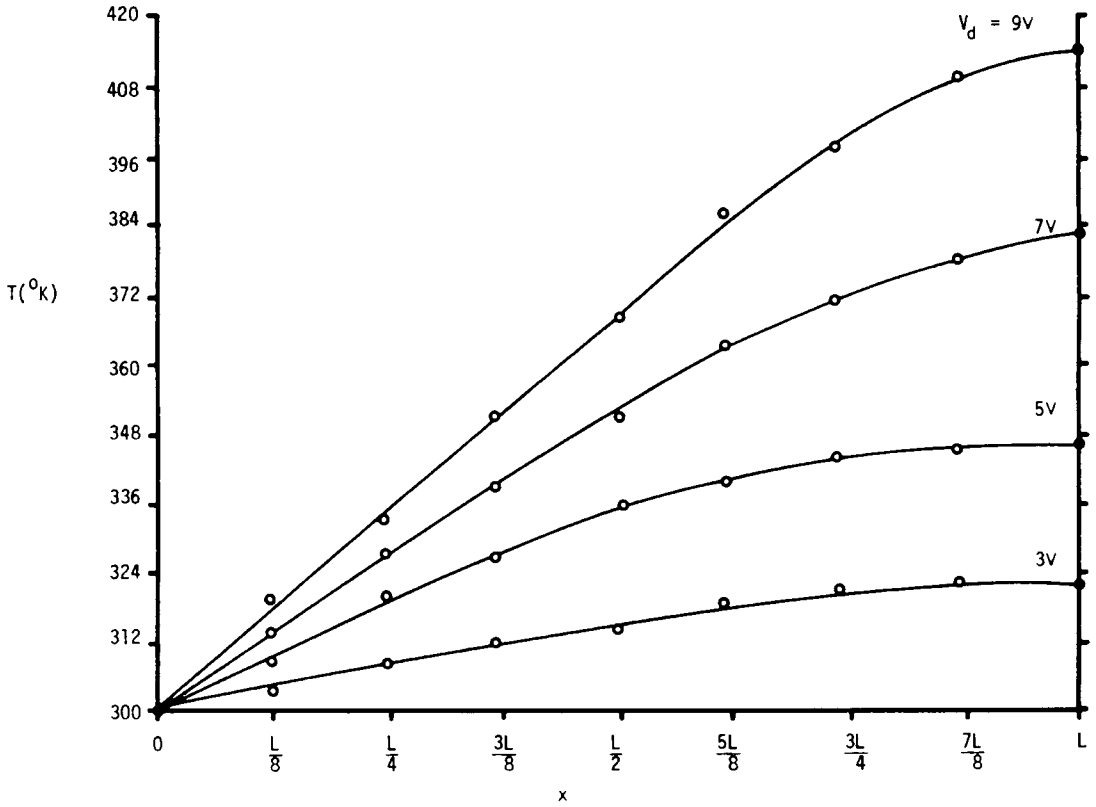


Figure 5. $T$ vs. $x$ for a constant gate voltage, $V_g = 13$ V, and various drain voltages

temperature distribution is needed. A constant temperature model was developed for this purpose. Holding the temperature constant in the non-isothermal model, the $\phi_n$ distribution was compared with the constant temperature model and found to be equal. The temperature distribution may be verified analytically for low $V_d$ cases. For a constant thermal conductivity and small $V_d$, $E(x)$ may be approximated by

$$E(x) = ax + b \tag{31}$$

and equation (11) may be solved as

$$T(x) = -\frac{aJx^3}{6\kappa} - \frac{Jbx^2}{2\kappa} + \left(\frac{aJL^2}{2\kappa} + \frac{JbL}{\kappa}\right)x + T_s. \tag{32}$$

Figure 8 is a plot of the analytical solution and the computer generated solution. Similar analytical results have been obtained for large $V_d$, but the approximation of a linear $E(x)$ is not as valid and results in an overestimation of $T(x)$. The results obtained, however, do indicate that the temperature distribution generated by the finite difference method is consistent with expected results.

The one-dimensional form of Gauss's law used in obtaining equation (1) implies that the charge within the inversion player of the MOSFET produces an electric field only in the $y$ direction. This approximation is only partially correct and is only appropriate when $E(y) \gg E(x)$. As previously discussed, a two-dimensional analysis at least is required to represent this condition properly. Not all the inversion charge in the channel contributes to $E(y)$; there are a small number of charges within the channel that are not associated with the gate voltage. This small number of charges produce the electric field, $E(x)$, required to maintain the current continuity between the source and drain.

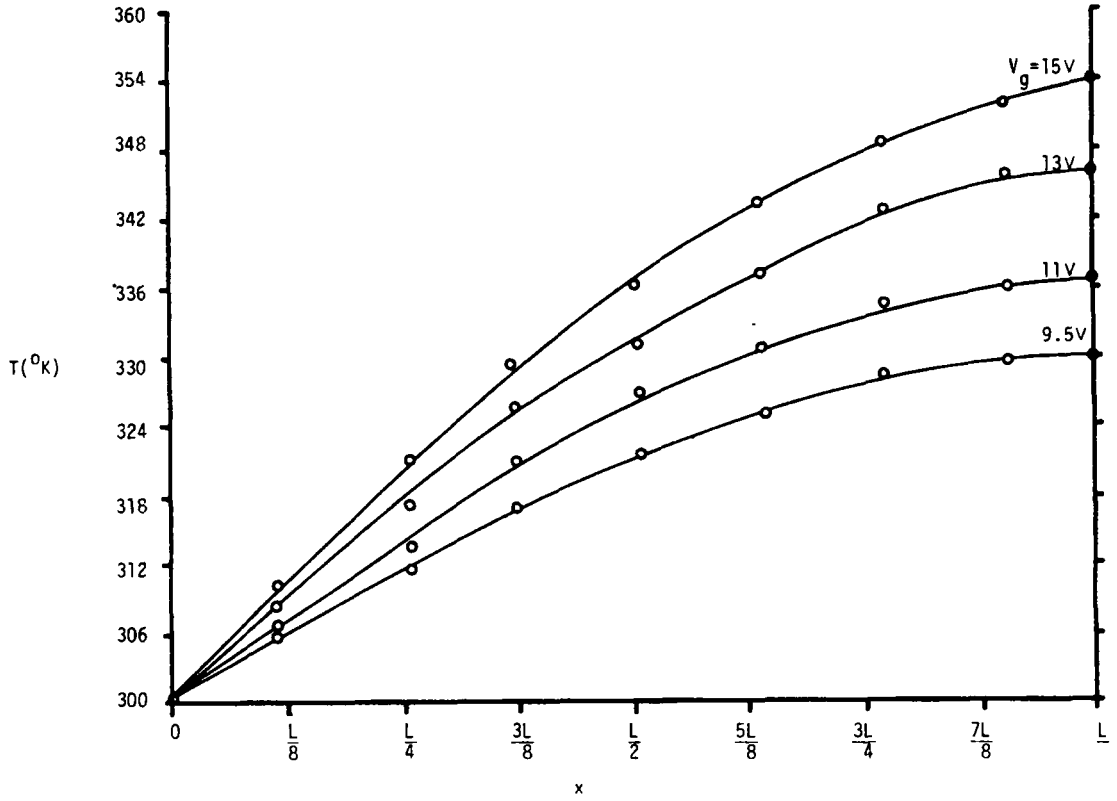


Figure 6. $T$ vs. $x$ for a constant drain voltage, $V_d = 5$ V, and various gate voltages

Therefore, as $V_d$ becomes larger for constant $V_g$, $\phi_n(x)$ becomes larger, especially in the vicinity of the drain, and because of the relationship between $\phi_n(x)$ and $E(x)$, the electric field in the $x$ direction becomes significant with respect to $E(y)$. The gradual channel approximation is no longer valid, and the carrier density, equation (1), is no longer an accurate representation for the carrier density in the channel.[6]

The spatial variations of the other variables and coefficients of interest are shown in Figure 9, $E(x)$, Figure 10, $\mu_n(x)$ and Figure 11, $\kappa(x)$.

Figure 12 compares the static curves for the non-isothermal model with the constant temperature model. For the strong inversion case, the values obtained from the non-isothermal model are lower than those for the constant temperature model; this is probably the result of reduced carrier mobility in the channel because of heating.

The time used for the computation is not very significant since this number varies with the applied voltages and number of mesh points. On an IBM 360-65J for 100 mesh point with $V_g = 13$ V and $V_d = 9$ V, the GO step took 30·53 sec of CPU time.

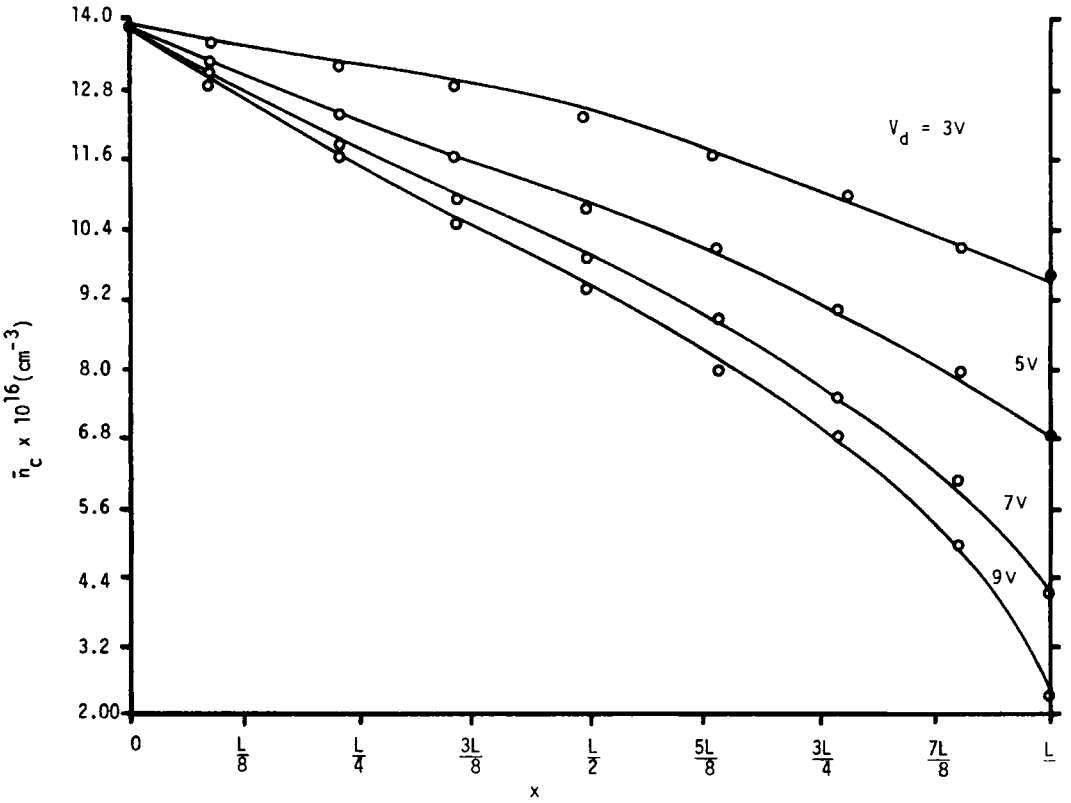


Figure 7. $\bar{n}_c$ vs. $x$ for a constant gate voltage, $V_g = 13$ V, and various drain voltages

## CONCLUSIONS

A non-isothermal numerical scheme was developed to simulate a long channel MOSFET device. Numerical solutions were obtained for the basic majority carrier transport equations, the nonlinear heat conduction equation and a number of auxiliary differential equations. The rate of convergence of the numerical scheme was accelerated by using underrelaxation of both the quasi-Fermi potential and the temperature distribution. Steady-state solutions were found throughout the conducting channel of the device for the electron quasi-Fermi potential, $\phi_n(x)$, temperature $T(x)$, electric field, $E(x)$, and channel density, $\bar{n}_c(x)$. Terminal characteristics were compared with a constant temperature model.

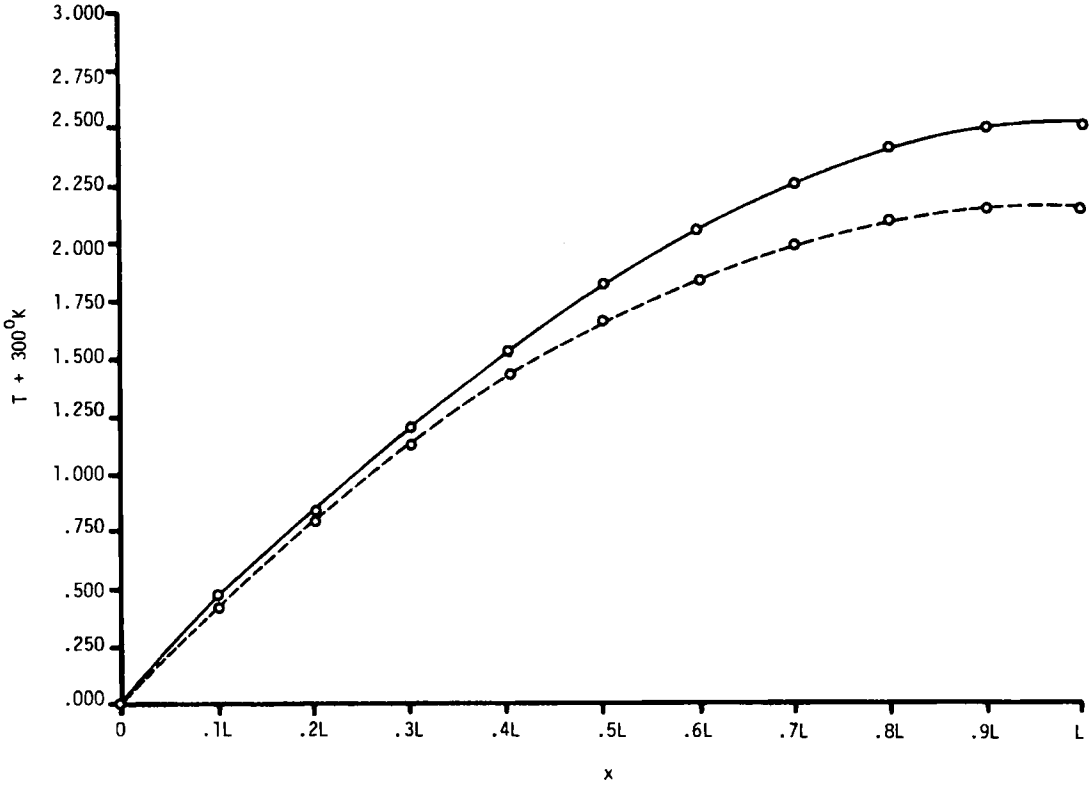


Figure 8. Comparison of the temperature profile for an analytical solution (solid line), and non-isothermal computer model (dashed line)

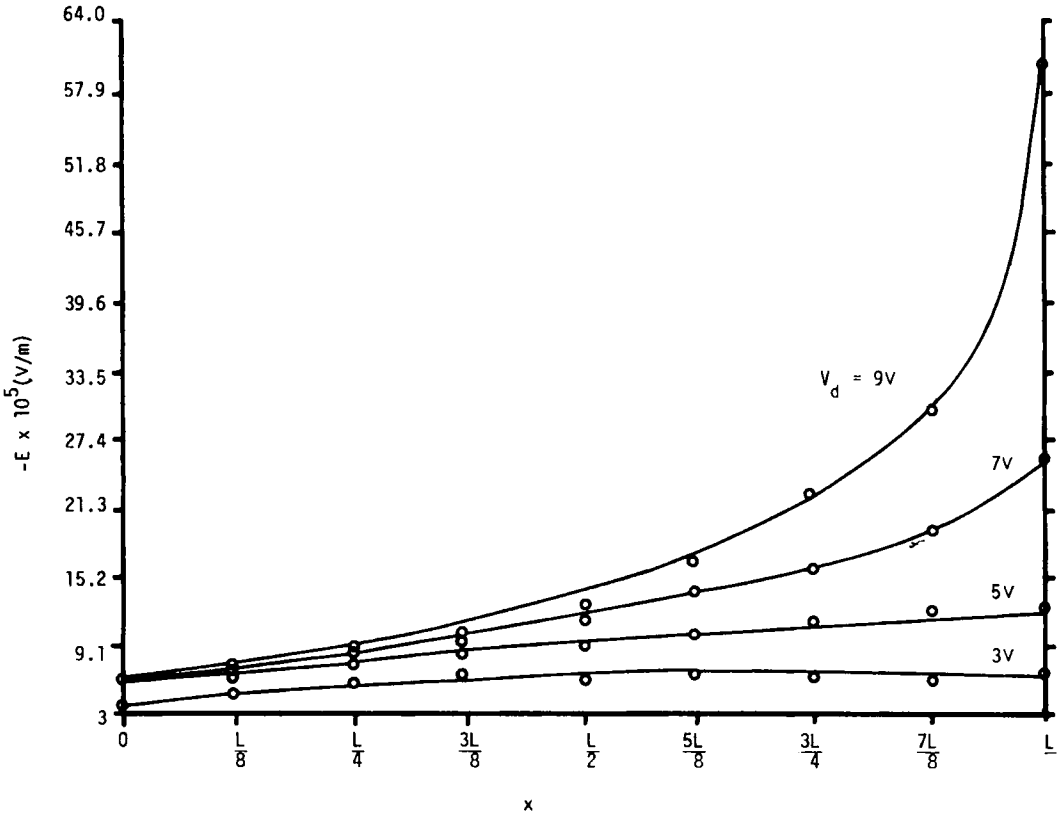


Figure 9. $E$ vs. $x$ for a constant gate voltage, $V_g = 13$ V, and various drain voltages

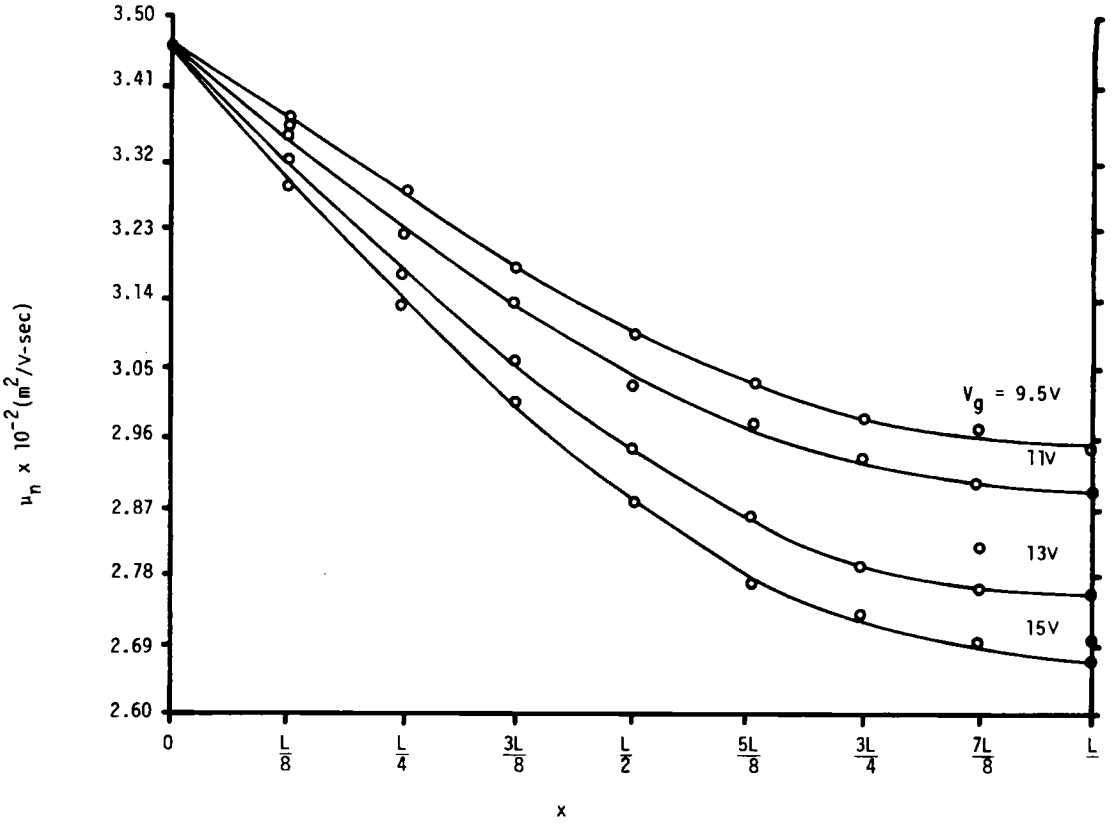


Figure 10. $\mu_n$ vs. $x$ for a constant drain voltage, $V_d = 5$ V, and various gate voltages

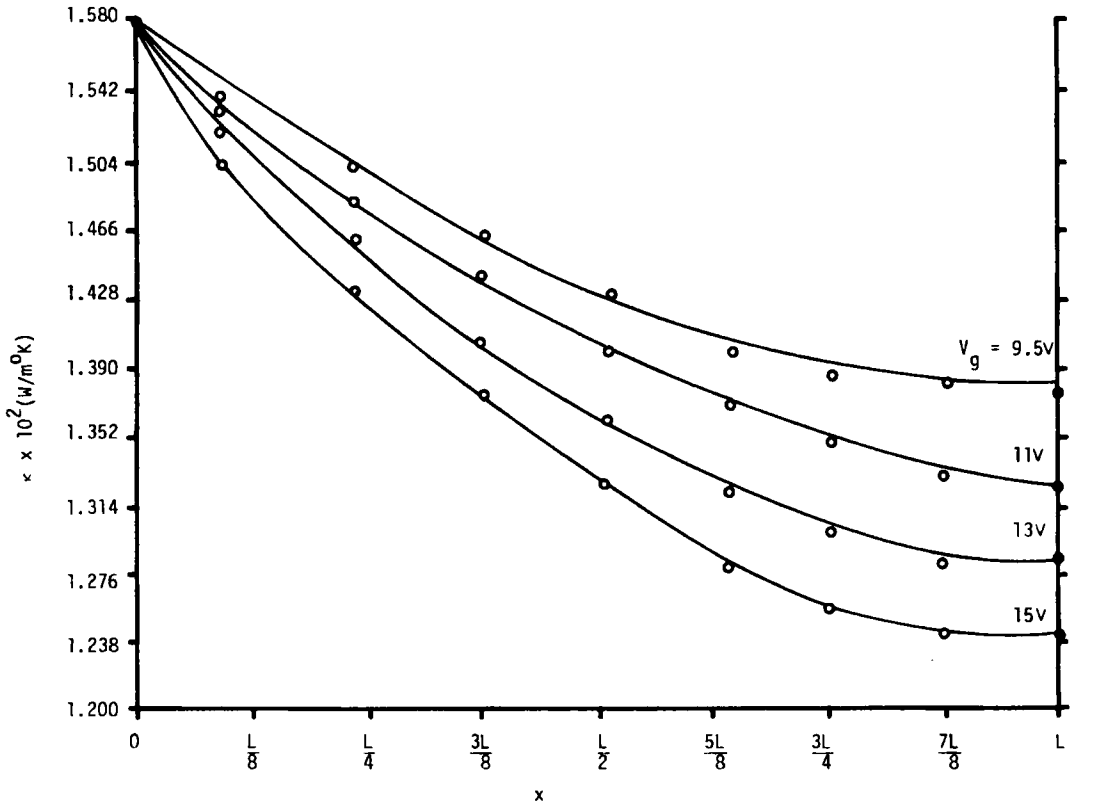


Figure 11. $\kappa$ vs. $x$ for a constant drain voltage, $V_d = 5$ V, and various gate voltages

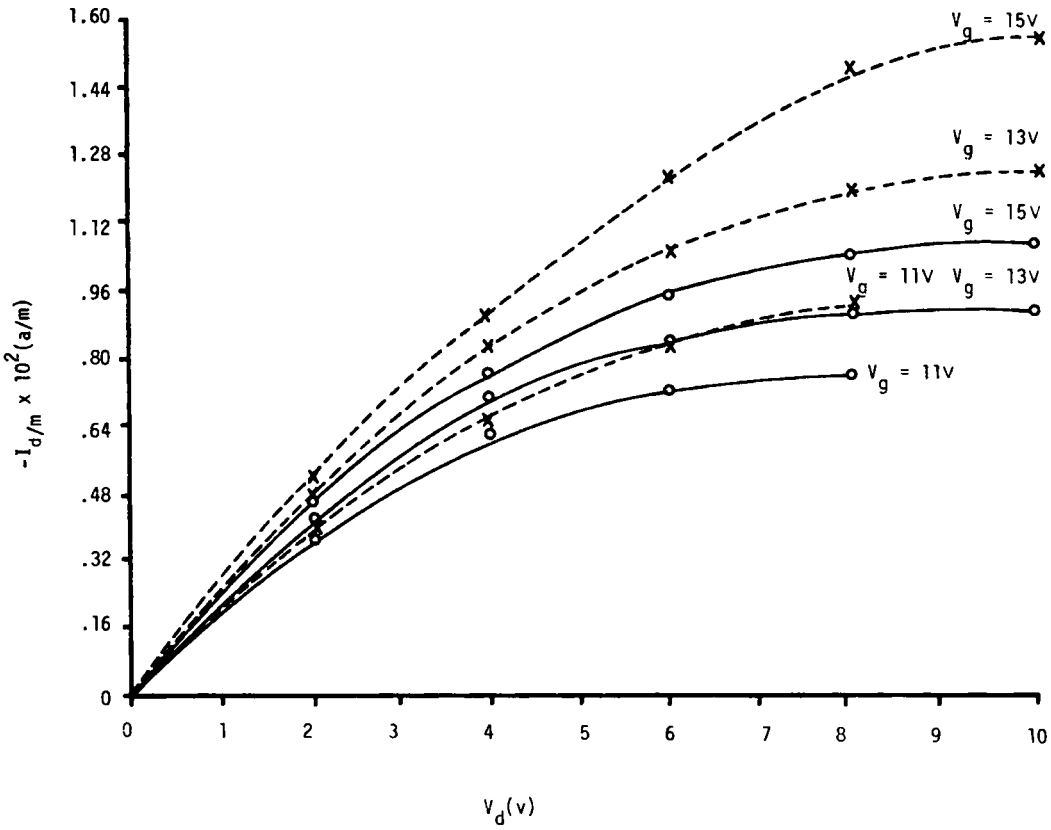


Figure 12. Comparison of a constant temperature (dashed line), and non-isothermal computer model (solid line)

One important point must be made made with respect to these calculations. A Boltzmann form of the carrier density was employed throughout. Although the concept of the quasi-Fermi level was originally introduced[15] to permit the use of the Boltzmann form, its validity is limited to non-equilibrium situations in which voltage variations are small, and temperature is essentially constant. When significant voltage and temperature excursions exist, the basic assumptions respecting relaxation processes in the material may be violated, and, in order to obtain a satisfactory distribution function, from which carrier density can be calculated, the Boltzmann equation, at last, must be solved for the non-equilibrium condition. For very large temperature and voltage excursions, the entire concept of a relaxation time may be questioned. It is not possible to perform calculations based on first principles under these conditions. The work performed here does not include the large voltage or temperature excursion case.

## APPENDIX: NOTATION

$C_1$ = Capacitance per unit area of the oxide.
$D_n$ = Diffusion constant.
$d_o$ = Oxide thickness.

$E_c$ = Critical electric field.
$E(x)$ = Longitudinal electric field.
$E(y)$ = Transverse electric field.
$i$ = Spatial index (integer).
$J_n$ = Electron current density.
$k$ = Boltzmann's constant.
$L$ = Channel length.
$n$ = Electron density.
$N^+$ = Source and drain density.
$N_A$ = Acceptor density.
$n_c$ = Electron channel density.
$\bar{n}_c$ = Average electron channel density.
$n_i$ = Intrinsic density.
$p$ = Hole density.
$q$ = Electronic charge.
$Q_n$ = Inversion charge per unit area.
$T$ = Absolute temperature.
$V_d$ = Drain voltage.
$V_g$ = Gate voltage.
$y_c$ = Channel depth.
$\varepsilon_o$ = Permittivity of oxide.
$\varepsilon_s$ = Permittivity of silicon.
$\varepsilon_1, \varepsilon_2$ = Error bounds.
$\kappa$ = Thermal conductivity
$\phi_B$ = Bulk potential.
$\phi_n$ = Electron quasi-Fermi potential.
$\phi_p$ = Hole quasi-Fermi potential.
$\psi_s$ = Surface potential.
$\mu_{ave}$ = Average $n$-type surface mobility.
$\mu_n$ = Surface electron mobility.

## REFERENCES


1. B. Hoeneisen and C. A. Mead, 'Fundamental limitations in microelectronics—1 MOS technology', *Solid-St. Electron*, **15**, 819–829 (1972).
2. F. H. Gaensslen, 'Geometry effects of small MOSFET devices', *Digest of IEDM*, 512–515 (1977).
3. M. Reiser, 'Large-scale numerical simulation in semiconductor device modelling', *Comp. Meth. Appl. Mech. Engng*, **1**, 17–38 (1972).
4. J. J. Barnes and R. J. Lomax, 'Two-dimensional finite element simulation of semiconductor devices', *Electron. Lett.* **10**, 341–343 (1974).
5. D. P. Kennedy and P. C. Murley, 'Steady state mathematical theory for the insulated gate field effect transistor', *IBM 5, Res. Develop.*, Jan., 2–12 (1973).
6. S. M. Sze, *Physics of Semiconductor Devices*, Wiley–Interscience, New York, 1969.
7. G. Baccarani, A. M. Mazzone and C. Morandi, 'The diffuse scattering model of effective mobility in the strongly inverted layer of MOS transistors', *Solid-St. Electron*, **17**, 785–789 (1974).
8. G. Merckel, J. Borel and N. Z. Cupcea, 'An accurate large-signal MOS transistor model for use in computer-aided design', *IEEE Trans. Electron Devices*, **19**, 681–690 (1972).
9. D. M. Caughey and R. E. Thomas, 'Carrier mobilities in silicon empirically related to doping and field', *Proc. IEEE*, **55**, 2192–2193 (1967).
10. H. F. Wolf, *Silicon Semiconductor Data*, Pergamon Press, Oxford, 1969.
11. M. S. Mock, 'A two-dimensional mathematical model of the insulated-gate field effect transistor', *Solid-St. Electron*, **16**, 601–609 (1973).

12. D. Von Rosenberg, *Method for the Numerical Solution of Partial Differential Equations*, Elsevier, New York, 1969.
13. R. W. Hornbeck, *Numerical Methods*, Quantum, New York, 1975.
14. A. DeMari, 'An accurate numerical steady-state one-dimensional solution of the P–N junctions', *Solid-St. Electron*, **11**, 33–58 (1968).
15. W. Shockley, *Electrons and Holes in Semiconductors*, Van Nostrand, Princeton, N.J., 1950.