Modeling MOS Transistors in the Avalanche-Breakdown Regime

A. Schütz
Siemens Research Laboratories
Otto-Hahn-Ring 6
D-8000 Munich, Federal Republic of Germany

S. Selberherr and H. W. Pötzl
Department of Electrical Engineering & Electronics
Technical University of Vienna
Gusshausstrasse 27-29
A-1040, Vienna, Austria

ABSTRACT
Impact ionization has a very disadvantageous effect for many semiconductor devices because it may cause avalanche breakdown and destruction of the device. For an MOS circuit design an understanding of avalanche-induced breakdown effects is desirable when the dimensions of transistors are reduced. This requires an accurate solution of the semiconductor equations and a good description of the ionization rates. Such a model is presented here and applied to calculate the electronic potential and carrier density distribution in MOS transistors operating near breakdown.

1. INTRODUCTION
Several multi-dimensional numerical models of semiconductor devices have been developed in recent years [5,13,15,24,31] as a response to rapid progress in the development of integrated circuits. The aim of these models is to help in understanding device physics. This has become absolutely necessary when shrinking the dimensions of the device [11]. To enhance the efficiency of such models, i.e., to reduce computer costs and to lower the program complexity, certain simplifying assumptions are usually made. In general, computing costs and program complexity increase when more physical accuracy is required. Generation-recombination mechanisms have been treated as second-order effects for MOS transistors and, therefore, have been often neglected in MOS models [32]. As a consequence, such models cannot be used to investigate avalanche breakdown. This normally undesirable effect is originated by impact ionization of hot carriers in a high field.
region. Avalanche breakdown in MOS transistors is associated with substrate current which, as indicated by experiments, increases rapidly when the drain voltage approaches the breakdown voltage.

The importance of breakdown analysis increases with the increasing degree of miniaturization. Therefore, exact modeling of impact ionization is required and generation-recombination mechanisms must be considered. For these reasons a consistent model, including both electrons and holes, impact ionization and various recombination mechanisms, has been developed and is presented below.

2. FUNDAMENTAL EQUATIONS

The numerical models of semiconductor devices are based on the basic semiconductor equations first given by van Roosbroeck [40]. These equations are repeated here for the reader’s convenience.

Poisson’s equations:

\[ \nabla \cdot \varepsilon \nabla \Psi = -q \left( p - n + N_D^* - N_A^- \right) \]

(1)

Continuity equations:

\[ \nabla \cdot J_n = -q \left( G - R \right) \]

(2)

\[ \nabla \cdot J_p = q \left( G - R \right) \]

when the current relations:

\[ J_n = -q \left( \mu_n n \nabla \Psi - D_n \nabla n \right) \]

(3)

\[ J_p = -q \left( \mu_p p \nabla \Psi + D_p \nabla p \right) \]

The right-hand sides of Equations (2) are the generation and recombination rates which are usually negligible in non-avalanche regions. When explicitly considering avalanche, however, electron-hole pair generation can no longer be neglected since avalanche breakdown is completely governed by generation. Recombination must not be neglected either because the high level of ionization in the high field regions can lead to a drastic increase of carrier densities throughout the device, thus rendering recombination more important. Therefore, the \( G - R \) term consists of thermal and surface generation/recombination, Auger recombination, and avalanche generation:

\[ (G - R) = (G - R)_{\text{th}} + (G - R)_s + (G - R)_{\text{Aug}} + G_a, \]

(4)

with:

\[ (G - R)_{\text{th}} = \frac{n_i^3 - p \cdot n}{\tau_n \left( p + p_i \right) + \tau_p \left( n + n_i \right)} \]

(5a)

\[ (G - R)_s = \frac{n_i^3 - p \cdot n}{\left( p + p_i \right) / s_n + \left( n + n_i \right) / s_p} \cdot \delta(y) \]

(5b)

*See pp. 12-13 for list of notations.
where \( \delta(y) \) is the Dirac-delta generalized function (delta-distribution) \( y = 0 \) denotes the interface,

\[
(G - R)_{\text{Aug}} = (n_i^2 - p \cdot n) \quad (C_n n + C_p p)
\]  

(5c)

Equations (5a) and (5b) are the commonly known Shockley-Read-Hall terms for thermal bulk and surface recombination processes. Equation (5c) describes Auger recombination as given in [8]. The avalanche generation term, \( G_a \), is defined as follows

\[
G_a = \frac{|J|}{q} \alpha_n(E) + \frac{|J|}{q} \alpha_p(E)
\]  

(5d)

Equation (5d) states that the impact generation rate, \( G_a \), is proportional to the current density, \( J \), times an ionization probability \( \alpha(E) \). Other sources [4] suggest slightly different formulation:

\[
G_a = n \cdot |v_n(E)| \cdot \alpha_n(E) + p \cdot |v_p(E)| \cdot \alpha_p(E)
\]  

(5e)

The difference between Equations (5d) and (5e) is based on the treatment of the diffusion current. The question whether a diffusion current contributes to impact ionization has been discussed in the literature but could not be answered satisfactorily [3,4].

A simplified method to model the avalanche breakdown was chosen by Toyabe [36] and Kotani [17]. First, the drain current is calculated neglecting impact ionization. This can be performed by solving Poisson's equation, (1), with one continuity equation, (2). Avalanche breakdown effect is then introduced by multiplying the drain current by a factor, \( M_n > 1 \), determined using the formula

\[
\frac{1}{1 - M_n} = \int_0^L \alpha_n \exp \int_0^x (\alpha_p - \alpha_n) \, dx' \, dx
\]  

(6)

where \( L \) is the channel length.

The simplified treatment allows for some savings in computational costs; however, the influence of an increased hole density cannot be treated. In particular, this simplified method is not able to account for the inherent feedback mechanism which becomes apparent by the snap-back phenomenon.

### 3. Models for \( \alpha_{n,p} \)

This section treats field dependence of impact ionization probabilities.

It is a well-known fact that the ionization rates strongly depend on the electric field. Early measurements [21,22,23] showed that they could satisfactorily be fitted by Chynoweth's law [6]

\[
\alpha_{n,p}(E) = A_{n,p} \exp \left( - \frac{B_{n,p}}{|E_r|} \right).
\]  

(7)
Where $A_{n,p}$ and $B_{n,p}$ are constants, and $E_r$ is the field strength. Later measurements [10, 12, 18, 25, 26, 39] support Chynoweth's law; however, some of them show that two or three sets of values for $A_{n,p}$ and $B_{n,p}$ should be used in order to fit the expressions of type (7) to the experimental data over a broad range of the electric field.

Figures 1 and 2 compare experimental results for the ionization rates of electrons and holes, respectively. Taking into account that these figures are drawn in a logarithmic scale, the differences are rather large, especially for the holes. In addition to the experimental data, the theoretical results of Baraff [2] with material constants from Crowell and Sze [9] are also drawn in Figures 1 and 2. These results have been obtained by solution of Boltzmann's equation with several simplifications concerning the band structure. Using the energy conservation considerations Okuto and Crowell [27] have given an upper limit for the ionization rates which implies that all the energy a carrier obtains from the electric field will be lost by impact ionization; thus

$$\alpha(E) = \frac{q|E_r|}{W_i} \quad (8)$$

---

**Figure 1.** Ionization rates for electrons ($\log \alpha_n$) versus reciprocal field strength obtained by various authors.
Figure 2. Ionization rates for holes (log \( \sigma_h \)) versus reciprocal field strength obtained by various authors.

where \( W_f \) is the mean energy loss per ionization and is assumed to be 1.6 eV and 1.8 eV for electrons and holes, respectively (Figures 1 and 2).

Grant's data and van Overstraeten's data show good agreement, but they differ considerably from Lee's data. On the other hand, Lee's data agree with Baraff's results which cannot be reasonably fitted to Grant's data [12]. Recent theoretical results [34] obtained by solution of Boltzmann's equation using Monte-Carlo techniques and realistic band structures [7] support both Lee's and Baraff's data.

4. IONIZATION THRESHOLD ENERGY

So far the ionization rates were considered to be functions of the electric field strength only. Let us consider a situation with a rather abrupt change from a low-field region to a high-field region. According to Chynoweth's law, Equation (6), the ionization rates would immediately follow the electric field. Energy conservation, however, requires that the ionizing carrier has an energy which at least is equal to the band gap. Therefore, the ionization rates cannot immediately follow the electric field but the carriers have to travel a minimum distance within the electric field to gain energy enough for the first ionization. That distance where no ionization can occur is commonly called "dark space." Another name related to the dark space
effect is "non-local effects" as the ionization rates are no more functions of field strength alone but also functions of the electron trajectories in the electric field.

In addition to the energy conservation, the momentum must also be conserved during an ionization process. This additional condition may require the minimum energy sufficient for ionization (ionization threshold energy) to become larger than the band gap. For a direct semiconductor with equal effective masses of electrons and holes, the ionization threshold energy becomes 1.5 times the band gap [33]. For realistic band structures the calculation of the threshold energy becomes quite different. Anderson [1] calculated the ionization energy for electrons to be 1.1 eV, which is just the band gap. Hauser [14] obtained $W_I = 1.3$ eV. An experimental value has been found by van Overstraeten [39] ($W_I = 1.8$ eV). Thornber has found another one [35] ($W_I = 3.6$ eV).

5. CHARACTERISTICS OF MOS TRANSISTOR AVALANCHE BREAKDOWN

In MOS transistors impact ionization is most prominent when the device operates in the pentode region, i.e., at large $V_{DS}$ and moderate $V_{GS}$. In this part of the $I$-$V$ characteristic the applied drain-to-source voltage does not drop gradually along the channel but drops mainly within the small pinch-off region near drain. This results in a sharply localized, very high electric field. Figure 3 shows the potential distribution for such operating conditions. If the peak field strength exceeds approximately 200 kV/cm impact ionization is considerable. The electric field splits the generated electron-hole pairs: electrons travel to the drain and holes flow into the substrate giving rise to substrate current strongly dependent on $V_{DS}$ [18, 37]. Measured and calculated substrate currents in a transistor with 2 $\mu$m effective channel length are shown in Figure 4.

The hole density is very much affected by impact ionization, as demonstrated by comparison of Figures 5 and 6. In simulation creating Figure 5 the avalanche effect was included; in Figure 6 it was neglected. Below the source and under the channel no depletion of holes can be observed in Figure 5 — a contrast to the situation represented in Figure 6. The additional space charge of the ionization-generated holes and the voltage drop due to the substrate current across the parasitic substrate resistor force more electrons to be injected from the source. This additional electron injection is not only restricted to the surface but also extends into the substrate. This is corroborated by Figure 7 showing electron distribution. A deep electron channel similar to a punch-through channel [16, 19, 38] can be seen in this plot. Both the punch-through channel and the avalanche-induced channel are caused by the source-to-substrate junction. To show that the deep electron channel in Figure 7 is basically caused by avalanche generation rather than by punch-through, we present the electron distribution without impact ionization in Figure 8. Indeed, a deep electron channel is not observed in this figure.

As mentioned above, the parasitic substrate resistor is in part responsible for the barrier lowering. The value of this resistor may be approximated, as shown in [29], by the expression

$$R_{sub} = \frac{q}{2(W - L)} \left( \ln \left( \frac{L + 2d}{L + 2d_{dep}} \right) - \ln \left( \frac{W + 2d}{W + 2d_{dep}} \right) \right)$$
Figure 3. Potential distribution of an N-channel MOSFET \( L_{\text{eff}} = 2 \, \mu\text{m}, V_{DS} = 8 \, \text{V}, V_{GS} = 2 \, \text{V}, V_{BS} = -1 \, \text{V} \).

Figure 4. Substrate current (log plot) versus \( V_{DS} \). \( L_{\text{eff}} = 2 \, \mu\text{m}, W = 100 \, \mu\text{m} \).
**Figure 5.** Hole density distribution (log plot), same conditions as in Figure 3.

**Figure 6.** Same as Figure 5, impact ionization neglected.
**Figure 7.** Electron density distribution (log plot), same conditions as in Figure 3.

**Figure 8.** Same as Figure 7, impact ionization neglected.
The symbols $L$ and $W$ represent the channel length and width, respectively, $\rho$ is the resistivity of the substrate, and $d$ and $d_{dep}$ are the thickness of the substrate and depletion width (extension of the hole depletion into the substrate, Figure 5). Typical values for $R_{sub}$ are 1 k$\Omega$ to 5 k$\Omega$. The voltage drop $R_{sub} \cdot I_{sub}$ increases the internal substrate potential and lowers the source to substrate barrier. If this voltage drop is large enough the parasitic bipolar transistor (source-substrate-drain) may be turned on [20]. This results in a rapid rise of the drain current with increasing drain voltage as demonstrated in Figure 9. Once the parasitic bipolar has been turned on, its current gain strongly influences the breakdown voltage and snap-back may be caused under certain conditions [28,30].

![Graph](image)

**Figure 9.** Drain current versus $V_{DS}$. $L_{eff} = 2 \mu m$, $W_{eff} = 100 \mu m$.

The snap-back branch of the $I-V$ curve is characterized by a negative differential resistance. Due to this effect numerical calculation is very difficult in the snap-back region. Nevertheless, the carrier densities are only weakly affected by the snap-back effect provided that the operating point is stable. To show carrier densities similar to a snap-back operating point Figures 10 and 11 are presented for a transistor with $L_{eff} = 1 \mu m$ operating near its breakdown voltage (although without snap-back). Drain and substrate current are very large and, consequently, the carrier densities have been increased considerably compared to Figures 6 and 8. In this situation the carrier density product greatly exceeds $n_i^2$ and the Shockley-Read-Hall recombination, (5a), is very prominent. The current obtained by spatially integrating the recombination may be considered as base current of the parasitic bipolar transistor.
Figure 10. Electron density distribution (log plot) for $L_{\text{eff}} = 1 \ \mu\text{m}, V_{DS} = 7 \ \text{V}, V_{GS} = 2 \ \text{V}, V_{RS} = -1 \ \text{V}$.

Figure 11. Hole density distribution (log plot) for $L_{\text{eff}} = 1 \ \mu\text{m}, V_{DS} = 7 \ \text{V}, V_{GS} = 2 \ \text{V}, V_{RS} = -1 \ \text{V}$.
6. CONCLUSIONS

Two-dimensional simulation of MOS transistors using the model which includes a complete description of impact ionization as well as all relevant recombination process provides important insight into device operation. Examples of the internal distributions of selected physical quantities are given. The computed currents are compared with the experimental data. The agreement in the case of bulk current, which is exclusively due to the impact ionization, is especially noteworthy. This correspondence confirms the validity of the model used in the simulation. The role of hole space charge enhancing the avalanche breakdown was displayed and discussed, thus demonstrating the model's usefulness.

ACKNOWLEDGMENTS

This work was supported by the “Fonds zur Förderung der wissenschaftlichen Forschung” under project S 22/11. The authors would like to thank the “Interuniversitäres Rechenzentrum der T.U. Wien” for providing a generous amount of computer time. Essential help in providing MOS-devices by Siemens AG, Munich is very appreciated.

NOTATION

\[ \varepsilon \]
permittivity constant

\[ = 11.7 \cdot 8.8541878 \cdot 10^{-12} \text{ AsV}^{-1}\text{m}^{-1} \text{ for silicon} \]

\[ = 3.9 \cdot 8.8541878 \cdot 10^{-12} \text{ AsV}^{-1}\text{m}^{-1} \text{ for silicon dioxide} \]

\[ \Psi \]
electrostatic potential

\[ q \]
elementary charge \( = 1.6021892 \cdot 10^{-19} \text{ As} \)

\[ n \]
electron concentration

\[ p \]
hole concentration

\[ N_D^+ \]
ionized donor concentration

\[ N_A^- \]
ionized acceptor concentration

\[ J_n \]
electron current density

\[ J_p \]
hole current density

\[ \mu_n \]
mobility of electrons

\[ \mu_p \]
mobility of holes

\[ D_n \]
diffusivity of electrons

\[ D_p \]
diffusivity of holes

\[ G - R \]
et generation recombination

\[ n_i \]
intrinsic number

\[ \tau_n \]
lifetime of electrons

\[ \tau_p \]
lifetime of holes

\[ s_n \]
surface recombination velocity of electrons

\[ s_p \]
surface recombination velocity of holes

\[ n_1 \]
electron concentration imposed by the trap energy level

\[ p_1 \]
hole concentration imposed by the trap energy level

\[ C_n \]
Auger coefficient of electrons

\[ C_p \]
Auger coefficient of holes
REFERENCES


Date received: June 20, 1983; Date revised: October 7, 1983.