

Non Local Impact Ionization Effects in Semiconductor Devices

DUILIO MEGLIO, CORRADO CIANCI, ALDO DI CARLO and PAOLO LUGLI*

INFN-Department of Electronic Engineering, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome, Italy

Impact ionization processes define the breakdown characteristics of semiconductor devices. An accurate description of such phenomenon, however, is limited to very sophisticated device simulators such as Monte Carlo. A new physical model for the impact ionization process is presented, which accounts for dead space effects and high energy carrier transport at a Drift Diffusion level. Such model allows to define universal impact ionization coefficients which are device-geometry independent. By using available experimental data these parameters have been calculated for $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$.

Keywords: semiconductors, devices, simulations, impact ionization, $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$

INTRODUCTION

Device simulation has become a fundamental part in the development and optimization of electronic devices. Next to standard approaches such as the Drift Diffusion method [1], which has lead to several commercial software packages, newer techniques are also available, ranging from hydrodynamic approaches to particle-based simulations, such as the Monte Carlo or the Cellular Automaton methods [2, 3].

In general, one pays the enhanced physical content of the latter approaches with a heavier computational load. Thus is particularly true in the presence of high electric fields. While physical simulators require an accurate description of high energy effects, Drift Diffusion methods can rely on simplified phenomenological approaches.

The aim of this paper is to show the inadequacy of some of these approaches, by focusing on the process

which is the main responsible for the breakdown of devices, namely impact ionization (II). Furthermore, we will indicate improvements which allow us to keep the speed advantages of Drift Diffusion methods while incorporating the main physical features of the II process.

IMPACT IONIZATION IN DRIFT DIFFUSION AND MONTE CARLO SIMULATORS

II is the phenomenon by which a hot carrier colliding with a valence electron creates a new pair of carriers, both available for conduction. A careful model of II must account for the high energy behaviour of carriers, since the threshold for the process is at least equal to the energy gap of the semiconductor under consideration. The best treatment in terms of physical accuracy is provided by the Monte Carlo method, which

* Corresponding author. Tel: +39 6 72594469. Fax: +39 6 2020519. E-mail: lugli@roma2.infn.it

typically includes II by associating to each carrier ionization probability per unit time, dependent on the carrier status (energy, velocity, momentum etc.). Several degrees of depth are possible, based on simplified bands and phenomenological II rates, or on complete bands and microscopical II rates. The latter approach has been applied mainly to bulk semiconductors, due to its numerical complexity, while the former one can be applied to devices. Recently, a Monte Carlo analysis has been presented of an AlGaAs/GaAs Heterojunction Bipolar Transistor (HBT) [4,5], which perfectly reproduced available experimental results both on the multiplication factor as well as on the electroluminescence spectra of the device in a near-breakdown regime [6,7]. There, a three valley non-parabolic model was implemented for electrons and holes, together with the Kane model for II. The main conclusions of that work have been :

1. The average energy and the ionization coefficients reach their maximum not at the base-collector junction, where the electric field reaches its maximum, but rather inside the collector. Such effect is referred to as "dead space" effect [8-11].
2. Electrons and holes gain considerable energy in the collector due to the presence of very high electric fields. As a result, the electron and hole distribution functions are very hot, leading to strong ionization processes and to radiative transitions within the conduction and the valence band, respectively [6], responsible for the observed electroluminescence.

Although the Monte Carlo simulation is physically very accurate, it is also extremely time consuming. It would therefore be preferable, in the context of device modeling (and especially in the presence of ionization phenomena), to use faster numerical tools such as those based on the Drift Diffusion algorithm. There, the following equations are used :

$$\frac{d}{dx} \left(\epsilon \frac{dV}{dx} \right) = -\rho \quad (1)$$

$$J_p = qp\mu_p F - qD_p \frac{dp}{dx}; \quad J_n = qn\mu_n F + qD_n \frac{dn}{dx} \quad (2)$$

$$\frac{dJ_p}{dx} = q(G - R) = -\frac{dJ_n}{dx} \quad (3)$$

In Eqs. (3), II is accounted for by means of the generation term, G . Unfortunately such approaches do not correctly describe hot carrier and non local effects, unless ad hoc phenomenological corrections are implemented. In the following, we will briefly revise the widely accepted way to deal with II in conjunction with Eqs. (1-3).

LOCAL MODEL: DESCRIPTION AND FAILURE

The most common method to deal with II phenomena in Drift Diffusion simulations, is represented by the Local Model (LM) [12-15]. Within such model the probability to generate a II event in $(x, x + dx)$ for a carrier moving in the $+x$ direction depends only on the local electric field. Under this hypothesis, II can be fully characterized by the mean free path between ionizing collisions, for both electrons, $\langle l_n \rangle$, and holes, $\langle l_p \rangle$.

The II generation rate is linked to the $\langle l_i \rangle$ parameters and to the currents by the equations

$$qG = \alpha_L J_n + \beta_L J_p = \frac{dJ_p}{dx} = -\frac{dJ_n}{dx} \quad (4)$$

$$\alpha_L = \frac{1}{\langle l_n \rangle}; \quad \beta_L = \frac{1}{\langle l_p \rangle}. \quad (5)$$

Here α_L and β_L represent the ionization coefficients of the LM for electrons and holes respectively. Equations (4) are usually introduced referring to a constant electric field, but can be extended to a generic field shape by assuming the II coefficients to be function of the local electric field.

Within the LM, the II process in a given semiconductor under a specified electric field is fully characterized by a pair of real numbers, namely the ionization coefficients.

Equations (4-5) can be used to link the II coefficients (i.e. the inverse of the carriers mean free path between ionizing collision) to macroscopically observable quantities, such as the multiplication fac-

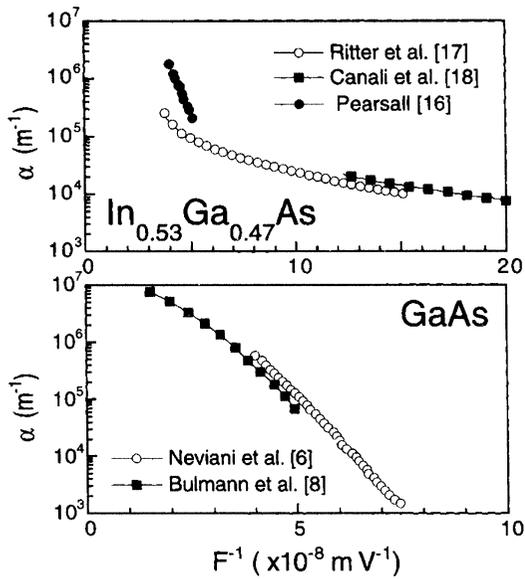


FIGURE 1 Dispersion of the LM ionization coefficients measured by different authors for electrons on (a) $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ and (b) GaAs

tor or the breakdown length [12-14]. Such equations can, in turn, be applied to extract the II coefficients as a function of the electric field from experimental data [12, 13]. As a matter of example, we report in Fig. 1a a comparison of the InGaAs electron ionization coefficients obtained by application of the LM from different authors [16-18]. The same quantity is plotted in Fig. 1b for GaAs [6, 8]. The observation of Fig. 1 reveals that the ionization coefficients are dependent on the particular device under examination. Particularly, their inverse cannot be taken as a reliable estimate of the mean free path between ionizing collisions.

MONTE CARLO ANALYSIS: THE DEAD SPACE CONCEPT

The dispersion in the experimental determinations of the II coefficients shown in Fig. 1 arises from the inadequacy of the LM. In order to clarify this point we refer to a Monte Carlo simulated experiment, by considering an electron current entering a region where a strong electric field ($F=5 \cdot 10^7 \text{ V/m}$) is applied. For simplicity, only electron ionizations are allowed

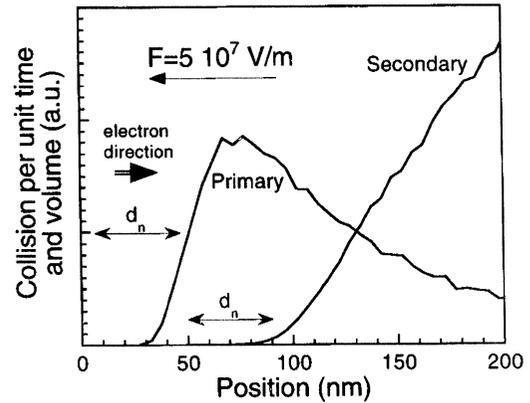


FIGURE 2 Number of ionizing collisions for unit time predicted with Monte Carlo simulation both for primary and secondary carriers. The quasi-exponential decay of the number of ionizing collisions due to the primary allows to define dead spaces and DM ionization coefficients

in the simulation. In Fig. 2, the number of ionizing collisions (per unit time and volume) is plotted as a function of position. Two types of processes are considered: the first ('primary') collisions of the carriers entering the high field region, and all other ('secondary') collisions, either of the primary carriers or of those that are generated within the field region.

Figure 2 clearly illustrates a fundamental physical feature of II process, that is the "dead-space" effect: a carrier must travel a certain distance before reaching the threshold energy for II. The "cold" injected primary electrons are not immediately available for II. Thus, a dead-space zone (d_n) where no ionizing collisions occur is present near the contact. Equivalently, the dead-space effect can be described by the associated energy ($E_{th,n}$) gained from the carriers over the d_n length. For a constant electric field, threshold energy and dead-space are related by the equations

$$E_{th,p} = q \cdot |F| \cdot d_p; \quad E_{th,n} = q \cdot |F| \cdot d_n \quad (6)$$

for holes and electrons respectively. It is often more convenient to speak of threshold energy because such quantity is generally less sensitive to the field value than d_i , and often, as a first approximation, it is considered to be constant for a given semiconductor. $E_{th,i}$

represents the energy that a carrier must receive from the field to appreciably initiate II.

Secondary carriers come into existence only after the primary ones collide, that is not before $x = d_n$. As it can be seen on the figure, they also cannot ionize before another dead-space, that is around $x = 2d_n$.

The reason of such behaviour can be understood in the light of the microscopy of the ionizing event. When the primary carrier impinges on the valence electron, it release its energy almost completely to generate a hole-electron pair roughly in rest condition. Accordingly, all secondary carriers can be assumed to start their motion close to zero kinetic energy. This behaviour, totally neglected in the LM, must be taken into account to adequately reproduce the II process. The mean free path between ionizing collision is only a first order description of the process. The dead space concept gives further information and, if considered in a suitable form, can add in accuracy to an adopted model, as we will see shortly.

DELAY MODEL: DESCRIPTION AND IMPLEMENTATION

We will refer here to a region with a step-constant, positive electric field, F . We further assume, for the sake of simplicity, that carriers are approximately in the same rest condition after suffering an ionizing collision, or after being generated by an ionizing collision.

Under these hypothesis a very general procedure for II modeling can start by considering the spatial probability density of the ionization event, $P_n(x)$ ($P_p(x)$), for a rest electron (hole) entering the field region at $x = 0$ [10]. Clearly, $P_n(x) = 0$, $x > 0$ and $P_p(x) = 0$, $x < 0$. Field edge inclusion is easily accomplished with the further assumption of zero II probability soon after the field falls to 0. Under the previous assumption Eqs. (5) must be substituted with the more general

$$\langle l_n \rangle = \overline{|P_n(x)|}; \quad \langle l_p \rangle = \overline{|P_p(x)|} \quad (7)$$

where overbar indicates the expected value calculation. Different choices for the $P_i(x)$ functions would

lead to different degrees of accuracy of the model and for the associated computational effort. If the $P_i(x)$ functions are taken as simple exponentials of the form

$$\begin{aligned} P_n(x) &= \begin{cases} \alpha_L e^{+\alpha_L x}, & x \leq 0 \\ 0, & x > 0 \end{cases} \\ P_p(x) &= \begin{cases} \beta_L e^{-\beta_L x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \end{aligned} \quad (8)$$

the LM is recovered.

As we have previously shown, most of the limitations implicit in the LM arise from the fact that the dead space effect is neglected. The simplest assumption that allows us to overcome such difficulty is to consider the P_i functions to be delayed exponential functions of the type

$$\begin{aligned} P_n(x) &= \begin{cases} \alpha_D e^{+\alpha_D (x+d_n)}, & x \leq -d_n \\ 0, & x > -d_n \end{cases} \\ P_p(x) &= \begin{cases} \beta_D e^{-\beta_D (x-d_p)}, & x \geq d_p \\ 0, & x < d_p \end{cases} \end{aligned} \quad (9)$$

that is, to assume that the probability of a hole (electron) initiated II event occurring between x and $x + dx$ ($x - dx$) is 0 until the carrier travels a dead space length, and $\beta_D dx$ ($\alpha_D dx$) afterwards [10]. The reason for such an assumption can be supported by the observation that the shape of the curve referred to as "primary" in Fig. 2 is proportional to $P_n(-x)$. Indeed, it is very close to the delayed form of Eqs. (9), and α_D and d_n can be extracted from a fitting procedure. Similar arguments hold for holes.

It is possible to show that the previous assumptions, which define the Delay Model (DM) of II, lead to the following expressions for the generation rate [9] (see also [19] where carriers are used instead of currents)

$$\frac{dJ_p}{dx} = qG = \alpha_D J_n^H + \beta_D J_p^H = -\frac{dJ_n}{dx} \quad (10)$$

$$J_n^H = J_n(x+d_n) - \alpha_D \int_x^{x+d_n} J_n^H dx' \quad (11)$$

$$J_p^H = J_p(x-d_p) - \beta_D \int_{x-d_p}^x J_p^H dx'. \quad (12)$$

Equations (10-12) are valid if II is the only relevant generation-recombination mechanism. The appropriate boundary conditions are

$$\begin{aligned} J_p^H(x) &= 0, & x < d_p \\ J_n^H(x) &= 0, & x > W - d_n \\ J_p(0) &= J_{p0} \\ J_n(W) &= J_{nW} \end{aligned} \quad (13)$$

The DM is actually equivalent to the model developed in [10, 20] to study the mean gain of avalanche photodiodes. Equations (10-12) can be easily inserted in Drift Diffusion simulators, whose solution can still be accomplished via an iterative scheme, as proposed by Gummel [21].

Within the DM, the II characteristic of a given semiconductor under a specified electric field are expressed by two pairs of parameters: the ionization coefficients and the threshold energies for electrons and holes, respectively. Those parameters are linked to the mean free path between ionizing through Eqs. (7) which takes the form

$$\langle l_p \rangle = \frac{1}{\beta_D} + d_p; \quad \langle l_n \rangle = \frac{1}{\alpha_D} + d_n. \quad (14)$$

The d_i terms are usually negligible in the lower field range, but their contribution increases with field strength. Similar results were reported for the first time in [19] by Y. Okuto and C. R. Crowell, but up to now they have not been received considerable attention.

It should be noticed that Eqs. (10) - (13) have been introduced referring to constant fields, but they can be easily extended by assuming both threshold energies and II coefficients to be dependent on the local electric field.

PARAMETERS EXTRACTION FOR DELAY MODEL

A practical way to obtain the DM parameters is through general techniques of error minimization. Let us suppose that N different experimental values are available, such as the multiplication factors (M_i) for

electrons and holes, for a certain number of devices under a specified bias condition. First, the field shape, $F_i(x)$ corresponding to each value has to be determined. Then, a sample set of the unknown functions ($\alpha_D(F)$, $\beta_D(F)$, $E_{th,n}(F)$, $E_{th,p}(F)$) is chosen, in correspondence to an arbitrary number (P) of field values, F_h ($h \in (1, P)$), and a guess is made about the range of variation of the $4 \cdot P$ unknown samples $\alpha_D(F_h)$, $\beta_D(F_h)$, $E_{th,n}(F_h)$, $E_{th,p}(F_h)$.

The extraction procedure simply consists of two further steps : in the first we randomly choose the value of the $4 \cdot P$ unknown samples in their relative range of variation. For each choice we simulate all the N structures using a suitable function to interpolate the samples, and we calculate an error function which could be just the simple mean root square of the differences between experimental and calculated $\log(M_i)$ values. By tracking the sample set which gave the minimum value, we reduce the searching range and update its bound to maintain the central value over the best samples set. When the error function decreases under a prescribed quantity, the last step starts, consisting of a simple gradient search for the minimum value of the error function until the zero value is reached.

A series of improvements can be thought of in order to enhance the method convergence and accuracy, depending on the particular device that provided the experimental data. The fundamental point is that, thanks to this method, we can successfully extract a set of highly reproducible parameters that can be used to describe II process in a given semiconductor. If we restrict our attention to structures with low overall multiplication factor ($M - 1 \ll 1$), many simplifications can be done on the full DM, that allow a faster parameter extraction.

We have applied the described algorithm to experimental data measured on npn InP/In_{0.53}Ga_{0.47}As HBTs [17,18]. In such transistors, the electron current injected from the base grows in the collector region because of II processes, the strength of multiplication depending on collector bias. The electron multiplication factor, M_n , is defined as the ratio between the col-

lector current and the electron current injected from the base into the collector

$$M_n = \frac{J_{n,C}}{J_{n,B}}$$

With a suited procedure [18,22] the M_n dependence on collector polarization for a given device can be extracted with great sensitivity. Both structures considered in [17, 18] are operating in the range of collector biases that satisfy the relation $M_n - 1 \ll 1$.

For sake of simplicity, we have made the widely accepted assumption to consider $E_{th,n}$ independent on the local electric field (at least for the range of fields considered). The procedure to extract an estimate of the $\alpha_D(F)$ and $d_n(F)$ functions proceeds by applying the algorithm described above. We calculated the field shapes in the collector region for the device reported in [18] by means of a Drift Diffusion simulator, using doping data obtained with C-V measurements [22]. For the structure reported in [17] we referred to the nominal doping value, reported in the original work.

In Fig. 3 we plot the universal electron impact ionization coefficient for $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ extracted from our algorithm at the threshold energy of 1.25 eV in the considered field range. It is worth reminding that, if the LM was to be used (as actually done in [17,18], see Fig. 1), the ionization coefficients for the two

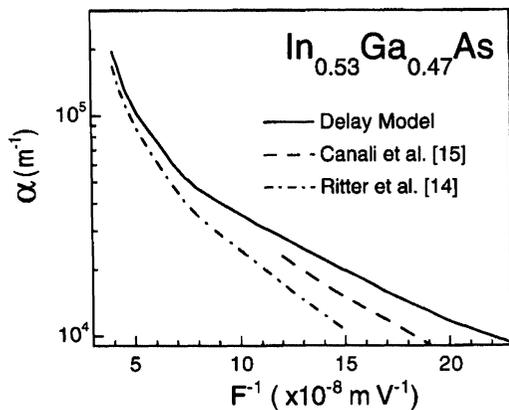


FIGURE 3 Universal DM ionization coefficient for electrons on $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ compared with previous LM determinations. The calculated values, used in DM Drift Diffusion simulator, allow a perfect calculation of the multiplication factors in both HBTs

structures would differ from one another. The calculated values on Fig. 3, used in DM-based Drift Diffusion simulator, allow a perfect calculation of the multiplication factors in both HBTs.

In conclusion, we have shown that it is possible to extend Drift Diffusion algorithms to include non-local effects of impact ionization, thus recovering the accurate physical description provided by Monte Carlo simulations.

The work was partially supported by Piano Nazionale "Materiali Innovativi e Avanzati" of the Italian Ministry for University and Research

References

- [1] S. Selberherr, "Analysis and Simulation of Semiconductor Devices", Springer Verlag, Wien (1984)
- [2] C. Jacoboni, P. Lugli, "The Monte Carlo Method for Semiconductor Device Simulation", Springer Verlag, Wien (1989)
- [3] K. Kometer, G. Zandler and P. Vogl, Physical Review B, **46**, 1382, (1992); G. Zandler, A. Di Carlo, K. Kometer, P. Lugli, P. Vogl and E. Gornik, IEEE Electron Device Letters, **14**, 77, (1993)
- [4] A. Di Carlo and P. Lugli, IEEE Electron Device Letters, **14**, 103, (1993)
- [5] C. Canali, P. Pavan, A. Di Carlo, P. Lugli, M. Manfredi, A. Neviani, L. Vendrame, E. Zanoni and G. Zandler, IEEE Transaction on Electron Devices, to be published (1995)
- [6] A. Neviani, C. Tedesco, E. Zanoni, M. Manfredi, C. Canali, A. Di Carlo and P. Lugli, Alta Frequenza, **5**, 179, (1993); C. Canali, F. Capasso, R.J. Malik, A. Neviani, P. Pavan, C. Tedesco, E. Zanoni, IEEE Electron Device Letters, **9**, 354, (1994)
- [7] G. Zandler, A. Di Carlo, P. Vogl and P. Lugli, Semiconductor Science and Technology, **9**, 666, (1994)
- [8] G.E. Bulman, V.M. Robbins and G.E. Stillman, IEEE Transaction on Electron Devices, **11**, 2454, (1985)
- [9] M. Claassen, H. Grothe, R. Pierzina and W. Harth, AEÜ, **41**, 380, (1987)
- [10] M.M. Hayat, B.E.A. Saleh, M.C. Teich, IEEE Transaction on Electron Devices, **39**, 546, (1992)
- [11] J.S. Marshland, R.C. Woods and C.A. Brownhill, IEEE Transaction on Electron Devices, **39**, 1129, (1992)
- [12] S.L. Miller, Physical Review, **99**, 1234, (1955).
- [13] A.G. Chynoweth, Physical Review, **109**, 1537, (1958).
- [14] C.A. Lee, R.A. Logan, R.L. Batdorf, J.J. Kleimack and W. Wiegmann, Physical Review, **134**, A761, (1964).
- [15] S.M. Sze, "Physics of Semiconductor Devices" - 2nd Edition, John Wiley & Sons, NY (1981), p. 45
- [16] T.P. Pearsall, Applied Physics Letters, **36**, 218, (1980).
- [17] D. Ritter, R.A. Hamm, A. Feyngenson and M.B. Panish, Applied Physics Letters, **60**, 3150, (1992).
- [18] Canali, C. Forzan, A. Neviani, L. Vendrame, E. Zanoni, R.A. Hamm, R.J. Malik, F. Capasso, S. Chandrasekhar, Applied Physics Letters, **66**, 1095, (1995).
- [19] Y. Okuto and C.R. Crowell, Physical Review B, **10**, 4284, (1974)
- [20] M.M. Hayat, W.L. Sargeant and B.E.A. Saleh, IEEE Journal of Quantum Electronics, **28**, 1360, (1992)

- [21] H.K. Gummel, IEEE Transaction on Electron Devices, **11**, 455, (1964)
[22] C. Canali, Private communication

Biographies

Duilio Meglio was born in Rome in 1966. He received the “Laurea degree” *cum laude* in Electrical Engineering from the University of Rome “La Sapienza”, Italy, in 1993.

He is currently enrolled in the doctorate program of Microelectronics and Telecommunications at the University of Rome “Tor Vergata”, Italy. E-mail: meglio@venere.eln.utovrm.it

Corrado Cianci was born in Rome in 1963. He received the “Laurea degree” in Physics in 1990 and the Ph.D. degree in Microelectronics in 1995 at the University of Rome “Tor Vergata”, Italy, where he is currently working. He is involved in the numerical simulation of heterojunction electron devices based both on GaAs and InP technology.

E-mail: cianci@roma2.infn.it

Aldo Di Carlo graduated in Physics at the University of Rome “La Sapienza”, Italy, in 1991. He obtained the Ph.D. degree in Physics at the Walter Schottky Institute of the Technical University of Munich (Germany) in 1995. The research work of Dr. Di Carlo focuses on numerical simulation of semiconductor devices and optical properties of III-V compounds. E-mail: dicarlo@venere.eln.utovrm.it

Paolo Lugli graduated in Physics at the University of Modena, Italy, in 1979. In 1981 he joined Colorado State University, Fort Collins, CO, where he received the M.S. degree in 1982 and a Ph.D. degree in 1985, both in electrical engineering. Since 1993 is Full Professor of optoelectronics at the University of Rome “Tor Vergata”. His current research involves the numerical simulation of semiconductor devices and the theoretical study of ultrafast phenomena in polar semiconductor.

Prof. Lugli is the author of more than a hundred scientific publications and co-author of the book “The Monte Carlo Modeling for Semiconductor Devices Simulation”. E-mail: lugli@roma2.infn.it



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

