3 Numerical Modelling and Simulation

3.1 INTRODUCTION

Modern silicon technologies, based on highly complex systems, require rigorous methods to predict the behaviour of electronic devices and the computational approach becomes essential. In the past decades, the applied mathematics branch called numerical computation has closely followed the evolution of information technologies, refining its techniques according to the complexity of the systems to be described and benefiting from the upgrade in the computational hardware performances. In this view, one of the most significant achievements is represented by the introduction of the Finite Element (FE) analysis, based on a particular solving strategy for partial differential equations (PDE), which has been developed from the 30s of the last century.

FE analysis is a mathematical tool able to offer a considerable saving of computational costs and the advantage of setting a proper approximation level with respect to the exact solution. The user can choose the optimal balance between the precision of the result and the time required to process data.

The important ingredients of any physical modelling of semiconductor devices are presented in the following sections. Before focusing our attention on the Drift-Diffusion (DD) framework, one of the most commonly used techniques both in the industry and in the R&D field, the microscopic description of charge carriers transport, as well as its numerical treatment, will be provided. To this aim, how these models are discretized before being implemented in a calculator will be briefly described. Finally, the case study of a UFSD-based detector is presented at the end of the chapter to show how theory applies to real life. In this section, standard carrier statistics and transport is coupled with advanced semiconductor physics (quantum models and radiation effects) to highlight the properties and issues of simulating the electrical behaviour and the operating performances of a real device.

3.2 PHYSICAL MODELLING OF SEMICONDUCTOR DEVICES

As anticipated in the introduction, the ingredients of physics-based modelling are now introduced. The first ingredient is the relationship between the electric field and the charge density, described by the Poisson's equation. The second ingredient is how the charge carriers, i.e., electrons and holes, *react* to the applied field or, in other words, the carrier dynamics. This step needs a dedicated transport model (TM). Here, two TMs models are detailed: the Boltzmann Transport Equation (BTE) and the Drift-Diffusion (DD).

3.2.1 ELECTROMAGNETIC MODEL: THE POISSON'S EQUATION

The *Poisson's equation* connects the electric field acting within a semiconductor material $\mathscr{E}$ and the overall charge density ρ (the sum of the densities of positive, negative and fixed charges):

$$\frac{\mathrm{d}^2 U_x}{\mathrm{d}x^2} = q\frac{\rho}{\varepsilon}, \tag{3.1}$$

where U_x is the potential energy along an hypothetical direction x and ε is the dielectric constant of the semiconductor. To make Eq. (3.1) explicit with respect to the electric field, it is sufficient to remember that the potential energy and the electrostatic potential φ are linked by

$$U_x = -q\varphi_x \tag{3.2}$$

and that

$$\mathscr{E}_x = -\frac{\mathrm{d}\varphi_x}{\mathrm{d}x}, \tag{3.3}$$

so it can be written

$$\frac{\mathrm{d}\mathscr{E}_x}{\mathrm{d}x} = \frac{\rho}{\varepsilon}, \tag{3.4}$$

which is the most common way to represent the Poisson's equation for semiconductors.

Depending on the system to be modelled, Eq. (3.4) has to be rewritten as a function of two or three spatial variables, transforming the derivative into a divergence and the density ρ into a distribution of charges $\rho(\mathbf{r})$ in the generic space vector $\mathbf{r}$. In the following section of this chapter, it will be shown how to write Eq. (3.1) in a more general form and, especially, how to solve it within a FE framework.

3.2.2 TRANSPORT MODELS

The simplest Transport Model suitable for electronic devices modelling in the semi-classical approach is the so-called *Boltzmann Transport Equation* (BTE). Written for a generic distribution function $f(\mathbf{k},\mathbf{r},t)$, which describes a population of charge carriers, the BTE is

$$\frac{\partial f}{\partial t} + \mathbf{v}(\mathbf{k})\cdot\nabla_{\mathbf{r}} f + \frac{\mathbf{F}}{\hbar}\cdot\nabla_{\mathbf{k}} f = \left.\frac{\mathrm{d}f}{\mathrm{d}t}\right|_{\mathrm{coll}}, \tag{3.5}$$

where $\mathbf{v}(\mathbf{k})$ is the fermion group velocity. In the equation above, the term

$$\mathbf{F}\left(n, p, \varphi, \nabla, \frac{\partial}{\partial t}\right) \tag{3.6}$$

is a hypothetical force acting on the system, written as a function of the electron-hole carrier densities n and p, the electrostatic potential φ, and the gradient and derivative operators.

The last term of the BTE is called *collision term*, and describes the dynamics of the perturbation induced by the force $\mathbf{F}$. Under the *relaxation time* approximation, it can be written

$$\left.\frac{\mathrm{d}f}{\mathrm{d}t}\right|_{\text{coll}} = \frac{|f(\mathbf{k},\mathbf{r},t) - f_0(\mathbf{k})|}{\tau(\mathbf{k})}, \tag{3.7}$$

where $f_0(\mathbf{k})$ is the distribution function at equilibrium and $\tau(\mathbf{k})$ the time required to restore such equilibrium (relax) after the initial perturbation. Following the statistical theory,

$$f(\mathbf{k},\mathbf{r},t)\,\Delta\mathbf{k}\,\Delta\mathbf{r}\,\Delta t \tag{3.8}$$

represents the number of electrons/holes having momentum $\mathbf{p} = \hbar\mathbf{k}$, at position $\mathbf{r}$ and time t. Similarly, integrating Eq. (3.8) with respect to $\mathbf{r}$, the number of electrons/holes having momentum $\mathbf{p} = \hbar\mathbf{k}$ at time t is found and, finally, by integrating with respect to $\mathbf{k}$, the number of carriers at position $\mathbf{r}$ and time t is determined. Introducing now an opportune function $\lambda(\mathbf{k})$ such that

$$\lambda(\mathbf{k}) = \sum_{j=0}^{N} a_j \mathbf{k}^j, \tag{3.9}$$

with

$$a_0 = 0, \quad a_1 = \hbar\mathbf{k}, \; a_2 = \frac{\hbar^2\mathbf{k}^2}{2m^*}, \; \ldots \tag{3.10}$$

then a set of N *moments* M_j of the distribution f having the general form

$$\mathsf{M}_j = \int \lambda_j(\mathbf{k})\, f(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k} \tag{3.11}$$

are found. These moments assume a noticeable relevance since they provide information about the properties of the system. For instance, applying the 0th-order moment M_0 to the BTE written for the electrons, the following equation

$$\frac{\partial n(\mathbf{r},t)}{\partial t} + \nabla_{\mathbf{r}}\left(\langle v_n\rangle\, n(\mathbf{r},t)\right) = \left.\frac{\mathrm{d}n(\mathbf{r},t)}{\mathrm{d}t}\right|_{\text{coll}} \tag{3.12}$$

is obtained, which represents the *continuity equation* (i.e., at the same time, a charge conservation law and a transport equation for electrons), since

$$\mathsf{M}_0 = \int f(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k} = n(\mathbf{r},t). \tag{3.13}$$

Note that, in order to ensure that Eq. (3.12) is valid, it is assumed that

$$\langle v_n\rangle = \frac{\int \mathbf{v}(\mathbf{k})\, f(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k}}{\int f(\mathbf{k},\mathbf{r},t)\,\mathrm{d}\mathbf{k}} \tag{3.14}$$

is the average electron velocity or, to simplify, the electron *drift velocity* v_n. Since the electron current density can be expressed as

$$J_n = -q v_n n(\mathbf{r},t), \tag{3.15}$$

and assuming that

$$\left.\frac{\mathrm{d}n(\mathbf{r},t)}{\mathrm{d}t}\right|_{\mathrm{coll}} = -U_n(\mathbf{r},t) \,, \tag{3.16}$$

where the term U_n will be defined shortly, the BTE for electrons becomes

$$\frac{\partial n(\mathbf{r},t)}{\partial t} = \frac{1}{q}\nabla_{\mathbf{r}} J_n(\mathbf{r},t) - U_n(\mathbf{r},t) \tag{3.17}$$

and, similarly for holes

$$\frac{\partial p(\mathbf{r},t)}{\partial t} = -\frac{1}{q}\nabla_{\mathbf{r}} J_p(\mathbf{r},t) - U_p(\mathbf{r},t) \,. \tag{3.18}$$

What has been done so far is to start from the generic expression of the BTE and, introducing the method of moments, rewrite it in a form more suitable to describe the transport of free charges in semiconductors. Such a procedure can be improved by adding a couple of other considerations. First of all, recall that in solid-state physics the current density of charge carriers $J_{n,p}$ has a contribution driven by the electric field (called drift current) and a second component due to the gradient of charge density (the diffusion part). So, in one space dimension, it can be written as

$$\begin{aligned} J_n &= q\mu_n n\mathscr{E} + qD_n\frac{\partial n}{\partial x} \\ J_p &= q\mu_p p\mathscr{E} - qD_p\frac{\partial p}{\partial x} \end{aligned} \tag{3.19}$$

where $\mu_{n,p} = v_{n,p}/\mathscr{E}$ are the electron-hole mobilities and

$$D_{n,p} = \mu_{n,p} k_{\mathrm{B}} T \tag{3.20}$$

the Einstein diffusion coefficients, functions of the material-dependent mobilities $\mu_{n,p}$, the Boltzmann constant k_{B} and the absolute temperature T. Finally, the term $U_{n,p}$ is the so-called net generation-recombination (GR) rate, i.e., the net number of interband energy transitions given by the electrons relaxed into the valence band (recombination rate R_n) minus the electrons promoted into the conduction band (generation rate G_n) – or viceversa for the holes – per unit volume per second. Then, if the BTE for electrons and holes are combined with Eq. (3.4), the following system is obtained

$$\begin{aligned} \frac{\partial n}{\partial t} &= \frac{1}{q}\frac{\partial J_n}{\partial x} - (R_n - G_n) \\ \frac{\partial p}{\partial t} &= -\frac{1}{q}\frac{\partial J_p}{\partial x} - (R_p - G_p) \\ \frac{\partial \mathscr{E}}{\partial x} &= \frac{\rho}{\varepsilon} \end{aligned} \tag{3.21}$$

which represents the so-called Drift-Diffusion (DD) model (written for clarity in one space dimension). The first two equations, deriving from the standard BTE, are the electron-hole continuity equations, while the last one is the Poisson's equation.

It is interesting to highlight that in the so-called *lifetime approximation*, equivalent to the relaxation time approximation for the BTE collision term, it is possible to write

$$\begin{aligned} U_n &= R_n - G_n \\ &\approx \frac{n - n_0}{\tau_n} = \frac{n'}{\tau_n} \end{aligned} \tag{3.22}$$

and

$$\begin{aligned} U_p &= R_p - G_p \\ &\approx \frac{p - p_0}{\tau_p} = \frac{p'}{\tau_p}, \end{aligned} \tag{3.23}$$

where n_0 and p_0 are the carrier densities at equilibrium, n' and p' the excess carrier densities (out of equilibrium) and $\tau_{n,p}$ the (doping- and temperature-dependent) recombination lifetimes, that change according to the semiconductor material and to the particular GR process considered. These assumptions – and, thus, the entire DD model – are acceptable only if the system dynamics is sufficiently slower than the lifetimes $\tau_{n,p}$, as always occurs in traditional semiconductors devices. A more detailed description of the possible GR mechanisms in silicon sensors will be provided in the following sections of this chapter.

3.3 NUMERICAL TREATMENT OF MODELS

The equation of the DD model needs to be modified in order to be implemented in a computer program. The appropriate tool to solve a transport equation is the partial differential equation (PDE). The complexity of the system in Eq. (3.21), two non-linear and one linear equation in four variables, has to be reduced. The usual strategy is to discretize the PDE both in the space and time domains, and transform them into ordinary differential equations (ODE). To achieve this target, the device geometry is divided into a grid of nodes and the dynamic transitions are treated as a sequence of quasi-stationary states. These transitions can be solved with the discretized version of the DD equations, i.e., in each node of the grid, and where all physical quantities are expressed as functions of their nodal values.

At the end of this section, properties of the solving method used to calculate the discretized DD model will be shown, providing information about the possible numerical issues.

3.3.1 METHODS OF SPATIAL DISCRETIZATION

The schemes commonly used to discretize the system are essentially two: the Finite Element (FE) and Finite Differences (FD). To explore the main differences between the FE and FD schemes, the Poisson's equation is used as an example. In its generic form, the Poisson's equation is written as

$$\nabla_{\mathbf{r}}^2 \varphi = f(\mathbf{r}), \tag{3.24}$$

where $f(\mathbf{r})$ is the known term of this PDE. The potential needs to be expressed as a set of basis functions defined in each j-th element of the space such that

$$\varphi \sim \sum_j \phi_j w_j(\mathbf{r}), \tag{3.25}$$

where ϕ_j is the nodal value of the potential and $w_j(\mathbf{r})$ are opportune weighting functions. So, the Poisson's equation can be written as

$$\sum_j \phi_j \nabla^2_{\mathbf{r}} w_j(\mathbf{r}) = f(\mathbf{r}). \tag{3.26}$$

Now, Eq. (3.26) can be integrated on a given volume of space, called Ω, obtaining

$$\sum_j \phi_j \int_\Omega w_k(\mathbf{r}) \nabla^2_{\mathbf{r}} w_j(\mathbf{r}) \, \mathrm{d}\mathbf{r} = \int_\Omega w_k(\mathbf{r}) f(\mathbf{r}) \, \mathrm{d}\mathbf{r}. \tag{3.27}$$

This is a matrix equation in the form

$$\mathbf{A} \cdot \Phi = \mathbf{f}, \tag{3.28}$$

where

$$A_{jk} = \int_\Omega w_k(\mathbf{r}) \nabla^2_{\mathbf{r}} w_j(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{3.29}$$

are the elements of the sparse matrix $\mathbf{A}$ and where Φ and $\mathbf{f}$ are the column vector of, respectively, the discretized potential and the known term. Now, assuming that Ω is the triangular region defined by three nearby nodes of the grid, Eq. (3.28) becomes the discretized Poisson's equation within the FE scheme, where the unit-elements are the triangular control regions.

Similarly, the grid (with a certain criteria) can be divided into boxes with area S_j that are surrounding each j-th node. If γ_j is the path around the box, the Gauss theorem can be exploited to write Eq. (3.24) as

$$\int_{S_j} \nabla \cdot (\nabla \varphi) \, \mathrm{d}\mathbf{r} = \oint_{\gamma_j} \nabla \varphi \cdot \mathbf{n} \, \mathrm{d}\ell = \int_{S_j} f(\mathbf{r}) \, \mathrm{d}\mathbf{r}. \tag{3.30}$$

By applying the 1st-order Taylor expansion, the following expression is obtained

$$\oint_{\gamma_j} \frac{\partial \varphi}{\partial \mathbf{n}} \, \mathrm{d}\ell \simeq \sum_{\text{sides}} \ell_i \frac{\varphi_j - \varphi_i}{d_{ij}} = S_j f(r_j), \tag{3.31}$$

where ℓ_i denotes the length of the side of the box around the jth-node, located between the adjacent nodes j and i, d_{ij} is the distance between these two nodes, S_j the area of the jth-box and $\mathbf{n}$ is a unit vector normal to the box side. Notice that Eq. (3.31) is, as in the previous case, a matrix equation

$$\mathbf{A} \cdot \Phi = \mathbf{S} \cdot \mathbf{f}, \tag{3.32}$$

where $\mathbf{S}$ is a diagonal matrix. The relation in (3.32) is the discretized version of the Poisson's equation written according to the FD scheme (that, in 2D, is usually called Finite Boxes (FB) method). The core of this scheme, and the reason for its name, is in the term $\varphi_j - \varphi_i$, which takes into account the potential difference between two adjacent nodes.

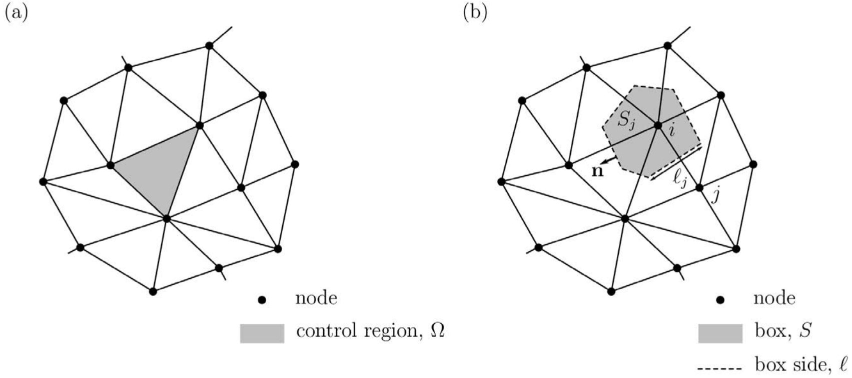

Figure 3.1 Bidimensional representation of *left*, the Finite Element (FE) and *right*, Finite Boxes (FB) discretization schemes. In the first case, the characteristic element (control region) to which all the physical quantities refer to, is the triangle described by three nearby nodes, whereas in the latter case is the box around a node.

As it was done for the Poisson's equation, the FB method can be exploited to discretize also the continuity equations. Before applying the Gauss theorem, one has to note that both the continuity equations included in the DD model have the form

$$\frac{\partial f}{\partial t} + \nabla \cdot \mathbf{F} = s\,, \tag{3.33}$$

where f and $\mathbf{F}$ are, respectively, a scalar and a vector unknown and where s is a (scalar) source term.

The domain of the box, having area S and perimeter Γ, can now be integrated over. The Gauss theorem applied to Eq. (3.33) yields to

$$\frac{\partial}{\partial t}\int_{S_j} f\,\mathrm{d}S + \oint_{\Gamma_j} \mathbf{F}\cdot\mathbf{n}\,\mathrm{d}\Gamma = \int_{S_j} s\,\mathrm{d}S\,, \tag{3.34}$$

that becomes

$$\frac{\mathrm{d}f_j(t)}{\mathrm{d}t} S_j + \sum_{\text{sides}} \ell_i \langle \mathbf{F}\cdot\mathbf{n}\rangle_i \approx s_j S_j\,. \tag{3.35}$$

The (3.35) applied to the electron continuity equation gives

$$\frac{\mathrm{d}n_j(t)}{\mathrm{d}t} S_j - \frac{1}{q}\sum_i \ell_{ij}\langle \mathbf{J}_n \cdot \mathbf{n}\rangle_{ij} = -U_{n,j} S_j\,, \tag{3.36}$$

where i is a index to identify all the possible box sides around the node j. Most of the terms appearing in Eq. (3.36) can quite easily be managed. What may induce ambiguity is the mean value of $(\mathbf{J}_n \cdot \mathbf{n})$ carried out over all the neighboring nodes. To proceed further, the twofold nature of the current density (drift and diffusion components) needs to be considered. In this way, the second term on the left-hand side of Eq. (3.36), except for the sum, is rewritten as

$$\frac{1}{q}\langle \mathbf{J}_n \cdot \mathbf{n}\rangle_{ij} = -\langle n(r,t)\,\mu_n \nabla_{\mathbf{r}}\varphi \cdot \mathbf{n}\rangle_{ij} + \langle D_n \nabla_{\mathbf{r}} n(r,t)\cdot \mathbf{n}\rangle_{ij}\,. \tag{3.37}$$

Now two assumptions are made. (i) Between the nodes i and j the potential φ is a linear function with values, respectively, φ_i and φ_j. This means that, along that segment, the electric field $\mathscr{E}_{ij} = (\varphi_j - \varphi_i)/d_{ij}$ can be considered a constant. (ii) The current densities are constant between the two nodes (i.e., $\langle \mathbf{J}_n \cdot \mathbf{n}\rangle_{ij} = J_{ij}$ is a constant).

For this reason, Eq. (3.37) becomes

$$\frac{1}{q}\langle \mathbf{J}_n \cdot \mathbf{n}\rangle_{ij} \approx -n(r,t)\,\mu_n(\mathscr{E}_{ij})\frac{\varphi_j - \varphi_i}{d_{ij}} + D_n(\mathscr{E}_{ij})\frac{\partial n(r,t)}{\partial r} = \frac{J_{ij}}{q}, \tag{3.38}$$

and the electron continuity equation is

$$\frac{\mathrm{d}n_j(t)}{\mathrm{d}t}S_j + \sum_i \ell_{ij}\, n(r,t)\,\mu_n(\mathscr{E}_{ij})\frac{\varphi_j - \varphi_i}{d_{ij}} - \sum_i \ell_{ij} D_n(\mathscr{E}_{ij})\frac{\partial n(r,t)}{\partial r} = -U_{n,j}S_j\,, \tag{3.39}$$

with the unknowns n, φ and $\mathscr{E}$ (a dual equation also holds for holes).

Since the solution of Eq. (3.39) may generate stability issues, several approaches have been developed in the past years. One of the most robust and suitable methods (but not the only one) for the implementation in a software program is the so-called Scharfetter-Gummel (SG) solving scheme. Being the stability mainly due to the behaviour of the function $n(r,t)$ (as well as of the hole density), the SG approach makes some particular assumptions on the carrier density and its trend between the nodes that allows reducing the continuity equations to the form

$$\begin{aligned}
\frac{\mathrm{d}n_j(t)}{\mathrm{d}t} &\overset{\text{SG}}{\approx} \sum_i \frac{D_n(\mathscr{E}_{ij})}{\ell_{ij}S_j}\Big[n_j(t)\,\mathsf{B}(\Delta\varphi) - n_i(t)\,\mathsf{B}(-\Delta\varphi)\Big] - U_{n,j}\\
\frac{\mathrm{d}p_j(t)}{\mathrm{d}t} &\overset{\text{SG}}{\approx} -\sum_i \frac{D_p(\mathscr{E}_{ij})}{\ell_{ij}S_j}\Big[p_j(t)\,\mathsf{B}(\Delta\varphi) - p_i(t)\,\mathsf{B}(-\Delta\varphi)\Big] - U_{p,j}
\end{aligned} \tag{3.40}$$

where

$$\mathsf{B}(\Delta\varphi) = \frac{\Delta\varphi}{\exp(\Delta\varphi) - 1} \tag{3.41}$$

is a Bernoulli function and where $\Delta\varphi \equiv \varphi_j - \varphi_i$.

At the end of the calculations, a formalism that allows converting all the continuous physical quantities into nodal values, which depend on the discretization scheme chosen to simplify the system, has been obtained. Equation (3.40) differs from its original form since the PDEs are now a set of ODEs. As stated in the introduction of the present chapter, this represents the ideal scenario for the implementation of the transport model into a numerical solver.

3.3.2 THE ITERATIVE SOLUTION OF THE EQUATIONS

A physical model described in analytical form cannot be handled easily by a numerical solver: it needs to be rewritten in a discretized formalism. This paragraph shows this procedure for the Poisson's equation, while leaving the treatment of the whole DD model to the reader's interest. The simplest expression of the Poisson's equation is written as

$$\nabla^2 \varphi(\mathbf{r}) = -\frac{q}{\varepsilon}\left[n(\varphi(\mathbf{r})) + N_\mathrm{A}(\varphi(\mathbf{r})) - p(\varphi(\mathbf{r})) - N_\mathrm{D}(\varphi(\mathbf{r}))\right]. \tag{3.42}$$

Two assumptions are necessary: (*i*) the effective densities of donors and acceptors (N_D and N_A) are kept constant with respect to the electrostatic potential (and coordinate $\mathbf{r}$); (*ii*) the potential can be written as

$$q\,\varphi(\mathbf{r}) \equiv u(\mathbf{r}), \tag{3.43}$$

where $u(\mathbf{r})$ is a given analytical function.

It follows that, in one space dimension x, Eq. (3.42) becomes

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\varepsilon\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right) = q^2\left[N_\mathrm{D} - N_\mathrm{A} + p\left(u(x)\right) - n\left(u(x)\right)\right]. \tag{3.44}$$

Suppose now to evaluate Eq. (3.44) for the unknown $u(x)$ only in a set of equally spaced $(N+1)$ points x_i, where $i = 0,\ldots,N$. First, one has to impose a solution at the edges of the domain x_0 and x_N. These values are the boundary conditions (BC) of the problem. With the BC set, the solution is restricted to only $(N-1)$ points. Then, the Poisson's equation is evaluated in a generic node x_i through the finite difference scheme, where the interval $\left[x_{i-1/2}\,;x_{i+1/2}\right]$ was used as a control region to perform the calculations. From what has been said, in x_i it holds that

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\mathrm{d}}{\mathrm{d}x}\left(\varepsilon\frac{\mathrm{d}u(x))}{\mathrm{d}x}\right)\mathrm{d}x = q^2 \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left[N_\mathrm{D} - N_\mathrm{A} + p\left(u(x)\right) - n\left(u(x)\right)\right]\mathrm{d}x. \tag{3.45}$$

Thanks to the properties of definite integrals it follows

$$\begin{aligned}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{\mathrm{d}}{\mathrm{d}x}\left(\varepsilon\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right)\mathrm{d}x &= \varepsilon_{x_{i+\frac{1}{2}}}\left.\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right|_{x_{i+\frac{1}{2}}} - \varepsilon_{x_{i-\frac{1}{2}}}\left.\frac{\mathrm{d}u(x)}{\mathrm{d}x}\right|_{x_{i-\frac{1}{2}}} \\ &= \varepsilon_{x_{i+\frac{1}{2}}}\frac{u(x)(x_{i+1}) - u(x_i)}{\Delta x} - \varepsilon_{x_{i-\frac{1}{2}}}\frac{u(x_i) - u(x_{i-1})}{\Delta x},\end{aligned} \tag{3.46}$$

where, in the second line, the definition of difference quotient has been used to evaluate the two derivatives. Assuming now that N is sufficiently high to have a dense set of nodes, then $u(x)$ can be approximated by a linear function around the point x_i and, more in general, within the whole control volume.

This hypothesis leads to

$$q^2 \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} [N_\mathrm{D} - N_\mathrm{A} + p - n] \approx q^2 [N_\mathrm{D} - N_\mathrm{A} + p - n]_{x_i} \Delta x , \tag{3.47}$$

where Δx is defined as the spacing between nodes (which is constant over the entire domain x). Combining equations (3.46) and (3.47), a new discretized version of the Poisson's equation is obtained:

$$\begin{aligned} &\varepsilon_{x_{i+\frac{1}{2}}} u(x_{i+1}) - \left(\varepsilon_{x_{i+\frac{1}{2}}} + \varepsilon_{x_{i-\frac{1}{2}}} u(x_i) \right) + \varepsilon_{x_{i-\frac{1}{2}}} u(x_{i-1}) \\ &\quad - (\Delta x)^2 q^2 [N_\mathrm{D} - N_\mathrm{A} + p(u(x_i)) - n(u(x_i))] = 0 , \end{aligned} \tag{3.48}$$

which has three unknowns: $u(x_{i-1})$, $u(x_i)$, and $u(x_{i+1})$.

The expression obtained is non-linear with respect to the unknowns since the charge carriers n and p are, in turn, non-linear in the potential term. To solve this equation in all the nodes, a numerical strategy that overcomes such an issue needs to be applied. One of the most used formalism is the iterative Newton's method. Besides the BC, the Newton's method also requires opportune Initial conditions (IC) of the system. In this case, the *charge neutrality law* at equilibrium, consisting in

$$N_\mathrm{D} - N_\mathrm{A} + p(u(x)) - n(u(x)) = 0 , \tag{3.49}$$

can be chosen for this aim. Simplifying, for each node x_i an equation of the form

$$f_i(u_{i-1}, u_i, u_{i+1}) = 0 , \tag{3.50}$$

is found, where the notation has been relaxed such that now $u_i \equiv u(x_i)$. The goal of the Newton's method is to provide an approximate solution of Eq. (3.50) starting from the IC and through subsequent iterations k, each one having a guess solution to be achieved within a certain tolerance. The iterative method requires a maximum precision $(\Delta u)_\mathrm{max}$ as input parameter (automatic or user-defined) and assumes, for all nodes x_i, that

$$u^k = u^{k-1} + \Delta u^k , \tag{3.51}$$

with u^k the solution at the kth-iteration and Δu^k the difference between two consecutive outcomes. Here, Δu^k represents the progressive correction factor of the method towards the final solution.

By applying the scheme (3.51) to the Eq. (3.50), the following expression is obtained:

$$f_i(u_{i-1}^k, u_i^k, u_{i+1}^k) = f_i \left(u_{i-1}^{k-1} + \Delta u_{i-1}^k, u_i^{k-1} + \Delta u_i^k, u_{i+1}^{k-1} + \Delta u_{i+1}^k \right) = 0 , \tag{3.52}$$

which is, finally, a system of linear equations in u where each solution at k depends on the solution found at the $(k-1)$th iteration. At the step $k = 1$ the value u^1 is a function

of u^0, the so-called *initial guess* of the iterative scheme. Each solver estimates this term through different techniques, depending on the application field. The final form of the Poisson's equation is obtained by rewriting Eq. (3.52) as a first-order Taylor expansion:

$$f_i\left(u_{i-1}^{k-1}+\Delta u_{i-1}^{k}, u_i^{k-1}+\Delta u_i^{k}, u_{i+1}^{k-1}+\Delta u_{i+1}^{k}\right) \approx f_i(u_{i-1}^{k}, u_i^{k}, u_{i+1}^{k})+ \\ + \left.\frac{\partial f_i}{\partial u_{i-1}}\right|_{u_{i-1}^{k}} \Delta u_{i-1}^{k} + \left.\frac{\partial f_i}{\partial u_i}\right|_{u_i^{k}} \Delta u_i^{k} + \left.\frac{\partial f_i}{\partial u_{i+1}}\right|_{u_{i+1}^{k}} \Delta u_{i+1}^{k} = 0\,. \tag{3.53}$$

Equation (3.53) can be cast into matrix form:

$$\begin{pmatrix} \bullet & \bullet & & & & & 0 \\ \bullet & \bullet & \bullet & & & & \\ & \bullet & \bullet & \bullet & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \bullet & \bullet & \bullet & \\ & & & & \bullet & \bullet & \bullet \\ 0 & & & & & \bullet & \bullet \end{pmatrix} \begin{pmatrix} \\ \\ \\ \Delta \mathbf{u}^k \\ \\ \\ \\ \end{pmatrix} = - \begin{pmatrix} \\ \\ \\ \mathbf{f} \\ \\ \\ \\ \end{pmatrix} \tag{3.54}$$

which is composed by a tri-diagonal matrix and the column vectors $\Delta\mathbf{u}^k$ and $\mathbf{f}$, that are, respectively, the correction and the residual vector.

The iterative method proceeds until at least one of the following requirements is satisfied: $\left\|\Delta\mathbf{u}^k\right\| < \delta$ or $\|\mathbf{f}\| < \delta$, where $\delta \equiv (\Delta u)_{\max}$ is the tolerance of the Newton's method. For any given x_i and k, if a solution is found within a finite number of iterations, then the method converges, otherwise the procedure does not converge.

Reason determining a non-converging system are (not exhaustive): (*i*) inadequate boundary conditions, (*ii*) poor discretization scheme (extremely dense or coarse mesh nodes), (*iii*) too small tolerance, and (*iv*) low computational power or (*v*) badly conditioned problems (for instance, due to a high number of charges or a large domain to be simulated).

In order to minimize the risk that non-convergence occurs, the Russian mathematician B. N. Delaunay developed in 1934 a robust triangulation procedure that makes use of non-obtuse triangles [13]. In combination with a particular domain tessellation that identifies the finite boxes by connecting the three bisector lines of each Delaunay triangle, a discretization scheme providing an even more stable solution, particularly suitable for simulating semiconductor devices, is found.

3.4 UFSD IMPLEMENTATION AND MODELLING

The aim of this section is to present the simulation of a UFSD-based particle detector and, at the same time, to show a real application of the concepts exposed in the previous paragraphs. Given the complexity of a silicon detector, only a model based on a numerical approach provides a reliable description. In this regard, the Technology

Computer-Aided Design (TCAD) is the most used solution. The TCAD implementation of the DD model requires a set of mathematical handles in order to solve the equations through the iterative method, for example the tolerance and the maximum number of iterations. In addition, for any structure to be modelled, all the geometrical and physical properties have to be taken into account. The materials and all the topological elements necessary to correctly impose the BC (as in the case of contacts or external edges), need to be defined. This step includes the definition of doping implants and profiles, and the presence of structural defects in the lattice having relevance in the physics of carrier transport, such as energy traps or fixed charges due to some fabrication processes. TCAD also allows simulating ion implantation, thermal annealing, crystal growth or material deposition. These features can provide more realistic scenarios in the event that such technological details can make a difference. Besides the electromagnetic and transport equations just derived, it is necessary to activate in the simulations the models of all physics processes deemed essential to the operation of the device. These are the processes that can alter the current flow in the device under test and, for such reason, must be implemented into the DD model as appropriate GR terms.

So far, only the core of the numerical modelling has been described. The following step is to define the target of the simulation (for example a voltage ramp or transient phenomena). The space domain is divided into a grid and the discretized version of the DD equations is solved in each node of the grid, where all the physical quantities are expressed as functions of their nodal values. A solid procedure is to use variable node spacing: where the relevant quantities or their gradients are expected to be particularly high, a finer mesh is necessary to allow the numerical solving. In case of transients, the time domain has to be discretized too so that a sweep of a physical quantity or the simulation of a time-dependent process is treated as a sequence of quasi-stationary states. As for the space domain, the density of steps must properly follow the time scale of the process: the quasi-stationary approximation is based on the assumption that the interval between two steps must be significantly shorter than the time of the transient.

Once the simulation domains are set up, the user selects the tolerance, the maximum number of iterations as well as all the BC and IC of the problem. Then the tool introduces in each node the initial guess for the Poisson's equation at equilibrium ($t = 0$), which is the electrostatic potential φ_0'. After solving the Poisson's equation, a numerical estimation φ_0 of the potential is obtained. This solution can be used as the initial guess to solve the whole DD system in the first time step, $t = 1$. Using the iterative scheme, TCAD estimates all the DD unknowns (charge densities and potential and, in turn, also the electric field and current densities), until the simulation goal is achieved. This happens only if the algorithm converges for each time step.

3.4.1 GENERATION-RECOMBINATION MECHANISMS

The physical processes occurring in silicon sensors are mostly defined in terms of a GR rate, representing the net number of generated or recombined electrons/holes, depending on which of the two competing processes is dominating. Before

introducing the physics used to model the UFSD-based silicon detectors, the minimum settings of any simulation is briefly discussed here.

First, all the energy transitions between the valence and conduction band that are assisted by those lattice impurities (like defects of dopants) acting as energy traps have to be considered. This kind of processes goes under the name of Shockley-Read-Hall (SRH) GR mechanisms [120], from the name of the scientists who developed this formalism. Out of equilibrium, the net GR rates $U_{n,p}$ are the algebraic sum of the terms $(R_{n,p} - G_{n,p})$. The statistical nature of the emission and capture processes is defined by introducing the *capture* and *emission coefficients* for electrons and holes $c_{n,p}$ and $e_{n,p}$. Both terms are a constant property (in s^{-1}) of a given trapping process. Thus, the SRH recombination rates (in $\mathrm{cm}^{-3}\mathrm{s}^{-1}$) is defined as

$$\begin{aligned} R_n &= c_n N_t\,(1 - f(E_t))\,n \\ R_p &= c_p N_t\, f(E_t)\, p \end{aligned} \tag{3.55}$$

where N_t is the total number of SRH trap states, E_t its energy, $f(E_t)$ the probability that a trap with energy E_t is occupied, and $1 - f(E_t)$ the probability to find that trap unoccupied.

Symmetrically, the generation rates (in $\mathrm{cm}^{-3}\mathrm{s}^{-1}$) are

$$\begin{aligned} G_n &= e_n N_t\, f(E_t) \\ G_p &= e_p N_t\,(1 - f(E_t)) \end{aligned} \tag{3.56}$$

being the recombination a capture process of an electron coming from the conduction band and the generation an emission that excites an electron from the trap state to the conduction band (see Fig. 3.2). It can be demonstrated that, in steady-state conditions, the electron and hole net rates are equal, giving

$$c_n N_t \left[(1 - f(E_t))\, n - f(E_t)\, n_0\,(E_\mathrm{F}) \right] = c_p N_t \left[f(E_t)\, p - (1 - f(E_t))\; p_0\,(E_\mathrm{F}) \right] \tag{3.57}$$

where $n_0\,(E_\mathrm{F})$ and $p_0\,(E_\mathrm{F})$ are the equilibrium electron-hole concentrations when the trap has energy $E_t \equiv E_\mathrm{F}$ (with E_F the Fermi level). The solution to this equation is

$$f(E_t) = \frac{c_n\, n + c_p\, p_0\,(E_\mathrm{F})}{c_n\,(n + n_0\,(E_\mathrm{F})) + c_p\,(p + p_0\,(E_\mathrm{F}))}\,. \tag{3.58}$$

Combining equations (3.57) and (3.58), the net rates are obtained:

$$U_n = U_p = \frac{n\,p - n_i^2}{\tau_p\,(n + n_0\,(E_\mathrm{F})) + \tau_n\,(p + p_0\,(E_\mathrm{F}))}\,, \tag{3.59}$$

where the *mass action law*

$$n\,p = n_0\,(E_\mathrm{F})\; p_0\,(E_\mathrm{F}) = n_i^2 \tag{3.60}$$

was used, being n_i the intrinsic carrier concentration, and with the assumption that

$$\tau_{n,p} = \frac{1}{c_{n,p} N_t} \tag{3.61}$$

are the electron-hole lifetimes.

The formula written in Eq. (3.59) represents the most used expression of the net SRH generation-recombination rate. In order to be included in the DD problem, it must be plugged into the right-hand side of both continuity equations and then self-consistently solved with the Poisson's equation.

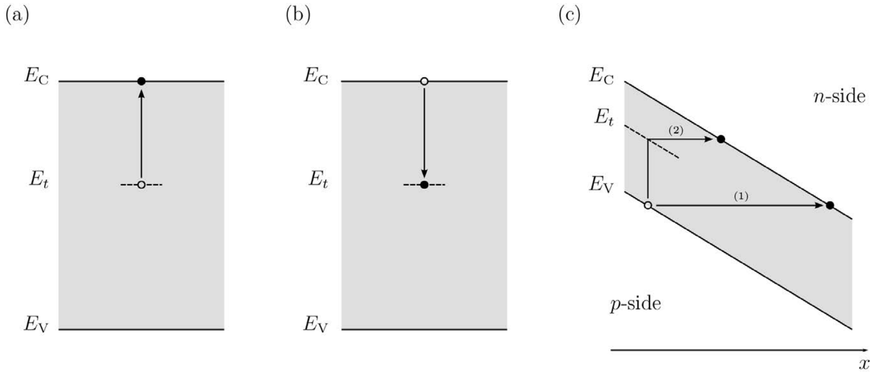

Figure 3.2 Band diagrams showing three different processes: (a) electron emission from a trap state into the conduction band (generation); (b) electron capture (recombination); (c) band-to-band (1) and trap-assisted (2) tunneling (generation) mechanisms in a reversely biased *pn* junction. Labels E_C, E_V, and E_t represent the conduction band, valence band, and trap energy, respectively.

The other important family of GR processes is represented by the tunneling mechanisms. They differ from the SRH ones by the fact that tunneling is a transition in space and not (necessarily) in energy. Again, trap-assisted tunneling processes (TAT) are found, starting or finishing with a trap capture/emission, or direct band-to-band tunneling processes (BTBT), taking place without any intermediate level. They may also occur in UFSD-based detectors, when the band bending far from the equilibrium determines a distance encompassing the conduction and valence band edges which is comparable to the wavelength of carriers, or when the presence of traps is such that this path is physically reduced, even with lower applied field. For this kind of mechanisms, net rates have the form [88, 89]

$$U_{\mathrm{TAT}} = \frac{n\,p - n_i^2}{\tau_p(\mathscr{E})\left(n + n_i\,\mathrm{e}^{\frac{E_t - E_{\mathrm{F},i}}{k_{\mathrm{B}}T}}\right) + \tau_n(\mathscr{E})\left(p + p_i\,\mathrm{e}^{\frac{E_{\mathrm{F},i} - E_t}{k_{\mathrm{B}}T}}\right)} \tag{3.62}$$

and [93]

$$U_{\mathrm{BTBT}} = A\,\mathscr{E}^2\,\mathrm{e}^{-B/\mathscr{E}}, \tag{3.63}$$

where $E_{\mathrm{F},i}$ is the energy of the intrinsic Fermi level (at the mid-gap), and the coefficients A and B are material-dependent constants. Both these two latter expressions represent theoretical models deriving from *ab initio* calculations based on,

respectively, the multiphonon emission theory and on the $\mathbf{k} \times \mathbf{p}$ approximation of bands [99].

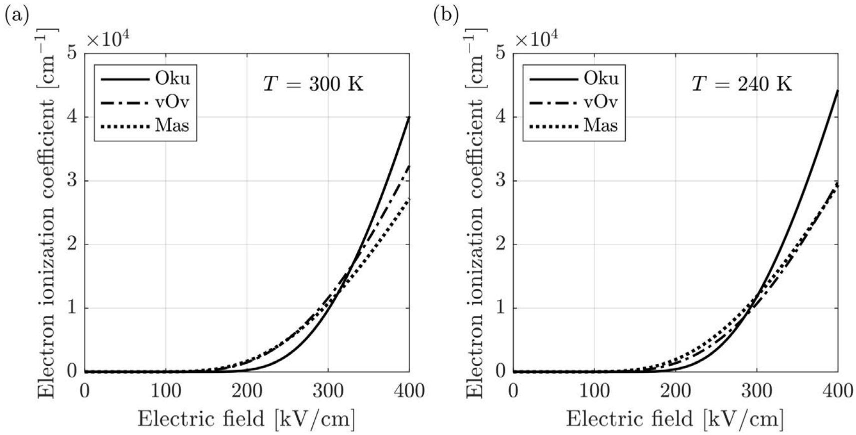

Figure 3.3 Trend of the electron ionization coefficient α_n as a function of the applied field, according to the three different avalanche models and for two temperatures: (a) at $T = 300$ K and (b) $T = 240$ K.

What characterizes the simulation of a UFSD device is the presence of the gain layer. This implies, from the numerical standpoint, that a model describing the avalanche of charge multiplication needs to be implemented. All the avalanche processes usually have a net rate of the form

$$U_{\text{aval}} = \alpha_n(\mathscr{E})\, n\, v_n + \alpha_p(\mathscr{E})\, p\, v_p\,, \tag{3.64}$$

in which $v_{n,p}$ are the carrier drift velocities and $\alpha_{n,p}(\mathscr{E})$ are the so-called electron-hole *ionization coefficients* (in cm^{-1}), corresponding to the inverse of the mean free path between two subsequent scattering events producing secondary charges. The various models developed in the past years differ in the form of the ionization coefficients. Three formalisms, that follow the Chynoweth law [7], are normally employed. The first one is the van Overstraeten-de Man model [133]:

$$\alpha_{n,p}^{\text{vOv}}(\mathscr{E}) = \gamma A_{n,p}^{\text{vOv}} \exp\left(-\gamma \frac{B_{n,p}^{\text{vOv}}}{\mathscr{E}}\right), \tag{3.65}$$

whose parameters are

$$\begin{aligned} A_n^{\text{vOv}} &= 7.030 \cdot 10^5 \text{ cm}^{-1} \\ B_n^{\text{vOv}} &= 1.231 \cdot 10^6 \text{ V/cm}\,, \end{aligned} \tag{3.66}$$

$$
\begin{aligned}
A_p^{\mathrm{vOv}} &= 1.582 \cdot 10^6\,\mathrm{cm}^{-1} \quad (\mathscr{E} \text{ below } 4 \cdot 10^5 \text{ V/cm}) \\
B_p^{\mathrm{vOv}} &= 2.036 \cdot 10^6\,\mathrm{V/cm} \quad (\mathscr{E} \text{ below } 4 \cdot 10^5 \text{ V/cm}) \\
A_p^{\mathrm{vOv}} &= 6.710 \cdot 10^5\,\mathrm{cm}^{-1} \quad (\mathscr{E} \text{ above } 4 \cdot 10^5 \text{ V/cm}) \\
B_p^{\mathrm{vOv}} &= 1.693 \cdot 10^6\,\mathrm{V/cm} \quad (\mathscr{E} \text{ above } 4 \cdot 10^5 \text{ V/cm})
\end{aligned} \tag{3.67}
$$

and

$$
\gamma = \frac{\tanh\left(\frac{\hbar\omega_{\mathrm{op}}}{2k_{\mathrm{B}}300\mathrm{K}}\right)}{\tanh\left(\frac{\hbar\omega_{\mathrm{op}}}{2k_{\mathrm{B}}T}\right)}, \tag{3.68}
$$

with $\hbar\omega_{\mathrm{op}} = 0.063$ eV the optical phonon energy in silicon.

The second formalism, proposed by Y. Okuto and C.R. Crowell [105], reads

$$
\begin{aligned}
\alpha_{n,p}^{\mathrm{Oku}}(\mathscr{E}) = & A_{n,p}^{\mathrm{Oku}}\left(1 + (T-300)\,C_{n,p}^{\mathrm{Oku}}\right)\mathscr{E} \\
& \times \exp\left[-\left(\frac{B_{n,p}^{\mathrm{Oku}}\left(1 + (T-300)\,D_{n,p}^{\mathrm{Oku}}\right)}{\mathscr{E}}\right)^2\right],
\end{aligned}
$$

where

$$
\begin{aligned}
A_n^{\mathrm{Oku}} &= 0.426\ \mathrm{V}^{-1} \\
A_p^{\mathrm{Oku}} &= 0.243\ \mathrm{V}^{-1} \\
B_n^{\mathrm{Oku}} &= 4.81 \cdot 10^5\ \mathrm{V/cm} \\
B_p^{\mathrm{Oku}} &= 6.53 \cdot 10^5\ \mathrm{V/cm}
\end{aligned} \tag{3.69}
$$

and

$$
\begin{aligned}
C_n^{\mathrm{Oku}} &= 3.05 \cdot 10^{-4}\ \mathrm{K}^{-1} \\
C_p^{\mathrm{Oku}} &= 5.35 \cdot 10^{-4}\ \mathrm{K}^{-1} \\
D_n^{\mathrm{Oku}} &= 6.86 \cdot 10^{-4}\ \mathrm{K}^{-1} \\
D_p^{\mathrm{Oku}} &= 5.67 \cdot 10^{-4}\ \mathrm{K}^{-1}.
\end{aligned} \tag{3.70}
$$

Finally, in the Massey model [100], the ionization coefficients are written as

$$
\alpha_{n,p}^{\mathrm{Mas}}(\mathscr{E}) = A_{n,p}^{\mathrm{Mas}} \exp\left(-\frac{B_{n,p}^{\mathrm{Mas}}(T)}{\mathscr{E}}\right), \tag{3.71}
$$

with parameters

$$
\begin{aligned}
A_n^{\mathrm{Mas}} &= 4.43 \cdot 10^5\ \mathrm{cm}^{-1} \\
A_p^{\mathrm{Mas}} &= 1.13 \cdot 10^6\ \mathrm{cm}^{-1}
\end{aligned} \tag{3.72}
$$

$$
\begin{aligned}
B_n^{\mathrm{Mas}}(T) &= C_n^{\mathrm{Mas}} + D_n^{\mathrm{Mas}} \cdot T \\
B_p^{\mathrm{Mas}}(T) &= C_p^{\mathrm{Mas}} + D_p^{\mathrm{Mas}} \cdot T,
\end{aligned} \tag{3.73}
$$

$$C_n^{\text{Mas}} = 9.66 \cdot 10^5 \text{ V/cm}$$
$$C_p^{\text{Mas}} = 1.71 \cdot 10^6 \text{ V/cm}, \tag{3.74}$$

and

$$D_n^{\text{Mas}} = 4.99 \cdot 10^2 \text{ V cm}^{-1} \text{ K}^{-1}$$
$$D_p^{\text{Mas}} = 1.09 \cdot 10^3 \text{ V cm}^{-1} \text{ K}^{-1}. \tag{3.75}$$

The Okuto-Crowell and Massey models have a more pronounced dependence of the ionization coefficients on temperature with respect to the van Overstraeten-de Man one, as it can be observed by comparing the two plots of Fig. 3.3.

3.4.2 RADIATION DAMAGE MODELLING

Typically, TCAD tools do not include built-in functions accounting for specific bulk radiation damage models in silicon particle detectors. To overcome this fact, *ad-hoc* models are added in the numerical framework. As seen in Section 1.3, a simple formulation describing the production of acceptor-like defects and the deactivation of acceptor dopants in UFSD-based detectors is [48]

$$N_{\text{A}}(\Phi, x) = g_{\text{eff}}\Phi + N_{\text{A}}(0, x)\,\mathrm{e}^{-\Phi \cdot c(N_{\text{A}}(0,x))}, \tag{3.76}$$

where N_{A} is the acceptor density in silicon, Φ is the fluence (in $\text{n}_{\text{eq}}/\text{cm}^2$), g_{eff} a coefficient determining the effective acceptor states production, and c is an appropriate function of the acceptor density before irradiation $N_{\text{A}}(0, x)$. This function indicates how strongly radiation deactivates the acceptor atoms. Notice that N_{A} is also a function of the position x inside the device. This means that both the acceptor removal process, described by the last term of Eq. (3.76), and the function c change according to the initial local density.

When performing parametric UFSD simulations where the unknowns are functions of the irradiation level, Eq. (3.76) is implemented in the system in such a way that the variable x maps each node of the discretization grid. To this aim, two steps are required: first of all, the user has to recompute off-line all the p-type profiles included in the detector according to the acceptor removal-creation law. Secondly, this profile is discretized and plugged into the Poisson's and continuity equations so that the initial conditions can include the new doping when solving the DD model. Usually, TCAD programs automatically adapt the profile discretization created by the user to the mesh, which is a remarkable advantage.

A different procedure can be exploited to define radiation-induced trap states in oxides or at interfaces between different materials, where the physical models governing their creation are typically simpler than Eq. (3.76). Since most of the TCAD tools predispose the implementation of defect levels or bands, the user should introduce their characteristic parameters (i.e., the energy and their eventual energy distribution, in case of band states, the initial concentration and the scattering or GR cross-sections with electrons and holes). If the experiments suggest that these defects acting like traps vary with the fluence in terms of density or energy, the user

needs to specify the parametric law through which the dependence takes place and TCAD automatically recomputes them material-wise or region-wise by reason of the radiation dose. One of the most robust frameworks in the literature describing how surface damage behaves with fluence is the so-called Perugia model [28, 29], a phenomenological set of parameters which introduces, besides the oxide charges, also acceptor- and donor-like interface trap states. In particular, acceptors uniformly occupy a band included between $E_C - 0.56$ eV and E_C, while donors are distributed in a band of width 0.6 eV, starting in correspondence of the valence band energy E_V. The concentrations of these defects, as well as of the charges in the oxide, are fluence-dependent and may slightly change according to the foundry producing the devices.

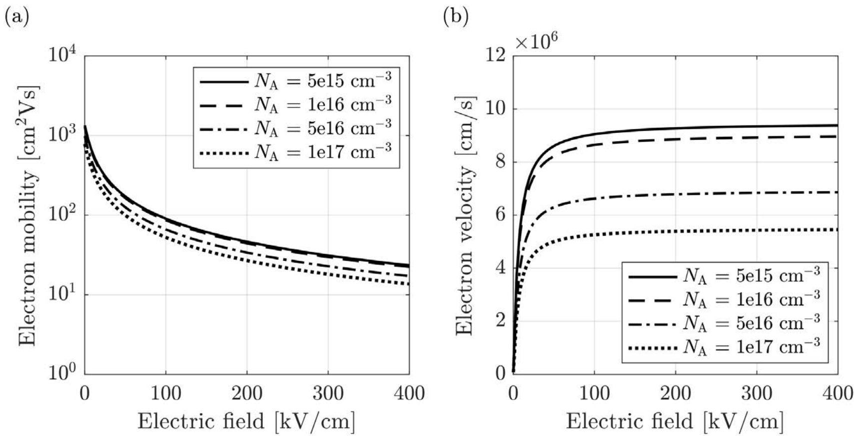

Figure 3.4 Trend of the electron mobility, *left*, and velocity in silicon *right*, as a function of the parallel component (with respect to the drift lines) of the electric field, calculated analytically for four acceptor atoms concentrations N_A, and at room temperature.

When the number of free charges changes due to an effective doping variation, also two fundamental solid-state quantities driving the operation of a silicon device change: the mobility and velocity of carriers (see, for instance, Ref. [3]). This effect is shown, for electrons, in Fig. 3.4, where the variation of both quantities as a function of the electric field has been calculated considering a silicon sample homogeneously doped with four different acceptor concentrations N_A. As one may see, when N_A increases, both quantities decrease for a given field. The curves here reported have a trend well-known in the literature and refer to a specific doping level. This means that they are valid locally, and the global behaviour of the whole device is given by the contribution of all the different doping concentrations. Since the effective doping concentration also depends on the fluence, in order to simulate the impact of radiation damaging on carrier mobility and velocity at the device-level, it is necessary to compute the new physical quantities in each node of the discretization mesh. Moreover, if the acceptor states generated by the radiation are explicitly declared as traps, the solver will also treat them as further scattering centers in which

calculate appropriate GR rates for trapping processes. This feature provides a more realistic simulation, especially if the goal is to study in detail the signal shape due to a charged particle passing through the detector. Generally, the radiation has the overall effect of decreasing both mobility and velocity in UFSD-based detectors.

3.4.3 OTHER PHYSICAL MODELS

Due to the crucial role of dopants in the operation of silicon devices, two additional effects, both function of the doping concentration, may be required in UFSD simulations. The first mechanism is the so-called *band-gap narrowing*, consisting – as the name suggests – in a slight reduction of the forbidden energy gap of silicon. This is due to the orbital overlapping of dopant atoms when they exceed a critical density N_{crit}. The energy bands generated by these new states may be sufficiently shallow and wide (the band is larger than E_{crit}) that they enter the conduction or valence band, with the result of narrowing the gap E_g. The gap reduction at $T = 300$ K is of the order of few tens of meV when the dopants are about 10^{18} cm^{-3}, and slightly decreases with temperature. The higher the total doping density N_{tot}, the larger the band-gap narrowing ΔE_g is. One of the most used formalisms is the Slotboom model [95, 123, 124, 125], according to which

$$\Delta E_g = E_{\text{crit}} \left(\ln\left(\frac{N_{\text{tot}}}{N_{\text{crit}}}\right) + \sqrt{\frac{1}{2} + \left[\ln\left(\frac{N_{\text{tot}}}{N_{\text{crit}}}\right)\right]^2} \right) , \tag{3.77}$$

where, in silicon, $N_{\text{crit}} = 1.3 \cdot 10^{17}$ cm^{-3} and $E_{\text{crit}} = 6.92 \cdot 10^{-3}$ eV.

So far, we always referred to the dopants without distinguishing between implanted dose and effective doping concentration. The second important phenomenon involving both the doping and the temperature is the capability of the dopant atoms to provide free charges (negative for donors and positive for acceptors). Since the ion implantation mostly generates interstitial impurities, a thermal process – called annealing – is performed after the implantation to activate the dopants, that become substitutional. During the annealing, besides the activation, also the lattice repair takes place, since the ion implantation has locally induced dislocations, vacancies, point defects or stacking faults. As the annealing time or temperature becomes higher, the lattice rearranges and dopants diffuse. Usually, this process can activate only a fraction of the nominal quantity of implanted atoms, which comes from a delicate balance between applying the minimum energy needed to reduce the lattice defects and keeping under control the diffusion of dopants. So, the thermal cycles applied to the implanted sample must follow a precise recipe, where the process temperature and duration have to be accurately determined. Even supposing to activate all the impurities, once the annealing has been performed, also the operating temperature of the device, in principle, can act on the capability to provide free charge carriers. In fact, each dopant is characterized by a ionization energy, given by the energy difference between the impurity level and the corresponding band edge (the bottom of the conduction band for donors or the top of the valence band for acceptors). When the thermal energy of the system – related to the device

operating temperature – is lower than the ionization energy of dopants the resulting fraction of active impurity atoms participating to the conduction is lower than 100%. This effect is known as *incomplete ionization* and, besides the temperature, depends also on the nominal concentration of dopants as well as on their chemical nature.

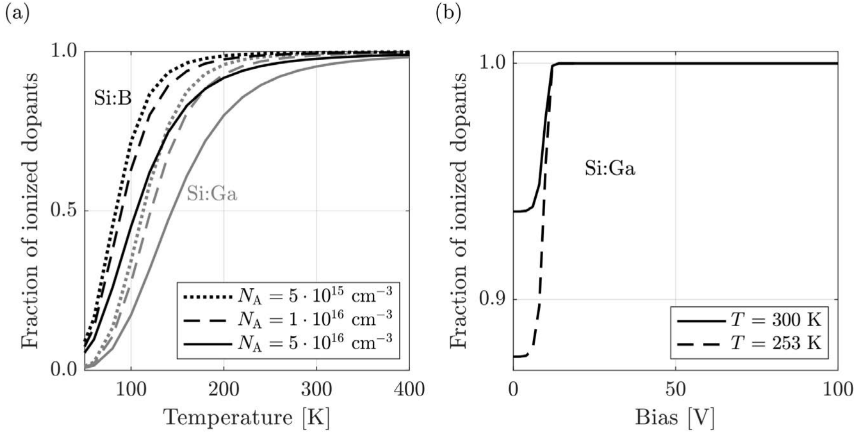

Figure 3.5 *Left*: fraction of activated acceptors in silicon, calculated for boron and gallium at equilibrium, as a function of the absolute temperature and for three different nominal acceptor atoms concentrations. *Right*: fraction of activated acceptors, in a UFSD-based structure, with gallium peak dose $N_A \sim 5 \cdot 10^{16}$ cm^{-3}, plotted at 300 K and 253 K, as a function of the applied reverse bias.

The *left* panel of Fig. 3.5 shows the fraction of the active acceptor dopants as a function of the simulation temperature in p-type silicon for three different nominal concentrations of boron and gallium. These calculations have been performed by using the incomplete ionization law

$$N_{A,0}(\mathscr{E},T) = N_A \left(1 + g_A \exp\left(\frac{E_A - E_{F,p}(\mathscr{E})}{k_B T}\right)\right), \tag{3.78}$$

which essentially is the Fermi-Dirac distribution for a population of acceptors (similarly for donors) with ionization energy E_A and density of active dopant atoms N_A. The terms $E_{F,p}$ and g_A are, respectively, the quasi-Fermi level of holes and the degeneracy factor, an integer number changing according to the dopant element and equal to 2 in case of gallium or boron. Equation (3.78) is applied until $N_{A,0}$ is lower than an effective concentration, above which we usually consider all dopants completely ionized (in Si:B, such threshold value is 10^{22} cm^{-3}). As the plot shows, the higher the nominal concentration, the lower the ionization fraction at fixed temperature, for both boron and gallium atoms.

Since the energy difference $(E_{F,p} - E_V)$ changes with the applied field, it is also possible to study how the activation evolves out of equilibrium with the external

bias. To this purpose, a UFSD-like structure with a gallium-doped multiplication layer has been implemented in TCAD. The *right* panel, (b), of Fig. 3.5 demonstrates that, for a peak density $\sim 5 \cdot 10^{16}$ cm^{-3} of the gallium implant and two different temperatures, the acceptors are almost completely activated even at low bias, far below the operating point of a standard UFSD-based detector.

3.5 SIMULATING ULTRA-FAST SILICON DETECTORS

In this section, several examples of Ultra-Fast Silicon Detectors simulations, and their comparison with experimental data, will be presented. In the following, if not otherwise specified, the TCAD results are obtained by solving the DD model, which includes SRH generation-recombination, avalanche multiplication, BTBT, TAT, band-gap narrowing and, when necessary, also proper radiation damage models.

The present section is divided into two paragraphs, the first one focuses on the use of numerical simulations to replicate the detector leakage currents and internal electric fields, while the second one on predicting the signal properties in both not irradiated and irradiated detectors.

3.5.1 STATIC CHARACTERISTICS AND ELECTRIC FIELD

As seen when introducing the numerical implementation of physical models, the sweep of an electrical quantity – e.g., the applied bias – corresponds to a sequence of steps, each representing a quasi-stationary state. By solving the DD equations in the whole space domain, it is possible to predict the trend of the total current versus the applied potential, the $I(V)$ characteristics. In order to have realistic results for both the leakage current and the breakdown voltage, the simulations must include as many technological details as possible. Among them, the doping profiles (and also their eventual lateral diffusion spreadings, when simulating in 2D-3D), the intergap defect levels, the interface or oxide charges, and all the geometrical features characterizing the device like the spacing among different implants and structures, or the thickness of each layer.

Figure 3.6 shows in the *left* panel the comparison between the measured and simulated $I(V)$ characteristics of UFSD2 detectors, differing in type and dose of acceptors in the gain implant. The increase of the current is more and more pronounced as the applied reverse voltage is raised due to the effect of charge multiplication. This is a clear example of why a correct modelling of the avalanche mechanism is so crucial in UFSD-based detectors. The *right* panel, instead, shows two $C(V)$ characteristics of two devices from the same UFSD production. The numerical calculation matches well the experimental data and can accurately predict the depletion voltage of both the gain layer and the active substrate. In these calculations, the van Overstraeten–de Man model has been adopted and a quasi-1D simulation domain has been used, having implemented just a silicon slice including the UFSD junction. It is worth noting that this kind of numerical characterizations can be used not only to adapt and calibrate the models and their parameters with respect to the specific

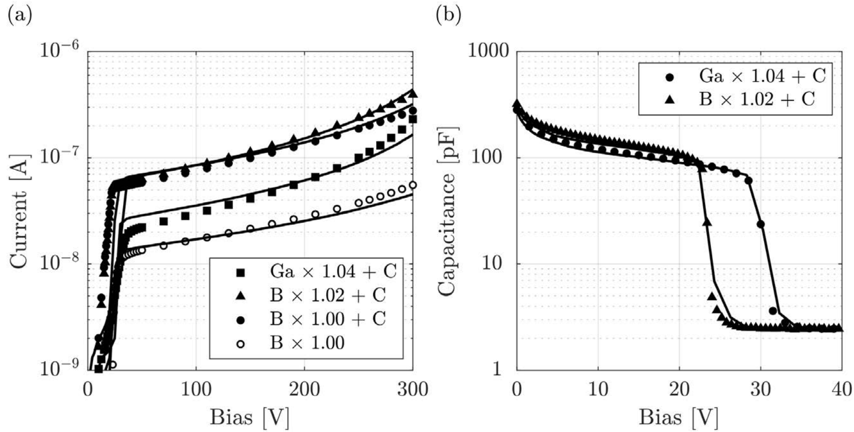

Figure 3.6 *Left*: measured and simulated current-voltage $I(V)$ characteristics. *Right*: measured and simulated capacitance-voltage $C(V)$ characteristics. The curves are for several UFSD2 samples having different dose (from 1.00 to 1.04) and species (gallium and boron) of acceptors, with or without the co-implantation of carbon atoms. Symbols are laboratory measurements, while lines are TCAD simulations [51, 54].

fabrication technology, but also – and especially – to predict the behaviour of the detectors or, in other words, to design them. In this regard, it should be highlighted that the $I(V)$ and $C(V)$ characteristics can help in defining the optimal gain implant profile, for example through the study of breakdown voltages. The same can be said for the gain curves, analyzed in the next section. Besides ensuring that the internal field is well distributed at the periphery, the designer has to assure electrical isolation between nearby active regions (gain implants) in correspondence of the inter-pad regions.

Through a 2D or 3D implementation of the UFSD under study, it is possible to infer important conclusions about the trend of the field and drift lines. Figure 3.7 shows a simulated cross-sectional view of the inter-pad separating two adjacent active regions (the p-gain implants). On both sides, there is a n-type implant, the junction termination extension (JTE), while in the middle a p-stop structure is implanted. The JTEs are used to confine the high fields produced in reverse bias by the gain layer, and to prevent particles crossing the detector in the inter-pad from generating out-of-time signals (see Section 2.8). The p-stop implants, instead, avoid that the electrons inversion layer, due to the oxides and interface charges, short-circuits two adjacent pads. These isolation structures determine – as it will become more clear in the next paragraph – a performance drop because where there is no multiplication the detector gain tends to be close to one. The calculated drift lines help in determining the width of the no-gain region (see, for instance, the lines marked in white), since they allow to predict which is the volume of silicon affected by the lower charge collection. By properly tuning the technological parameters of the isolation implants (such as dose and depth) it is possible to minimize their impact on the detector performance.

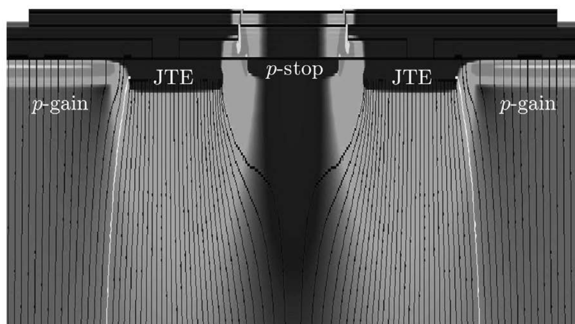

Figure 3.7 Cross-section of a generic UFSD device in the inter-pad region between two neighboring active areas, showing the electron drift lines and the electric field map (in grey tones, where brighter areas have a higher field intensity). White lines are the external drift lines enclosing the JTE implants. They highlight the collection volume of the *p*-gain.

Other considerations may also concern the chemical nature and thickness of the oxide and passivation layers deposited on the top of the detector that, as well as the use of proper interface defects, represents an important feature, especially for irradiated devices.

Besides the implantation properties, simulations are also used to understand the effects on the internal field distribution of two important key-elements: the oxide thickness and the metal extensions on the device surface. Both concur in determining the device operation, either of the detector periphery, the cut-line region at the physical edge of the device or between active areas. To predict their impact on the breakdown voltage in correspondence of the isolation implants, the simulations proposed in Fig. 3.8 have been performed. The figure reports, in the *top* part, two cross-sections of the UFSD inter-pad region between active areas showing the field intensity map (in grey tones) simulated at the breakdown voltage. *Top left* panel refers to a structure with a very short metal overhang (with respect to the n^+-contact) deposited on top of the device, while the *top right* panel concerns the case of implementing a field plate. Furthermore, the *bottom* panels (c) and (d) show the $I(V)$ characteristics of the devices sketched, respectively, in subfigures (a) and (b). It's is evident that the use of a metal field plate fully covering the JTE implants is beneficial to keep under control the electric field, which relaxes in silicon. Indeed, the region with the highest field intensity moves from the junctions of the inter-pad implants to the oxide, in correspondence with the metallization edge, allowing to reach the breakdown at a voltage approximately 200 V higher.

Similar reasoning can be made if we want to characterize the field in the detector periphery, for example, in order to design more robust protection structures which

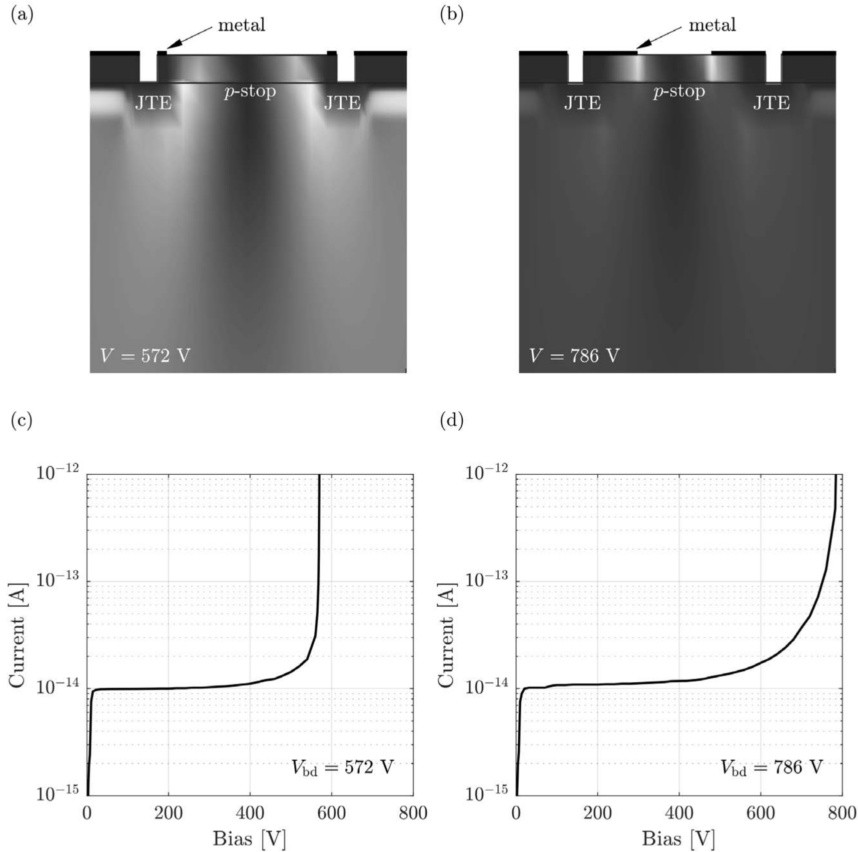

Figure 3.8 *Top* : simulated electric field intensity map (in grey tones - brighter areas have a higher intensity) at the breakdown voltage V_{bd}, for two UFSD with different inter-pad regions, short (a) or long (b) metal overhang (black pattern on the device surface). *Bottom*: corresponding $I(V)$ curves.

allow higher breakdown voltages. Overall, TCAD simulation offers a powerful tool to have relatively fast and reliable feedback on the physics driving each part of the system.

3.5.2 TRANSIENT PROCESSES

This section deals with one of the most interesting and crucial targets of simulating UFSDs: the signal formation. The UFSD operation mechanism is based on the multiplication by a certain gain factor (which is bias-dependent) of the primary *e-h* pairs produced when a charged particle crosses the detector.

The numerical implementation consists of a time-dependent process during which a heavy ion or a laser beam simulates a particle crossing the detector, with

a custom trajectory and a well-defined energy released in the silicon lattice. Since this procedure, in principle, holds for each particle in any kind of semiconductor, it is important to characterize the detection process for a minimum ionizing particle (MIP) in silicon. A possible approach can be represented by the calibration of the collected charges in a detector without the gain implant. Here, the absence of multiplication allows to properly tune the energy released by the ion or the laser, allowing to accurately describe the gain in UFSD at low applied voltages, i.e., below the avalanche onset. Moreover, as explained in Section 2.10, after high values of fluence, the electric field in a PIN can be raised so much that charge multiplication happens in the sensor bulk. This aspect is important in the simulation of heavily irradiated UFSD.

Figure 3.9 shows two significant examples coming from the calibration campaign carried out on two independent productions of PIN devices, one by CNM and the second by HPK [36]. The number of charges has been obtained by integrating the signal response in time, both in simulations and in the laboratory measurements. In order to accurately fit the experimental data, the curves presented in the *left* panel have been simulated by setting a MIP releasing an energy equivalent to the production of $\sim$ 60 electron-hole pairs per crossed micron whereas in the *right* panel we set $\sim$ 70 pairs. Both devices are irradiated (respectively, with neutrons at a fluence of $3{\cdot}10^{15}$ n_{eq}/cm^2 and with pions at $1.5{\cdot}10^{15}$ n_{eq}/cm^2) so these simulations also accounted for the phenomenological model written in Eq. (3.76) for the acceptor creation/deactivation with fluence. The plots indicate that both Massey and van Overstraeten-de Man models are adequate in reproducing the collected charge while the Okuto-Crowell predicts a higher bias value for the onset of the multiplication in the bulk.

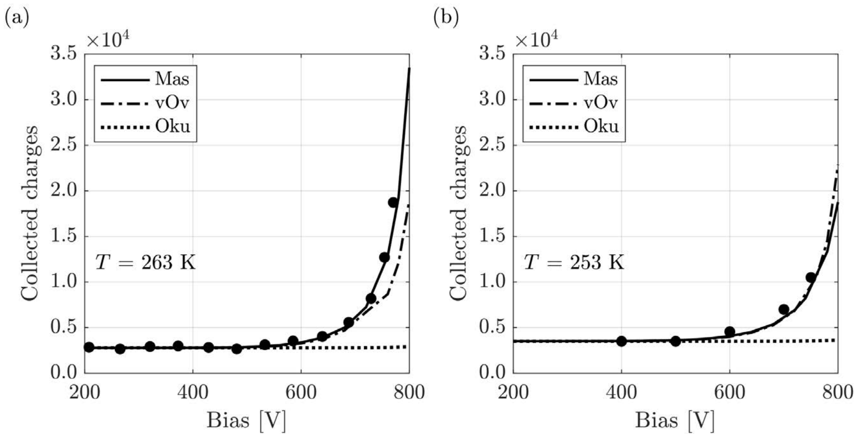

Figure 3.9 Measured (symbols) and simulated (lines) number of collected charges in 50 μm-thick PIN diodes manufactured by (a) CNM and (b) HPK. Simulations have been performed using three different avalanche models [52, 54].

It is worth stressing that, even with different formalism, the same result is obtained simulating either the injection of heavy ions or that of an infrared laser beam (for the wavelength 1064 nm) releasing the energy of 310 W/cm^2. The main difference between the two numerical approaches is that the simulation with ions is more suitable for a comparison with results obtained with particle while the simulation with laser, allowing to define the illumination over a window with finite width, is more appropriate with results obtained with the Transient Current Technique, TCT (see Section 4.3 and Ref. [74]).

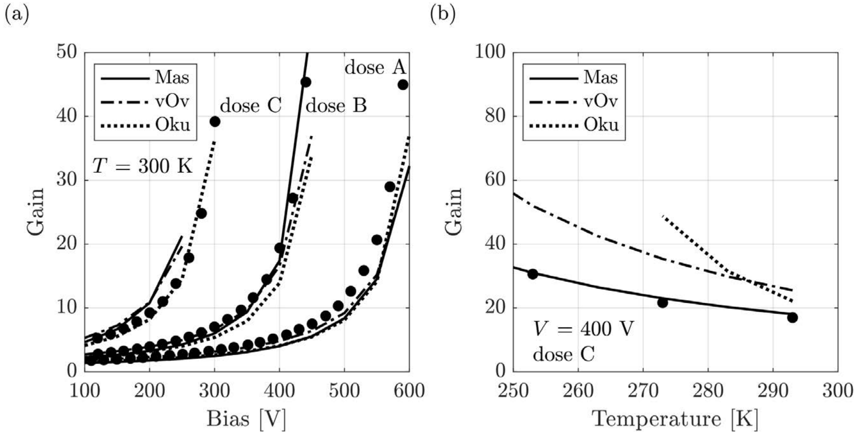

Figure 3.10 Measured and simulated gain value versus (a) bias and (b) temperature, in 50 μm-thick UFSDs with different gain implant dose (increasing from A to C) [52, 117].

Once the avalanche models have been optimized in terms of charges generated without multiplication layer, the following step is to compare TCAD simulations and measurements of the gain versus voltage $G(V)$ characteristics in UFSDs. The gain is calculated as the ratio between the charges generated in the UFSD and those in the corresponding PIN diode. For this reason, both devices have to be implemented in the TCAD solver and subjected to the passage of a heavy ion or a laser beam. Then, the gain curve is obtained by computing at each bias step the ratio of charges $Q_{\text{UFSD}}/Q_{\text{PIN}}$.

In Fig. 3.10, results obtained with a 50 μm-thick UFSD are reported. The *left* panel shows three different room temperature gain curves coming from devices with, from left to right, a decreasing boron concentration in the gain implant. The *right* panel refers to the device with the highest dose (as the leftmost $G(V)$ curve). These plots demonstrate that the three multiplication models are highly competitive in reproducing the gain curve of these structures at a fixed temperature and for all the gain implant doses considered. For what concerns the gain increase with temperature, the Massey model is the most reliable. The slope of $G(T)$ is quite satisfactory also for the van Overstraeten-de Man avalanche, while the Okuto model predicts a much steeper dependence.

The *left* panel of Fig. 3.11 presents a second comparison between simulated and measured $G(V)$ curves at room temperature, using sensors from the UFSD2 production. The gain implant has been obtained with boron or gallium at different doses, with/without the co-implantation of carbon atoms. For simplicity, only results obtained with the van Overstraeten-de Man model have been reported. The plot on the *right* panel reports the amplitude seen by two adjacent pads during a TCT position scan (see Section 4.3) obtained – both experimentally and numerically – by moving the spot of an infrared laser beam along the detector surface in correspondence of the inter-pad region.

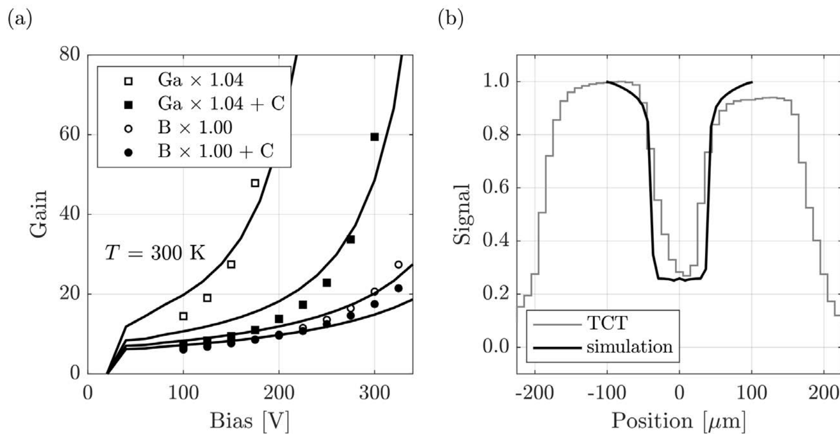

Figure 3.11 *Left*: measured and simulated (using the van Overstraeten-de Man avalanche) room temperature gain curves in 50 μm-thick UFSD differing in the gain implant dopants. *Right*: laser scan of the inter-pad region showing the normalized signal intensity versus the position of the laser spot on the detector surface [51].

The next step in the UFSD simulation is the calibration of the parameters characterizing the gain implant. Figure 3.12 shows the simulated effects of varying the gain implant dose and depth. The *left* panel reports how the gain boron dose (normalized at 1 μm) has to be scaled when implanting at different depths in order to have gain $G = 20$ at the reverse bias $V = 200$ V. As explained in Section 2.1.1, when the gain implant is deeper, the doping has to be decreased. The *right* panel of Fig. 3.12 shows how the gain changes varying the doping density of the gain implant. For a fixed depth of the boron profile, 2 μm, the implant dose has been varied from 100% to 85%. As the acceptor density becomes lower, the gain decreases. These predictions quantitatively depend on the avalanche model: in this example, the Massey (black curves) model gives higher gain with respect to the Okuto-Crowell model (grey curves).

One important ingredient of the UFSD design is simulating the effects of radiation on the detector performances. To this aim, it is necessary to account for the doping creation and deactivation, both in the gain implant and in the bulk, and the

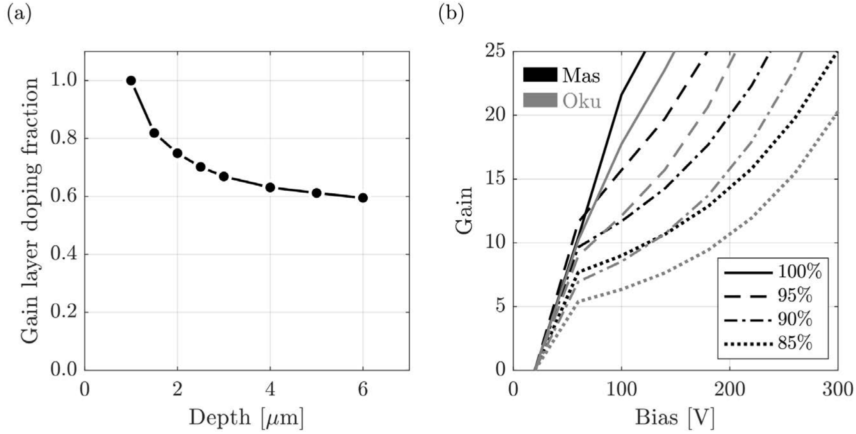

Figure 3.12 *Left*: gain implant dose versus gain implant depth, needed to have $G = 20$ at $V = 200$ V. *Right*: gain as a function of the bias voltage, calculated with the Massey (black) and Okuto-Crowell model (grey), for a gain implant depth of 2 μm and four different doses of gain implant doping.

additional generation mechanism due to surface/oxide radiation damage described. This can be done, for instance, through the Perugia model. By using the van Overstraeten-de Man formalism for the avalanche, it has been possible to reproduce the experimental data shown in the *left* panel of Fig. 3.13, where the collected charge in a UFSD device is plotted as a function of the bias, before irradiation and at two fluences. The numerical prediction is quite satisfactory, also in consideration of the changes induced by a temperature variation

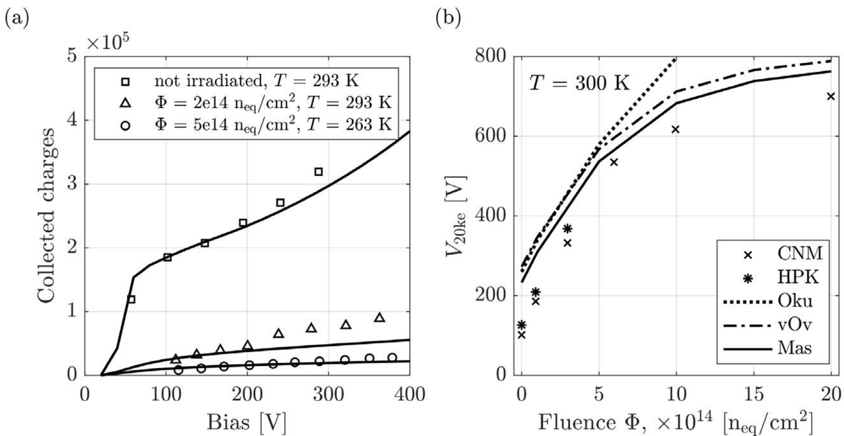

Figure 3.13 *Left*: number of collected charges versus bias for a 300 μm-thick UFSD before and after irradiation. *Right*: bias values needed to collect $2 \cdot 10^4$ electrons as a function of the fluence in 50 μm-thick UFSDs from different vendors [52].

(293 K vs 263 K), confirming the robustness of the TCAD approach in terms of radiation-hardness. Finally, in the *right* panel of Fig. 3.13, the measured and simulated bias to collect $2 \cdot 10^4$ electrons versus fluence is reported for several 50 μm-thick UFSDs. Here all the three avalanche models have been tested, obtaining that – as in the simulation of gain as a function of temperature – the Okuto–Crowell model is the worst choice to reproduce this set of measurements.

3.6 RESISTIVE AC-COUPLED SILICON DETECTORS DESIGN

As explained above, the UFSD design is optimized to achieve the best temporal resolution. Consequently, the signals generated by particles hitting in the inter-pad region should not be amplified to avoid out-of-time signals (see Section 2.8). The no-gain region decreases the detector fill factor and leads to the use of multiple staggered layers to achieve hermetic coverage. A possible strategy to overcome this issue is the AC-LGAD [61] paradigm, which consists in the implementation of a continuous gain implant that achieves 100% fill factor, as shown in Fig. 3.14. In order to

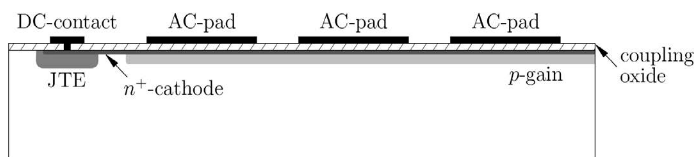

Figure 3.14 Schematic cross-section (not to scale) of an RSD (AC-LGAD) sensor.

reconstruct the hit position without segmentation, AC-LGADs make use of two key design elements: (*i*) a coupling oxide and (*ii*) a resistive n^+-cathode. For the signal to be visible on the AC-pads, both elements need to be correctly engineered so that the lowest impedance path to ground for the signal is via the read-out electronics. The signal discharges with an *RC* time constant given by the product of the AC-pad capacitance and the n^+-cathode resistivity. Thus, since the target is to have a discharge time long enough for the signal to be seen by the pads, the *RC* time constant must be chosen to be longer than the signal formation time ($\sim 1-2$ ns). However, to avoid pile-up effects, the *RC* should be sufficiently short to allow a prompt return to the baseline. The most important parameters are the oxide composition and thickness, determining the coupling strength, and the n^+ implant dose and profile, which instead directly modify the resistivity of the cathode. The geometrical configuration of the AC-pads (e.g., dimension and pitch) is also decisive since their geometrical dimension directly impacts the *RC* time constant. Given the large number of parameters concurring in the AC-LGADs design, it is crucial to have a reliable numerical tool for the prediction of signal formation and read-out properties.

In the last few years, several research groups began to develop and also fabricate AC-LGADs. In this chapter, a specific design, produced by INFN in Torino (Italy) and FBK, called Resistive AC-Coupled Silicon Detectors (RSD), is explained. The

name, RSD, refers to its two main features: the resistive implant and the coupling oxide. Here, some parametric simulations referring to standard values are shown in order to demonstrate how TCAD can help in the design of such devices. To this aim, the impact of varying the RSD design parameters on the signal waveform is analyzed. Several simulated signals produced by a MIP in an RSD are reported in Fig. 3.15. In these results, the detector is composed of a row of three AC-pads, and the particle is always crossing the device in the center of the first pad and black, grey, and light grey lines refer, respectively, to the signal seen in the hit pad, its first and second neighbor pad. The preliminary observation coming from the simulations is that the current as a function of time has a bipolar behaviour, an intrinsic feature of the AC-coupled read-out paradigm.

The first lobe is generated by the coupling with the AC-pads, when the signal is collected by the resistive n^+-cathode, while the opposite lobe is due to their subsequent discharge to ground, which takes place through the DC-contact (see Fig. 3.14). The discharge characteristics – such as the amplitude and duration – depend on the *RC* constant of the equivalent read-out circuit, so, on the coupling capacitance of the oxide layer and on the sheet resistance of the cathode. The second fundamental observation is that the signal is shared among several pads. Such effect, besides the 100% fill factor, represents the most important difference between the RSD paradigm and standard UFSDs, where the signal is visible only on one pad. This difference leads to a clear advantage in the accuracy of position reconstruction: combining the information of many pads, the impact position can be identified with the precision of a few microns [55, 58]. As for the other figures of merit related to the signals, also the charge sharing depends on the properties of both the coupling oxide and the resistive implant.

The *top left* and *top right* panels of Figure 3.15 report the signals simulated with a 2D TCAD implementation, injecting 1 MIP in a 50 μm-pitch RSD device with 45 μm pad size. First of all, consider only the pad hit directly by the particle (black curves). The solid lines show the signals obtained implementing the standard values of the FBK technological parameters (low oxide thickness and low n^+-cathode dose). The dotted lines, instead, refer to an increase of either the oxide thickness (*top left*) or cathode resistivity (*top right*). The simulations show that both the peak amplitude and the discharge duration increase when the oxide is thinner (higher capacitance) or the n^+-cathode dose decreases (higher resistivity). The signals seen on the first and second neighboring pad, respectively in grey and light grey, have the same behaviour. The simulation, therefore, indicates that if the *RC* time constant is too short, the signal discharges before being fully formed.

The *bottom left* panel of Fig. 3.15 shows a different scenario. In this plot, the same RSD geometry of the previous simulations is compared to a modified version, where the pitch is doubled and the pad size is 95 μm. Moreover, the fabrication technology has been slightly changed, a 50% thicker oxide has been used. Focusing again on the signals coming from the pad crossed by the particle (black lines), it is evident that the bigger pad has a larger signal. Considering the adjacent pads, the first and second neighboring pads in the 100 μm-pitch RSD produce a signal which is lower in amplitude and broadened in time with respect to the 50 μm-pitch case.

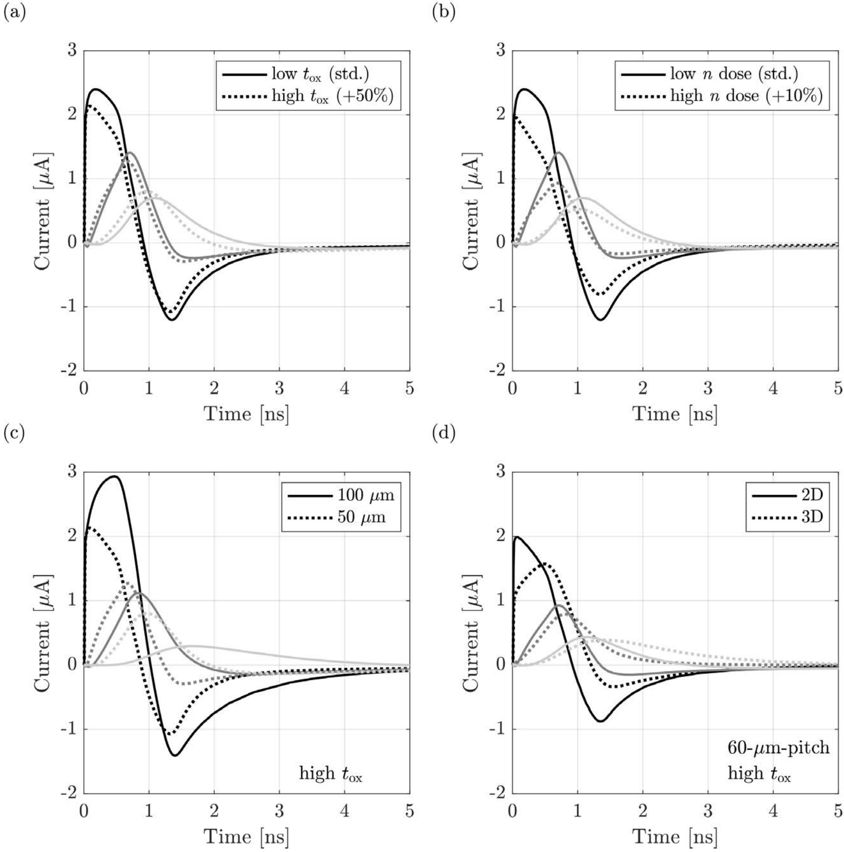

Figure 3.15 Simulated signals in a three-pads RSD detector, biased at 300 V, at room temperature. Black lines represent the signal coming from the hit pad, while the different tones of grey refer to the first and second neighboring pad. *Top left*: signals in two RSDs differing for the thickness of the coupling oxide. *Top right*: signals in two RSDs differing for the n^+-cathode resistivity. *Bottom left*: signals in two RSDs differing for the sensor pitch. *Bottom right*: signals generated using either a 2D or 3D simulation [54].

Another important result stressing the importance of numerical simulations is presented in the *bottom right* panel of Fig. 3.15. In this plot, the 2D and 3D implementations of a 60 µm-pitch RSD are compared. Here the resistive n^+-cathode dose is the standard one, while the oxide has been chosen in its thick version. In these simulations, all the pads have in common the same behaviour: the signals computed with the 3D structure are slightly lower and longer with respect to the 2D case. Even if in the 2D condition the tool emulates the 3D scenario projecting the third dimension by a customizable factor, the full-3D simulation takes into account additional

built-in volume-related physical effects that originate the observed differences in the signals.

This example is very instructive from the numerical standpoint since it well represents the importance of choosing the most appropriate framework according to the target of the simulation. If the 2D geometry is sufficient to reproduce the electrical behaviour of a UFSD, or its gain curve, the 3D implementation is necessary to study the transient phenomena generating induced signals in RSD detectors.

To conclude this short overview of the RSD design, the effects of the termination structures present at the borders of the RSD, in correspondence of the DC-contact (as shown in Fig. 3.14), are presented. To understand the edge effect in RSD, consider

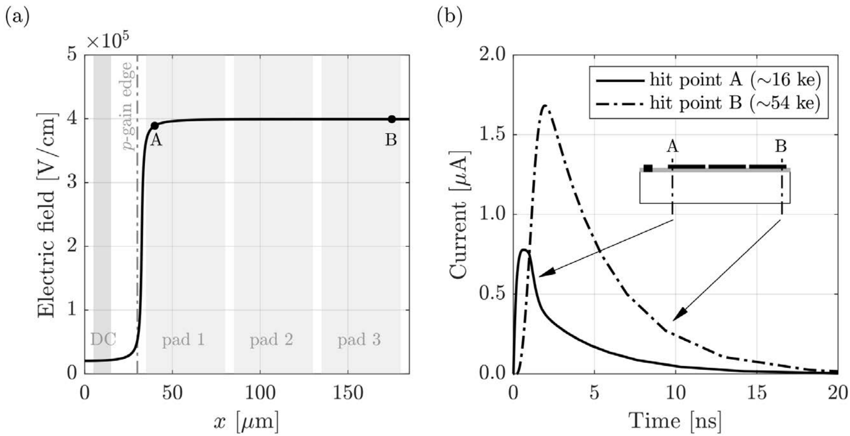

Figure 3.16 *Left*: electric field in a 1×3 array of a 50 μm-pitch RSD, with 45 μm pad size, biased at 160 V. The MIP has been injected in positions A and B. *Right*: simulated signals seen on the DC contact of the same device as a function of the distance from the sensor edge [54].

the simulation of a 1×3 array with 50 μm-pitch and 45 μm pad size through a 2D implementation again. A MIP is injected in the first pad, 5 μm from its left edge, and in the third one, 5 μm before its end. The electric field beneath the silicon/oxide interface as a function of position is reported in the *left* panel of Fig. 3.16. At small x values, where the DC-contact is located, the field is of the order of few tens of kV. Then, the field experiences a steep increase, starting at the edge of the multiplication implant, located at 30 μm from the origin (the greater the lateral spread of this implant, the earlier the field increase). Past this sharp rise, the field increases slowly: at the point A, at 40 μm, is 394 kV/cm while at the point B, at 175 μm, is 404 kV/cm. The *right* panel reports the simulated signals collected by the DC-contact at 160 V, when the MIP crosses the device in A and B. The inset drawing shows the cross-sectional view of the simulated RSD, as well as the two tracks crossing the detector perpendicularly to the surface. The signal from A arrives earlier than that from B, since the discharge path from the hit point to the contact is shorter, however, the number of collected charges is smaller (respectively, 1.6 against $5.4 \cdot 10^4$ electrons)

since the multiplication field is lower in A than in B. This simulation highlights that termination structures may affect the detector response, creating an area of lower gain at the periphery of the active region. This is why numerical simulations are essential in designing the RSD and, in particular, in defining specific layout rules, such as the minimum distance between the peripheral pads and the gain implant edge.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com