Scanning Microscopy

Volume 9 | Number 2 Article 4

5-25-1995

Cathodoluminescence Contrast of Localized Defects Part I. Numerical Model for Simulation

K. L. Pey
Institute of Microelectronics, Singapore

D. S. H. Chan
National University of Singapore

J. C. H. Phang
National University of Singapore

Scanning Microscopy, Vol. 9, No. 2, 1995 (Pages 355-366) 0891-7035/95$5.00+.25
Scanning Microscopy International, Chicago (AMF O'Hare), IL 60666 USA

CATHODOLUMINESCENCE CONTRAST OF LOCALIZED DEFECTS PART I. NUMERICAL MODEL FOR SIMULATION

K.L. Pey[1], D.S.H. Chan* and J.C.H. Phang

Centre for Integrated Circuit Failure Analysis and Reliability, Faculty of Engineering
National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511
[1]Institute of Microelectronics, 11 Science Park Road, Science Park II, Singapore 0511

(Received for publication June 1, 1994, and in revised form May 25, 1995)

Abstract

A three-dimensional model has been developed for cathodoluminescence contrast of localized defects in semiconductors. The numerical model incorporates electron-solid interaction effects, charge transport phenomena and optical losses. Electron-solid interaction is modelled by a Monte Carlo method. Three-dimensional continuity equation and derivative boundary conditions are discretized by a central-difference quotients scheme. Localized defects are represented by regions of enhanced non-radiative recombination. The discretized linear difference equations of the boundary value problem are solved by the successive-over-relaxation method. A method for avoiding the divergence problem during the successive-over-relaxation calculation is illustrated. The solutions of the model are compared with the analytical results of several established models.

Key Words: Cathodoluminescence, cathodoluminescence contrast, defects, scanning electron microscope, numerical model, Monte Carlo, semiconductors.

*Address for correspondence:
D.S.H. Chan, Department of Electrical Engineering
National University of Singapore
10 Kent Ridge Crescent, Singapore 0511
Telephone number : (65) 772-2132
FAX number : (65) 779-1103

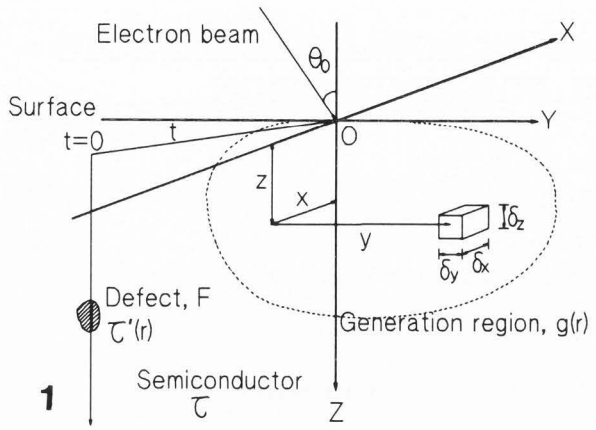
Introduction

Cathodoluminescence (CL) has been widely used to investigate electrically active lattice imperfections in optoelectronic materials. A number of theoretical methods for calculating CL signals from localized defects have been proposed [5, 6]. There have, however, been only limited attempts at incorporating a more realistic and accurate carrier generation function into CL models. For example, Löhnert and Kubalek [6] and Jakubowicz [5] analysed the contrast of dislocations with a uniform generation sphere or a point source. The approximations may fail either when the source is sufficiently close to the defect or when the excitation region is comparable to the defect size. Jakubowicz's model also neglected the fraction of carriers generated inside the dislocation cylinder. A method similar to Jakubowicz's model [5] was proposed by Pasemann and Hergert [11] which considered a dislocation lying parallel to the surface at a particular depth. Czyżewski and Joy [1] analysed the ratio of CL contrast to EBIC contrast for isolated dislocations using the Monte Carlo method to represent the source of e-h pairs. This proposed model is based on Donolato's Born approximation for excess carrier density in the presence of a point defect.

In this paper, an approach incorporating realistic e-h pair generation obtained from Monte Carlo calculations into a three-dimensional model for semiconductors with localized defects is described. The three-dimensional carrier diffusion equation is represented by a set of finite difference equations. This approach offers the flexibility of investigating different kinds of defect structures in the bulk by simply reducing the non-radiative recombination lifetime in the region of influence.

Formulation of Model

Cathodoluminescence is the emission of light as the result of electron bombardment. In the case of semiconductors, only radiative recombination contributes to photon generation. The generated photons propagate in all direction within the material, but only a fraction of them



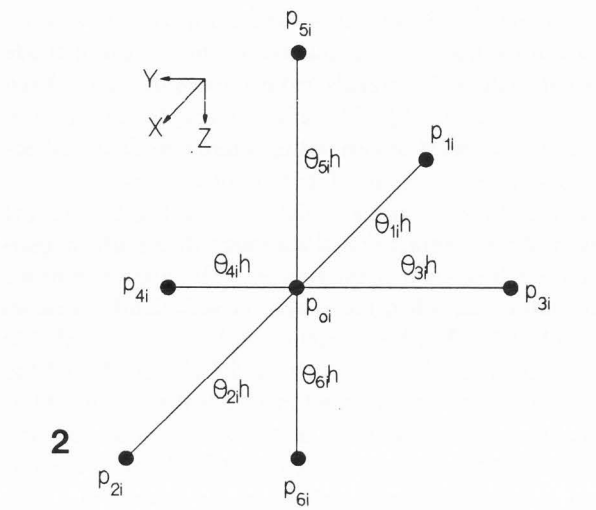


Figure 1. Schematic diagram of the model used for deriving the CL image of a localized defect.

Figure 2. A seven-point system arranged in an unequal-armed star for deriving the derivatives of the continuity equation.

---

emerges from the surface, giving rise to CL emissions. In steady state, the total CL intensity is the integral of radiative recombination rate over the entire sample volume corrected by the optical loss function of the generated photons. To analyse the generation of CL signal, it is important to know the excess minority carrier density accurately. The solution of the continuity equation for an arbitrary generation function presents a challenging problem, since no analytical expression for this function is available.

Electron-solid interaction and minority carrier diffusion equation in the presence of a defect

The Monte Carlo procedure described by Phang *et al.* [12] is used to evaluate the rate of energy dissipation of the electron beam on its way through the sample. The sample is assumed to be semi-infinite, bounded only by the top surface at z = 0, and divided into volume elements with dimensions δx, δy and δz. An electron beam of energy E is assumed to be incident on the surface at an angle θ_o with respect to the surface normal at the origin O of Figure 1. The three-dimensional spatial energy dissipation δE of all electrons traversing the sample volume is calculated and stored as a matrix of δE versus x, y and z using the Nearest-Grid-Point method [4]. Due to the small values chosen for δx, δy and δz, each cube in Figure 1 can be treated as a point source located at the centre of that volume [12], exciting e-h pairs in the semiconductor at a rate g(r) per unit volume.

For a homogeneous semiconductor with minority carrier diffusion coefficient D, lifetime τ, and a surface recombination velocity v_s, a localized defect under the surface can be represented by a bounded region of space F, where the minority carrier lifetime τ' is lower than that in the rest of the semiconductor. The right-handed, orthogonal, coordinate system shown in Figure 1 is used. The Z-axis is defined to be normal to the surface and the positive sense is into the specimen. In the SEM CL operation, the quantity used to form an image is the total photon flux collected by the detector as a function of the position t of the electron beam relative to the defect.

Following the definition used in [2], if r is in F, then $\tau(r)$ is equal to $\tau'(r)$. Elsewhere $\tau(r)$ equals τ. Introducing $L = (D\tau)^{1/2}$ and $L'(r) = (D\tau'(r))^{1/2}$, the three-dimensional continuity equation describing the diffusion process of the minority carriers is

$$\nabla^2 p(r) = -\frac{1}{D}g(r) + \frac{1}{L^2}p(r) + \gamma(r)e(r)p(r) \tag{1}$$

where p(r) is the excess minority carrier concentration,

$$e(r) = \begin{cases} 1 \text{ for } r \text{ inside } F \\ 0 \text{ elsewhere} \end{cases} \tag{2}$$

and

$$\gamma(r) = \frac{1}{D}\left[\frac{1}{\tau'(r)} - \frac{1}{\tau}\right] = \frac{1}{L'(r)^2} - \frac{1}{L^2} \tag{3}$$

is defined as the 'strength' of the defect [2]. If $L'(r) << L$, i.e., $\tau'(r) << \tau$, $\gamma(r) \approx 1/L'(r)^2$, then the defect strength becomes independent of the diffusion length or the lifetime of the host material.

The top surface boundary is characterized in eq. 4:

Table 1. List of normalization coefficients for the quantities of interest for GaAs materials [13].

Description	Normalized quantities	Normalization coefficients	
		Symbol	Numerical value
position coordinates	x,y,z	$L_{Di}=(\epsilon v_t/q^2\eta_{io})^{1/2}$	$0.3222*10^4\mu m$
carrier concentration	p(r)	η_{io}	$1.79*10^6 cm^{-3}$
carrier diffusion coefficient	D	D_o	$1*10^8\mu m^2 sec^{-1}$
generation rate	g(r)	$D_o\eta_{io}/L_{Di}^2$	$0.1723*10^{-4}\mu m^{-3} sec^{-1}$
defect strength	γ(r)	$1/L_{Di}^2$	$0.9628*10^{-7}\mu m^{-2}$
carrier lifetime	τ(r)	L_{Di}^2/D_o	$0.1039 sec^{-1}$
surface recombination velocity	v_s	D_o/L_{Di}	$0.3103*10^5\mu m.sec^{-1}$
grid spacing	h	L_{Di}	$0.3222*10^4\mu m$

ϵ = Permissiveness of the material v_t = Thermal voltage q = Electronic charge
L_{Di} = Intrinsic Debye length η_{io} = Intrinsic carrier concentration

---

$$D\frac{\partial p(r)}{\partial Z}\bigg|_{Z=0} = v_s\, p(r) \tag{4}$$

Discretization of diffusion equation and boundary conditions

By employing central-difference quotients [3], the boundary-value problem described can be solved by replacing the differential equation in eq. (1) with a finite difference equation. Consider the general case of a group of seven points whose spacings are non-uniform and arranged in an unequal-armed manner as shown in Figure 2. Each distance is represented by θ_{ai}h, where θ_{ai} is the fraction of the standard spacing h that the particular distance represents. Replacing the derivative in eq. (1) by its difference quotients as in eq. (I.8) in Appendix I, and rearranging eq. (1), the minority carrier density at point i = (x_i, y_i, z_i) becomes

$$\tilde{p}_{0i}^{(m+1)} = \frac{1}{A_i}(B_iP_{1i}^{(m+1)} + C_ip_{2i}^{(m)} + D_ip_{3i}^{(m+1)} + E_ip_{4i}^{(m)} + F_ip_{5i}^{(m+1)} + G_ip_{6i}^{(m)} + \frac{g_{0i}}{D} - \frac{p_{0i}^{(m)}}{L_{0i}^2} - \gamma_{0i}e_{0i}p_{0i}^{(m)}) \tag{5}$$

where the superscripts represent the number of iteration. This set of equations can be solved using Gauss-Seidal iteration [10].

The boundary values given by eq. (4) involve derivatives which require the domain to be extended beyond z = 0, i.e., negative z in the finite difference scheme. These fictitious exterior points [3] are nodes located at a row of the extended network. The derivative boundary condition is used to write difference quotients that permit the elimination of the fictitious points at z < 0.

Using eq. (I.10) in Appendix I, the set of equations from eq. (5) can be solved after eliminating the fictitious points. This set of equations at z_i = 0 is

$$\tilde{p}_{0i}^{(m+1)} = \frac{1}{A_i+F_iQ_i}(B_iP_{1i}^{(m+1)} + C_ip_{2i}^{(m)} + D_ip_{3i}^{(m+1)} + E_ip_{4i}^{(m)} + (F_i + G_i)p_{6i}^{(m)} + \frac{g_{0i}}{D} - \frac{p_{0i}^{(m)}}{L_{0i}^2} - \gamma_{0i}e_{0i}p_{0i}^{(m)}) \tag{6}$$

For semiconductor regions that are far away from the sources, it can be assumed that there are no excess minority carriers. Within a three-dimensional space, the excess carrier density decreases as (1/r)exp(-r/L) [2],

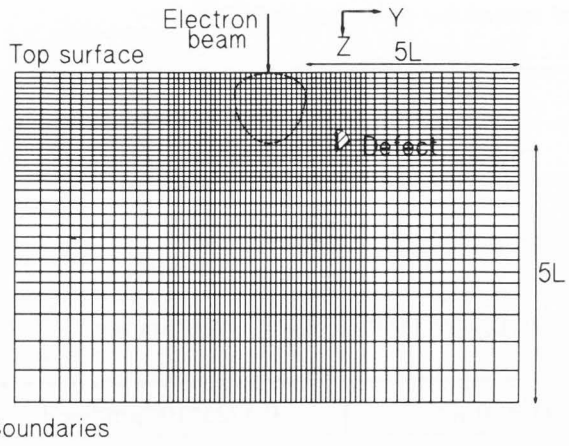


Figure 3. Schematic illustration of a typical grid system used for simulation.

---

where r is the radial distance from a source. Using this as a guide, the other boundaries are fixed at about five minority carrier diffusion lengths away from the edge of the generation volume. The error introduced by using this cut-off criterion for p(r) is much less than 0.7%.

A normalization scheme that is similar to that devised by de Mari [7] has been used to normalize the variables of the discretized semiconductor equations. The normalized coefficients are listed in Table 1.

Cell Partitioning Criteria

A scheme has been developed for allocating the nodes for the simulation. The main factors determining the choice of the partitioning scheme are:

(1) The size of the electron scattering volume which is directly related to the electron beam energy and its incident angle;

(2) The positions where variables are expected to change drastically; for example, locations within the generation volume, near the top surface and defect region etc.;

(3) The bulk diffusion length;

(4) The desired accuracy of the solutions.

Each of these factors plays an important role in node allocation. No general rules or empirical formulae have evolved regarding cell partitioning owing to the fact that each simulation model is unique in its own area of interest. In this model, the size of the generation volume and the bulk diffusion length are directly related to the physical dimensions of the sample volume of interest. The distance from the edge of the generation volume to the sample boundaries would be approximately five carrier diffusion lengths to ensure that the distributed excess carrier density is sufficiently small to be ignored. Hence, the size of the Monte Carlo profile and the value of L can provide a quick guide in estimating the size of the sample volume required for the simulation.

In the actual programme flow, the choice of grid spacings and the size of simulation volume is based on the following criteria:

(1) Fine grid spacings are employed within and near the generation volume, and they become progressively coarse as the nodes move away from the generation edge. A good estimation of the fine grid size is less than L/10;

(2) Fine grid spacings should also be used near the top surface and in the vicinity of the defect location;

(3) Each of the boundary conditions mentioned previously should be at least five diffusion lengths away from the nearest edge of the generation volume; and

(4) The selection of a set of grid spacings is dependent on the desired accuracy. This will be discussed in a later section.

Figure 3 shows a typical choice of grid spacings and the size of the sample volume based on the above-mentioned imposed conditions.

Formation of the cathodoluminescence signal

Photons are generated within the material when the excess minority carriers recombine radiatively. It is assumed that the luminescent intensity of light produced by radiative recombination at a node defined in the numerical model is proportional to the integral of the density of the excess carriers over its volume minus the optical losses.

Basically there are three main mechanisms causing the light output reduction, namely total internal reflection, material self-absorption and Fresnel loss [12]. The discrete model formulated by Phang *et al.* [12] for calculating the fraction of radiative loss of generated photon flux through the three mechanisms has been employed. The net CL emission from the surface is obtained by considering each angle of light propagation, up to the critical angle, from a particular position in the simulation model and summing up the total contributions. This calculation is then done for each position at which significant excess carriers are present.

Solving Linear System by Successive-Over-Relaxation

As the number of linear equations in the three-dimensional system is large, the successive-over-relaxation (SOR) method has been adopted to solve this set of equations iteratively. When implemented in SOR, eqs. (5) and (6) become

$$p_{0i}^{(m+1)} = P_{0i}^{(m)} + \omega(\tilde{p}_{0i}^{(m+1)} - p_{0i}^{(m)}) \qquad (7)$$

where ω is an over-relaxation factor. The optimum value of ω always lies between 1.0 and 2.0.

Criterion for convergence

Since SOR is a special case of the Gauss-Seidal iteration, the convergence criteria for the Gauss-Seidal method have to be followed for the proper implementation of the SOR method. In the above formulation, eq. (1) can be rewritten as

$$MP = F(p,x,y,z,g,e,\gamma) \tag{8}$$

where **M** is a matrix containing the coefficients of the sets of linear equations,

$$\begin{aligned} P = [\ & p_{01} \quad p_{02} \quad \cdots\cdots \quad p_{0m} \\ & p_{0m+1} \quad p_{0m+2} \quad \cdots\cdots \quad p_{0mxn} \\ & p_{0mxn+1} \quad p_{0mxn+2} \quad \cdots\cdots \quad p_{0r}\ ]^T \end{aligned} \tag{9}$$

$$\begin{aligned} & F(p,x,y,z,g,e,\gamma) = \\ & [\ F(p_{01},x_1,y_1,z_1,g_{01},e_{01},\gamma_{01}) \\ & F(p_{02},x_2,y_1,z_1,g_{02},e_{02},\gamma_{02}) \\ & \cdots\cdots\ F(p_{0m},x_m,y_1,z_1,g_{0m},e_{0m},\gamma_{0m}) \\ & F(p_{0m+1},x_1,y_2,z_1,g_{0m+1},e_{0m+1},\gamma_{0m+1}) \\ & F(p_{0m+2},x_2,y_2,z_1,g_{0m+2}e_{0m+2},\gamma_{0m+2}) \\ & \cdots\cdots\ F(p_{0mxn},x_m,y_n,z_1,g_{0mxn},e_{0mxn},\gamma_{0mxn}) \\ & F(p_{mxn+1},x_1,y_1,z_2,g_{0mxn+1},e_{0mxn+1},\gamma_{0mxn+1}) \\ & F(p_{mxn+2},x_2,y_1,z_2,g_{0mxn+2},e_{0mxn+2},\gamma_{0mxn+2}) \\ & \cdots\cdots\ F(p_{0r},x_m,y_n,z_q,g_{0r},e_{0r},\gamma_{0r})\ ]^T \end{aligned} \tag{10}$$

and r = m x n x q, where m, n and q are the number of nodes in the X-, Y- and Z-directions respectively.

A sufficient condition for the Gauss-Seidal method to converge is that **M** must be strictly diagonally dominant, i.e.,

$$|A_i| > |B_i| + |C_i| + |D_i| + |E_i| + |F_i| + |G_i| \qquad i = 1,2,3.....r \tag{11}$$

When this is true, p_{oi} will converge to the solution no matter what initial vector is used.

However, it can be deduced from eqs. (I.6)-(I.8) that

$$|A_i| = |B_i| + |C_i| + |D_i| + |E_i| + |F_i| + |G_i| \qquad i = 1,2,3.....r \tag{12}$$

This implies that **M** is not strictly diagonally dominant, and therefore the SOR method may not converge.

Linearization to obtain convergence

In order to obtain convergence, it is necessary to modify the formulation so as to make the coefficient matrix **M** strictly diagonally dominant. Let

$$p_{0i} = \bar{p}_{0i} + \Delta p_{0i} \tag{13}$$

where $\bar{p}_{oi}$ is the current best available approximation for p_{oi}, and Δp_{oi} is the difference required to make p_{oi} exact. Applying eq. (13) to eq. (1), the left-hand side (LHS) and right-hand side (RHS) of eq. (1) respectively become,

$$\begin{aligned} \frac{\partial^2 p_{0i}}{\partial x^2} + \frac{\partial^2 p_{0i}}{\partial y^2} + \frac{\partial^2 p_{0i}}{\partial z^2} = & \frac{\partial^2 \bar{p}_{0i}}{\partial x^2} + \frac{\partial^2 \bar{p}_{0i}}{\partial y^2} + \frac{\partial^2 \bar{p}_{0i}}{\partial z^2} \\ & + \frac{\partial^2 \Delta p_{0i}}{\partial x^2} + \frac{\partial^2 \Delta p_{0i}}{\partial y^2} + \frac{\partial^2 \Delta p_{0i}}{\partial z^2} \end{aligned} \tag{14}$$

and

$$-\frac{g_{0i}}{D} + \frac{p_{0i}}{L_{0i}^2} + \gamma_{0i} e_{0i} p_{0i} = -\frac{g_{0i}}{D} + \left[\frac{1}{L_{0i}^2} + \gamma_{0i} e_{0i}\right] \bar{p}_{0i} + \left[\frac{1}{L_{0i}^2} + \gamma_{0i} e_{0i}\right] \Delta p_{0i} \tag{15}$$

Based on the expression in eq. (I.8) for $\nabla^2 p_{oi}$, a similar expression can be derived for $\nabla^2 \bar{p}_{oi}$ and $\nabla^2 \Delta p_{oi}$ respectively as follows,

$$\nabla^2 \bar{p}_{0i} = -A_i \bar{p}_{oi} + B_i \bar{p}_{1i} + C_i \bar{p}_{2i} + D_i \bar{p}_{3i} + E_i \bar{p}_{4i} + F_i \bar{p}_{5i} + G_i \bar{p}_{6i} \tag{16}$$

and

$$\nabla^2 \Delta p_{0i} = -A_i \Delta p_{oi} + B_i \Delta p_{1i} + C_i \Delta p_{2i} + D_i \Delta p_{3i} + E_i \Delta p_{4i} + F_i \Delta p_{5i} + G_i \Delta p_{6i} \tag{17}$$

Now, substituting eqs. (16) and (17) into eqs. (14) and (15), and rearranging, Δp_{oi} becomes

$$\Delta p_{0i} = \frac{-T1\ \bar{p}_{0i} + T2 + T3 + \dfrac{g_{0i}}{D}}{T1} \tag{18}$$

where

$$T1 = A_{iHx} + \frac{1}{L_{0i}^2} + \gamma_{0i} e_{0i} \tag{19}$$

$$T2 = B_i \bar{p}_{1i} + C_i \bar{p}_{2i} + D_i \bar{p}_{3i} + E_i \bar{p}_{4i} + F_i \bar{p}_{5i} + G_i \bar{p}_{6i} \tag{20}$$

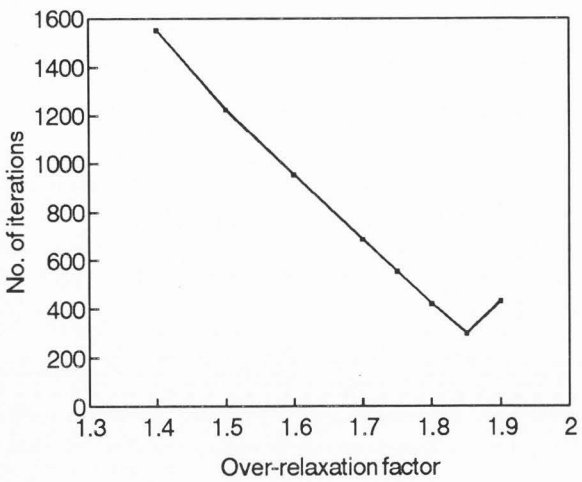


Figure 4. Effect of the successive-over-relaxation factor on the speed of convergence of the numerical model. Other relevant parameters are L = 1 μm, θ_o = 0°, $V_s \rightarrow \infty$, $\gamma(r) = 0\ \mu m^{-2}$ and $Z_m = 0\ \mu$m. The radius of the beam $\delta = 0\ \mu$m and D = $1*10^9$ cm^2/sec.

$$T3 = B_i \Delta p_{1i} + C_i \Delta p_{2i} + D_i \Delta p_{3i} + E_i \Delta p_{4i} + F_i \Delta p_{5i} + G_i \Delta p_{6i} \quad (21)$$

If f_r represents the RHS of eq. (18), then for each node, the steps for iteration are

(1) $\Delta\tilde{p}_{oi} = f_r$;

(2) $\Delta p_{oi} = \omega\ (\Delta\tilde{p}_{oi})$; and

(3) $p_{oi} = p_{o1} + \Delta p_{oi}$

p_{oi} is updated after every Δp_{oi} is calculated. This updating has the objective of making Δp_{1i} to Δp_{6i} zero [9]. Therefore f_r in step 1 could be simplified by removing all factors containing Δp_{1i} to Δp_{6i}. This formulation also makes the coefficient matrix diagonally dominant since

$$\left| A_i + \frac{1}{L_{0i}^2} + \gamma_{0i} e_{0i} \right| > |B_i| + |C_i| + |D_i| + |E_i| + |F_i| + |G_i| \quad (22)$$

and L_{oi} and $\gamma_{oi} > 0$, and $e_{oi} \geq 0$.

Results

When the above formulation was implemented on a digital computer, the solution converged readily. The relaxation factor ω was determined by trial and error. Figure 4 illustrates some of the results showing how the successive-over-relaxation method can speed up the convergence in the three-dimensional calculation of the excess carrier density generated by a 20 keV electron beam. The number of grid points used in the X-, Y- and Z-directions are 34, 34 and 36 respectively. The grid spacings start from 0.1 μm and are progressively increased to 0.4 μm and then to 0.8 μm. For this example, the iteration continued until the maximum change represented by TOL in any component of $p_{oi}^{(m+1)}$ was less than 1% variation of $p_{oi}^{(m)}$.

Accuracy

Discretization errors

The discretization error is a good measure of the accuracy of the solution of the continuity equation. A method for determining this error for the approximated function of eq. (1) is described in Appendix II.

It can be observed from eq. (II.5) that the discretization error is a function of the grid spacings and the higher derivatives of the function. This means that the larger the grid size, the less accurate the approximation will be. The higher derivatives give an indication of how fast the function is varying spatially. The faster the change, the closer is the grid spacing required to maintain a given accuracy [9]. In addition, since the discretization error is solely determined by the particular numerical solution procedure selected and is independent of computing equipment characteristics, it is a good indicator for selecting a set of optimum grid sizes of a system for a prespecified error criterion.

The actual program uses the normalized discretization error ($Nd\epsilon_{oi}$), which is evaluated using eq. (II.7) in Appendix II. Assuming the defect is not present in the system, i.e., $\gamma(r) = 0\ \mu m^{-2}$, eq. (1) is simplified to $\nabla^2 p(r) = -g(r)/D + p(r)/L^2$. The third and fourth derivatives of p(r) required in the calculation of $Nd\epsilon_{oi}$ can therefore be expressed in terms of the first and second derivatives of the RHS of the above expression [9]. This is implemented in the program to avoid the formidable task of having to numerically evaluate the third and fourth derivatives of p(r) using the finite difference scheme, as dictated by eq. (II.5).

To show the effect of the grid spacing on the maximum $Nd\epsilon_{oi}$ and on the number of loops for iteration, two different grid systems, one with grid spacings half the value of the other, were used for a 20 keV electron beam impinging on a defect free GaAs material with L = 1 μm, $V_s \rightarrow \infty$, $\gamma(r) = 0\ \mu m^{-2}$ and $Z_m = 0\ \mu$m. Other relevant parameters are $\theta_o = 0°$, $\omega = 1.85$ and TOL = 0.1%. The results are tabulated in Table 2.

The simulation was performed on a NEC-SX1A supercomputer. Double precision accuracy was used throughout this work. Table 3 shows the $Nd\epsilon_{oi}$ recorded

Table 2. Discretization errors of the test grid systems.

Grid system			1			2		
NX, NY, NZ			40,42,44			20,21,22		
Grid directions	X	Node numbering	0 -28	29 -34	35 - 40	0 - 14	15 - 17	18 -20
		Size (μm)	0.1	0.4	0.8	0.2	0.8	1.6
	Y	Node numbering	0 - 30	31 - 36	37 - 42	0 - 15	16 - 18	19 - 21
		Size (μm)	0.1	0.4	0.8	0.2	0.8	1.6
	Z	Node numbering	0 - 32	33 - 38	39 - 44	0 - 16	17 - 19	20 - 22
		Size (μm)	0.1	0.4	0.8	0.2	0.8	1.6
No. of iterations			349			293		
Maximum $\lvert Nd\epsilon_{oi} \rvert$			3.88%			9.10%		

NX,NY,NZ = Number of nodes in the X-, Y- and Z-directions respectively

Table 3. Discretization errors at some of the strategic locations.

Nodal point	Location (x,y,z)(μm)	$\lvert Nd\epsilon_{oi} \rvert$ (%)	
		NX, NY, NZ 20, 21, 22	NX, NY, NZ 40, 42, 44
1	0.0, 0.0, 0.4	1.06	0.39
2	0.0, 0.2, 0.2	0.84	0.21
3	0.0, 0.4, 0.2	0.87	0.11
4	0.0, 0,2, 0.4	0.07	0.03

at some of the strategic locations of the simulation model for the same conditions as in Table 2. The four points are identical in the two sets of equations. In principle, since the error in approximating the derivatives by central difference quotients is $O(h^3)$, which is the order relation signifying that the error is proportional to h^3 as $h \rightarrow 0$ [3], the anticipated errors in the solution through sets of difference equations would also vary as h^3. The results in Table 3 reflect this behaviour with reasonably good agreement.

Comparison with Analytical Solutions

In this section, the solutions of the developed numerical formulation for the diffusion equation, hereafter named as CL Model, are compared with the analytical answers of some models.

Test Model I

This is a model used to measure steady-state photoconductivity. Assuming that the sample is in the form of a rectangular solid of dimensions x_o, y_o, and z_o,

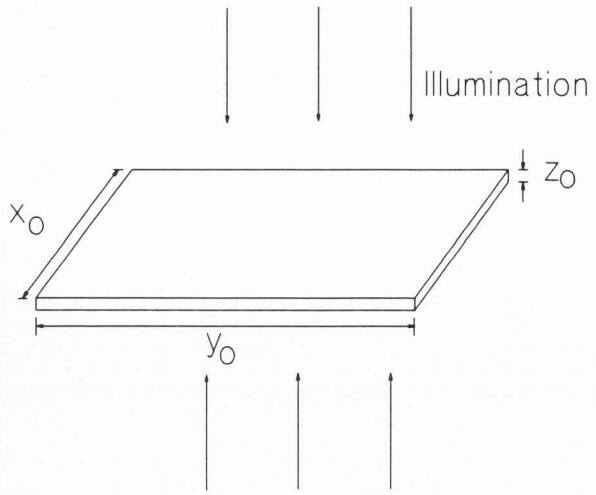


Figure 5. Geometry used to measure dc photoconductivity.

---

where the thickness z_o is much less than the other dimensions x_o and y_o. The illumination is assumed to be incident along the Z-direction, as shown in Figure 5. It is also assumed that both the top and bottom surfaces have the same surface recombination velocity. If the absorption coefficient is relatively small, the light intensity may lead to a generation rate g′ of excess carriers which is constant throughout the sample. This situation can be modelled by a one-dimensional continuity along the Z-direction. Putting $\gamma(r) = 0\ \mu m^{-2}$ and $g(r) = g'$, eq. (1) takes the form

$$\frac{d^2p}{dz^2} = -\frac{g'}{D} + \frac{p}{L^2} \qquad (23)$$

The solution of eq. (23), subject to the top and bottom boundary conditions, is [8]

$$p(z) = g'\tau\left[1 - \frac{v_s \cosh\frac{z}{L}}{v_s \cosh\frac{z_o}{2L} + \frac{D}{L}\sinh\frac{z_o}{2L}}\right] \qquad (24)$$

The geometry shown in Figure 5 was implemented in CL Model. The value of z_o was chosen to be 1 μm. To simulate constant generation throughout the sample, the input function g(r) of the set of numerical equations at each of the nodes was set to g′. To ensure that x_o and y_o were much greater than z_o, the boundaries at the planes $x = x_o/2$, $x = -x_o/2$, $y = y_o/2$ and $y = -y_o/2$ were implemented by simply assuming that all the nodes in these planes also have the generation rate g′, and $p_{0(xo/2)} = p_{0(xo/2-\theta 1h)}$, $p_{0(-xo/2)} = p_{0(-xo/2+\theta 2h)}$, $p_{0(yo/2)} = p_{0(yo/2-\theta 3h)}$ and $p_{0(-yo/2)} = p_{0(-yo/2+\theta 4h)}$. Constant grid sizes of 0.5 μm, 0.5 μm and 0.1 μm were used in the X-, Y- and Z-directions respectively in this model. Other relevant parameters are L = 1.0 μm, Z_m = 0 μm, $\gamma(r) = 0\ \mu m^{-2}$ and D = 1 x 10^9 μm^2/sec. The concentration profiles p(z) of the simulated and analytical answers are shown in Figure 6 for several values of the normalized v_s, i.e., V_s. In all the cases, the numerical approximations converged to their solutions without problem and a maximum discrepancy of less than 1% has been observed for the case of $V_s \rightarrow \infty$.

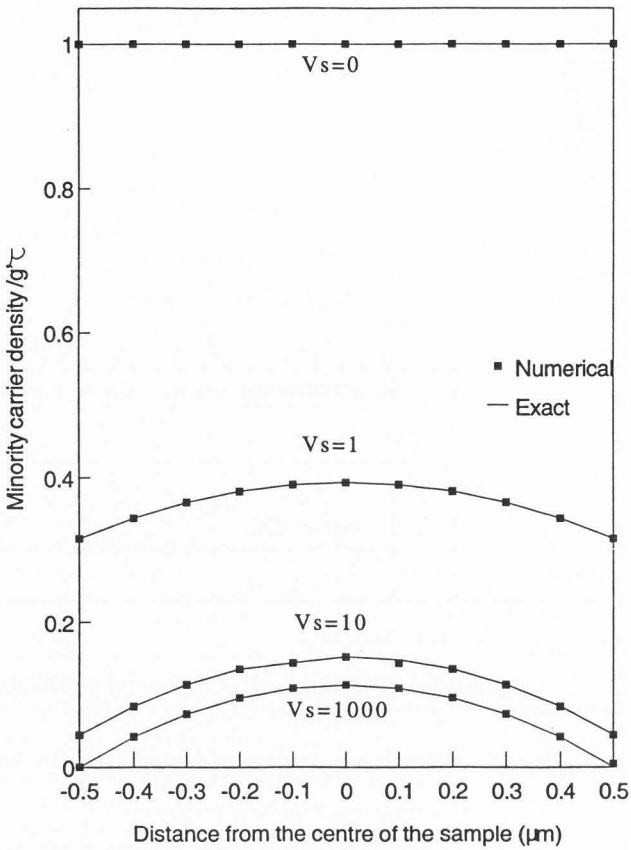


Figure 6. Comparison of concentration profiles calculated using eq. (24) and CL Model within the sample, assuming to be GaAs, illustrated in Figure 5 for different values of V_s.

Comparison between CL Model and the CL model of Phang *et al.* [12]

This test was performed by considering a homogeneous semiconductor in both CL Model and the CL model proposed by Phang *et al.* [12]. A generation point source of identical strength was assumed to be located at

Table 4. Comparison of CL Model and the CL model by Phang *et al.* [12].

V_s	Location (x,y,z) (μm)	$p(x,y,z)/g'\tau$		| error | (%)
		Phang *et al.*'s CL model [12]	CL Model	
0	0.0, 0.0, 2.5	0.0642532	0.0643698	0.18
	0.0, -2.5, 2.5	0.0021453	0.0021823	1.72
	0.0, 5.0, 5.0	0.0003538	0.0003487	1.45
$\rightarrow \infty$	0.0, 0.0, 2.5	0.0634492	0.0635628	0.18
	0.0, -2.5, 2.5	0.0015546	0.0015949	2.59
	0.0, 5.0, 5.0	0.0002368	0.0002336	1.33

(0,0,2.5 μm) in both models. Simulations were carried out in GaAs with L = 3 μm, D = 1 x 10^9 μm^2/sec and a grid spacing of 0.5 μm in all the directions for V_s = 0 and $V_s \rightarrow \infty$. The distributed carrier density at three identical locations in the two models are tabulated in Table 4. Both models show identical performance in the evaluation of p(r).

Conclusions

A three-dimensional numerical cathodoluminescence model has been developed for the study of defects in semiconductors. The representation of the continuity equation by its numerical equivalence is useful not only in incorporating realistic electron-hole pair generation obtained from Monte Carlo calculations but also in simulating defect properties. The linearization technique for obtaining convergence during the successive-over-relaxation calculation works satisfactorily. The accuracy tests showed that the numerical scheme employed is accurate, reliable and stable in evaluating the carrier distribution which is important for calculating cathodoluminescence emissions.

Acknowledgements

Research funding from the National University of Singapore under Research Project RP900614 is acknowledged. One of the authors, KL Pey, gratefully acknowledges the support given by National University of Singapore through a Research Scholarship.

References

[1] Czyżewski Z and Joy DC (1990) Monte Carlo simulation of CL and EBIC contrasts for isolated dislocations. Scanning **12**, 5-12.

[2] Donolato C (1978) On the study of SEM charge-collection imaging of localized defects in semiconductors. Optik **52**, 19-36.

[3] Gerald GF and Wheatley PO (1984) Numerical solution of elliptic partial differential equation. In: Applied Numerical Analysis, Addison-Wesley, Singapore. pp. 399-447.

[4] Hockney RW and Eastwood JW (1981) The particle-mesh force calculation. In: Computer Simulation Using Particles. McGraw-Hill, New York. pp. 120-165.

[5] Jakubowicz A (1986) Theory of cathodoluminescence contrast from localized defects in semiconductors. J. Appl. Phys. **59**, 2205-2209.

[6] Löhnert K and Kubalek E (1984) The cathodoluminescence contrast formation of localized non-radiative defects in semiconductors. Phys. Stat. Sol. **a83**, 307-314.

[7] dé Mari A (1968) An accurate numerical steady-state one-dimensional solution of the *P-N* junction. Solid-State Electronics **11**, 33-58.

[8] McKelvey JP (1966) Excess carrier in semiconductors. In: Solid-State and Semiconductor Physics. Harper and Row, London. pp. 320-368.

[9] Ong VKS (1988) Two-dimensional Modelling of the Solar Cell, M. Eng. Thesis, National University of Singapore.

[10] Ortega JM and Poole Jr WG (1981) The course of dimensionality. In: An Introduction to Numerical Methods for Differential Equations, Pitman Pub. Inc., Marshfield, MA. pp. 850-851.

[11] Pasemann L (1981) A contribution to the theory of the EBIC contrast of lattice defects in semiconductors. Ultramicroscopy **6**, 237-250.

[12] Phang JCH, Pey KL and Chan DSH (1992) A simulation model for cathodoluminescence in the

scanning electron microscope. IEEE Trans. Electron Devices **ED-39**, 782-791.

[13] Sze SM (1981) LED and semiconductor lasers. In: Physics of Semiconductor Devices (2nd ed.), John Wiley & Sons, New York. pp. 681-739; also, Appendix H - Properties of Ge, Si, and GaAs. pp. 850-851.

Discussion with Reviewers

J.F. Breese: Does your model take into account the recombination due to the reabsorption of the emitted light which may have an influence on the density of minority carriers? From this point of view, the boundary limits are taken as 5 times the diffusion length, are they enough?

Authors: The present treatment does not take into account the recombination due to the reabsorbed recombination radiation at the moment. According to Von Roos (1983), the influence of the reabsorbed recombination radiation on carrier transport is unimportant at low doping levels and the effect, therefore, can be incorporated into the model without significant modification of the formulation. The boundary limits may then have to be set at more than five times the diffusion length.

J.F. Breese: Your Monte Carlo calculations give an electron range which is underestimated as compared to values given by Kanaya's formula (K. Kanaya, S. Okoyama. J. Phys. D, 3, 43, (1972)) which is in agreement with experimental measurements. Does this difference modify your estimated values?

Authors: Although Kanaya and Okoyama's formulae give better prediction for the electron range, the calculated depth dose functions (see Fig. 12 of Kanaya and Okoyama) do not provide information on the energy dissipation in the radial direction and, more importantly, they do not agree well with the experimental data quantitatively. On the other hand, Monte Carlo calculations are able to determine the radial energy dissipation distribution. Contours of equal energy dissipation determined experimentally and calculated by Monte Carlo method have been shown to be in good agreement (Shimizu *et al.*, 1975). Our simulation results have shown that the CL contribution from the depths near and beyond the maximum penetration range is relatively insignificant, especially for materials with high self-absorption. As a result, we believe that the difference in the electron range between the KO and our model would not significantly modify the calculated values.

S. Myhajlenko: How readily can a spatially varying bulk lifetime be incorporated into the numerical formulation? In Part II, you effectively do this with dot-halo contrast by varying the defect strength $\gamma(r)$ at and away from the defect: I was thinking of the case of SI-GaAs where lifetime difference has been determined within cells and cell boundaries.

Authors: Theoretically, the CL contrast of a localized defect can be calculated using Eq. (6) in Part II which is dependent on the lifetime of the recombination processes. In this numerical model, using Eq. (3) in Part I and Eq. (5) in Part II, a spatial variation in bulk lifetime can be implemented easily as different lifetimes can be set for each node.

Additional References

Shimizu R, Ikuta T, Everhart TE and DeVore WR (1975). Experimental and theoretical study of energy dissipation profiles of keV elctrons in polymethylmethacrylate. J. Appl. Phys. **46**, 1581-1584.

Von Roos O (1983) Influence of radiative recombination on the minority-carrier transport in direct band-gap semiconductors. J. Appl. Phys. **54**, 1390-1398.

Appendices

I. Discretization of the second order derivatives and derivative boundary conditions

Considering the X-direction, the first derivatives are approximated by:

$$\frac{\partial}{\partial x}p_{0i\,1i} = \frac{p_{0i} - p_{1i}}{\theta_{1i}h}; \qquad \frac{\partial}{\partial x}p_{2i\,0i} = \frac{p_{2i} - p_{0i}}{\theta_{2i}h} \tag{I.1}$$

Since

$$\frac{\partial^2 p}{\partial x^2} = \frac{\partial}{\partial x}\left[\frac{\partial p}{\partial x}\right]$$

this gives

$$\frac{\partial^2}{\partial x^2}p_{0i} = \frac{2}{h^2(\theta_{1i}+\theta_{2i})}\left[\frac{p_{2i}-p_{0i}}{\theta_{2i}} - \frac{p_{0i}-p_{1i}}{\theta_{1i}}\right] \tag{I.2}$$

Similarly, for the Y- and Z-directions,

$$\frac{\partial^2}{\partial y^2}p_{0i} = \frac{2}{h^2(\theta_{3i}+\theta_{4i})}\left[\frac{p_{4i}-p_{0i}}{\theta_{4i}} - \frac{p_{0i}-p_{3i}}{\theta_{3i}}\right] \tag{I.3}$$

$$\frac{\partial^2}{\partial z^2}p_{0i} = \frac{2}{h^2(\theta_{5i}+\theta_{6i})}\left[\frac{p_{6i}-p_{0i}}{\theta_{6i}} - \frac{p_{0i}-p_{5i}}{\theta_{5i}}\right]$$

Combining eqs. (I.2)-(I.4), the second derivative can be expressed as

$$\begin{aligned}\nabla^2 p_{0i} &= \frac{\partial^2 p_{0i}}{\partial x^2} + \frac{\partial^2 p_{0i}}{\partial y^2} + \frac{\partial^2 p_{0i}}{\partial z^2} \\ &= \frac{2}{h^2(\theta_{1i}+\theta_{2i})}\left[\frac{p_{2i}-p_{0i}}{\theta_{2i}} - \frac{p_{0i}-p_{1i}}{\theta_{1i}}\right] + \\ &\quad \frac{2}{h^2(\theta_{3i}+\theta_{4i})}\left[\frac{p_{4i}-p_{0i}}{\theta_{4i}} - \frac{p_{0i}-p_{3i}}{\theta_{3i}}\right] + \\ &\quad \frac{2}{h^2(\theta_{5i}+\theta_{6i})}\left[\frac{p_{6i}-p_{0i}}{\theta_{6i}} - \frac{p_{0i}-p_{5i}}{\theta_{5i}}\right]\end{aligned}$$

Substituting

$$B_i = \frac{2}{h^2\theta_{1i}(\theta_{1i}+\theta_{2i})};$$

$$C_i = \frac{2}{h^2\theta_{2i}(\theta_{1i}+\theta_{2i})};$$

$$D_i = \frac{2}{h^2\theta_{3i}(\theta_{3i}+\theta_{4i})};$$

$$E_i = \frac{2}{h^2\theta_{4i}(\theta_{3i}+\theta_{4i})};$$

$$F_i = \frac{2}{h^2\theta_{5i}(\theta_{5i}+\theta_{6i})};$$

$$G_i = \frac{2}{h^2\theta_{6i}(\theta_{5i}+\theta_{6i})};$$

and

$$A_i = B_i + C_i + D_i + E_i + F_i + G_i$$

into eq. (I.5), the second derivative becomes

$$\begin{aligned}\nabla^2 p_{0i} &= -A_i p_{oi} + B_i p_{1i} + C_i p_{2i} \\ &\quad + D_i p_{3i} + E_i p_{4i} + F_i p_{5i} + G_i p_{6i}\end{aligned}$$

With reference to Figure 2 and using the central difference quotient, the approximated derivative boundary condition can be expressed as

$$\frac{\partial}{\partial z}p_{0i} = \frac{p_{6i}-p_{5i}}{h(\theta_{6i}+\theta_{5i})}$$

This approximation is written at $z_i = 0$ and p_{5i} is the fictitious point located at point $z_i < 0$. Using the difference quotient, the fictitious p_{5i} values can be expressed in terms of points within the network. Substituting eq. (I.9) into eq. (4), the fictitious point p_{5i} can be written as

$$\begin{aligned}p_{5i} &= p_{6i} - \frac{v_s\, p_{0i} h}{D}(\theta_{6i}+\theta_{5i}) \\ &= p_{6i} - Q_i p_{0i}\end{aligned}$$

where

$$Q_i = 2\frac{v_s h}{D}\theta_{6i} \text{ with } \theta_{6i} = \theta_{5i}$$

II. Derivation of discretization errors

Using the notations in Figure 2 for the three-dimensional system

$$\begin{aligned}p''(x_{0i}, y_{0i}, z_{0i}) &= \frac{\partial^2 p_{0i}}{\partial x^2} + \frac{\partial^2 p_{0i}}{\partial y^2} + \frac{\partial^2 p_{0i}}{\partial z^2} \\ &= \frac{2}{h^2(\theta_{1i}+\theta_{2i})}\left[\frac{p_{2i}-p_{0i}}{\theta_{2i}} - \frac{p_{0i}-p_{1i}}{\theta_{1i}}\right] - d\varepsilon x_{0i} + \\ &\quad \frac{2}{h^2(\theta_{3i}+\theta_{4i})}\left[\frac{p_{4i}-p_{0i}}{\theta_{4i}} - \frac{p_{0i}-p_{3i}}{\theta_{3i}}\right] - d\varepsilon y_{0i} + \\ &\quad \frac{2}{h^2(\theta_{5i}+\theta_{6i})}\left[\frac{p_{6i}-p_{0i}}{\theta_{6i}} - \frac{p_{0i}-p_{5i}}{\theta_{5i}}\right] - d\varepsilon z_{0i}\end{aligned} \quad \text{(II.1)}$$

where $d\varepsilon x_{0i}$, $d\varepsilon y_{0i}$ and $d\varepsilon z_{0i}$ are the discretization errors in the X-, Y- and Z-directions respectively.

Consider a typical one-dimensional case, along the X-axis, by a Taylor series expansion,

$$p(x_{0i}+h_R) = p(x_{0i}) + p'(x_{0i})h_R + \frac{p''(x_{0i})}{2!}h_R^2 + \frac{p'''(x_{0i})}{3!}h_R^3 + \frac{p''''(x_{0i})}{4!}h_R^4 + d1 \quad \text{(II.2)}$$

where

$$d1 = \sum_{i=5}^{\infty} \frac{p^{(i)}(x_{0i})}{i!} h_R^i$$

and $h_R = \theta_{2i}h$.
Similarly

$$p(x_{0i}-h_L) = p(x_{0i}) - p'(x_{0i})h_L + \frac{p''(x_{0i})}{2!}h_L^2 - \frac{p'''(x_{0i})}{3!}h_L^3 + \frac{p''''(x_{0i})}{4!}h_L^4 + d2 \quad \text{(II.3)}$$

where

$$d2 = \sum_{i=5}^{\infty} (-1)^i \frac{p^{(i)}(x_{0i})}{i!} h_L^i$$

and $h_L = \theta_{1i}h$.

Substituting eqs. (II.2) and (II.3) into the one-dimensional derivative,

$$p''(x_{0i}) = \frac{2p(x_{0i}-h_L)}{h_L(h_R+h_L)} + \frac{2p(x_{0i}+h_R)}{h_R(h_R+h_L)} - \frac{2p(x_{0i})}{h_R h_L} - d\varepsilon x_{0i} \quad \text{(II.4)}$$

where

$$d\varepsilon x_{0i} = \frac{p'''(x_{0i})}{3}(h_R-h_L) + \frac{p''''(x_{0i})}{12}(h_L^2 - h_R h_L + h_R^2) + \frac{2}{h_L+h_R}\sum_{i=5}^{\infty} \frac{p^i(x_{0i})}{i!}(h_R^{i-1} + (-1)^i h_L^{i-1}) \quad \text{(II.5)}$$

The same approach can be applied to derive $d\epsilon x_{0i}$, $d\epsilon y_{0i}$ and $d\epsilon z_{0i}$ in the three-dimensional system. For example, the second partial derivative with respect to x, $\partial^2 p_{0i}/\partial x^2$ is derived by holding y and z constant and evaluating the function at x equals x_{0i}, $x_{0i}+h_R$ and $x_{0i}-h_L$. The partial derivatives $\partial^2 p_{0i}/\partial y^2$ and $\partial^2 p_{0i}/\partial z^2$ are similarly computed, holding x and z, and x and y constant respectively. Therefore, expressions for $d\epsilon y_{0i}$ and $d\epsilon y_{0i}$ can also be obtained easily using eq. (II.5).

Putting the derived discretization errors into eq. (II.1) and rearranging,

$$p_{0i} = [-p''(x_{0i},y_{0i},z_{0i}) + B_i p_{1i} + C_i p_{2i} + D_i p_{3i} + E_i p_{4i} + F_i p_{5i} + G_i p_{6i} - d\epsilon_{0i}] / A_i \quad \text{(II.6)}$$

where $d\epsilon_{0i} = d\epsilon x_{0i} + d\epsilon y_{0i} + d\epsilon z_{0i}$.

The normalized discretization error is expressed as

$$Nd\varepsilon_{0i} = \frac{(d\varepsilon_{0i})/A_i}{p_{0i}} \quad \text{(II.7)}$$