SIMULATION OF SEMICONDUCTOR DEVICES AND PROCESSES Vol. 5
Edited by S. Selberherr, H. Stippel, E. Strasser - September 1993

Newton-GMRES Method for Coupled Nonlinear Systems Arising in Semiconductor Device Simulation

C. Simon[‡†], M. Sadkane[†], and S. Mottet[‡]

[‡]CNET Lannion B, France Telecom
Route de Trègastel, BP 40, F-22301 Lannion Cédex, FRANCE
[†]IRISA
Campus de Beaulieu, F-35042 Rennes Cédex, FRANCE

Abstract

We are interested in computing the solution of a system of coupled nonlinear PDE's which describes the electrical behaviour of semiconductor devices. This set of nonlinear equations is solved via a nonlinear version of the GMRES method [2]. This method consists in solving the linear system, that arises in Newton's method, by an iterative scheme, which constructs an orthonormal basis of a Krylov subspace, and minimizes the residual, over the current Krylov subspace. An advantage of this method over the classical ones is that the Jacobian is not stored and that little storage is required since the method restarts periodically whenever the size of the Krylov subspace reaches a maximum value fixed by the user.

1. Introduction

The classical steady state isothermal drift-diffusion conduction model in semiconductors is basically described by a set of three nonlinear PDE's (Poisson's equation, electron and hole continuity equations) :

$$div(\epsilon\nabla(\varphi)) = \rho \; ; \quad -\frac{1}{q}div(J_n) = -U \; ; \quad \frac{1}{q}div(J_p) = -U \tag{1}$$

The symbols have their usual meaning of the semiconductor device theory. ϵ is the dielectric constant, φ the electrostatic potential, $U \equiv U(\varphi,\ \varphi_n,\ \varphi_p)$ the net recombination-generation rate where φ_n and φ_p are the electrochemical potentials of electron and hole. The expression of the charge density $\rho \equiv \rho(\varphi,\ \varphi_n,\ \varphi_p)$ is : $\rho = n - p - dop$, where dop is the residual doping level, $n \equiv n(\varphi,\ \varphi_n)$ and $p \equiv p(\varphi,\ \varphi_p)$ are the electron and hole free carrier densities. The drift-diffusion current densities are : $J_n = -q.n.\mu_n.\nabla(\varphi_n)$ and $J_p = -q.p.\mu_p.\nabla(\varphi_p)$ where $\mu_n \equiv \mu_n(\varphi,\ \varphi_n)$ and $\mu_p \equiv \mu_p(\varphi,\ \varphi_p)$ are the electron and hole mobilities.

The numerical solution of system (1) is carried out by discretizing the equations on a mesh. We have implemented a Flux Conservative Box Method (FCBM) scheme [6], which is a variant of the so called "Box Method". This transforms the system of PDE's (1) into a nonlinear algebraic system. We consider that the potentials φ,

φ_n and φ_p are the fundamental entities which characterize the device behaviour. So, this leads to a natural choice of unknowns $\varphi, \varphi_n, \varphi_p$, and the set of equations can be written as :

$$\begin{cases} \nabla.\,(\epsilon\nabla\varphi) - \rho = 0 \\ \nabla.\,(n\mu_n\nabla\varphi_n) + U = 0 \\ \nabla.\,(p\mu_p\nabla\varphi_p) - U = 0 \end{cases} \tag{2}$$

with appropriate boundary conditions. The domain of definition Ω is a meshed bounded domain belonging to $\mathbf{R}^l$ (l=1,2,3). The boundary of the domain, $\Gamma = \partial\Omega$, is divided into classes, each of them corresponding to a given boundary condition type (ohmic contact, Schottky contact, insulating boundary...).

For the sake of conciseness, the expressions of the coefficients in (2) are not reported. However, we can point out that, depending on the expressions used to describe the mobility law, the recombination-generation term (Shockley-Hall-Read, Auger, spontaneous band to band, impact ionization ...) and the statistics (Maxwell-Boltzmann or Fermi-Dirac) of the free carrier densities, this general model is adequate for a wide range of devices including silicon and III-V optoelectronic hetero-junction structures.

2. Newton-GMRES algorithm

We are interested in solving the nonlinear system (2) using a coupled method where the three unknowns φ, φ_n and φ_p are sought simultaneously. The system (2) can be written as a nonlinear mapping :

$$F(\Phi) = 0 \quad \text{where} \quad \Phi = (\varphi,\ \varphi_n,\ \varphi_p)^t \in \mathbf{R}^{3N}. \tag{3}$$

where N stands for the number of nodes of the mesh. Let us denote by $J(\Phi)$ the Jacobian matrix $\nabla F(\Phi)$ computed at vector Φ. Recall that Newton's method for solving the nonlinear equation (3) can be described as follows :

1. Choose an initial guess Φ_1 ;
2. **for** $k = 1, 2, \cdots$ **do**
   (i.) Solve the linear system : $J(\Phi_k)\delta_k = -F(\Phi_k)$;
   (ii.) Correct the approximate solution $\Phi_{k+1} = \Phi_k + \delta_k$;

end for

The Newton-GMRES method consists in solving the linear system $J(\Phi_k)\delta_k = -F(\Phi_k)$ in step (2.i.) using GMRES method [1, 2, 3], which builds iteratively an orthonormal basis of a Krylov subspace, whose dimension m is very small compared to the size of the matrix J, and minimizes the residual over it. Unlike the methods based on Gaussian elimination, this method only requires to compute, at each iteration k, the action of the Jacobian J on a vector. This is accomplished without explicit storage of the Jacobian matrix, using the approximation :

$$J(\Phi_k)\Psi_k \approx \frac{F(\Phi_k + \sigma\Psi_k) - F(\Phi_k)}{\sigma} \tag{4}$$

where σ is a small quantity chosen according to [4]. Moreover, in order to improve the conditioning of the linear system (2.i.), a simple scaling by row is used.

3. Numerical results

We have focused the study on two points which may have an impact on the quality of the method. The first one is the size m of the Krylov subspace, the second one is the parameter σ which appears in approximation (4).

- Concerning the size m of the Krylov subspace, the basic method consists in choosing a value m and then restarting the GMRES loop periodically every m iterations. This strategy is, in general, not recommended because it does not take care of the excessive storage and its numerical instability effect on the algorithm. On the other hand, if m is small enough, the GMRES algorithm restarts more frequently, and this results in a high execution time. We are still working on this point. So far, we have implemented an adaptive method as suggested in [3] and which may be described in the following : let us denote by RESL the norm of the linear residual given by GMRES, and RESN the norm of the nonlinear residual $||F(\Phi_k)||$ where Φ_k stands for the current approximate solution. We fixe a maximum size m_{max} and start with a small value, say $m_{min} \leq 5$, the size m of the Krylov subspace is then updated at each restart according to the following conditions :

  if $(RESN \leq 1.5 * RESL)$ **then** $m := min(m + k, m_{max})$ **endif**
  if $(RESN \geq 5 * RESL)$ **then** $m := max(m - k, m_{min})$ **endif**

  where k is a small integer,$(k \leq 5)$, by which the size m is either increased or decreased. Moreover, within the GMRES loop, we use two different criterions for the periodical restart : either the number of iterations has reached the maximum size m_{max} of the krylov subspace or $RESL < \epsilon$ where ϵ is a precision parameter depending on the machine precision ϵ_{mach} and the norm of the first non linear residual. Although the results obtained with this strategy are rather good, an improvement we are trying to find is a more general autoadaptive method which would adapt automatically the size m.

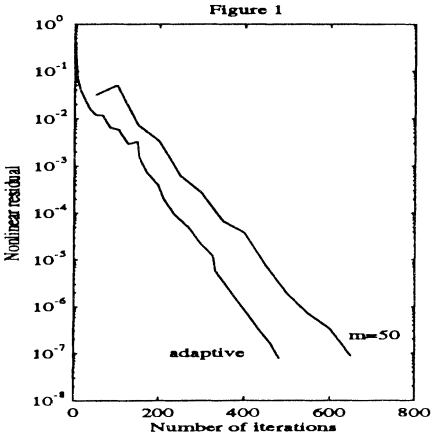
- Concerning the parameter σ in approximation (4), we have compared the choice suggested in [2], that is $\sigma = \sqrt{\epsilon_{mach}} \times ||F(\Phi_k)||$ (method A), with the adaptive one (method B) which has been successfully used in the context of CFD [4]. Both of the methods use the adaptive choice for the size m of the Krylov subspace. The cost of computing σ is equivalent to the cost of two residual evaluations. Consequently, the cost is minor when compared to the overall computation. The method could be improved because it seems not necessary to re-evaluate σ at each restart but rather every 5 or 10 restarts.

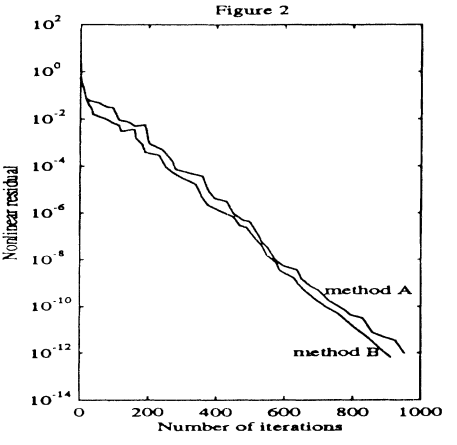
- The simulations have been performed, using the coupled method implemented in the 3D version of the simulator CARMES [5]. Two devices have been simulated. The first one is a classical Silicon pn junction with a $30 \times 30 \times 30$ rectangular mesh. Figure 1 shows the comparison between the static and the adaptive choice of the size of the Krylov subspace m. These results have been obtained under 5 Volts reverse bias (the size of the Jacobian matrix J is $3N = 81000$). We can observe that the adaptive method needs less iterations. The second device is a Buried Heterostructure (BH) used in a Semiconductor Optical Amplifier (SOA). It is mainly composed of a GaInAsP (p doped) active layer inserted into an InP (n doped) buffer layer and an InP (p doped) protection layer. The

simulations are performed with a $20 \times 25 \times 30$ rectangular mesh (the size of the Jacobian matrix is $3N = 45000$), under .95 Volt direct bias. Figure 2 shows the comparison between method A and method B. Here again, the latter one needs less iterations than the former one.

4. Conclusion

We have investigated a Newton-GMRES method for solving nonlinear equations coming from semiconductor device simulation. We have shown that the method gives reasonable results if some of the parameters are chosen adequately. There are two important points left to be done. The first one concerns a further study about the adaptive choice of the Krylov subspace, and the second one concerns the finding of a preconditioner of the Jacobian matrix which would take into account some physical considerations.





References

[1] Y. Saad and M. Schultz. GMRES: A Generalized Minimal Residual Algorithm for solving nonsymmetric linear systems. SIAM J. Sci. and Stat. Comp. Number 7. pp:856-869, 1986.

[2] P. Brown and Y. Saad. Hybrid Krylov methods for solving nonlinear systems of equations. SISSC. vol.11,no.3,pp 450-481,1990.

[3] T. Kerhoven and Y. Saad. On acceleration methods for coupled nonlinear elliptic. Numer. Math.60,525-548 (1992).

[4] Z. Johan. Data parallel finite element techniques for large scale computational fluid dynamics. PhD Thesis. Stanford University. July 1992.

[5] C. Simon, S. Mottet, J. E. Viallet. Autoadaptive Mesh Refinement. SISDEP91. Vol.4, pp225-233. W. Fichtner and D. Aemmer editors.

[6] C. Simon, S. Mottet. FCBM : Flux Conservative Box Method, a new discretization strategy of the semiconductor equations, in preparation.