Application of Matrix Transformation Methods in Three-Dimensional Device Simulation

Ke-Chih Wu, Ze-Yi Wang†, Robert F. Lucas, Robert W. Dutton

Integrated Circuits Laboratory
AEL 231F, Stanford University, Stanford, CA94305
† Visiting Scientist from Tsinghua University, Beijing, China

In semiconductor device simulation, both Gummel and Newton methods are used to solve the nonlinearly coupled system of Poisson and current continuity equation(s)[1]. In comparing with Gummel's method, Newton's method has the advantage of superlinear convergence when the solution is near. To avoid the large memory penalty associated with using a full Newton method, block-Newton methods have been proposed. However, performance has been reported to be not competitive with the combination of Gummel and full Newton methods[2]. Encouraged by the success of the Modified-Singular Perturbation (MSP) scheme in speeding up the convergence of the Gummel method for a 3-D one-carrier device solver, [3] we have now applied both the MSP and the Alternate-Block-Factorization (ABF)[4] schemes in conjunction with the block-Newton method. With a block matrix transformation, the system of equations for a new update is as follows :

$$A\,x = -F \equiv \begin{bmatrix} A_{\psi\psi} & D_{\psi\Phi_n} \\ A_{\Phi_n\psi} & A_{\Phi_n\Phi_n} \end{bmatrix} \begin{bmatrix} \delta\psi \\ \delta\Phi_n \end{bmatrix} = - \begin{bmatrix} F_\psi \\ F_{\Phi_n} \end{bmatrix}$$

where $A_{\psi\psi}$ and $A_{\Phi_n\Phi_n}$ are the main matrices of Poisson and current continuity equation respectively while $D_{\psi\Phi_n}$ and $A_{\Phi_n\psi}$ are the corresponding coupling matrices. The MSP scheme applies a transformation matrix T_{MSP} to the both sides of the equations

$$T_{MSP}\,A\,x = -\,T_{MSP}\,F$$

with

$$T_{MSP} = \begin{bmatrix} I & 0 \\ -A_{\Phi_n\psi}D_{\psi\psi}^{-1} & I \end{bmatrix}.$$

The ABF scheme applies T_{ABF} to the right of A thereby leaving the right hand side of the equations unchanged as shown below

$$(AT_{ABF})\,(T_{ABF}^{-1}\,x) = -\,F$$

with

$$T_{ABF} = \begin{bmatrix} I & -D_{\psi\psi}^{-1}D_{\psi\Phi_n} \\ -D_{\Phi_n\Phi_n}^{-1}D_{\Phi_n\psi} & I \end{bmatrix}.$$

Given the highly nonlinear nature of the semiconductor equations, it is essential to adopt an appropriate update damping scheme with the Newton method so as to ensure the global convergence. With our improved initial guess scheme[3], the task of finding such a scheme becomes less demanding. Still, according to our experience, it is essential for the scheme to keep a tight control over the maximum error in Poisson's equation and to adaptively limit the magnitude of ψ update. To evaluate the performances of these new schemes, calculations are performed on a coarsely grided one-micron MOSFET using these two schemes as well as the traditional scheme which employs no matrix transformation. Table 1 compares the average CPU time per bias point (averaged from ten bias points with Vds stepping from 0.5V to 5 V) using the traditional block-Newton (BN) method and MSP and ABF modified block-Newton methods at three values of Vgs. Consistent with the improvement obtained by the MSP-Gummel method over the traditional Gummel method[3], both schemes significantly improve the convergence rate in the strong inversion regime (high Vgs) of MOSFET calculations. At Vgs = 5V, the speed-ups achieved by the new schemes are 7.5 and 5.2 respectively. Furthermore, the use of these schemes has significantly extended the convergence range of the block-Newton method. Although the traditional scheme has

difficulty to converge beyond a Vds step of one volt, the new schemes converge with an initial bias as large as Vds = 5 V for all three Vgs values. Calculations are also made to compare the convergence performance of the MSP modified Gummel method and block-Newton methods modified by MSP and ABF. Table 2 shows the average CPU time per bias point as a function of Vds step for the three schemes. Although the difference between the various schemes is not very significant, MSP modified block-Newton method consistently outperforms the other two schemes.

Parallel computers are cost-effective alternatives to supercomputers in solving large systems of equations such as those encountered in 3-D device simulation. The 3-D device solver is, therefore, developed on such a machine (Intel iPSC2^{TM} hypercube, a distributed-memory parallel computer). With each processor (node) has 8M bytes of memory, the solver can handle more than 130K nodes (50 cubes in each dimension) in our 16-node system. In terms of CPU time per bias point, we have observed a performance of about every 2K nodes needing one CPU minute in a machine of about 1.5 MFLOP computation power. Having overcome the huddles in programming the more difficult distributed-memory machine, our 3-D device solver can also be adapted to shared-memory machines with minimum amount of overhead.

References

1. S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag, Vienna, 1984.
2. C. S. Rafferty, M. R. Pinto, and R. W. Dutton, "Iterative Methods in Semiconductor Device Simulation," *IEEE Trans. Electron Devices*, vol. ED-32, no. 10, pp. 2018 - 2027, Oct. 1985.
3. K.-C. Wu, R. F. Lucas, Z.-Y. Wang, and R. W. Dutton, "New Approaches in A Parallel 3-D One-Carrier Device Solver," *Workshop on Numerical Modeling of Processes and Devices for Integrated Circuits -- NUPAD II: Digest of Technical Papers*, May 9-10 1988.
4. R. E. Bank and R. K. Smith, "The Alternate Block Factorization Procedure for Systems of Partial Differential Equations," *to appear*.

Table 1
Effects of Matrix Transformation Schemes

	Average CPU (sec.)			CPU speedup	
Vgs	BN	MSP-BN	ABF-BN	MSP-BN	ABF-BN
1V	112	75.5	85.8	1.5	1.3
3V	661	134.	164.	4.9	4.0
5V	981	130.	189.	7.5	5.2

Table 2
Comparison of Gummel and Block-Newton Schemes

	Average CPU (sec.)		
Vds step	MSP-Gum.	MSP-BN	ABF-BN
0.5V	145	113	146
1.0V	159	132	171
2.0V	200	175	224
5.0V	286	238	289