SAND2008-1742C

Parallel Block-Oriented Preconditioners for FEM Modeling of Semiconductor Devices

SIAM PP08
Atlanta, GA
March 12, 2008

Paul Lin, John N. Shadid,
Robert Hoekstra, Gary Hennigan
Sandia National Labs

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Motivation and Goals

- High fidelity solutions for large scale semiconductor device problems on massively parallel platforms
- Block preconditioners for semiconductor problems
  - Developing scalable multilevel preconditioners for coupled systems with mixed parabolic / elliptic characteristics is often difficult
  - Block preconditioners can be used to produce sub-block systems that are more amenable to optimal multilevel methods
  - Memory usage can also be reduced to allow larger problems to be solved

Governing Equations

Drift-Diffusion Equations

Electric potential

$$\lambda^2 \nabla \cdot (\epsilon_r \mathbf{E}) = p - n + C \qquad \mathbf{E} = -\nabla \psi$$

Current conservation

$$\nabla \cdot \mathbf{J}_n = \frac{\partial n}{\partial t} + R \qquad \mathbf{J}_n = \mu_n n \mathbf{E} + D_n \nabla n$$

$$-\nabla \cdot \mathbf{J}_p = \frac{\partial p}{\partial t} + R \qquad \mathbf{J}_p = \mu_p p \mathbf{E} - D_p \nabla p$$

"constitutive" relation

- ψ: electric potential
- n: electron density
- p: hole density
- C: doping profile
- R: generation-recombination term

Modifications for Defect Physics

Each additional species adds an additional equation

$$\nabla \cdot (\mu_i X_i \nabla \psi + D_i \nabla X_i) = \frac{\partial X_i}{\partial t} + R_{X_i}, \qquad \mu_i = \frac{q_i D_i}{kT}$$

Modified Equation for Electric Potential

$$\varepsilon \nabla^2 \psi = -q(p - n + C) + \sum_{i=1}^{n} q_i X_i$$

- X_i : species concentration
- q : charge of electron
- q_i : integer charge for species

- Immobile and near immobile species makes problem very stiff (7 orders magnitude difference in diffusivity)
- Immobile species add ODEs rather than PDEs due to lack of spatial operators
- Large number of defect species leads to systems with large block sizes

Difficulties with Semiconductor Device Modeling

- Nonlinear Poisson equation is singularly perturbed
  - Doping can be greater than 10^20 and the coefficient of the Laplace operator is very small
- Strong gradients
  - Junctions between n- and p-doped materials
  - At boundaries for high voltage bias
- Type of transport can be very different in different parts of the domain
  - Strongly convective in regions of high gradients
  - Diffusion dominated in other regions
- Dynamic range of variables
  - Electron and hole densities can vary by 18 orders of magnitude
- Nonlinear coupling between Poisson equation for electric potential and equations for continuity of electron and hole currents
- Additional species makes the system of equations very stiff
  - Radiation defect species (charged and neutral) add equations with substantially lower mobilities than the carriers

Sandia National Laboratories

Implementation

- FEM solver: Charon (Hennigan, et. al.)
  - Drift-diffusion equations, species equations
  - Stabilized finite element formulation
  - Fully-coupled Newton-Krylov solver
- NEVADA framework (unstructured meshes, I/O, etc.)
- Trilinos solvers (Heroux,…)
  - Multigrid preconditioner: ML (Tuminaro, Hu, Sala, Gee, Tong)
  - Nonlinear solver: NOX (Pawlowski, Kolda,…)
  - Linear system solver: AztecOO (Tuminaro, Shadid, Hutchinson, Heroux, ..)
  - Direct coarse solver: Amesos (Stanley, Sala) interface with KLU
  - ILU preconditioners: Ifpack (Sala, Heroux)
  - Matrix redistribution: Epetra (Heroux, …)
- Load balance: Zoltan (Devine, Hendrickson,…)
- Graph partitioning: METIS and ParMETIS (Karypis)
- Mesh generation: Cubit

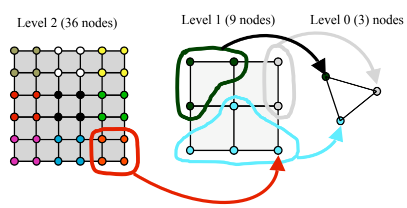
Sandia National Laboratories

ML library: Multilevel Preconditioners

(R. Tuminaro, J. Hu, M. Sala, M. Gee, C. Tong)

2-level and N-level Algebraic

- Generation of the aggregates to produce a coarser operator
  - Create graph where vertices are block nonzeros in matrix A_k
  - Edge between vertices i and j added if block $B_k(i,j)$ contains nonzeros
  - Decompose graph into aggregates
- Construction of restriction/prolongation operators
- Construction of A_{k-1} as $A_{k-1} = R_k A_k P_k$



- Nonsmoothed aggregation
- Domain decomposition smoothers (subdomain GS and ILU)

Sandia National Laboratories

New Preconditioner in ML for Nonsymmetric Systems

(M. Sala and R. Tuminaro)

- Typical choice for smoothing prolongator in smoothed aggregation

$$P_i = (I - \omega_i D^{-1} A)\hat{P}_i$$
$$R = P^T$$

$\hat{P}_i$: tentative prolongator
$D = diag(A)$
ω_i: damping parameter

- Smoothed aggregation preconditioner for nonsymmetric linear systems

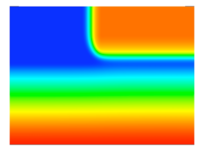
$$P_i = (I - \omega_i D^{-1} A)\hat{P}_i$$
$$R_i = \hat{P}_i^T (I - A D^{-1} \omega_i^{(r)})$$

  - Perform restriction smoothing
  - Restriction operator does not correspond to transpose of prolongator for nonsymmetric problems
  - Rather than use a single damping parameter, calculate values to minimize P_i and R_i

Sandia National Laboratories

Weak Scaling Study: Steady-State NPN BJT
1- and 3-level Preconditioners

- 2x1.5 micron BJT; steady-state 2D drift diffusion bias 0.3V; initial guess NLP solution
- 3-level preconditioner smoothers/solvers: ILU, ILU, KLU
- aggregation: METIS/ParMETIS; 125 nodes/aggregate
- "NSA": standard baseline nonsmoothed aggregation
- "EMIN" ("Energy Minimization"): new ML preconditioner
- Run on Sandia Red Storm machine (Cray XT3)



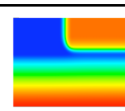
proc	fine grid size (elements)	fine grid unknowns	1-level ILU		3-level NSA agg125		3-level EMIN agg125	
			ave its per Newt step	time per Newt (s)	ave its per Newt step	time per Newt(s)	ave its per Newt step	time per Newt (s)
4	220x165	110058	68	3	63	3.5	51	3.4
16	440x330	437913	142	7.4	122	7.1	91	5.6
64	880x660	1.75M	287	21	198	14	128	8
256	1760x1320	6.98M	571	73	318	30	183	15
1024	3520x2640	27.9M	1145	289	475	73	249	24
4096	7040x5280	111.6M	2264	1088	677	173	326	65
8192	7040x5280	111.6M	2347	601	704	140	336	58

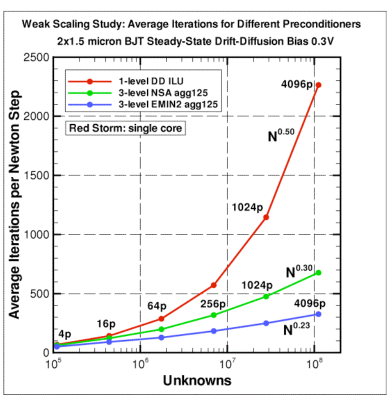
- New "EMIN" preconditioner is a significant improvement over baseline NSA
- "time" is time to construct preconditioner and perform linear solve

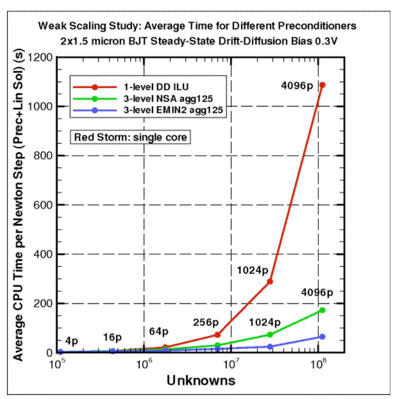
Sandia National Laboratories

Weak Scaling Study: 1-level and 3-level (with and without EMIN)
2x1.5 micron NPN BJT Steady-State Drift-Diffusion

- Charon FEM semiconductor device modeling code
- 2D steady-state drift-diffusion bias 0.3V
- 3-level AMG preconditioners (ML library): baseline NSA and EMIN







- "Time": construct preconditioner and perform linear solve

Sandia National Laboratories

Block Preconditioners for Drift-Diffusion System

- Block Preconditioners
  - Apply different methods to different sub-blocks
  - Applying multilevel method to full system often difficult
  - Can use multilevel designed for one unknown
  - ILU much cheaper on sub-blocks than full system
- Block preconditioner for fully-coupled Newton-Krylov method is an approximation of the Jacobian
  - Jacobian for Drift-Diffusion System

$$\mathbf{J} = \begin{bmatrix} \mathbf{D}_\psi & -\mathbf{M}_n & \mathbf{M}_p \\ \mathbf{F}_{n\psi} & \left(\frac{\mathbf{M}_n}{\Delta t} + \mathbf{C}_n + \mathbf{R}_{nn}\right) & \mathbf{R}_{np} \\ \mathbf{F}_{p\psi} & \mathbf{R}_{pn} & \left(\frac{\mathbf{M}_p}{\Delta t} + \mathbf{C}_p + \mathbf{R}_{pp}\right) \end{bmatrix}$$

$\mathbf{D}_\psi$ Diffusion Matrix
$\mathbf{C}_n$, $\mathbf{C}_p$ Convection-Diffusion Matrix
$\mathbf{M}_n$, $\mathbf{M}_p$ Mass Matrix
$\mathbf{R}_{nn}$, $\mathbf{R}_{np}$, $\mathbf{R}_{pn}$, $\mathbf{R}_{pp}$ Reaction Matrix
$\mathbf{F}_{n\psi}$, $\mathbf{F}_{p\psi}$ Drift Velocity Coupling Matrix

Sandia National Laboratories

Block Preconditioners

- Block Jacobi, Block Gauss-Seidel, Block SOR (1 unknown per block)
  - ILU or ML on sub-blocks

$\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{F}$
$\mathbf{D}$ is diagonal block
$\mathbf{E}$ is negative of blocks below diagonal
$\mathbf{F}$ is negative of blocks above diagonal

$$\mathbf{M}_{JAC} = \mathbf{D}$$
$$\mathbf{M}_{GS-F} = \mathbf{D} - \mathbf{E}$$
$$\mathbf{M}_{GS-B} = \mathbf{D} - \mathbf{F}$$
$$\mathbf{M}_{SOR-F} = \frac{1}{\omega}(\mathbf{D} - \omega\mathbf{E})$$
$$\mathbf{M}_{SOR-B} = \frac{1}{\omega}(\mathbf{D} - \omega\mathbf{F})$$

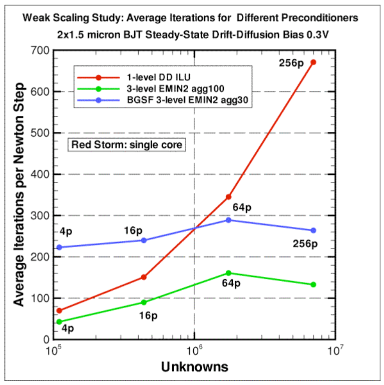
$$\mathbf{J} = \begin{bmatrix} \mathbf{D}_\psi & -\mathbf{M}_n & \mathbf{M}_p \\ \mathbf{F}_{n\psi} & \left(\frac{\mathbf{M}_n}{\Delta t} + \mathbf{C}_n + \mathbf{R}_{nn}\right) & \mathbf{R}_{np} \\ \mathbf{F}_{p\psi} & \mathbf{R}_{pn} & \left(\frac{\mathbf{M}_p}{\Delta t} + \mathbf{C}_p + \mathbf{R}_{pp}\right) \end{bmatrix}$$

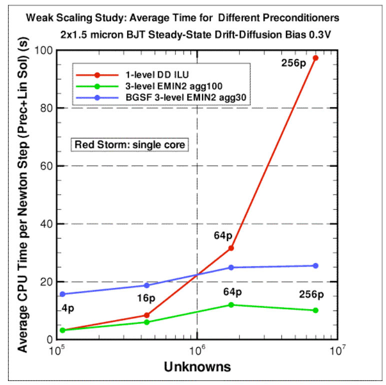
$$\mathbf{M}_{Jac} = \begin{bmatrix} \mathbf{D}_\psi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\frac{\mathbf{M}_n}{\Delta t} + \mathbf{C}_n + \mathbf{R}_{nn}\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \left(\frac{\mathbf{M}_p}{\Delta t} + \mathbf{C}_p + \mathbf{R}_{pp}\right) \end{bmatrix}$$

Sandia National Laboratories

Preliminary Results: Weak Scaling Study 2x1.5 micron NPN BJT Steady-State Drift-Diffusion

- Full Jacobian for preconditioner: 1-level DD ILU and 3-level EMIN2 agg100
- Block Gauss-Seidel (Forward): Sub-block iteration ML with EMIN and agg30
- Meshes: 220x165 to 1760x1320 elements (4-256 processors)





Sandia National Laboratories

Next Steps: Operator Split Approximation

- Operator split approximation to convection/diffusion/reaction system
  - 2x2 block system with n and p in second block
  - Apply block preconditioner (Jacobi for this example)

$$\mathbf{M} = \begin{bmatrix} \mathbf{D}_\psi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\frac{\mathbf{M}_n}{\Delta t} + \mathbf{C}_n + \mathbf{R}_{nn}\right) & \mathbf{R}_{np} \\ \mathbf{0} & \mathbf{R}_{pn} & \left(\frac{\mathbf{M}_p}{\Delta t} + \mathbf{C}_p + \mathbf{R}_{pp}\right) \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{D}_\psi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\frac{\mathbf{M}_n}{\Delta t} + \mathbf{C}_n + \mathbf{R}_{nn}\right) & \mathbf{R}_{np} \\ \mathbf{0} & \mathbf{R}_{np} & \left(\frac{\mathbf{M}_p}{\Delta t} + \mathbf{C}_p + \mathbf{R}_{pp}\right) \end{bmatrix}$$

Sandia National Laboratories

Next Steps: Operator-split Approximation

- Operator split approximation

$$\mathbf{M} = \begin{bmatrix} \mathbf{D}_\psi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\mathbf{M}_n^L}{\Delta t}\left(\mathbf{I} + \Delta t \left(\mathbf{M}_n^L\right)^{-1} \mathbf{C}_n\right) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\mathbf{M}_p^L}{\Delta t}\left(\mathbf{I} + \Delta t \left(\mathbf{M}_p^L\right)^{-1} \mathbf{C}_p\right) \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{I} + \Delta t \left(\mathbf{M}_n^L\right)^{-1} \mathbf{R}_{nn}\right) & \Delta t \left(\mathbf{M}_n^L\right)^{-1} \mathbf{R}_{np} \\ \mathbf{0} & \Delta t \left(\mathbf{M}_p^L\right)^{-1} \mathbf{R}_{pn} & \left(\mathbf{I} + \Delta t \left(\mathbf{M}_p^L\right)^{-1} \mathbf{R}_{pp}\right) \end{bmatrix}$$

- Inexpensive system to solve
  - AMG for diffusion solve for psi
  - Decoupled convection-diffusion solves for n and p (AMG)
  - Reorder to obtain a block diagonal 2x2 system

Next Steps

- Drift diffusion with defect species
  - Blocks: first block consists of drift-diffusion system, remaining blocks have one species per block

$$\mathbf{M} = \left[\begin{array}{c|cccc} \mathbf{DD} & & & & \\ \hline & \mathbf{X_1} & & & \\ & & \mathbf{X_2} & & \\ & & & \ddots & \\ & & & & \mathbf{X_i} \end{array}\right]$$

- Form Gummel iteration as a preconditioner
- Schur decomposition on 2x2 block system

Conclusions

- Block preconditioners seem to be a promising way to solve the semiconductor device equations
  - Easier to solve sub-blocks rather than full system
    - Applying multilevel method to full system often difficult
    - Can leverage previous work with AMG methods
  - Can handle different sub-blocks with different methods (methods which we hope to be optimal)

Sandia National Laboratories

Acknowledgements

- Marzio Sala, Ray Tuminaro, Roger Pawlowski, Jonathan Hu, Pavel Bochev
- Sandia Red Storm team
- Those that pay the bills: DOE MICS, ASC

Sandia National Laboratories

The End

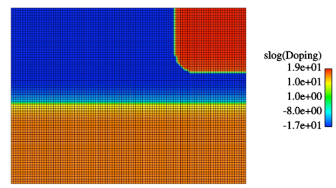
FEM for Semiconductor Device Modeling

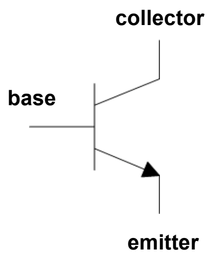
- Existing commercial codes use FVM
  - FVM approaches are fast and robust
    - Can use coarser meshes for the same accuracy
  - Most common approach: Scharfetter-Gummel
    - derived in 1D
    - believed to “introduce a potentially significant amount” of numerical diffusion for more than 1D (Kramer and Hitchon)
- FEM approach has advantages
  - Theoretical error bounds for FE approximations
  - Design error estimators for mesh adaptivity methods
  - Extension to higher order methods
  - Convergence properties for unstructured meshes
  - Possible advantages for multi-dimensional problems

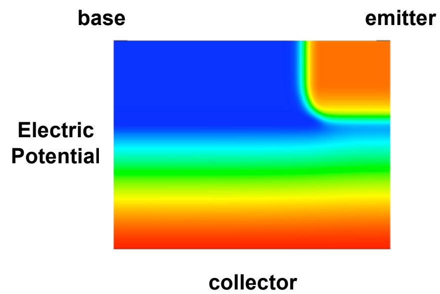
Example Problem: 2D NPN BJT

- 2D NPN Bipolar Junction Transistor
  - 2x1.5 microns
  - Three contacts
- Voltage bias: emitter at ground while base and collector voltage increased
- Steady-state

Example Doping







Sandia National Laboratories

Extremely Preliminary Results

- 100x75 elem 2x1.5 um BJT
- NLP->0.3V bias; serial runs
- For block preconditioners using almost a direct solve on the blocks

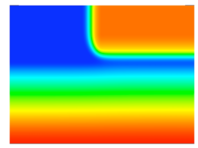
Preconditioner	ave its per Newt step
ILU(fill=2)	32
ILU(fill=1)	49
ILU(fill=0)	72
Block Jacobi	282
Block Gauss-Seidel-F	108
Block Gauss-Seidel-B	173
Block SOR-F:0.7	164
Block SOR-F:1.3	183
Block SOR-B:0.8	171
Block SOR-B:1.2	184

- Extremely preliminary parallel result
  - 2x1.5 um BJT; NLP->0.3V bias
  - 1760x1320 elem (7 M unknowns)
  - 256 procs on Red Storm
- 1-level ILU (baseline preconditioner)
- Block GS-F with ML as sub-block iteration
  - smoothers gs1/ILU/KLU
  - aggregate30 (unks: 2.3M/77K/2580)
  - Competitive with 1-level ILU; will be better for larger problems
- For sub-blocks, can use ML for one unknown
- For our system based aggressive coarsening ML as preconditioner, cannot use GS as smoother, must use ILU

Preconditioner	ave iter per Newt step	time per Newt (s)
ILU	668	107
Block GS-F (ML subblock)	602	101

Sandia National Laboratories

Weak Scaling Study: Steady-State NPN BJT 1- and 3-level Preconditioners



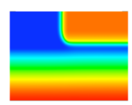
- 2x1.5 micron BJT; steady-state 2D drift diffusion bias 0.3V; initial guess NLP solution
- 3-level preconditioner smoothers/solvers: ILU, ILU, KLU
- aggregation: METIS/ParMETIS; 125 nodes/aggregate
- “NSA”: standard baseline nonsmoothed aggregation
- “EMIN” (“Energy Minimization”): new ML preconditioner
- Run on Sandia Red Storm machine (Cray XT3)

proc	fine grid size (elements)	fine grid unknowns	1-level ILU		3-level NSA agg125		3-level EMIN agg125	
			ave its per Newt step	time per Newt (s)	ave its per Newt step	time per Newt(s)	ave its per Newt step	time per Newt (s)
8	220x165	110058	73	1.7	67	2	55	2
32	440x330	437913	147	3.9	126	3.9	95	3.2
128	880x660	1.75M	297	13	231	10	147	6
512	1760x1320	6.98M	590	40	341	19	206	11
2048	3520x2640	27.9M	1181	158	490	41	265	18
8192	7040x5280	111.6M	2347	601	704	140	336	58

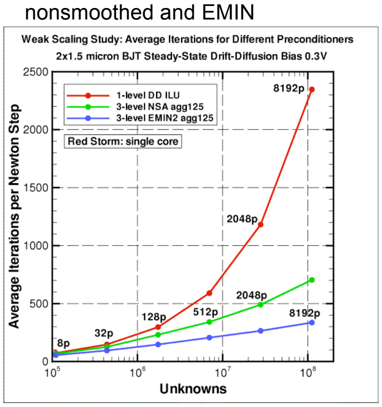
- “time” is time to construct preconditioner and perform linear solve
- New “EMIN” preconditioner is a significant improvement over baseline NSA

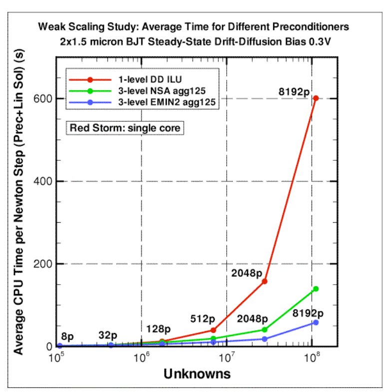
Sandia National Laboratories

Weak Scaling Study: 1-level and 3-level (with and without EMIN) 2x1.5 micron NPN BJT Steady-State Drift-Diffusion



- Charon FEM semiconductor device modeling code
- 2D steady-state drift-diffusion bias 0.3V
- 3-level AMG preconditioners (ML library): baseline nonsmoothed and EMIN





- “Time”: construct preconditioner and perform linear solve

Sandia National Laboratories