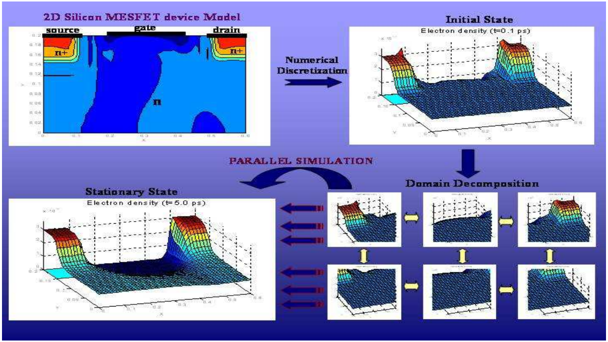
Simulación Paralela Determinista de dispositivos semiconductores 2D basada en esquemas WENO-Boltzmann

J. M. Mantas Ruiz, R. J. García
Depto. Lenguajes y Sistemas Inf.
Univ. Granada

M.J. Cáceres
Depto. Matemática Aplicada.
Univ. Granada

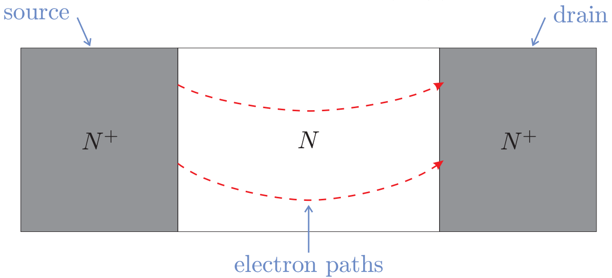


XVIII Jornadas de Paralelismo, Zaragoza, 11-14 Septiembre 2007

Contents

Introduction

$N^+ - N - N^+$ Device



Cross-section of a $N^+ - N - N^+$ device

WENO-Boltzmann scheme

Semiclassical Approximation. Boltzmann-Poisson System

- **Boltzmann Transport Equation for semiconductors:**

$$\frac{\partial f}{\partial t} + \frac{1}{\hbar}\nabla_{\mathrm{k}}\varepsilon \cdot \nabla_{\mathrm{x}} f - \frac{e}{\hbar}\mathrm{E} \cdot \nabla_{\mathrm{k}} f = Q(f)$$

- f is the electron probability density function in phase space (x, k) at each time t.
- **Poisson Equation:** To compute Electric field $\mathrm{E} = -\nabla_{\mathrm{x}} V$, the potential V satisfies:

$$\nabla_{\mathrm{x}} \left[\epsilon_r(\mathrm{x}) \nabla_{\mathrm{x}} V\right] = -q \left[\rho(t, \mathrm{x}) - N_D(\mathrm{x})\right],$$

$\rho(t, x)$=electron density, $N_D(x)$ is the doping profile and $\epsilon_r(\mathrm{x})$ is the dielectric constant.

WENO-Boltzmann scheme (*Carrillo et. al. 2003*)

Adimensionalization + pseudospherical change of variable for k

$$\Phi(t,x,y,w,\mu,\phi) = s(w) f(t,x,y,w,\mu,\phi).$$

Adimensional Boltzmann Eq. in 2D physical space

$$\frac{\partial \Phi}{\partial t} + \frac{\partial}{\partial x}(a_1 \Phi) + \frac{\partial}{\partial y}(a_2 \Phi) + \frac{\partial}{\partial w}(a_3 \Phi) + \frac{\partial}{\partial \mu}(a_4 \Phi) + \frac{\partial}{\partial \phi}(a_5 \Phi) = s(w) C(\Phi)$$

a_3, a_4, a_5 depends on $\mathbf{E} = [E_x, E_y] = [\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}]$.

2D Poisson Equation

$$\frac{\partial \left[\epsilon(x,y)\frac{\partial V}{\partial x}\right]}{\partial x} + \frac{\partial \left[\epsilon(x,y)\frac{\partial V}{\partial y}\right]}{\partial y} = -q\left[n(t,x,y) - n_D(x,y)\right].$$

WENO-Boltzmann scheme

$$\frac{\partial \Phi}{\partial t}+\frac{\partial}{\partial x}(a_1\Phi)+\frac{\partial}{\partial y}(a_2\Phi)+\frac{\partial}{\partial w}(a_3\Phi)+\frac{\partial}{\partial \mu}(a_4\Phi)+\frac{\partial}{\partial \phi}(a_5\Phi)=s(w)C(\Phi)$$

Approximation of derivatives with WENO5

- Fith order WENO (*Weighted Essentially Non-Oscillatory*, Shu98) finite difference schemes to approximate the fluxes (WENO5) in x, y, w, μ, ϕ:
  $$\frac{\partial}{\partial x}(a_1\Phi)+\frac{\partial}{\partial y}(a_2\Phi)+\frac{\partial}{\partial w}(a_3\Phi)+\frac{\partial}{\partial \mu}(a_4\Phi)+\frac{\partial}{\partial \phi}(a_5\Phi)$$
  - Nonoscillatory character near shocks or gradient regions.
  - Suitable to deal with very strong steep spatial gradients in these devices.
- Time Integration of $\frac{\partial \Phi}{\partial t}$ with 3rd order TVD (*Total Variation Diminishing*) Runge-Kutta.

WENO-Boltzmann scheme

WENO-Boltzmann scheme vs. Direct Simulation Monte Carlo (DSMC)

- **Advantages with respect DSMC**
  1. Suitable to describe almost empty regions of the device.
  2. Transient computation
  3. Explicit Management of the Probability density function.
- **Main disadvantage**: Great demands of computational power even in 2D physical space (5 dimensions + time).

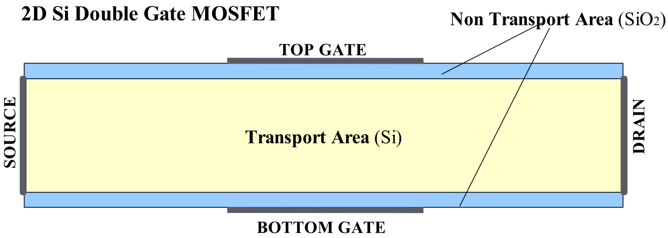
Approach to enable its use

- Development of flexible and efficient parallel versions of the scheme for parallel architectures.
- Low cost and widespread architecture: Clusters of SMPs.

Representation of Device information

Two areas are considered in the 2D physical space of the device:

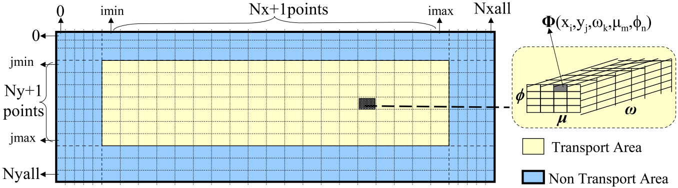
- **Area without charge transport**
- **Transport area**
  - $\Phi(t, x, y, w, \mu, \phi)$ is only defined for (x, y) in this area.
  - The computational work of this area is parallelized



Representation of Device information. Discretization

- **Transport Area**: 5D uniform grid $(x_i, y_j, w_k, \mu_m, \phi_n)$,
  $i = 0, .., N_x,\ j = 0, .., N_y,\ k = 1, .., N_w,\ m = 1, .., N_\mu,\ n = 1, .., N_\phi$.
  - $\mathbf{\Phi(t, x_i, y_j, w_k, \mu_m, \phi_n)} \Longrightarrow$ 5D array $\Phi(i, j, k, m, n)$.
  - $\mathbf{n(t, x_i, y_j)}$, $\mathbf{E_x(x_i, y_j)}$ and $\mathbf{E_y(x_i, y_j)} \Longrightarrow$ 2D arrrays.

- **Full spatial device** (2D): (x_i, y_j), $i = 0, \ldots, Nxall, j = 0, \ldots, Nyall$.
  - $\mathbf{V(x_i, y_j)} \Longrightarrow$ 2D array $V(i, j)$.



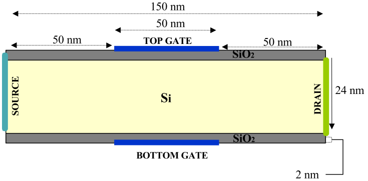
Representation of Device information

Description file for a DG-MOSFET device

```
SET SCALE (centimeter,volt,sec);
GEOMETRY
  XLength = 150.0e-7;  YLength = 28.0e-7;
  TRANSPORT[(0.0, 2.0e-7)->(150.0e-7, 26.0e-7)];
  SOURCE [(0.0     , 2.0e-7)->(0.0     , 26.0e-7)];
  DRAIN  [(150.0e-7, 2.0e-7)->(150.0e-7, 26.0e-7)];
  GATE   [(50.0e-7, 0.0)->(100.0e-7, 0.0)];
  GATE   [(50.0e-7, 28.0e-7)->(100.0e-7, 28.0e-7)];
Nxall=30; Nyall=56;  Nx=30; Ny=48; Nw=120; Nmu=12;
         Nphi=12;  NMULT   =2;
END
DOPING
                    FILE ("doping.dat");
END
POTENTIAL
  SOURCE=0.0;
  DRAIN   =0.5;
  GATE[0]=1.0;
  GATE[1]=1.0;
END
DIELECTRIC
   REGION [ (0.0, 0.0)->(15.0e-6, 2.0) ]                =3.9;
   REGION [ (0.0, 26.0e-7(->(150.0e-7, 28.0e-7) ]= 3.9;
   REGION  [ (0.0, 2.0e-7(->(150.0e-7, 26.0e-7) ] =11.9;
END
SIMULATION
  TEND    =3.5; NTOT    =9999999; CFL     =0.9;
END
```

doping.dat

```
0          2e-07    1.0e+20
5e-07      2e-07    1.0e+20
1e-06      2e-07    1.0e+20
1.5e-06    2e-07    1.0e+20
2e-06      2e-07    1.0e+20
2.5e-06    2e-07    1.0e+20
3e-06      2e-07    1.0e+20
3.5e-06    2e-07    1.0e+20
...
```



Numerical Schemes in Transport area

$$\left[\frac{\partial \Phi}{\partial t}\right]_{i,j,k,m,n} = L(\Phi)(i,k,k,m,n) = [s(w)C(\Phi)]_{i,j,k,m,n} - \left[\frac{\partial}{\partial x}(a_1\Phi) + \ldots + \frac{\partial}{\partial \phi}(a_5\Phi)\right]$$

Time integration of $\Phi(i,j,k,m,n)$: 3rd order TVD RK method

- Φ^n (in t^n) $\rightarrow \Phi^{n+1}$ (in $t^{n+1} = t^n + \Delta t$)
- 3 stages

$$\Phi^n = \Phi^{(0)} \xrightarrow{stage} 0\ \Phi^{(1)} \xrightarrow{stage} 1\ \Phi^{(2)} \xrightarrow{stage} 2\ \Phi^{(3)} = Phi^{n+1}$$

- In s-th stage, $L(\Phi^{(s)})(i,k,k,m,n)$ must be evaluated.

Computation of $C(\Phi)_{i,j,k,m,n}$ and $n(i,j)$

- Composite mid-point rule to approximate the integrals

Numerical Schemes in Transport area

Computation of spatial derivatives in $L(\Phi)_{i,j,k,m,n}$

Dimension by dimension approximation to the spatial derivatives using WENO5 : **Most costly computing phase**

- **Example**: WENO5 to approximate $\frac{\partial}{\partial x}(a_1\Phi)$ in $(x_i, y_j, w_k, \mu_m, \phi_n)$

$$\boxed{\frac{\partial}{\partial x} a_1\Phi(i,j,k,m,n) = \begin{cases} W(g_{i-3},\ldots,g_{i+2}), & \text{si } a_1 > 0 \\ W(g_{i+3},\ldots,g_{i-2}), & \text{si } a_1 \leq 0 \end{cases}}$$

$$g_i = a_1\Phi(i,j,k,m,n), \quad i = -3,\ldots,N_x+3,$$

- **Variable 6-point asymmetric stencil.**

Numerical Schemes in Full Spatial Device

Computation of Electric Potential $V(i,j)$

Numerical solution of 2D Poisson Eq. $\rightarrow$ Linear system

$$A_{n\times n} \cdot V = -q\left[n(i,j) - n_D(i,j)\right]$$

- $n = (Nxall + 1)(Nyall + 1)$, $Nxall < 300$ and $Nyall < 300$.
- A is a **constant** banded matrix with bandwidth $2Nxall + 3$.
- One initial Banded LU factorization.
- One Forward elimination and back substitution is made for each Runge-Kutta stage

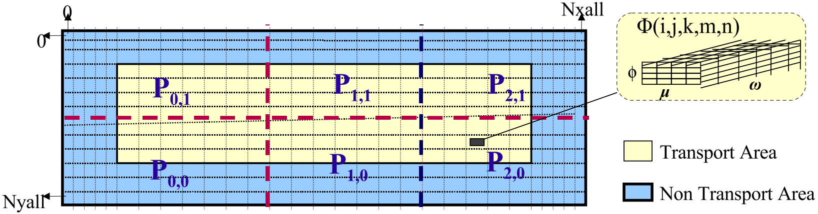
Computation of Electric field $E_x(i,j)$ and $E_y(i,j)$

$$E_x(i,j) = \frac{V(i+1,j) - V(i-1,j)}{x_{i+1} - x_{i-1}}, \quad E_y(i,j) = \frac{V(i,j+1) - V(i,j-1)}{y_{i+1} - y_{i-1}}$$

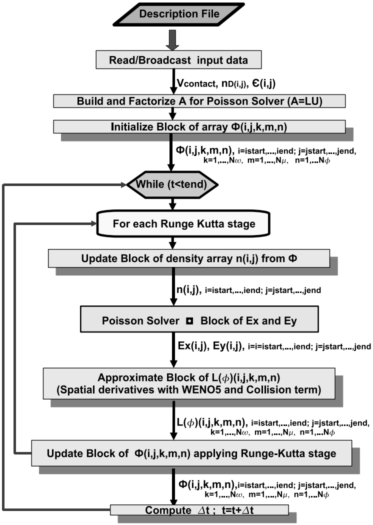
Parallel Algorithm. Decomposition Strategy

- $\Phi(i, j, k, m, n)$ is block distributed onto a 2D proc. grid by splitting (x_i, y_j)
- A similar distribution for 2D arrays in transport area
- Workload balance centred on transport area
- Satisfactory load balance, low communication costs and high reuse.

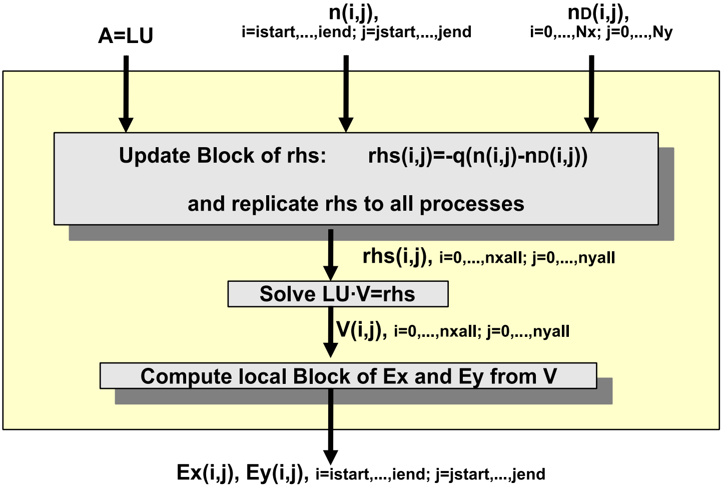
P=6 procs. , Nx=18, Ny=6 ⇒ 3×2 processor grid



Parallel Algorithm. General View of a process



Parallel Algorithm. Poisson Solver



Parallel Algorithm. Compute Block of $L(\Phi)(i,j,k,m,n)$

Input: $n(i,j)$, $E_x(i,j), E_y(i,j)$, $\Phi(i,j,k,m,n)$,
$i = istart, \dots, iend$; $j = jstart, \dots, jend$;
$k = 1, \dots, N_w$, $m = 1, \dots, N_\mu$, $n = 1, \dots, N_\phi$.

1. $L(i,j,k,m,n) = s(w_k, \mu_m) C(i,j,k,m,n)$

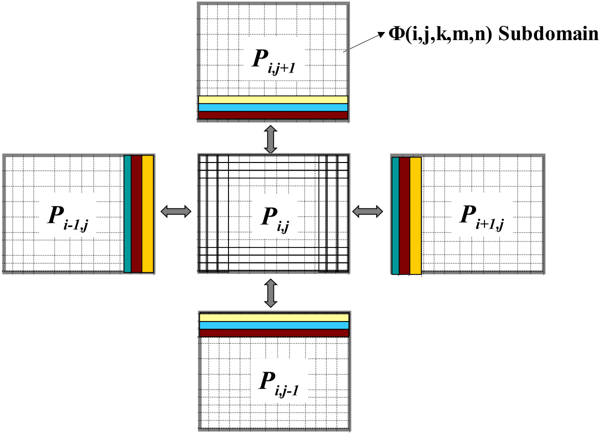
2. $$L_(i,j,k,m,n) - = \left(\frac{\partial}{\partial w}(a_3 \Phi) + \frac{\partial}{\partial \mu}(a_4 \Phi) + \frac{\partial}{\partial \phi}(a_5 \Phi) \right)_{i,j,k,m,n}$$

3. Exchange Φ boundaries in x direction
$L(i,j,k,m,n) - = \frac{\partial}{\partial x}(a_1 \Phi)_{i,j,k,m,n}$

4. Exchange Φ boundaries in y direction
$L(i,j,k,m,n) - = \frac{\partial}{\partial y}(a_2 \Phi)_{i,j,k,m,n}$

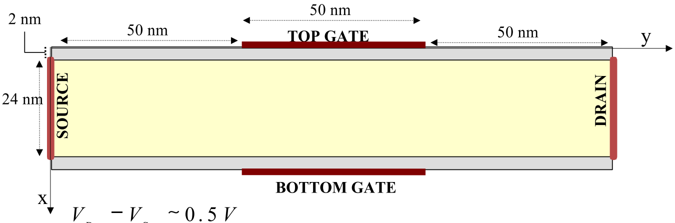
Parallel Algorithm. Communication pattern for WENO5

- Only the fluxes in x and y require interprocessor communication.
- A 7 point symmetric stencil is assumed to apply WENO5.



Experimental results

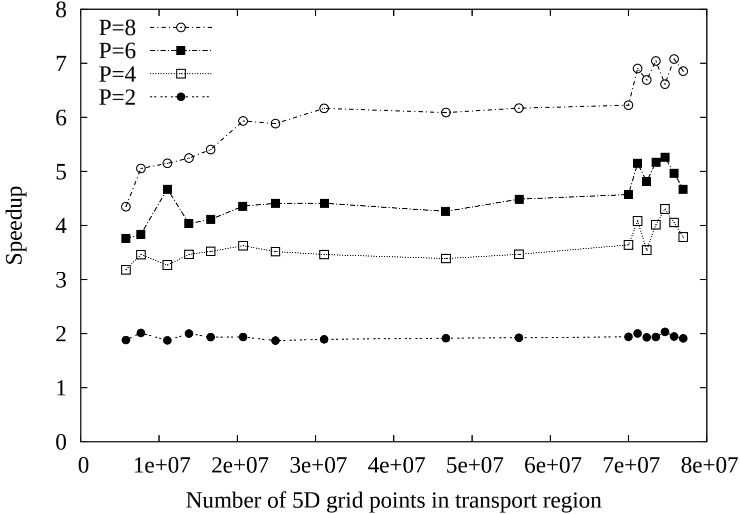
- Spatial grid for full device: $(Nxall, Nyall) = (30, 56)$.
- 5D grid for transport area: $(Nx, Ny, Nw, N\mu, N\phi) = (30, 48, 120, 12, 12)$.
- Cluster of 4 dual AMD Opteron 2.4 MHz via Gigabit Ethernet.
- C++ code + mpich + ACML Lapack
- Time integration from $0ps$ through $3.5ps$.



$$V_D - V_S \sim 0.5\ V$$

$$V_G - V_S \sim 1.0\ V$$

Experimental results. Parallel speedup



Further Work

- Development of a **quantum-deterministic model** for DG-MOSFET devices and integration in the parallel simulator.
- **Use of GPUs** to speedup the numerical simulation on CPU-GPU clusters.
- **Non Uniform Grids**: WENO with interpolation at subdomain interfaces (*Sebastian-Shu 2003*).
- Dynamic Load balancing strategy to exploit efficiently heterogeneous systems.

GRACIAS POR VUESTRA ATENCIÓN

Experimental results. Execution times