Mixed-Mode Device/Circuit Simulation

Tibor Grasser

Institute for Microelectronics
Gußhausstraße 27–29, A-1040 Wien, Austria
Technical University Vienna, Austria
http://www.iue.tuwien.ac.at
Outline

Circuit simulation and compact models

Numerical models instead of compact models

Challenges in numerical modeling

Mixed-mode device/circuit simulation

Examples

Conclusion
Circuit simulation fundamental

Development of modern IC
To understand and optimize the way a circuit works
Circuit Simulation

Circuit simulation fundamental

- Development of modern IC
- To understand and optimize the way a circuit works

For circuit simulation we need

- Lumped elements: R, C, L, etc.
- Current and voltage sources, controlled sources
- Semiconductor devices
- Thermal equivalent circuit (coupling and self-heating)
Circuit simulation fundamental
  Development of modern IC
  To understand and optimize the way a circuit works

For circuit simulation we need
  Lumped elements: R, C, L, etc.
  Current and voltage sources, controlled sources
  Semiconductor devices
  Thermal equivalent circuit (coupling and self-heating)

Electrical/thermal properties of semiconductor devices
  Characterized by coupled partial differential equations
Circuit simulation fundamental
Development of modern IC
To understand and optimize the way a circuit works

For circuit simulation we need
Lumped elements: R, C, L, etc.
Current and voltage sources, controlled sources
Semiconductor devices
Thermal equivalent circuit (coupling and self-heating)

Electrical/thermal properties of semiconductor devices
Characterized by coupled partial differential equations

For the simulation of large circuits we need compact models
Obtained from simplified solutions of these PDEs or empirically
Must be very efficient (compact!)
Compact Modeling

Derivation of compact models based on fundamental equations

Often the drift-diffusion framework is used
Simplifying assumptions on geometry, doping profiles, material parameters
⇒ Compact model

It is becoming increasingly difficult to extract main features
Compact Modeling

Derivation of compact models based on fundamental equations

- Often the drift-diffusion framework is used
- Simplifying assumptions on geometry, doping profiles, material parameters
  \[ \Rightarrow \text{Compact model} \]
- It is becoming increasingly difficult to extract main features

Ongoing struggle regarding

- Number of parameters
- Physical meaning of these parameters
- Predictiveness difficult to obtain, calibration required
Compact Modeling

Derivation of compact models based on fundamental equations

- Often the drift-diffusion framework is used
- Simplifying assumptions on geometry, doping profiles, material parameters
  \[
  \Rightarrow \text{Compact model}
  \]
- It is becoming increasingly difficult to extract main features

Ongoing struggle regarding

- Number of parameters
- Physical meaning of these parameters
- Predictiveness difficult to obtain, calibration required

Compact modeling challenges (ITRS)

- Quantum confinement
- Ballistic effects
- Inclusion of variability and statistics
Advantages of using compact models

Very fast execution (compared to PDEs)
Advantages of using compact models

- Very fast execution (compared to PDEs)

Disadvantages

- Many parameters
  - Physically motivated parameters
  - Fit parameters
- Parameter extraction can be quite cumbersome
- Device optimization via geometry and doping profile hardly possible
- Considerable model development effort
  - Limited model availability (DG, TriGate, FinFETs, GAAFETs, etc.)
- Scalability questionable
  - Quantum effects
  - Non-local effects
Instead of

Analytical expressions describing the device behavior (compact models)

Rigorous device simulation based on

Coupled partial differential equations!
Advantages of numerical device simulation

- Fairly arbitrary devices (doping, geometry)
- Realistic doping profiles from process simulation
- Natural inclusion of
  - 2D/3D effects
  - Non-local effects (via appropriate transport model)
  - Quantum mechanical effects (via simplified model or Schrödinger’s equation)
  - Temperature dependencies
- Sensitivity of device/circuit figures of merit to process parameters
- Better predictivity for scaled/modified devices
Advantages of numerical device simulation

Fairly arbitrary devices (doping, geometry)
Realistic doping profiles from process simulation
Natural inclusion of
  2D/3D effects
  Non-local effects (via appropriate transport model)
  Quantum mechanical effects (via simplified model or Schrödinger’s equation)
  Temperature dependencies
Sensitivity of device/circuit figures of merit to process parameters
Better predictivity for scaled/modified devices

Disadvantages of numerical modeling

Performance (don’t compare!)
Convergence sometimes costly/difficult to obtain
Realistic doping profiles from process simulation
Feature size approaches mean free path

Ballistic effects become important

*No ballistic transistor in sight, but still important effect*
Challenges in Device Simulation

Feature size approaches mean free path
  Ballistic effects become important
    *No ballistic transistor in sight, but still important effect*

Feature size approaches electron wavelength
  Quantum mechanical effects become important
  Transport remains classical
    *Critical gate length around 10 nm*
    *Modified transport parameters for thin channels*
Challenges in Device Simulation

Feature size approaches mean free path
- Ballistic effects become important
  - *No ballistic transistor in sight, but still important effect*

Feature size approaches electron wavelength
- Quantum mechanical effects become important
- Transport remains classical
  - *Critical gate length around 10 nm*
  - *Modified transport parameters for thin channels*

Exploitation of new effects
- Strain effects used to boost mobility
- Substrate orientation and channel orientation
Feature size approaches mean free path

Ballistic effects become important

*No ballistic transistor in sight, but still important effect*

Feature size approaches electron wavelength

Quantum mechanical effects become important

Transport remains classical

*Critical gate length around 10 nm*

*Modified transport parameters for thin channels*

Exploitation of new effects

Strain effects used to boost mobility

Substrate orientation and channel orientation

Exploitation of new materials

Strained silicon, SiGe, Ge, etc.

High-k dielectrics
Classical transport described by Boltzmann’s equation

Allows inclusion of sophisticated scattering models, quasi-ballistic transport
Device Simulation

Classical transport described by Boltzmann’s equation
  Allows inclusion of sophisticated scattering models, quasi-ballistic transport

Very time consuming
  Current resources do not allow us to look at circuits, no AC analysis
Classical transport described by Boltzmann’s equation
  Allows inclusion of sophisticated scattering models, quasi-ballistic transport

Very time consuming
  Current resources do not allow us to look at circuits, no AC analysis

Approximate solution obtained by just looking at moments of $f$
Device Simulation

Classical transport described by Boltzmann’s equation

Allows inclusion of sophisticated scattering models, quasi-ballistic transport

Very time consuming

Current resources do not allow us to look at circuits, no AC analysis

Approximate solution obtained by just looking at moments of $f$

Simplest moment-based model: the classic drift-diffusion model

$$\epsilon \nabla^2 \psi = q(n - p - C)$$

$$\nabla \cdot (D_n \nabla n - n \mu_n \nabla \psi) - \frac{\partial n}{\partial t} = R$$

$$\nabla \cdot (D_p \nabla p + p \mu_p \nabla \psi) - \frac{\partial p}{\partial t} = R$$

Requires models for physical parameters $D$, $\mu$, and $R$

These models capture fundamental physical effects

*Velocity saturation, SRH recombination, impact-ionization*

Models can be quite complex

Used to be basis for the derivation of compact models
Double-Gate MOSFETs

Drift-diffusion model inaccurate for short-channel devices
Double-Gate MOSFETs

Drift-diffusion model inaccurate for short-channel devices
Higher-order moment models available

Comparison of scaled DG-MOSFETs

Comparison with fullband Monte Carlo data

Transport parameters from FBMC
Double-Gate MOSFETs

Drift-diffusion model inaccurate for short-channel devices

Higher-order moment models available

Comparison of scaled DG-MOSFETs

Comparison with fullband Monte Carlo data

Transport parameters from FBMC

DD accurate down to 250 nm

No velocity overshoot
Drift-diffusion model inaccurate for short-channel devices

Higher-order moment models available

Comparison of scaled DG-MOSFETs

Comparison with fullband Monte Carlo data

*Transport parameters from FBMC*

DD accurate down to 250 nm

*No velocity overshoot*

ET accurate at 100 nm

*Maxwellian distribution function*
Double-Gate MOSFETs

Drift-diffusion model inaccurate for short-channel devices

Higher-order moment models available

Comparison of scaled DG-MOSFETs

Comparison with fullband Monte Carlo data

*Transport parameters from FBMC*

**DD** accurate down to 250 nm

*No velocity overshoot*

**ET** accurate at 100 nm

*Maxwellian distribution function*

**SM** accurate at 50 nm

*Non-Maxwellian effects*

*Low computational effort*

'TCAD' compatible
Mixed-Mode Simulation

Simulator coupling
Simple, straight forward solution
Two-Level Newton algorithm
SPICE-like damping algorithms usable
Many iterations for device equations needed
Parallelization straight-forward
Mixed-Mode Simulation

**Simulator coupling**
- Simple, straightforward solution
- Two-Level Newton algorithm
- **SPICE**-like damping algorithms usable
- Many iterations for device equations needed
- Parallelization straight-forward

**All-In-One solution (Full-Newton)**
- Circuit and device equations in one single matrix
- Full-Newton algorithm
- Complex convergence behavior
- Parallelization more complicated
Two-Level Newton

Device simulator is called for each circuit iteration

*Fixed set of contact voltages*

*Contact current response* \( I_C^k \)

*Problematic:* \( g_{\text{eq}}^k = \frac{\partial I_C}{\partial V_C} \bigg|_k \)

Device simulator iterates until convergence

Last iteration as initial-guess

*Linear prediction algorithm*
Simulator Coupling

Two-Level Newton

Device simulator is called for each circuit iteration

Fixed set of contact voltages

Contact current response $I_C^k$

Problematic: $g_{eq}^k = \frac{\partial I_C}{\partial V_C} |_k$

Device simulator iterates until convergence

Last iteration as initial-guess

Linear prediction algorithm

Quasi Full-Newton

Only one iteration of device simulator

Calculation of $I_C^k$ and $g_{eq}^k$
Simulator Coupling

Two-Level Newton

Device simulator is called for each circuit iteration

- Fixed set of contact voltages
- Contact current response \( I^k_C \)
- Problematic: \( g^k_{eq} = \frac{\partial I_C}{\partial V_C}|_k \)

Device simulator iterates until convergence

- Last iteration as initial-guess
  - Linear prediction algorithm

Quasi Full-Newton

- Only one iteration of device simulator

  - Calculation of \( I^k_C \) and \( g^k_{eq} \)

Advantages

- Straight-forward parallelization
- \texttt{SPICE}-like damping schemes can be applied
- Stable operating point computation
Simulator Coupling

Two-Level Newton

Device simulator is called for each circuit iteration

*Fixed set of contact voltages*

*Contact current response* $I_C^k$

*Problematic:* $g_{eq}^k = \frac{\partial I_C}{\partial V_C} |_k$

Device simulator iterates until convergence

Last iteration as initial-guess

*Linear prediction algorithm*

Quasi Full-Newton

Only one iteration of device simulator

*Calculation of* $I_C^k$ *and* $g_{eq}^k$

Advantages

Straight-forward parallelization

*SPICE*-like damping schemes can be applied

Stable operating point computation

Disadvantages

Considerable overhead
Device and circuit equations in one matrix

Simultaneous damping of device and circuit equations
Full-Newton Approach

Device and circuit equations in one matrix
  Simultaneous damping of device and circuit equations

No simulator communication overhead
  No input-deck generation, no temporary input and output files, etc.
Full-Newton Approach

Device and circuit equations in one matrix
  Simultaneous damping of device and circuit equations

No simulator communication overhead
  No input-deck generation, no temporary input and output files, etc.

Full-Newton equation system extremely sensitive to node voltages
Full-Newton Approach

Device and circuit equations in one matrix
   Simultaneous damping of device and circuit equations

No simulator communication overhead
   No input-deck generation, no temporary input and output files, etc.

Full-Newton equation system extremely sensitive to node voltages

Properties of the newton method
   Quadratic convergence properties for a good initial-guess (fast!)
   Initial-guess hard to construct
   Damping schemes
**Full-Newton Approach**

Device and circuit equations in one matrix
- Simultaneous damping of device and circuit equations

No simulator communication overhead
- No input-deck generation, no temporary input and output files, etc.

Full-Newton equation system extremely sensitive to node voltages

Properties of the newton method
- Quadratic convergence properties for a good initial-guess *(fast!)*
- Initial-guess hard to construct
- Damping schemes

Reliable DC operating point calculation of utmost importance
- Drift-diffusion solution as initial-guess for
  - Higher-order transport models
  - Electro-thermal solution
- Transient simulations better conditioned
Why is convergence hard to obtain?
Why is convergence hard to obtain?

Conventional boundary condition for numerical devices

\[ V_{C,i} (\text{device contact potential}) = \varphi_{C,i} (\text{node voltage}) \]

Carrier concentrations depend exponentially on the potential
Convergence

Why is convergence hard to obtain?

Conventional boundary condition for numerical devices

\[ V_{C,i} \text{ (device contact potential)} = \varphi_{C,i} \text{ (node voltage)} \]

Carrier concentrations depend exponentially on the potential

No pure voltage boundary conditions

Current flowing out of the contact affects node voltages

*System is extremely unstable at the beginning of the iteration*

Similar situation as with current boundary condition

Shifts in the DC offset require many iterations

*Distributed quantities provide 'internal state'*
Convergence

Why is convergence hard to obtain?

Conventional boundary condition for numerical devices

\[ V_{C,i} \text{ (device contact potential)} = \varphi_{C,i} \text{ (node voltage)} \]

Carrier concentrations depend exponentially on the potential

No pure voltage boundary conditions

Current flowing out of the contact affects node voltages

*System is extremely unstable at the beginning of the iteration*

Similar situation as with current boundary condition

Shifts in the DC offset require many iterations

*Distributed quantities provide 'internal state’*

Alternative boundary condition for numerical devices

\[ V_{C,i} = \varphi_{C,i} - V_{\text{ref}} \quad \text{with} \quad V_{\text{ref}} = \frac{1}{N_c} \sum_j \varphi_{C,j} \text{ (average potential)} \]

Average potential changes during the iteration and operation
Simple Methods

Limitation of node voltage update to $2V_T$

*Many iterations needed*

Initial guess close to the solution (experimental value: $\pm 0.2$ V)
Convergence – Damping Schemes

Simple Methods

- Limitation of node voltage update to $2V_T$
  - Many iterations needed
- Initial guess close to the solution (experimental value: $\pm 0.2$ V)

Traditional device simulation methods

- Damping after Bank and Rose (SIAM 1980)
- MINIMOS damping scheme

Standard damping schemes not suitable for mixed-mode problems
Convergence – Embedding Scheme

Shunt an iteration dependent conductance $G^k_S$ at every contact

Purely empirical expression

$$G^k_S = \max \left( G_{\text{min}}, \ G_0 \times 10^{-k/\kappa} \right)$$

$G_0 = 10^{-2} \ \text{S}$

$G_{\text{min}} = 10^{-12} \ \text{S}$

$\kappa = 1.0 \ldots 4.0$
Shunt an iteration dependent conductance $G^k_S$ at every contact

Purely empirical expression

$$G^k_S = \max\left(G_{\min}, G_0 \times 10^{-k/\kappa}\right)$$

- $G_0 = 10^{-2}\, S$
- $G_{\min} = 10^{-12}\, S$
- $\kappa = 1.0 \ldots 4.0$

Method works for small circuits

- Zero initial-guess for node voltages
- Charge neutrality assumptions for semiconductor devices
- Convergence within 20–50 iterations
- Comparable to SPICE with compact models
Examples

Five-stage CMOS ring oscillator
  Long-channel/short-channel behavior

Electro-thermal analysis of an operational amplifier ($\mu$A709)
Five-Stage CMOS Ring Oscillator
CMOS Ring Oscillators

Long-channel devices ($L_g = 2 \mu m$)

First timestep: $\varphi_{in} = 0$ V

Excellent agreement DD and ET

Non-local effects negligible

\[
\varphi_1 \quad \varphi_2 \quad \varphi_3 \quad \varphi_4 \quad \varphi_5
\]

\[
0 \quad 0.5 \quad 1 \quad 1.5
\]

\[
0 \quad 2 \quad 4 \quad 6 \quad 8 \quad 10
\]
CMOS Ring Oscillators

Long-channel devices \( (L_g = 2 \, \mu m) \)

First timestep: \( \varphi_{in} = 0 \, V \)

Excellent agreement DD and ET

Non-local effects negligible

Short-channel devices \( (L_g = 0.13 \, \mu m) \)

Significant difference DD and ET

Non-local effects important

Larger currents for ET

15% difference in delay time

Complexity of models can be increased

Higher-order transport models

More accurate quantum corrections

Different mobility models
Electro-Thermal Analysis of a $\mu$A709
Electro-Thermal Analysis of a $\mu$A709

Temperature Gradient
Thermal coupling modeled via a thermal circuit

Thermal coupling between individual devices

Thermal equations similar to Kirchhoff’s equations

*Formally derived from the discretized lattice heat-flow equation*
Simple thermal equivalent circuit

\[ P_1, G_1, G_{1,9} \rightarrow \vartheta_1 \]

\[ G_{1,15} \rightarrow \vartheta_1, \vartheta_9 \]

\[ G_{2,9} \rightarrow \vartheta_2 \]

\[ G_{2,15} \rightarrow \vartheta_2, \vartheta_{15} \]

\[ P_2, G_2, G_{15} \rightarrow \vartheta_2, \vartheta_{15} \]

\[ \vartheta_{\text{ref}} \rightarrow \vartheta_1, \vartheta_2, \vartheta_9, \vartheta_{15}, \vartheta_{\text{ref}} \]

\[ P_9, G_9 \rightarrow \vartheta_{\text{ref}}, \vartheta_9 \]

\[ P_{15}, G_{15} \rightarrow \vartheta_{\text{ref}}, \vartheta_{15} \]
Electrical simulation

All 15 transistors numerically simulated
System-size: 37177, simulation time: 1:08 hours (101 points, DC transfer)
Electro-Thermal Analysis of a $\mu$A709

**Electrical simulation**

- All 15 transistors numerically simulated
- System-size: 37177, simulation time: 1:08 hours (101 points, DC transfer)

**Electro-thermal simulation**

- Input and output stage with self-heating (4 Transistors)
- Thermal coupling effects
  - Thermal feedback from the output to the input stage
  - Thermal interaction between all 4 transistors
- Highly non-linear problem, complex convergence behavior
- System-size: 40449, simulation time: 3:08 hours
Electro-Thermal Analysis of a \( \mu A709 \)

**Electrical simulation**

All 15 transistors numerically simulated  
System-size: 37177, simulation time: 1:08 hours (101 points, DC transfer)

**Electro-thermal simulation**

Input and output stage with self-heating (4 Transistors)  
Thermal coupling effects  
   *Thermal feedback from the output to the input stage*  
   *Thermal interaction between all 4 transistors*  
Highly non-linear problem, complex convergence behavior  
System-size: 40449, simulation time: 3:08 hours

**Electro-thermal simulation with simplified self-heating model**

Same coupling effects as before  
Practically same results  
System-size: 38477, simulation time: 1:22 hours
Electro-Thermal Analysis of a $\mu$A709

DC Stepping

Gain $\approx$ 35000
$\Delta \varphi_{out} = 0.7$ V (101 points)
Critical point 0 V

Thermal feedback caused bumps

Input stage: $\Delta T$

$\Delta T \propto P$

$max(\Delta T) = -22$ mK

Input voltage difference
Electro-Thermal Analysis of a $\mu$A709

Open-loop voltage gain $|A_v|$

Optimistic thermal conductances

Stronger impact published

$|A_v|$ can even change sign

OpAmp can become unstable
Conclusions

For circuit design compact models are indispensable
Intermediate phase when devices structures is not established
   Mixed-mode circuit/device simulation can be used

Motivation for mixed-mode device-circuit simulation
   When compact models are inconvenient/not available
   Verification of compact models in a more realistic environment
   Optimization of devices
   Exploitation of new device designs

Examples have been simulated with Minimos-NT
   Go to http://www.iue.tuwien.ac.at and try it