Enablement of compact models for ultra-scaled CMOS technologies

D. Yakimets, P. Schuddinck, D. Jang, M. Garcia Bardon, N. Sharan, B. Parvais*, P. Raghavan, and A. Mocuta

IMEC, Kapeldreef 75, 3001 Leuven, Belgium
*also with Vrije Universiteit Brussel, Belgium

10th International MOS-AK Workshop – Dec. 6, 2017
**LOGIC ROADMAP VIEW**

Happy scaling era
- # transistors per area doubles every two years for same cost

Less happy scaling era
- Still doubles but device scaling provides diminishing returns

14nm: FinFET
- FinFET device saves the day

20nm: First sign of trouble
- Double patterning (cost !)
  - Planar device runs out of steam

10-7nm: More trouble
- Multi-patterning cost escalates

7-5nm: At last...
- EUV reduces cost

4-3nm:
- Nanowire/sheet FET to continue gate length scaling
- 2nm: High-NA EUV?

- Heterogenous scaling
  - 2D Material for power gating
  - Back Side PDN
  - MRAM
  - ...

Focus of process technology innovation is
- Scale device and wire
- Scale basic logic cells
- Scale (sub-)system functions

New Compute
- Machine learning
- Quantum computing
- Cryogenic electronics
SCALING IS ABOUT POWER-PERFORMANCE-AREA (PPA) AND COST (C)

- Recent scaling trends have involved cell height reduction and fins depopulation causing
  - FEOL structural changes (taller fins / nanosheets)
  - Relative increase in MOL / BEOL importance

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7nm</td>
<td>56</td>
<td>40</td>
<td>7.5T</td>
<td>3</td>
<td>1.00</td>
</tr>
<tr>
<td>5nm</td>
<td>48</td>
<td>28</td>
<td>6.0T</td>
<td>2</td>
<td>0.48</td>
</tr>
<tr>
<td>3nm</td>
<td>42</td>
<td>21</td>
<td>5.5T</td>
<td>2 / 1</td>
<td>0.29</td>
</tr>
</tbody>
</table>
**Scaling is about Power-Performance-Area (PPA) and Cost (C)**

- What device architecture / MOL scheme can enable required massive PP gains?
- Which performance boosters are needed?
  - Stressors
  - Low-k spacers
  - Contact improvements
- How to enable PPA with compact models? (TCAD is too slow to cover all experiments)
Flow for PPA assessment at imec
Quite common for the industry
FLOW FOR PPA ASSESSMENT AT IMEC
QUITE COMMON FOR THE INDUSTRY

Accurate compact models for both intrinsic transport and RC-parasitics are crucial elements of the flow.
OU TLINE

- Introduction

- Compact model structure:
  - Intrinsic part
  - Parasitics

- Example of RC-extraction for vertical transistor

- Summary
OUTLINE

- Introduction

- Compact model structure:
  - Intrinsic part
  - Parasitics

- Example of RC-extraction for vertical transistor

- Summary
HARDWARE-INFLUENCE ON CM DEVELOPMENT
EITHER DIRECT OR THROUGH TCAD

Lateral GAAFETs with inner spacers

Low contact resistivity on blanket wafers

S/D junction engineering

TCAD simulations

Compact model and RC-parasitics

3D solvers for RC extraction
FLOW FOR PPA ASSESSMENT AT IMEC
QUITE COMMON FOR THE INDUSTRY

1. INV layout
2. PEX deck
3. Annotated netlist (cell parasitics + device parameters)
4. Parasitics compact model
5. Core compact model (calibrated to TCAD and HW)
6. RO simulations
7. PPA
TCAD-BASED COMPACT MODEL FOR INTRINSIC TRANSPORT

**Extended BSIM-CMG**

**TCAD**

- Ballistic Ratio (Literature or advanced sims)
- Ballistic current (TCAD S-band)
- Electrostatics and charge (TCAD DD with QM-corrections)

**CM Fitting based on BSIM-CMG model**

1. **Current (IV) fitting:**
   \[ I_{DSAT}(L_g, \text{Stress}) = I_{BAL}(\text{Stress}) \times BR(L_g, \text{Stress}) \]
   - Full ballistic currents are independent of \( L_g \) and \( V_D \).
   - Quasi-ballistic current by applying \( BR(L_g, \text{Stress}) \).
   - CM is assumed to have no series resistance and ideal electrostatics (Turned off all related BSIM parameters).

2. **Capacitance (CV) fitting**
   - Capture QM charge centroid behavior
   - Decoupled from IV fitting (QMTCENCV=1)

3. **SS/DIBL fitting**

**BSIM parameters**

- \( U_0 \)
- \( V_{SAT} \)
- \( CGS(D)O \)
- \( CGS(D)L \)
- \( QM0 \)
- \( PQM \)
- \( ETAQM \)
- \( TOXP \)
- \( SS \)
- \( DIBL \)
- \( CDSC \)
- \( CDSCD \)
- \( DVT1SS \)
- \( CIT \)
- \( ETA0 \)
- \( DSUB \)
- \( DVTP0 \)
- \( DVTP1 \)
**Drive Current**

**Ballistic Current**

Full ballistic current provide good indication of intrinsic potential of various devices

May be easily simulated as a function of stress
Drive Current

Ballistic current times Ballistic ratio = Target Drive Current

Ballistic ratio is tricky to compute as various scattering mechanisms affect it

\[ I_{DSAT}(L_g, \text{Stress}) = I_{BAL}(\text{Stress}) \times BR(L_g, \text{Stress}) \]
HOW TO CAPTURE THIS IN BSIM-CMG?

BRIEF THEORETICAL BACKGROUND

- BSIM model has been described with a drift-diffusion model which is defined by carriers mobility ($\mu_{DD}$):
  - At low fields $\mu_{DD} \approx \mu_{low}$
  - At high fields (high $V_{DS}$ and/or short gates) $v_{sat}$ defines $\mu_{DD}$

- This results in two key parameters: $\mu_{low}$ and $v_{sat}$

- The $\mu_{low}$ is replaced by the apparent mobility due to ballistic mobility reduction and additional scattering mechanisms.

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_{long}} + \frac{\alpha \mu}{\mu_{sat}}$$

$\mu_{DD}(E) = \frac{\mu_{low}}{1 + \frac{\mu_{low}E}{v_{sat}}}$

[Graph showing mobility vs. gate length]
**Capacitance Fitting**

**Both Channel capacitance and overlap capacitance should be fitted**

Gate capacitance (channel & S/D overlap) is fitted to *S-device TCAD* for short-channel devices.
**How to Capture This in BSIM-CMG?**
**By Modifying Charge Centroid**

Confinement has been considered in CV fitting by using QMTCENCV=1 but this does not change IV results (i.e. decoupled).

\[ <\text{BSIM-CMG v107.0}> \]

**Charge Centroid Calculation for Inversion**

\[
T_4 = \frac{q_{ba} + ETAQM \cdot q_{ba}}{QM0} \quad (3.302)
\]

\[
T_5 = 1 + T_4^{\eta_{QM}} \quad (3.303)
\]

\[
T_{cen} = \frac{T_{cen0}}{T_5} \quad (3.304)
\]

If \( GEOMOD = 1 \) then

\[
W_{eff} = W_{eff0} - 4 \cdot QMTCENIV_i \cdot T_{cen} \quad (3.307)
\]

\[
W_{eff, CV} = W_{eff, CV0} - 4 \cdot QMTCENCNV_i \cdot T_{cen} \quad (3.308)
\]
Fitting Process for Electrostatics

The Last Step Towards Intrinsic Compact Model

L_g dependency of SS/DIBL at fixed V_{dd} / I_{off} (TCAD)

Tuning BSIM parameters for SS/DIBL separately

SPICE simulation at varied L_g and same I_{off} for single fin structure

Comparing between CM and TCAD

Done.

<BSIM-CMG v107.0>

Vth Roll-off, DIBL, and Subthreshold Slope Degradation

\[
\psi_m = 0.4 + PHIN_i + \Phi_B
\]

\[
C_{dxe} = \frac{0.5}{\cosh \left( DVT1SS_i \cdot \frac{L_{eff}}{\lambda} \right) - 1} \cdot (CDSC[N] + CDSCD_a \cdot V_{dxx})
\]

SS

\[
n = \begin{cases} 
\Theta_{SS} \cdot \left( 1 + \frac{C_{TSS} + C_{dxe}}{(2C_{TSS}) C_{dxe}} \right) & \text{if GEOMOD} \neq 3 \\
\Theta_{SS} \cdot \left( 1 + \frac{C_{TSS} + C_{dxe}}{C_{dxe}} \right) & \text{if GEOMOD} = 3 
\end{cases}
\]

\[
\Delta V_{th,SCB} = -\frac{0.5 \cdot DVT0_i}{\cosh \left( DVT1_i \cdot \frac{L_{eff}}{\lambda} \right) - 1} \cdot (V_{th} - \psi_m)
\]

Vt roll-off

\[
\Delta V_{th,DIBL} = -\frac{0.5 \cdot ETA_{0i}}{\cosh \left( DSUB_i \cdot \frac{L_{eff}}{\lambda} \right) - 1} \cdot V_{dxx} + DVTP0 \cdot V_{dxx}^{DVTP1}
\]

DIBL

In v107.0 BSIM-CMG, Vt roll-off and SS were decoupled by introducing the DVT1SS parameter which simplified fitting a lot.
Summary on Intrinsic Transport

Device model (CM)

- Electrostatics (S-device)
- Quasi-Ballistic (S-band + Ballistic ratio)
- FEOL parasitics

imec CM is calibrated with
- quasi-ballistic current
- quantum-mechanical CV
- electrostatics based on TCAD simulation
OUTLINE

- Introduction

- Compact model structure:
  - Intrinsic part
  - Parasitics

- Example of RC-extraction for vertical transistor

- Summary
FLOW FOR PPA ASSESSMENT AT IMEC
QUITE COMMON FOR THE INDUSTRY

INV layout → PEX deck → Annotated netlist (cell parasitics + device parameters)

PPA ← RO simulations ← Parasitics compact model

Core compact model (calibrated to TCAD and HW)
FLOW FOR PPA ASSESSMENT AT IMEC
QUITE COMMON FOR THE INDUSTRY

- INV layout
- PEX deck
- Annotated netlist (cell parasitics + device parameters)
- Parasitics compact model
- Core compact model (calibrated to TCAD and HW)
- RO simulations

Integration-dependent, hard to maintain in the analytical form even for FF/NW/NSh. Switched to finite element modelling for the RC calculations.
**Flow for PPA Assessment at IMEC**

Quite common for the industry

- **INV layout**
- **PEX deck**
- **RO simulations**
- **Annotated netlist** (cell parasitics + device parameters)
- **Parasitics compact model**
- **Core compact model** (calibrated to TCAD and HW)

Macro-model goes hand-in-hand with device recognition in the PEX deck

Integration-dependent, hard to maintain in the analytical form even for FF/NW/NSh. Switched to finite element modelling for the RC calculations

Hard to construct for advanced devices (VFET/CFET, etc.)
Device model covers parasitic components up to the self-aligned IM1 contact

Enforced boundaries result in capacitance loss, say from IM2 to fins

(this picture is a bit old, layer names changed, but the idea is correct)
Various resistances:
- Extension
- S/D Epi
- Contact
- M0A core material
- M0A liners

- Capacitances are computed with 3D finite element modeling software
- Resistances and capacitances are connected in a distributed network
**Edge Effects Become More and More Critical**  
(Due to Cell Height Reduction and Fin Depopulation)

Simplified analytic model on relaxed pitches

3D simulations critical at tighter pitches with actual designs from .gds accounting for all fringe capacitances
**Actual Designs From .gds**

**But still with the same distinction between macro and PEX**

Output is at M1, not captured during 3D RC extraction

Independent simulations might be needed if devices are not symmetrical

There are a lot of various devices in a full cell library → heavy .va look-up tables have to be generated
**Full TCAD Based DTCO**

**That is probably way too heavy**
Yet, the idea of RC-extraction from a cell is attractive (in the form of distributed RC-netlist)

- (Almost) no need in a macro model
- Lighter PEX deck because of simplified device recognition
- No missing capacitances because there would be no macro / PEX boundary

----------

- Can this be **fast** and to be used for full library characterization? As otherwise this solution is limited to the path-finding PPA on RO-level.
OUTLINE

- Introduction

- Compact model structure:
  - Intrinsic part
  - Parasitics

- Example of RC-extraction for vertical transistor

- Summary
An example of RC-analysis of a VFET device

Why vertical? -- Gate pitch budgeting is challenging

- Parasitic capacitance
- Reliability
- Electrostatics control
- Direct S/D tunneling
- Gate stack
- Access resistance
- S/D stressors
An example of RC-analysis of a VFET device

Why vertical? -- Gate pitch budgeting is challenging

- Parasitic capacitance
- Reliability
- Electrostatics control
- Direct S/D tunneling
- Gate stack
- Access resistance
- S/D stressors

FinFET → Vertical FET
A VARIANT OF VFET

2-TRANSISTORS PRIMITIVES

→ Transistors are chained by **pairs**

→ This ensures that all device terminals are accessible
A VARIATION OF VFET
NAND2 BECOMES THE SMALLEST CELL, NOT INV

Series connection is typically tricky with VFETs

→ NAND gate structurally reduces by 50% by the absence of dummies
A VARIANT OF VFET

SELF-ALIGNED GATE

$$D_{NW} = 10 \text{ nm}$$

$D_D = 20 \text{ nm}$

$D_{NW} = 10 \text{ nm}$

$D_S = 20 \text{ nm}$

Si VFET

Si & Si/SiGe VFET

Make S/D with SiGe instead of Si

Possibility to create a thinner core due to different etch rate (Si vs SiGe)
**KEY CAPACITANCES**

Electrodes are split into logical segments:
- overlap capacitance
- capacitance to via
- ...

![Graph showing capacitance values for different electrodes](image)

**Capacitance [aF]**

- G-Tov
- G-Bov
- G-Top
- G-Bot
- G-SiGe
- Top-Bot

**Dielectrics:**
- HfO$_2$
- SiCO
- Oxide

**Conductors:**
- Cobalt
- Silicon
- SiGe

![Diagram showing electrode configurations](image)
SENSITIVITIES TO KEY GEOMETRICAL PARAMETERS

CAPS ARE RATHER INSENSITIVE AND DOMINATED BY OVERLAP CAPACITANCE

Sensitivities are given around the middle point for every parameter

![Graphs showing sensitivities to key geometrical parameters.](image)
**Device Resistance**

Plug is used for pMOS, but not for nMOS

- Top contacts
- Top epi
- Bottom epi
- Bottom contact

- 10 nm
- 15..25 nm
**Resistance Sensitivities: Top**

The wider the extension, the better the resistance

\[ R_C = \frac{\rho_C}{A} \]

Holes mobility is smaller than electrons mobility in SiGe
RESISTANCE SENSITIVITIES: BOTTOM
CURRENT FLOW IS DIFFERENT FOR nMOS AND pMOS (NAND2 CELL)
**RO Performance**

**Key bottle necks may be identified**

- $L_g = 15$ nm,
- Bottom SiGe is 25 nm thick,
- Top SiGe is 10 nm thick
- Top to bottom via spacing is 7 nm

<table>
<thead>
<tr>
<th>VFET config</th>
<th>$\rho_C$ [$\Omega \cdot \text{cm}^2$]</th>
<th>$N_{SD}$ [cm$^{-3}$]</th>
<th>Ext. W [nm]</th>
<th>Cov [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1e^{-9}$</td>
<td>$1e20$</td>
<td>21</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>$5e^{-10}$</td>
<td>$1e21$</td>
<td>21</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>$5e^{-10}$</td>
<td>$1e21$</td>
<td>18</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>$5e^{-10}$</td>
<td>$1e21$</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

Overlap capacitance limits performance

Resistance is less of an issue due to small currents
OUTLINE

- Introduction
- Compact model structure:
  - Intrinsic part
  - Parasitics
- Example of RC-extraction for vertical transistor
- Summary
SUMMARY

- For pre-silicon PPA analysis of advanced nodes, CMs should be very flexible and, ideally, modules based:
  - stress,
  - electrostatics,
  - quasi-ballistic current...

- Analytical macro-model is hard to maintain as more and more device options appear + various edge effects become non-negligible.

- Device RC .va look-up tables may be well generated with 3D TCAD tools, but this approach is not sustainable because further library characterization would require exhaustive look-up tables.

- RC extraction from a full cell might be too time consuming.