Invited Talk – MOS-AK Workshop – New Delhi

Compact Model Application to Statistical Variability and Reliability Studies

Xing Zhou

School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

March 16, 2012

Email: exzhou@ntu.edu.sg

Outline

● Background and motivation on compact model application to statistical variations and the need for new paradigm shift

● Demonstration of ideas and preliminary results through SiNW example

● Interface-trap model in core CM for physics-based reliability modeling

● Summary and outlook
Moore’s Law: Essence of Technology Scaling

What really dictates technology scaling is the Economics (as “governed” by Moore’s law), and the essence of which is the “yield” which is mainly determined by variability/reliability.

Sources of Variations: Lots/Dies/Process/Layout/Etc

Due to statistical variations, the same (designed) transistors at different locations on different dies from different lots all have different characteristics!
Types of Variations: Temporal and Spatial

- **Temporal variations** — the *same device behaves differently at different times*
  - Random: transistor intrinsic noise (RTS, flicker, thermal, shot) — important for analog/RF circuits
  - Deterministic: coupling noise (cross-talk, EM interference, supply/substrate noise) — important for digital circuits
  - Deterministic: transistor aging (BTI, HCI) — important for reliability

- **Spatial variations** — *identically-designed devices at different locations* behave differently (i.e., no two devices are identical)
  - Systematic: parametric (inter/intra die variations) — important for process control
  - Stochastic: intrinsic variations (RDF, LER; mismatch – ‘spatial noise’) — important as a fundamental limit

- **Environment changes** — operating or parametric *temperature drift*
  - Deterministic: both temporal (self-heating) and spatial (hot-spot) variations

Visualizing Effects of Variation/Noise/Aging

- **Designed behavior** — *ideal characteristics*, as if drawing waveforms using an “infinitely” thin pencil
  - Deterministic variations: should be modeled physically
  - Random variations: modeled physically (noise) or stochastically (Monte Carlo sampling)

- **Spatial variations/temporal noise** — *real behaviors of an ensemble of devices or fluctuating waveforms*, as if drawing by a pencil with finite thickness
  - Variability/reliability/noise/aging: currently all modeled as an “after-thought” of ideal designs

- **New challenge of scaling** — *nominal design scaled but variations do not!*

How to design chips together with variations/noise/aging?
Ideal (Model) vs Real (Statistical) Device Characteristics

**Basic circuit analysis assumptions**

- **Performance** (FOM) — measured by SPICE circuit simulation
- **Technology** — represented by compact model parameters
- **Technology Variations** — given by model parameter statistical distributions

**Compact Model — Bias/Process/Geometry Variations**

- **Combined effect** due to bias/process/geometry variations

- **Bias variation**

- **Process variation**

- **Geometry variation**

- **Compact model can well capture intentional variations by design** — Spice circuit simulation

- **It can be extended to model unintentional variations by 'sampling'** — Spice MC simulation

- **Question: how to use a deterministic compact model in a stochastic manner?**
Observations and Motivation

- **CMOS-compatible SiNW/FinFET technology** — likely to be emerging after conventional planar bulk/SOI devices (thermal oxidation/diffusion in Si — better reproducible over 12” wafers; all other alternative choices are more difficult to meet Moore’s law on economics/yield’ in the foreseeable future)

- **A small number of physically-scalable model parameters** — a prerequisite for meaningful statistical studies (since statistical variations are captured by random sampling on a deterministic model)

- **Deterministic variations** (coupling noise, aging) need to be modeled physically, so do random variations (mismatch/variance, just as transistor noise modeling) — need to start from the same “core” compact model
  - SiNWs/FinFETs have less intrinsic variations — undoped [RDF] ultrathin body [SCE] and litho-independent gate width [LER] with multi-fingers

- **Challenges** — the need for new approaches and a paradigm shift to statistical compact modeling (“stat-CM”)

---

How Can a Compact Model “Capture” Statistics?

- **A model is a mental image of reality**
  - Different equations — different “images”
  - Different choice of parameters — correlations?

\[ \text{I}_D = f(V_{xy}, W/L, T, T_{ox}, N_A, X_j, \mu_n, V_t, \ldots) \]

- **Measured variability ET data (“reality”)** — includes ALL possible causes (“model” independent!)

- **Modeled statistics** (“mental image”) — due to a chosen set of parametric distributions of a given CM (model/parameter-dependent!)

  - How to separate various ‘causes’ (intrinsic vs. process)?
New Paradigm: Statistical Compact Model

\[ y = f(x_1, \ldots, x_n) \]

Assuming: \( y = I \) represents the transistor behavior (function/derivatives) accurately with respect to bias/geometry/structural parameters \((x_1, \ldots, x_n)\)

(where \( y \) can be \( I_{on}/I_{off}, g_m/g_d, V_t, \ldots \))

- **Statistical-CM** — Study FOM’s variations using CM equations, together with statistical data — derivation of transistor/gate FOM’s compact equations

<table>
<thead>
<tr>
<th>Mismatch</th>
<th>Variance</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \Delta y = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial x_i} \right) \Delta x_i ]</td>
<td>[ (\Delta y)^2 = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial x_i} \right)^2 (\Delta x_i)^2 ]</td>
<td>A small number of uncorrelated parameters</td>
</tr>
</tbody>
</table>

- **Validation** by comparison with Monte Carlo sampling [CM validation with TCAD]

- **Propagation** to gate-level (CSM) by providing the delay table (input-slope, load) and parameter sensitivity (partial derivative of charge w.r.t. parameters)

**SiNW Transistor “Statistical Compact Model”**

Pre-requisite for meaningful statistical modeling:
- A small number of physically scalable model parameters
- Identify FOM’s at different levels (transistor/circuit)
- Validating “stat-CM” with numerical corners and MC

**Analog FOM’s:** \( g_m/g_d \)

\[ g_m \equiv \left. \frac{dI_{ds}}{dV_s} \right|_{V_{ds}=V_{ds}^\text{ref}, V_{gs}=V_{gs}^\text{ref}} = f_3(V_{dd}; L_g, R, T_{ox}, \ldots) \]

\[ g_d \equiv \left. \frac{dI_{ds}}{dV_g} \right|_{V_{ds}=V_{ds}^\text{ref}, V_{gs}=V_{gs}^\text{ref}} = f_4(V_{dd}; L_g, R, T_{ox}, \ldots) \]

Compact model equations of FOM’s as functions of \( V_{dd}, L_g, R, T_{ox}, \ldots \)
SiNW Transfer Characteristics with Varying Radius

Model validation with various wire radius in comparison with numerical data

SiNW Saturation Characteristics

Model validation with various gate lengths in comparison with numerical data

I_{on} determines logic speed

I_{off} determines standby power
SiNW Output Characteristics

Model validation with various supply power in comparison with numerical data

\[ I_{on}(V_{dd}) \]

\[ g_d(V_{dd}) \]

\[ I_{on} \text{ and } g_d \text{ scaling with } V_{dd} \text{ (for speed/power trade-off)} \]

Two Variables with Gaussian/Uniform Distributions

- Evaluate CM target \((I_{on})\) vs. variables \((L_g, V_{dd})\)
- Run CM Monte Carlo with given distributions (e.g., Gaussian with 3\(\sigma\))
- Validate target variation with numerical corners
- Derive target stat-CM mean/variance
Three Variables: \(L_g, R, T_{ox}\) Gaussian Distributions

Mean \(L_g = 45\) nm, \(R = 10\) nm, \(T_{ox} = 2\) nm; Variations: +12%

Combined distributions (with correlations) validated by Medici best/worst corners (symbols).

SiNW Mismatch — Analytical “Stat-CM” At Any Bias

\[
\sigma_{I_{ds}}^2 = \left( \frac{1}{I_{ds}} \frac{\partial I_{ds}}{\partial R} \right)^2 \sigma^2 (R) \\
+ \left( \frac{1}{I_{ds}} \frac{\partial I_{ds}}{\partial V_{gs}} \right)^2 \sigma^2 (V_{gs}) \\
+ \left( \frac{1}{I_{ds}} \frac{\partial I_{ds}}{\partial \mu} \right)^2 \sigma^2 (\mu) \\
+ \rho (R, V_{gs}, \mu)
\]

where \(V_{gs} = V_g - V_{FB}\) (variance can be due to process \(V_{FB}\) or input voltage \(V_{gs}\)).

A compact statistical model readily applicable to major process variations (\(R, V_{FB}, \mu\)) and/or input (gate) voltage variations (\(V_{gs}\)).

A compact statistical model readily applicable to major process variations (\(R, V_{FB}, \mu\)) and/or input (gate) voltage variations (\(V_{gs}\)).

\(\sigma^2(\mu)\) can be similarly modeled with equations and more “independent” variables (\(N_A, T_{ox}, T, \ldots\))
CMOS Inverter Transfer Characteristics and Noise Margin

Mean $R = 10\,\text{nm}$, $L_g = 45\,\text{nm}$, $V_{dd} = 1.2\,\text{V}$; Variations: 12% Gaussian Distributions

Noise Margins with Gaussian Distributions
CMOS Inverter Transient and Delay/Power Definitions

Transient Characteristic: Delay Definition

\[ V_{\text{out}} (V) \]
\[ V_{\text{in}} (V) \]
\[ I(V_{dd}) \ (mA) \]

Input & Output Voltage, \( V_{\text{in}} \) & \( V_{\text{out}} \) (V)

\[ 50\% \ V_{dd} \]

Supply Current, \( I(V_{dd}) \) (mA)

Time, \( t \) (ns)

Propagation Delays with Gaussian Distributions

Mean \( R = 10 \) nm, \( L_g = 45 \) nm, \( V_{dd} = 1.2 \) V

Variations: 12%

Gaussian Distributions

Distribution (a.u.)

Propagation Delay (High-Low), \( t_{pHL} \) (ns)

Distribution (a.u.)

Propagation Delay (Low-High), \( t_{pLH} \) (ns)
## Gate-Level CSM for Parametric Variations

### Current Source Model (CSM) at TUM

Courtesy: C. Knoth (TUM)

\[
I_{dc,1} = f(v_1, v_2) \quad Q_1 = h(v_1, v_2) \quad I_{dyn,1} = Q_1 \n\]

\[
I_{dc,2} = g(v_1, v_2) \quad Q_2 = k(v_1, v_2) \quad I_{dyn,2} = Q_2
\]

dI/dP_x and dQ/dP_x can be obtained from CM.

---

## Physics-Based Interface-Trap Model

We have already developed a physics-based interface-trap model into the surface-potential based core compact model (Xsim). It will serve as the basis for incorporating physics-based reliability compact model in the proposed ‘rel-CM’.

### Surface Potential

- nMOSFET
  - \( V_{GS} = 0 \text{mV} \)
  - \( T_{OX} = 3.5 \text{nm} \)
  - \( N_x = 1 \times 10^{17} \text{cm}^{-3} \)
  - \( N_{ET} = 0 - 5 \times 10^{12} \text{cm}^{-2} \)

\( \phi_s (V) \) vs. \( V_{GB} (V) \)

\( \phi_s \) increasing with \( N_{ET} \)

Symbol: Xsim
Line: Numerical
Paired-linear traps

### Gate Capacitance

\( \frac{C}{C_{ox}} \) vs. \( V_{GB} (V) \)

- \( N_{ET} \): 5 \times 10^{12} - 3 \times 10^{13} \)

Symbol: Xsim
Line: Numerical
Paired-linear traps
Drain-Current & Transconductance Variation with Traps

**MOS-AK / India EEE / NTU**

Full bias-dependence in the current/charge model with $V_t$ and SS dependence on $N_{it}$ for physics-based aged device modeling and prediction.

- Next step: linking interface-trap density to stress time.

**Transfer $I_d$–$V_g$**

![Graph showing $I_d$ vs $V_g$](image)

**Transfer $g_m$–$V_g$**

![Graph showing $g_m$ vs $V_g$](image)

---

U-shaped Interface Traps into the Surface Potential

- Inclusion of interface-trap models into the surface-potential solution allows physical (rather than empirical) modeling and prediction of “aged” devices with a complete $\phi_s$-based core model.

**U-shaped energy distribution of neutral electron/hole interface traps**

- Donor-like
  - $E_C$
  - $E_F$
  - $E_V$
  - $-ve\ N_{it}$ at flatband

![U-shaped energy distribution](image)

- Acceptor-like
  - $E_C$
  - $E_F$
  - $E_V$
  - $+ve\ N_{it}$ in inversion

**Q_{it}[\phi_s(V_g)] built into the CM**

![Graph showing $Q_{it}$ vs $V_g$](image)

When the stress voltage is removed, the low energy portion of neutral interface traps, such as those near the conduction- and valence-band edges, would quickly lose charges and become neutral, which can account for the fast recovery of NBTI.

**$V_t$ recovery during NBTI**

![Graph showing $V_t$ vs time](image)

Summary and Outlook

- ** Thoughts on new approaches (paradigm shift) to statistical compact modeling — mismatch/variance (“spatial noise”) modeling based on a single core CM **

- ** Preliminary demonstration of ideas through SiNW example — extendable to other FOMs and more physical variables **

- ** Remaining challenges **
  - Approaches to mapping parameter variance to measured statistics
  - Gate-level FOM compact models (CM at the gate/cell level)

- ** Bridging over to gate-level variability/reliability modeling and designs **

- ** Ultimate goal — new approaches and tools for variation-aware technology/circuit co-design and prediction within a unified multi-level core model framework **