Intrinsic and Systematic Variability in Nanometer CMOS Technologies

Kedar Patel
Outline

• Introduction

• Robust method for estimation of LWR descriptors

• Comprehensive study of LWR for next-generation lithography processes

• Method to estimate FinFET device performance through the use of LWR descriptors

• Decomposition of Semiconductor Variability
Variation in a Semiconductor ICs

(a)

(b)

(c)

(d)

a) lot-to-lot variation
b) wafer-to-wafer variation
c) die-to-die or across-wafer variation
d) within-die variation

Schematic Courtesy: Kun Qian
Modeling Parameter

- **Process Parameter**
  - Examples: process temperature

- **Physical Parameter**
  - Examples: gate length

- **Electrical Parameter**
  - Examples: drive current

- **Circuit Parameter**
  - Examples: delay

Level of Abstraction:
- High
- Low
Intrinsic Process Variations in Nanometer Semiconductor ICs

- Random fluctuations in intrinsic physical parameters greatly impact circuit performance and yield
  - What you design ≠ what you get
- Huge impact on design optimization
  - Timing analysis (timing yield) affected by 20% [Orshansky, DAC02]
  - Leakage power analysis (power yield) affected by 25% [Rao, DAC04]
  - Circuit tuning: 20% area difference, 17% power difference [Choi, DAC04], [Mani DAC05]
- The three primary sources of intrinsic variability are:

  ![Line edge roughness (LER)](image1)
  ![Gate oxide thickness](image2)
  ![Random dopant fluctuations](image3)

Adapted from L. He (UCLA)
Origins of LER in Resist Processing
What is Line Edge Roughness (LER)?

\[ \sigma_{LWR}^2 = (N - 1)^{-1} \sum_{i=1}^{N} L_i - \bar{L} \]

\[ \sigma_{LWR}^2 = \sigma_L^2 + \sigma_R^2 - 2 \rho_X \sigma_L \sigma_R \]

\[ \sigma_L = \sigma_R \equiv \sigma_{LER} \]

\[ \sigma_{LWR}^2 = 2 \sigma_{LER}^2 (1 - \rho_X) \]

\[ \rho_X = 0 \]
\[ \sigma_{LWR}^2 = 2 \sigma_{LER}^2 \]

\[ \rho_X = 1 \]
\[ \sigma_{LWR}^2 = 0 \]

\[ \rho_X = -1 \]
\[ \sigma_{LWR}^2 = 4 \sigma_{LER}^2 \]
LWR Descriptors

\[ \tilde{\rho}(k) = \exp\left[-\left(\frac{k}{\xi}\right)^{2\alpha}\right] \]

\[ \sigma^2_{LWR,\infty} \]

\( \alpha \)

\( \xi \)
Device Performance

\[\sigma_d^2 = \sigma_I^2 \left( \frac{\partial d}{\partial I} \right)^2 = \left( \frac{\partial J(\bar{L})}{\partial \bar{L}} \right)^2 \sigma^2 \left( \frac{2\xi}{W} \right) \left( \frac{CV}{I^2} \right)^2\]
Outline

• Introduction

• Robust method for estimation of LWR descriptors

• Comprehensive study of LWR for next-generation lithography processes

• Method to estimate FinFET device performance through the use of LWR descriptors

• Decomposition of Semiconductor Variability
Characterization of LWR

\[ L = nd \]

\[ \bar{X}_j = \frac{1}{n} \sum_{i=1}^{n} X_{i,j} \]

\[ \hat{\sigma}_{LWR,j}^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_{i,j} - \bar{X}_j)^2 \]

\[ \bar{X}_j = \frac{1}{M} \sum_{j=1}^{M} \hat{\sigma}_{LWR,j}^2 \]

\[ \hat{\sigma}_{CD,local}^2 = (M-1)^{-1} \sum_{j=1}^{M} (\bar{X}_j - \bar{CD})^2 \]

\[ \bar{CD} = M^{-1} \sum_{j=1}^{M} \bar{X}_j \]
Generalized Framework

\( s = L \)

\( d \)

\( X_{is} \)

\( s = 1 \)

\( X_1 \)

\( X_2 \)

\( \ldots \)

\( X_i \)

\( \ldots \)

\( X_M \)

1 2 \ldots i \ldots M
Prior Work

- Effects of local CD non-uniformities are attributed to LWR

- What we are interested in is the variance in CD of a infinitely long line *purely* due to LWR

\[ \sigma^2_\infty = \sigma^2_{LWR}(L) + \sigma^2_{CD,\text{local}}(L) \]
Our Approach

1. Assume a model

\[ 2\gamma(h; \sigma^2, \theta) = 2\sigma^2 \left[ 1 - \exp \left( -\left| \frac{h}{\xi} \right|^{2\alpha} \right) \right], \quad h \in \mathbb{Z}. \]

2. Use Weighted Least Squares to estimating the parameters of model

\[
(\hat{\sigma}^2_{WLS}, \hat{\theta}) = \arg\min_{\sigma^2, \theta} \sum_{h=1}^{\ell_0} w(h) \left[ 2\hat{\gamma}(h) - 2\gamma(h; \sigma^2, \theta) \right]^2
\]

\[
w(h) = \frac{1}{\text{Var}(2\hat{\gamma}(h))}
\]

Bootstrap!
Bootstrap

Real World

\[ F \rightarrow X = \{X_1, X_2, \ldots, X_n\} \]

Bootstrap World

\[ \hat{F} \rightarrow X^* = \{X_1^*, X_2^*, \ldots, X_n^*\} \]

Unknown Distribution

Observed Data

Empirical Distribution

Bootstrap Sample

\[ h(X) \]

\[ \hat{\theta} \]

Statistic of Interest

Bootstrap Estimate

\[ h(X) \]

\[ \hat{\theta}^* \]

Block of Blocks Bootstrap

Original Data

\[ 2\hat{\gamma}(h) = \frac{1}{M(L-h)} \sum_{i=1}^{M} \sum_{p=1}^{L-h} \left( Y_{ip,1} - Y_{ip,(1+h)} \right)^2 \]

Bootstrapped Data

\[ 2\hat{\gamma}^*(h) = \frac{1}{M(L-h)} \sum_{i=1}^{M} \sum_{p=1}^{L-h} \left( Y_{ip,1}^* - Y_{ip,1+h}^* \right)^2 \]
Block of Blocks Bootstrap

\[
\tilde{\text{Var}}(2\hat{\gamma}(h)) = [w(h)]^{-1} = \frac{1}{B-1} \sum_{b=1}^{B} \left[2\gamma^*_b(h) - 2\tilde{\gamma}^*(h)\right]^2.
\]

WLS Fitting

\[
\tilde{\text{Var}}(\hat{\sigma}^2_{BBB}) = \frac{1}{B-1} \sum_{b=1}^{B} \left[(\sigma^2_{BBB})^*_b - (\bar{\sigma}^2_{BBB})^*\right]^2, \quad \{(\sigma^2_{BBB})^*_b : b = 1, \ldots, B\}
\]

\[
\tilde{\text{Var}}(\hat{\alpha}) = \frac{1}{B-1} \sum_{b=1}^{B} \left[\alpha^*_b - \bar{\alpha}^*\right]^2, \quad \{\alpha^*_b : b = 1, \ldots, B\}
\]

\[
\tilde{\text{Var}}(\hat{\xi}) = \frac{1}{B-1} \sum_{b=1}^{B} \left[\xi^*_b - \bar{\xi}^*\right]^2, \quad \{\xi^*_b : b = 1, \ldots, B\}
\]
Comparison of Methods
Outline

• Introduction

• Robust method for estimation of LWR descriptors

• Comprehensive study of LWR for next-generation lithography processes

• Method to estimate FinFET device performance through the use of LWR descriptors

• Decomposition of Semiconductor Variability
LWR (3σ)
LWR as % of CD

<10% final LWR for all NGL technologies
Correlation Length ($\xi$)

8-24nm; Meets ITRS roadmap
Roughness Exponent ($\alpha$)

DSA is closest to being ideal.
Transfer of LWR (1/3)

Example:
Self-Aligned Double Patterning (SADP) a.k.a “spacer lithography”
Transfer of LWR (2/3)
Transfer of LWR (3/3)
Outline

• Introduction

• Robust method for estimation of LWR descriptors

• Comprehensive study of LWR for next-generation lithography processes

• Method to estimate FinFET device performance through the use of LWR descriptors

• Decomposition of Semiconductor Variability
LER in FinFET

- Gate LER
- Fin LER
- Front and Back Gates

Parameters:
- $t_{fin}$
- $h_{fin}$
- $L_{bg}$
- $L_{fg}$
## Simulation Details

### TABLE 1: 2-D DEVICE SIMULATION PARAMETERS

<table>
<thead>
<tr>
<th>Electrical/Doping</th>
<th>Structural</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{dd}=0.9V$</td>
<td>$L_g=13\text{nm}$</td>
</tr>
<tr>
<td>$\phi_m=4.62\text{eV}$</td>
<td>$t_{ox}=6\text{A}$</td>
</tr>
<tr>
<td>$N_B=1\times10^{15} \text{ cm}^{-3}$</td>
<td>$L_{sp}=7.2\text{nm}$</td>
</tr>
<tr>
<td>$N_{s/d}=1\times10^{20} \text{ cm}^{-3}$</td>
<td>$t_{fin}=7.5\text{nm}$</td>
</tr>
<tr>
<td>$\sigma_{s/d}=4\text{nm/dec}$</td>
<td>$t_{poly}=13\text{nm}$</td>
</tr>
</tbody>
</table>

### TABLE 2: 2-D NOMINAL DEVICE PERFORMANCE PARAMETERS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{t,\text{sat}}$</td>
<td>mV</td>
<td>210</td>
</tr>
<tr>
<td>SS</td>
<td>mV/dec</td>
<td>69</td>
</tr>
<tr>
<td>DIBL</td>
<td>mV/V</td>
<td>30</td>
</tr>
<tr>
<td>$g_{m,\text{sat}}$</td>
<td>mA/V</td>
<td>6.75</td>
</tr>
<tr>
<td>$I_{d,\text{sat}}$</td>
<td>mA/µm</td>
<td>2.48</td>
</tr>
<tr>
<td>$I_{\text{off}}$</td>
<td>pA/µm</td>
<td>94.4</td>
</tr>
</tbody>
</table>
Model Description

Auto-correlation function

For a resist-defined gate electrode,

\[ \sigma^2_\delta = 2\sigma^2_{LER} \left( 1 - \rho_A(t_{\text{fin}}) \right) \]

\[ \rho(y) = e^{-\left(\frac{|y|}{\xi}\right)^{2\alpha}} \quad \text{Auto-correlation function} \]

For a spacer-defined gate electrode,

\[ \sigma^2_{\Delta L} = 4\sigma^2_{LER} \left[ 1 - \rho_A(t_{\text{fin}}) \right]. \]

\[ \sigma^2_{\Delta L} = 0 \]
Spacer v. Resist

$\sigma_{V_{t,\text{sat}}} [\text{mV}]$

$\sigma_{\text{LWR}} [\text{nm}]$

$\sigma_{V_{t,\text{sat}}} [\text{mV}]$

$\xi [\text{nm}]$

Correlated Edges
$\rho_X = 1$

Un-correlated Edges
$\rho_X = 0$

(a) Spacer Defined

(b) Resist Defined

$\sigma_{\text{LWR}}^2 = 2\sigma_{\text{LER}}^2 \left(1 - \rho_X \right)$

$\sigma_L = \sigma_R \equiv \sigma_{\text{LER}}$
Outline

• Introduction

• Robust method for estimation of LWR descriptors

• Comprehensive study of LWR for next-generation lithography processes

• Method to estimate FinFET device performance through the use of LWR descriptors

• Decomposition of Semiconductor Variability
Decomposition of Variability is NOT Unique

\[ P = \text{large-scale variation} + \text{small-scale variation} \]

\[ P = P_0 + P_{\text{inter-die}} + P_{\text{intra-die}} + P_e \]

\[ P_{\text{inter-die}} = P_{\text{lot-to-lot}} + P_{\text{wafer-to-wafer}}(\text{lot}) + P_{\text{die-to-die}}(\text{wafer}) \]

• Levels of sophistication
  
  – Lumped global model for inter-die
    
    \[ \tilde{P}_{\text{interdie}} \sim N \left( 0, \sigma^2_{\text{interdie}} \right) \]
    
    (Stine et al TSM’97)
    
    \[ \tilde{P}_{\text{intradie}} \sim \tilde{P}_{\text{intradie}}(x, y) \]

  – Average wafer positional dependence
    
    (Chang et al DAC’09)
    
    \[ \tilde{P}_{\text{interdie}} \sim \tilde{P}_{\text{interdie}}(X_{\text{wafer}}, Y_{\text{wafer}}) + N \left( 0, \sigma^2_{R,\text{interdie}} \right) \]
    
    \[ \tilde{P}_{\text{intradie}} \sim \tilde{P}_{\text{intradie}}(x, y) \]
Decomposition of Variability is Application-Specific

- The target end-user is the circuit designer.
- Our objective here is to create on-demand views of:
  - Global variation
  - Local variation
  - Spatial correlation
Proposed Hierarchical Model

\[ P_{ijkl} = \eta + \omega_{ijk} + \delta_l + \epsilon_{ijkl} \]

<table>
<thead>
<tr>
<th>Index</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>Lot</td>
</tr>
<tr>
<td>j</td>
<td>Wafer within a lot</td>
</tr>
<tr>
<td>k</td>
<td>Die within a wafer</td>
</tr>
<tr>
<td>l</td>
<td>Location within a die</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Term Estimator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\eta} = \bar{P} \ldots )</td>
<td>Global Average</td>
</tr>
<tr>
<td>( \hat{\omega}<em>{ijk} = \bar{P}</em>{ijk} - \bar{P} \ldots )</td>
<td>Global (inter-die) component</td>
</tr>
<tr>
<td>( \hat{\delta}<em>l = \bar{P}</em>{\ldots l} - \bar{P} \ldots )</td>
<td>Local (intra-die) systematic component</td>
</tr>
<tr>
<td>( \hat{\epsilon}<em>{ijkl} = P</em>{ijkl} - \bar{P}<em>{ijk} - \bar{P}</em>{\ldots l} + \bar{P} \ldots )</td>
<td>Residual; Local (intra-die) random component</td>
</tr>
</tbody>
</table>
\[ P_{ijkl} = \eta + \omega_{ijk} + \delta_l + \epsilon_{ijkl} \]

\[ \hat{\omega}_{ijk} = \bar{P}_{ijk} - \bar{P} \ldots \]

\[ \hat{\mu} = \frac{1}{M} \sum_{j=1}^{M} \omega_j \]

\[ \hat{\Sigma} = \frac{1}{M} \sum_{j=1}^{M} (\omega_j - \hat{\mu})(\omega_j - \hat{\mu})' \]
\[ P_{ijkl} = \eta + \omega_{ijk} + \delta_l + \epsilon_{ijkl} \]

\[ \hat{\delta}_l = \bar{P}_{..l} - \bar{P}_{..} \]

If within-die spatial location is available, look-up value \((x, y)\).

Otherwise, \((x, y)\) are randomly sampled on \(\delta_l\).

\[ \hat{z}(x, y) = X'c \]

\[ X = (1, x, y, x^2, y^2, \ldots, x^k, y^k)' \quad \text{and} \quad c = (c_0, c_1, c_2, \ldots, c_{k-1}, c_k)' \]

\[ AIC_c = \log \left( \frac{1}{n} \sum_{(x,y)} (z(x,y) - \hat{z}(x,y))^2 \right) + \frac{n + k}{n - k - 2} \]
Why is Spatial Correlation Important?

- Correlation of device parameters depends on spatial locations
  - Closeness of devices $\rightarrow$ higher probability of similarity
- Spatial variation is very important: 40~65% of total variation [Nassif, ISQED’00]

Adapted from L. He (UCLA)
\[ P_{ijkl} = \eta + \omega_{ijk} + \delta_l + \epsilon_{ijkl} \]

\[ \hat{\epsilon}_{ijkl} = P_{ijkl} - \bar{P}_{ijk} - \bar{P}_{...l} + \bar{P}_{...} \]

\[ C(0) = \sigma^2 \]

\[ C(h) = \sigma^2 \rho(h) \]
Variogram Re-visited

\[ \gamma(h) = \left(1 - \frac{1}{e}\right) \sigma^2 \]

\[ \sigma^2 \]

\[ c_0 \]

\[ \xi \]

\[ \phi \]

\[ \text{Range} \]

\[ \text{Sill} \]

\[ \text{Nugget} \]

\[ \text{Lag, } h \]

\[ \text{Var}(X(s) - X(s + h)) = \text{Var}(X(0) - X(h)) = 2\gamma(h) \]
Variogram Estimation Issue #1

• Covariogram is the end goal and not variogram, but

\[ 2 \hat{\gamma}(h) \neq 2 \left[ \hat{C}(0) - \hat{C}(h) \right] \]

• Solution:

\[ 2 \gamma(h; \theta) = 2c_0 + 2\sigma^2 \left[ 1 - \exp \left( - \left[ \frac{\|Ah\|_2}{\xi} \right]^{2\phi} \right) \right], \quad h \in \mathbb{R}^d \]

• Other forms also available.
Variogram Estimation Issue #2

• Matheron’s classical estimator is vulnerable to outliers

\[ 2\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{N(h)} \left( X(s_i) - X(s_j) \right)^2, \quad h \in \mathbb{R}^d, \]

• Solution: Robust Estimator (Cressie, 1980)

\[ 2\hat{\gamma}(h) = \frac{1}{0.457 + 0.494/N(h)} \left[ \frac{1}{N(h)} \sum_{N(h)} \left| X(s_i) - X(s_j) \right|^{1/2} \right]^4, \quad h \in \mathbb{R}^d, \]
Variogram Estimation Issue #3

- Insufficient density of within die observations

\[ P_{ijkl} = \eta + \omega_{ijk} + \delta_l + \epsilon_{ijkl} \]

\[ \hat{\epsilon}_{ijkl} = P_{ijkl} - \bar{P}_{ijk} - \bar{P}_{...l} + \bar{P} \]
Variogram Estimation Issue #3

• Solution:
  – Estimate **structure** of auto-correlation function at wafer-level
    \[ P_{ijkl} = \eta + \delta_l + \varepsilon_{ijkl} \]
    \[ \hat{\varepsilon}_{ijkl} = P_{ijkl} - \bar{P}_{...l} \]
  – Estimate the sill using die-level residuals
    \[ P_{ijkl} = \eta + \omega_{ijk} + \delta_l + \varepsilon_{ijkl} \]
    \[ \hat{\varepsilon}_{ijkl} = P_{ijkl} - \bar{P}_{ijkl} - \bar{P}_{...l} + \bar{P}.... \]
Weighted Least Squares Fit

\[
\hat{\theta} = \arg\min_{\theta} \sum_{h=1}^{h_0} N(h) \left[ \frac{2\hat{\gamma}(h)}{2\gamma(h; \theta)} - 1 \right]^2
\]
Wafer Selection

• Model: Garbage in, garbage out
• Wafers must be selected carefully
• Reject outliers → MVN assumption

• What happens when there are multiple across-wafer spatial patterns?
  – Cluster analysis using Least Angle Regression
Least Angle Regression
Cluster Analysis: Results

Cluster 1 = 117
Cluster 2 = 13
Cluster 3 = 31
Cluster 4 = 1
Cluster 5 = 8
Cluster 6 = 1
Cluster 7 = 1
Cluster 8 = 27
Cluster 9 = 96

Distance

C-Index

CH-Index

Heuristic Optimum
Conclusion

• A robust procedure was developed for estimating LWR parameters which greatly improved the estimation of $\sigma$.

• Many NGL processes were evaluated in terms of their LWR characteristics. Results indicate benefits of DSA and SADP processes.

• Possibility of incorporating LWR parameters in compact model was demonstrated using FinFET.

• Hierarchical decomposition of variability was performed to provide circuit designer on-demand views of global and local variability including spatial correlation.
Future Work

• Study the impact of gate oxide roughness in advanced technologies using the 2-D roughness characterization framework developed here.

• Demonstrate an application of the decomposition of variability using ISCAS’85 benchmark circuit and compare results with other approaches.