Width-dependent Statistical Leakage Modeling for Random Dopant Induced Threshold Voltage Shift

Jie Gu, Sachin Sapatnekar, Chris Kim

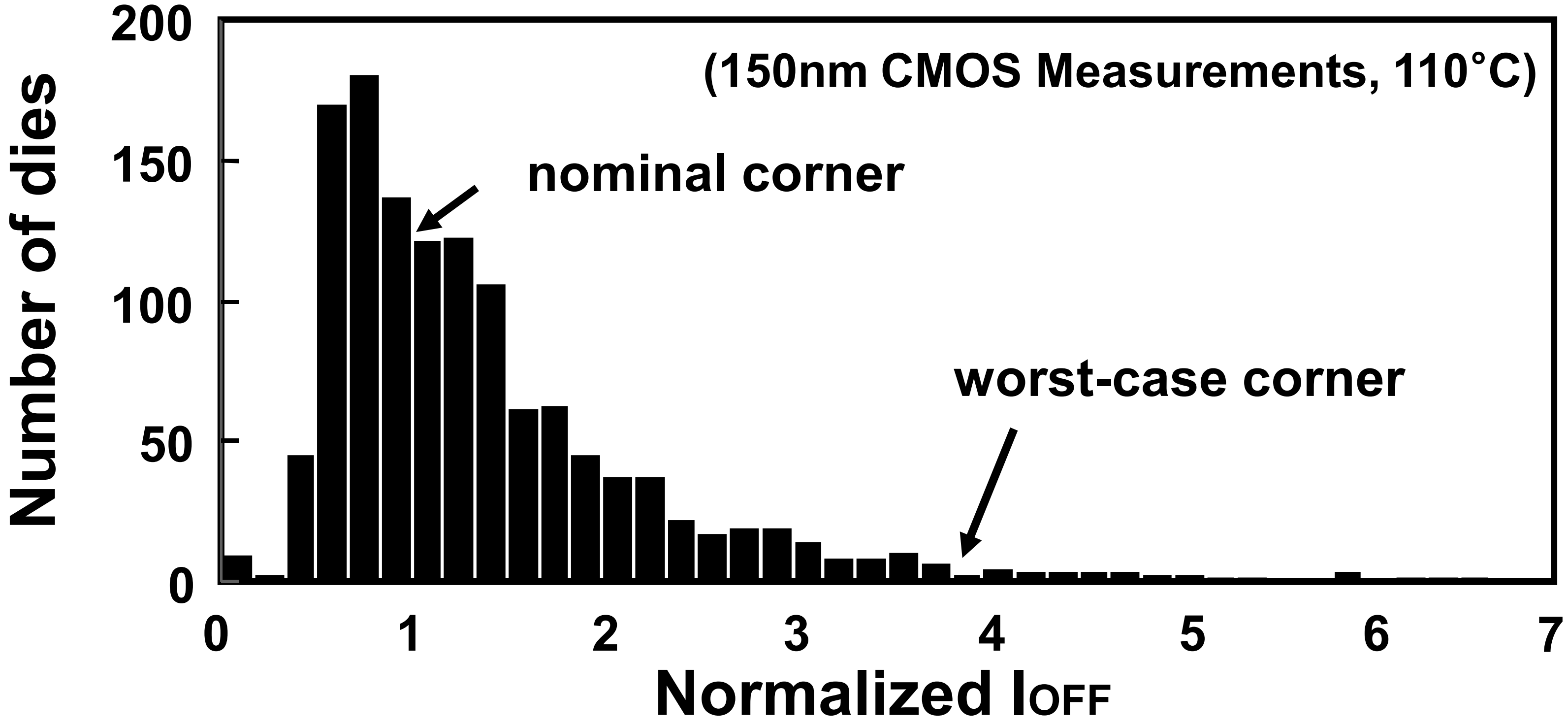
Department of Electrical and Computer Engineering
University of Minnesota

chriskim@umn.edu
www.umn.edu/~chriskim

Outline

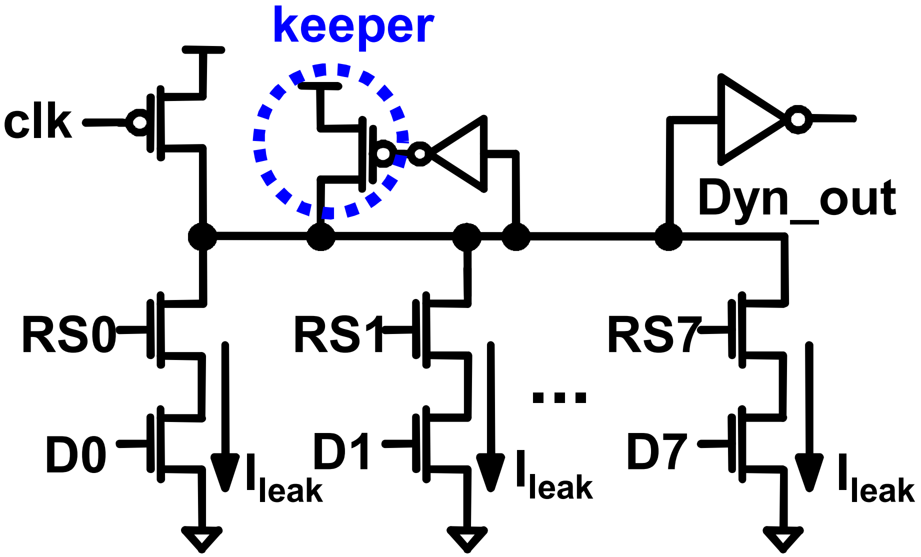
- Introduction
- Conventional statistical leakage modeling
- Proposed statistical leakage modeling
  - "Microscopic" random dopant fluctuation
- Experimental results
- Application to leakage sensitive circuits
  - Dynamic circuits
  - SRAM memory bitlines
- Conclusions

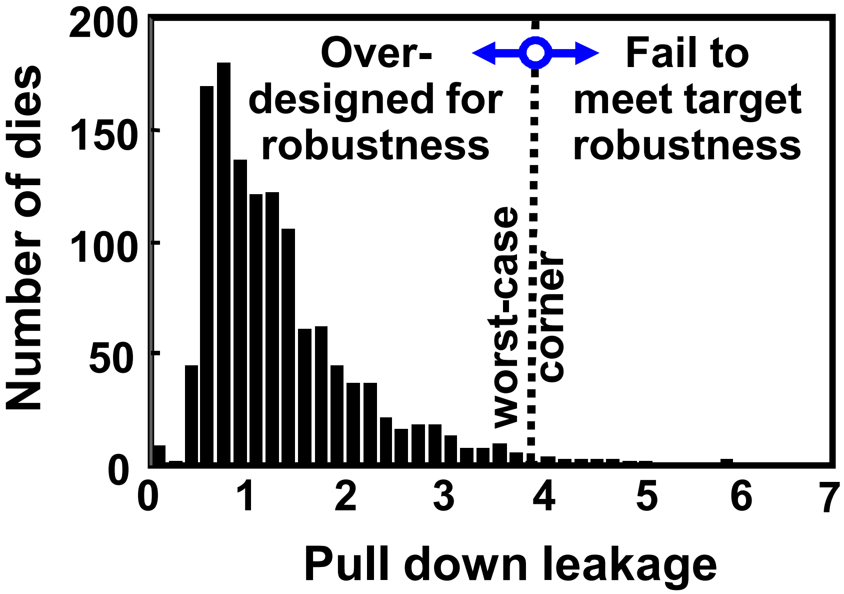
Motivation



- 4X variation between nominal and worst-case leakage
  - Channel length/width variation, line edge roughness, dopant fluctuation
- Performance determined at nominal leakage
- Power/robustness determined at worst-case leakage

Leakage Variation Impact: Dynamic Circuit Example



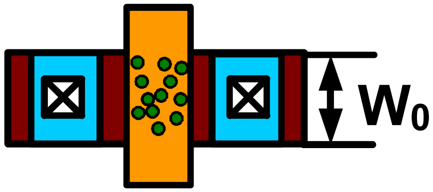


- Static keeper prevents the dynamic node droop
- Keeper has to be properly sized for sufficient noise margin
- Accurate leakage estimation is critical for meeting noise margin requirements

Conv. Statistical Leakage Modeling

Device parameters provided by fabs:

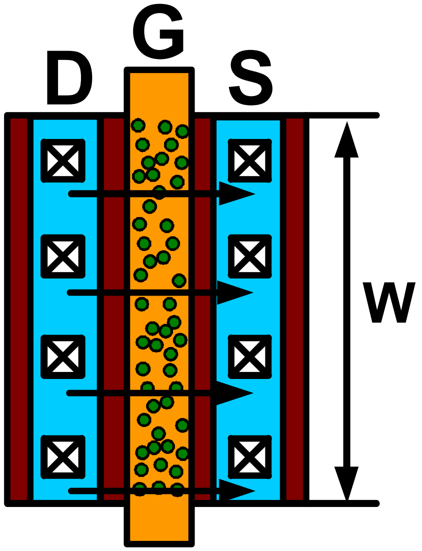
$$\mu_{V_{T0}}, \sigma_{V_{T0}}, W_0$$



Conventional approach (square-root method):

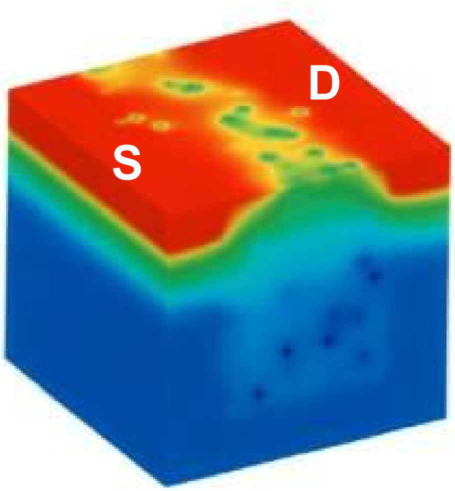
$$I_{leak} \propto W e^{-qV_T/mkT}$$

$$\begin{cases} \mu(V_T) = \mu_{V_{T0}} \\ \sigma(V_T) = \sigma_{V_{T0}} / \sqrt{\dfrac{W \cdot L}{W_0 \cdot L_0}} \end{cases}$$



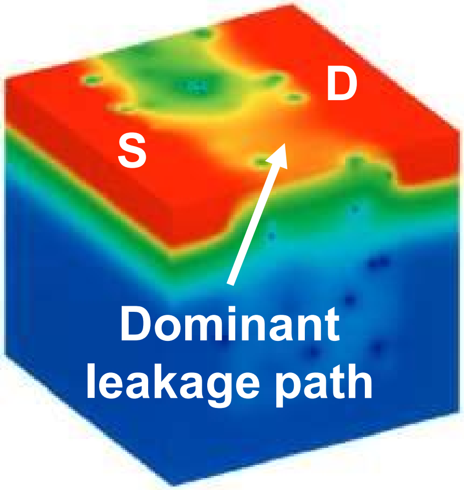
- **Statistically, larger devices have lesser variation**

μ-RDF Induced Threshold Voltage Shift



Evenly distributed dopants

V_T=0.78V, 130 dopants, L_{eff} = 30nm



Unevenly distributed dopants

V_T=0.56V, 130 dopants, L_{eff} = 30nm

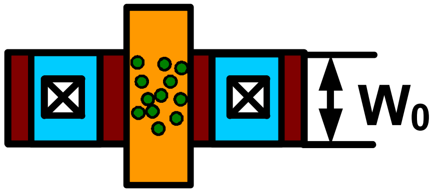
3D simulation results on surface potential A. Asenov, TED 1998

- **Threshold voltage depends on both**
  - **the # of dopants in the channel and**
  - **the "microscopic" random dopant placement**
- **30mV+ V_T shift has been reported by P. Wong (IEDM '93)**

Golden Statistical Leakage Modeling

Device parameters provided by fabs:

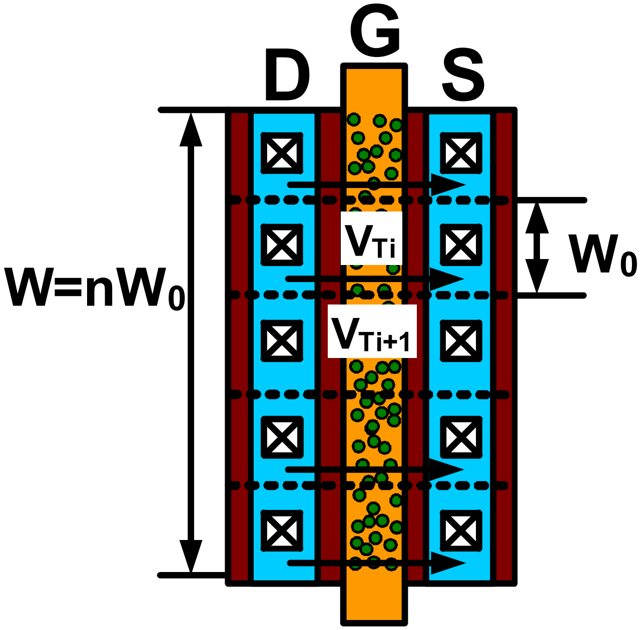
$$\mu_{V_{T0}},\ \sigma_{V_{T0}},\ W_0$$



Golden approach:

$$I_{leak} \propto \sum_{i=1}^{n} W_0 e^{-qV_{Ti}/mkT}$$

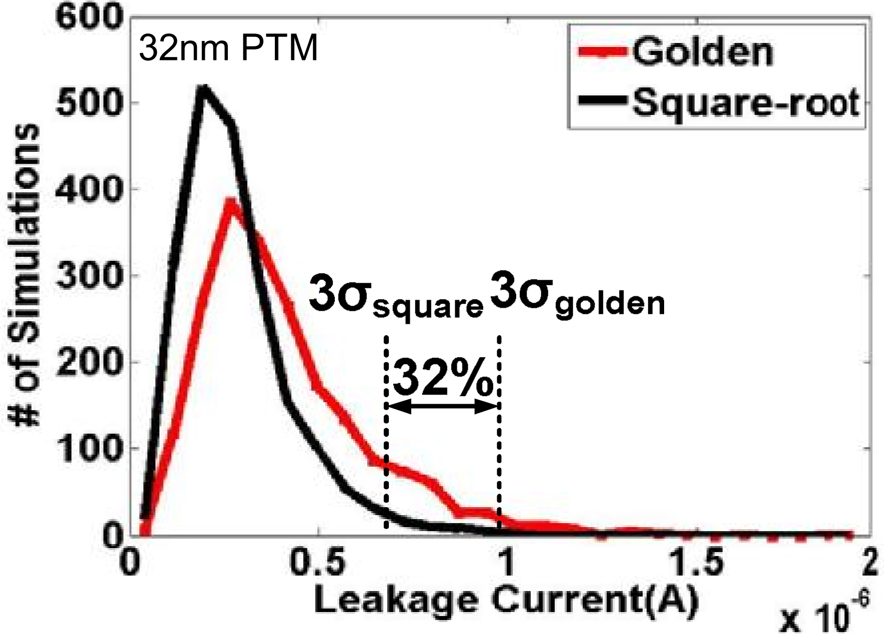
$$\begin{cases} \mu(V_{Ti}) = \mu_{V_{T0}} \\ \sigma(V_{Ti}) = \sigma_{V_{T0}} \end{cases}$$



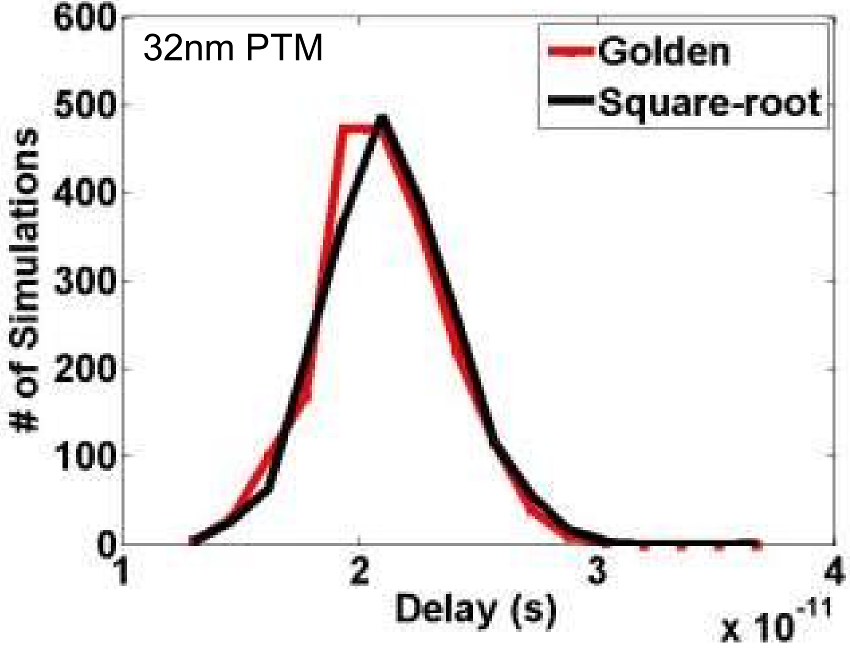
- Need to sum the leakage dist. of the sub-devices

Conventional versus Golden Method

Leakage distribution



Delay distribution

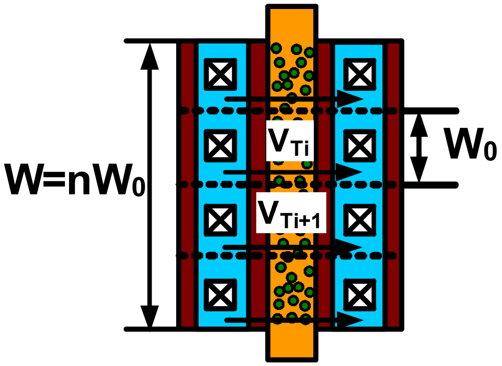
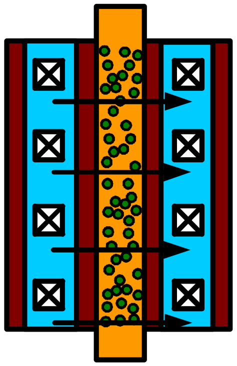
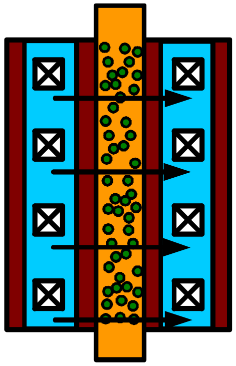
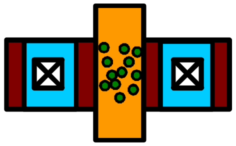


$$I_{leak} \propto \sum_{i=1}^{n} W_0 e^{-qV_{Ti}/mkT}$$

$$I_{on} \propto \sum_{i=1}^{n} W_0 (V_{dd} - V_{Ti})^{\alpha} \quad , \ \alpha \approx 1$$

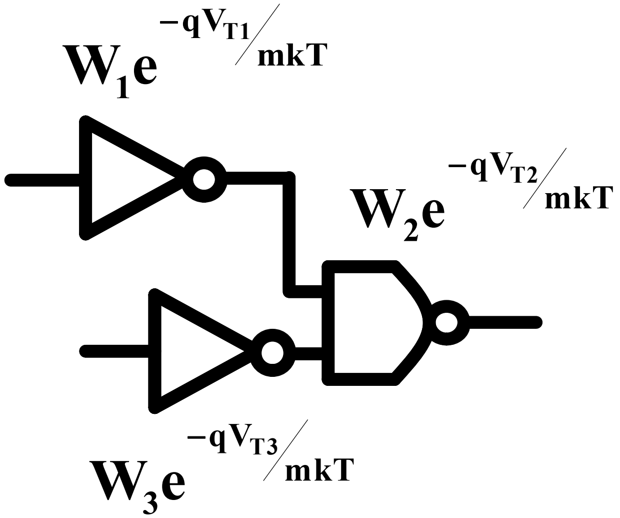
- **Leakage model shows 32% discrepancy btwn 3σ values**
  - **$\mu(V_T)$ reduces when adding lognormal distributions**
- **Delay model matches well with golden results**
  - **$\mu(V_T)$ does not change when adding normal distributions**

Statistical Leakage Estimation

Golden	Conventional	Proposed
 (V$_{Ti}$ varies due to RDF)		
Given reference device parameter $W_0, \mu_{V_{T0}}, \sigma_{V_{T0}}$		
$I_{leak} \propto \sum_{i=1}^{n} W_0 e^{-qV_{Ti}/mkT}$ $\mu(V_{Ti}) = \mu_{V_{T0}}$ $\sigma(V_{Ti}) = \sigma_{V_{T0}}$	$I_{leak} \propto We^{-qV_{T_{sq}}/mkT}$ $\mu(V_{T_{sq}}) = \mu_{V_{T0}}$ inaccurate $\sigma(V_{T_{sq}}) = \sigma_{V_{T0}} / \sqrt{\frac{W \cdot L}{W_0 \cdot L_0}}$	$I_{leak} \propto We^{-qV_{Teff}/mkT}$ $\mu(V_{Teff}) = f_\mu(W, \mu_{V_{T0}}, \sigma_{V_{T0}})$ $\sigma(V_{Teff}) = f_\sigma(W, \mu_{V_{T0}}, \sigma_{V_{T0}})$

- Effective V_T concept introduced to model width-dependency

Previous Work and Our Contribution



Previous work

$\mu(V_{T1}) = \mu(V_{T2}) = \mu(V_{T3})$

σ follows square – law

Proposed

$\mu(V_{Ti}) = f_{\mu}(W_i, \mu(V_{T0}), \sigma(V_{T0}))$

$\sigma(V_{Ti}) = f_{\sigma}(W_i, \mu(V_{T0}), \sigma(V_{T0}))$

- To the best of our knowledge, this is the first work to model the V_T dependency on device width
  - Previous work proposed by Ananthan (DAC06), Chang (DAC05), Narendra (ISLPED02) did not consider this
- Simple closed-form expression derived that can handle continuous width case

Calculation of Effective V_T

Given a reference device with W_x, μ_0, σ_0

- $W_y = nW_x$
  - Discrete width multiplication

$$I_{leak} = \sum_{i=1}^{n} W_x e^{\frac{-V_{Tx_i}}{mkT/q}} \approx W_y e^{\frac{-V_{T_y}}{mkT/q}}$$

  - Directly apply Wilkinson's method

- $W_y = \alpha W_x$ ($\alpha = n/m$)
  - Continuous width multiplication (α is any rational number)
  - Extend Wilkinson's method to handle continuous integration

Wilkinson’s Method

$$\sum_{i=1}^{n} We^{x_i} \cong nWe^{y}$$

- Sum of lognormals can be approximated as a single lognormal with a calculable mean and standard deviation
- y is the new Gaussian variable calculated by moments matching

First moment:

$$u_1 = E(S) = \sum_{i=1}^{n} e^{m_{x_i} + \sigma_{x_i}^2/2} = ne^{m_y + \sigma_y^2/2}$$

Second moment:

$$u_2 = E(S^2) = \sum_{i=1}^{n} e^{2m_{x_i} + 2\sigma_{x_i}^2} + 2\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} e^{m_{x_i} + m_{x_j}} e^{(\sigma_{x_i}^2 + \sigma_{x_j}^2 + 2r_{ij}\sigma_{x_i}\sigma_{x_j})/2} = n^2 e^{2m_y + 2\sigma_y^2}$$

Moment matching results:

$$\boxed{\begin{aligned} m_y &= 2\ln u_1 - 1/2\ln u_2 \\ \sigma_y^2 &= \ln u_2 - 2\ln u_1 \end{aligned}}$$

A. Abu-Dayya, IEEE Vehicular Technology Conference, 1994

Discrete Width Multiplication ($W_y=nW_x$)

Effective V_T Sub-device V_T

$$\mu_{V_{Ty}} = \mu_{V_{Tx}} - \frac{1}{2}\Delta / B$$

$$\sigma^2_{V_{Ty}} = \sigma^2_{V_{Tx}} - \Delta / B^2$$

where $\Delta = B^2\sigma^2_{V_{Tx}} - \ln\left(\frac{e^{B^2\sigma^2_{V_{Tx}}} + (n-1)\cdot e^{r_x B^2\sigma^2_{V_{Tx}}}}{n}\right) \geq 0$

$$B = {mkT}\big/{q} \quad , \quad r_x = \text{correlation coeff.}$$

- Both the mean and sigma of V_T decreases with larger device width
- The mean and sigma of V_T also decreases with smaller r_x

Spatial Correlation Coefficient

$$r\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx1_i}}, \sum_{i=1}^{n} W_x e^{-BV_{Tx2_i}}\right) = r\left(W_y e^{-BV_{Ty1}}, W_y e^{-BV_{Ty2}}\right)$$

$$r\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx1_i}}, \sum_{i=1}^{n} W_x e^{-BV_{Tx2_i}}\right)$$

$$= \frac{E\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx1_i}} \cdot \sum_{i=1}^{n} W_x e^{-BV_{Tx2_i}}\right) - E\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx1_i}}\right) \cdot E\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx2_i}}\right)}{\sigma\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx1_i}}\right) \cdot \sigma\left(\sum_{i=1}^{n} W_x e^{-BV_{Tx2_i}}\right)}$$

$$= \frac{e^{r_x B^2 \sigma_{V_{Tx}}^2} - 1}{\frac{e^{B^2 \sigma_{V_{Tx}}^2} + (n-1) e^{r_x \sigma_{V_{Tx}}^2}}{n} - 1}$$

$$r\left(W_y e^{-BV_{T_{y1}}}, W_y e^{-BV_{T_{y2}}}\right)$$

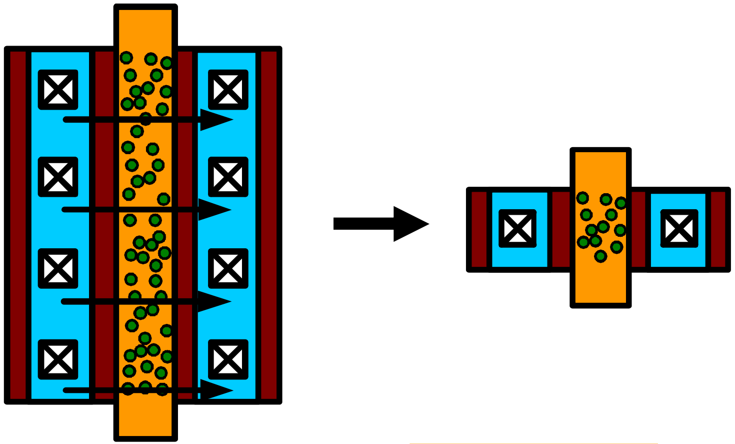
$$= \frac{E\left(W_y e^{-BV_{T_{y1}}} \cdot W_y e^{-BV_{T_{y2}}}\right) - E\left(W_y e^{-BV_{T_{y1}}}\right) \cdot E\left(W_y e^{-BV_{T_{y2}}}\right)}{\sigma\left(W_y e^{-BV_{T_{y1}}}\right) \cdot \sigma\left(W_y e^{-BV_{T_{y2}}}\right)}$$

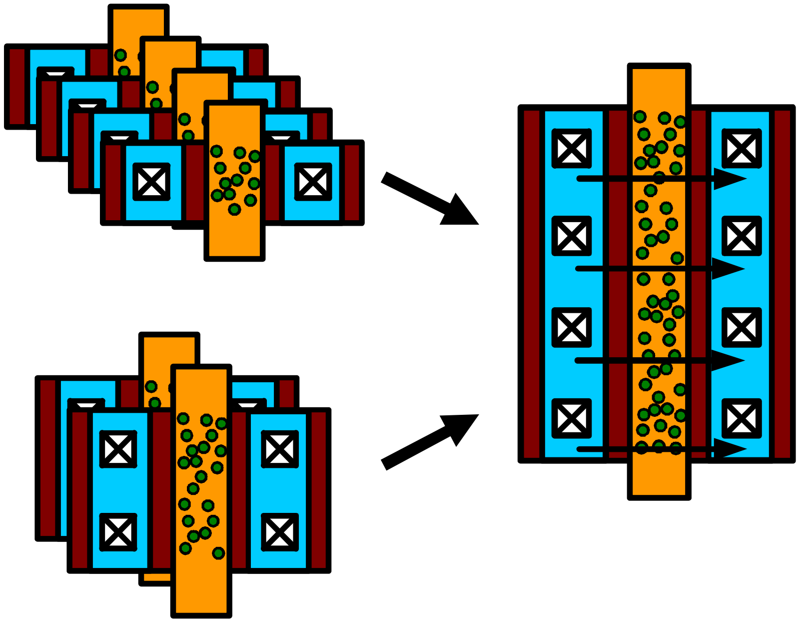
$$= \frac{e^{r_y B^2 \sigma_{V_{Ty}}^2} - 1}{e^{B^2 \sigma_{V_{Ty}}^2} - 1}$$

$$\boxed{r_y = \frac{B^2 \sigma_{V_{Tx}}^2}{\ln \frac{e^{B^2 \sigma_{V_{Tx}}^2} + (n-1) e^{r_x B^2 \sigma_x^2}}{n}} \cdot r_x = \frac{\sigma_{V_{Tx}}^2}{\sigma_{V_{Ty}}^2} \cdot r_x}$$

- r_y goes up as width increases ($r_x \le r_y \le 1$ and $r_y = 1$ iff $r_x = 1$)

Reference Device Size Independence





$$\mu_{V_{Tx}} = \mu_{V_{Ty}} - \frac{1}{2}\Delta'/B$$

$$\sigma^2_{V_{Tx}} = \sigma^2_{V_{Ty}} - \Delta'/B^2$$

$$\Delta' = B^2\sigma^2_{V_{Ty}} - \ln\left(\frac{e^{B^2\sigma^2_{V_{Ty}}} + (1/n-1)\cdot e^{r_y B^2\sigma^2_{V_{Ty}}}}{1/n}\right)$$

$$\boxed{r_y = \frac{\sigma^2_{V_{Tx}}}{\sigma^2_{V_{Ty}}}\cdot r_x}$$

- Given μ_y, σ_y of a large device with W_y, we can reverse the calculation to find out μ_x, σ_x of a smaller device with W_x
- Same results can be obtained independent of the reference device size

Continuous Width Multiplication ($W_y=\alpha W_x$)

Assume there exists a small virtual device that satisfies $W_x=mW_0$, $W_y=nW_0$. Applying Wilkinson's for both we get,

Effective V_T **Sub-device V_T**

$$\mu_{V_{Ty}} = \mu_{V_{Tx}} - \frac{1}{2}\Delta / B$$

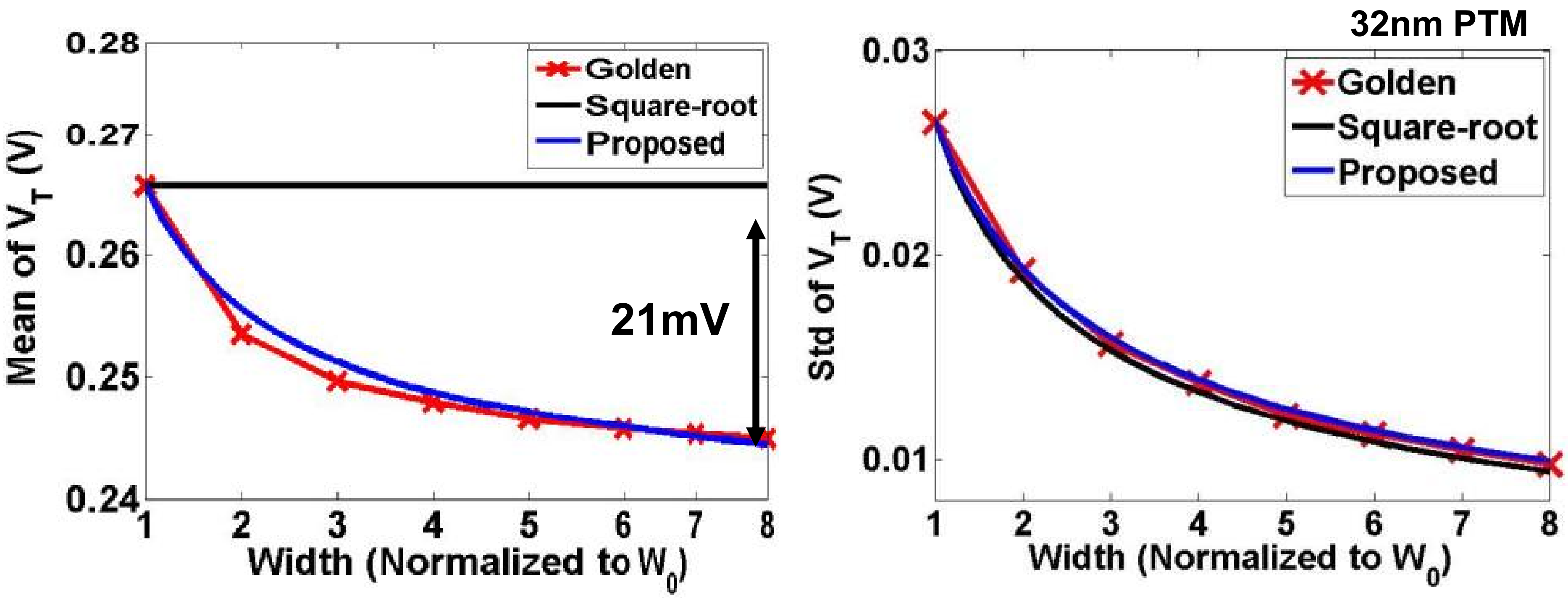
$$\sigma^2_{V_{Ty}} = \sigma^2_{V_{Tx}} - \Delta / B^2$$

where $\Delta = B^2\sigma^2_{V_{Tx}} - \ln\left(\frac{e^{B^2\sigma^2_{V_{Tx}}} + (\alpha - 1)\cdot e^{r_x B^2\sigma^2_{V_{Tx}}}}{\alpha}\right) \geq 0$

$$\boxed{r_y = \frac{\sigma^2_{V_{Tx}}}{\sigma^2_{V_{Ty}}}\cdot r_x}$$

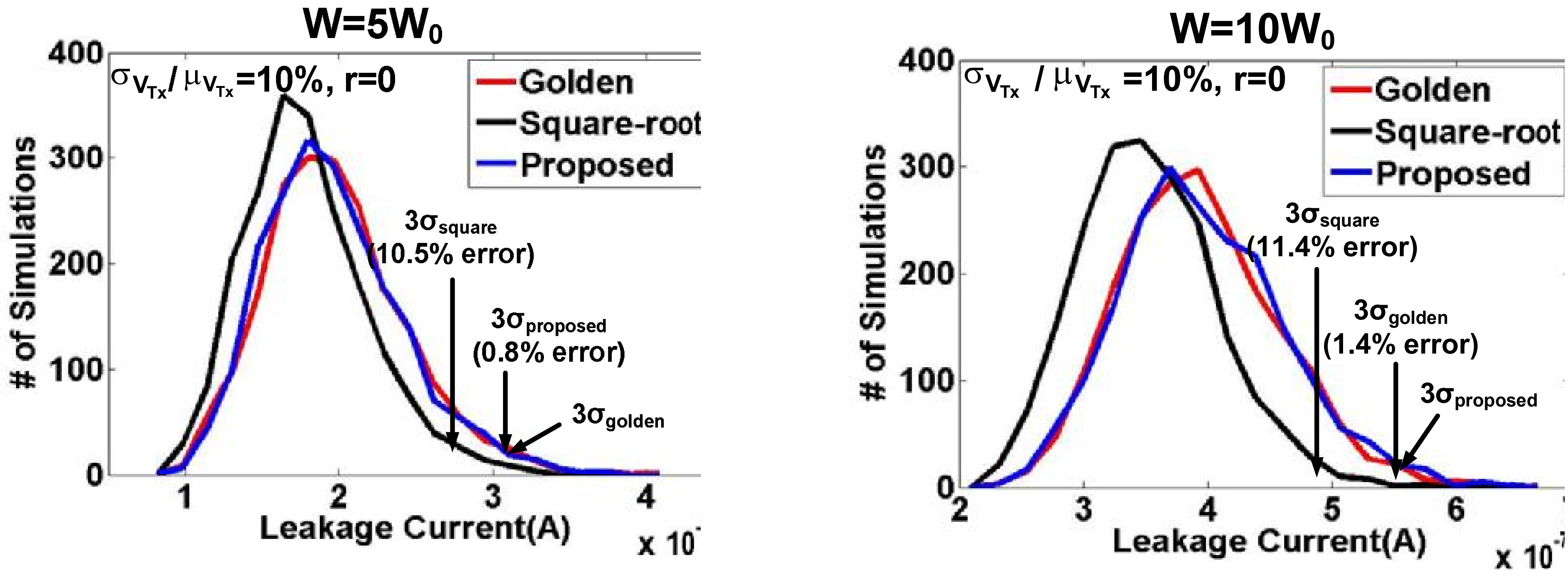
- **Expression identical to the discrete width multiplication case**

Width Dependent V_T Statistics



- **21mV difference in V_T mean between conventional and golden method**
- **No significant difference in V_T sigma between golden, square-root, and proposed method**

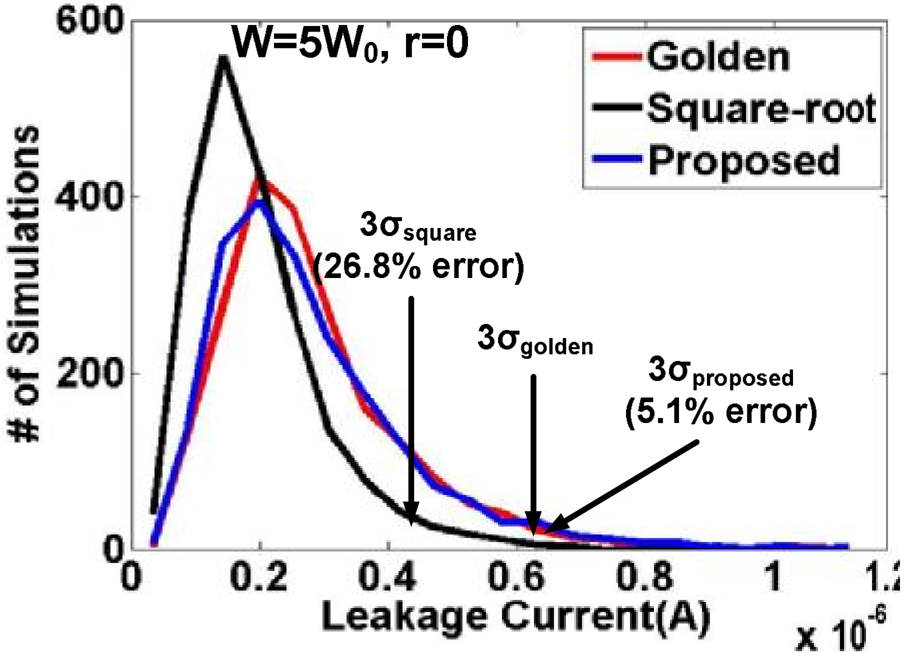
Leakage Distribution Comparison: Different Widths



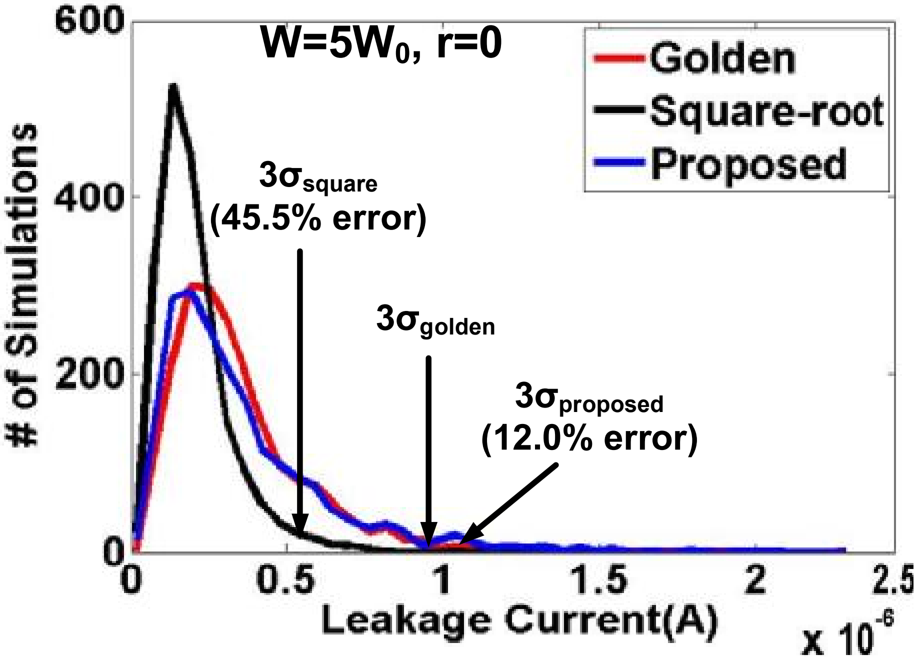
- Leakage estimation error (3σ point) reduced from 10.5% to 0.8% using the proposed model

Leakage Distribution Comparison: Different σ/μ's

$\sigma_{V_{Tx}} / \mu_{V_{Tx}} = 15\%$

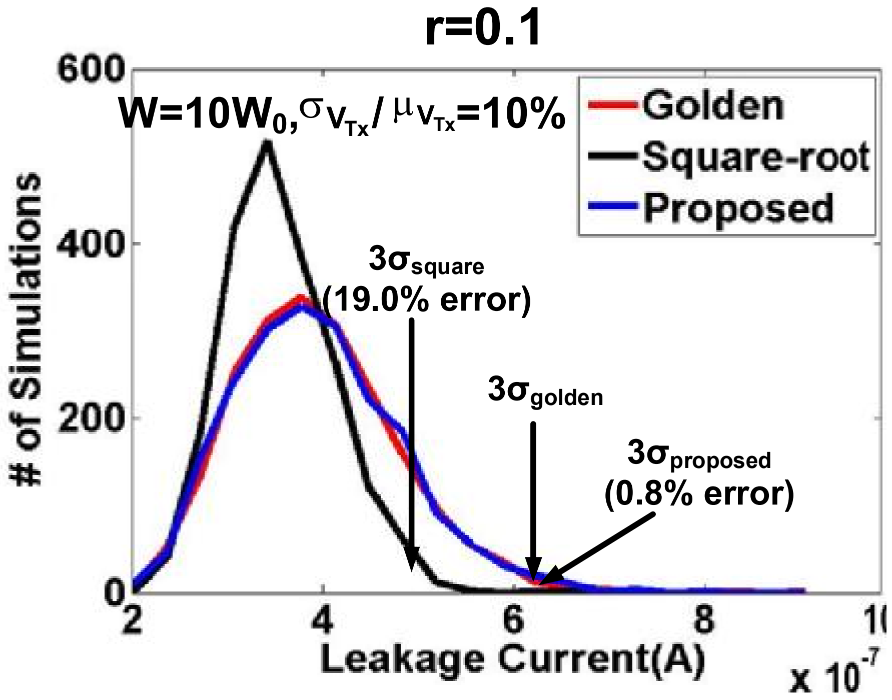


$\sigma_{V_{Tx}} / \mu_{V_{Tx}} = 25\%$

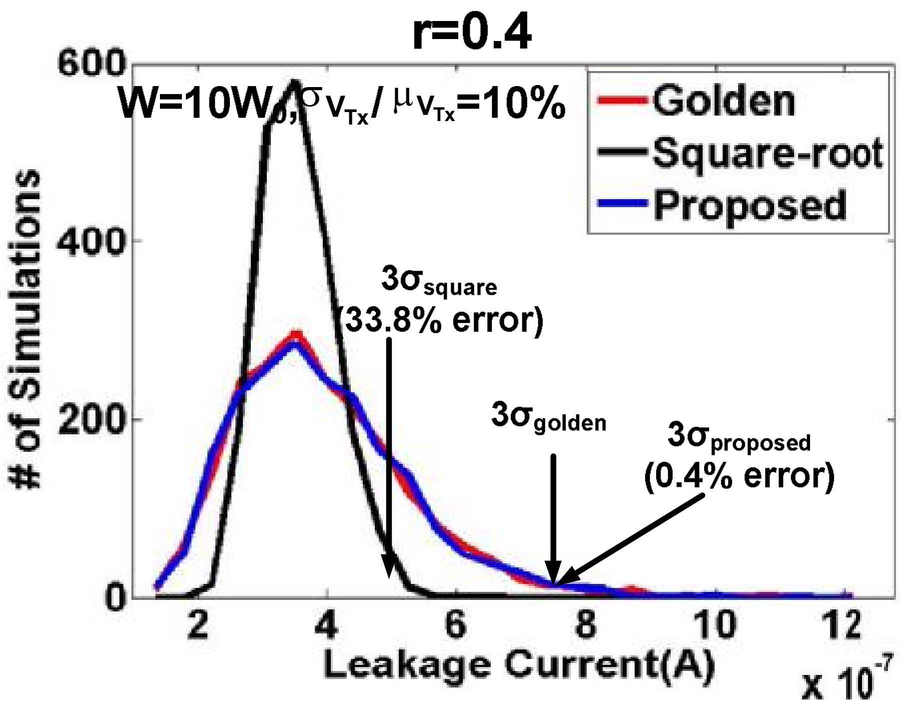


- Error of conventional approach increases to 45.5% for larger σ/μ due to larger variation between the sub-devices
- Proposed approach exhibits a smaller leakage estimation error (<12.0%) limited by the accuracy of the Wilkinson's formula

Leakage Distribution Comparison: Different Correlation Coefficients



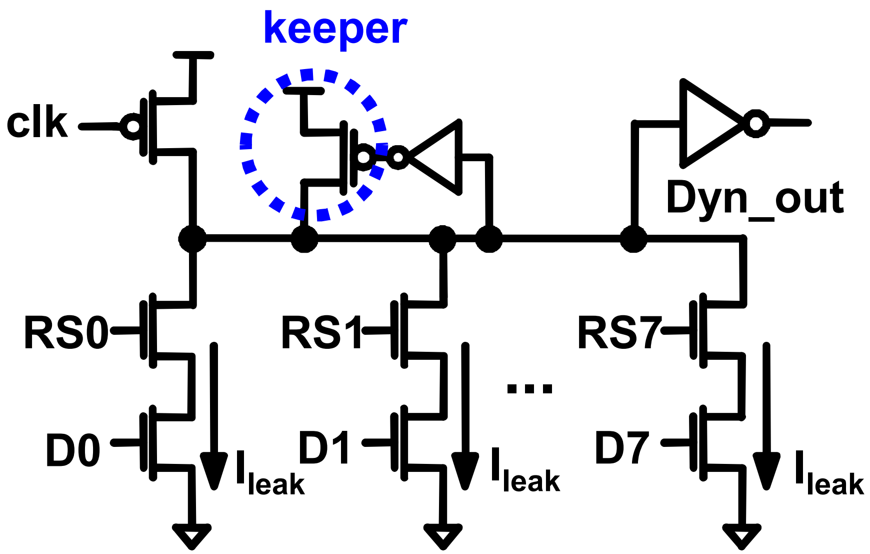
r=0.1

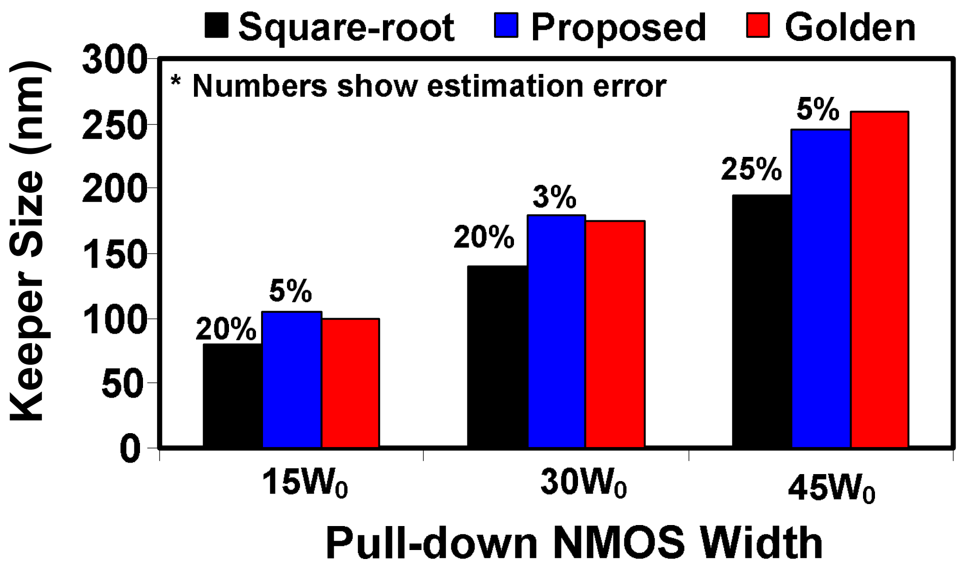


r=0.4

- **Conventional model does not consider spatial correlation (assumes that sub-devices are uncorrelated)**
- **Proposed work's estimation error is small for a wide-range of V_T correlation coefficients**

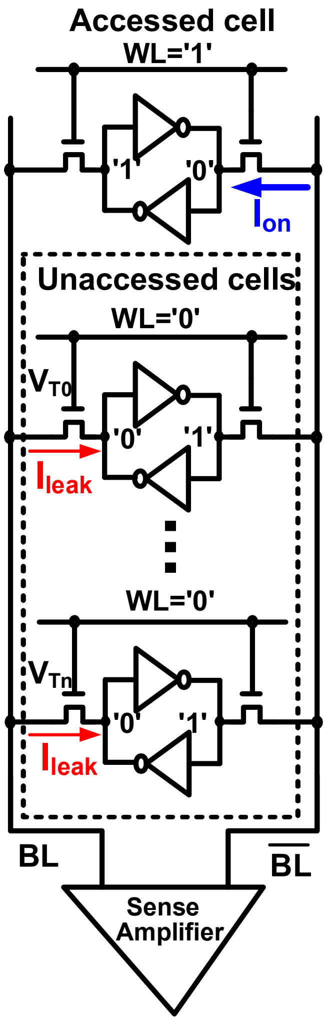
Design Example: Dynamic Circuit Keeper Sizing

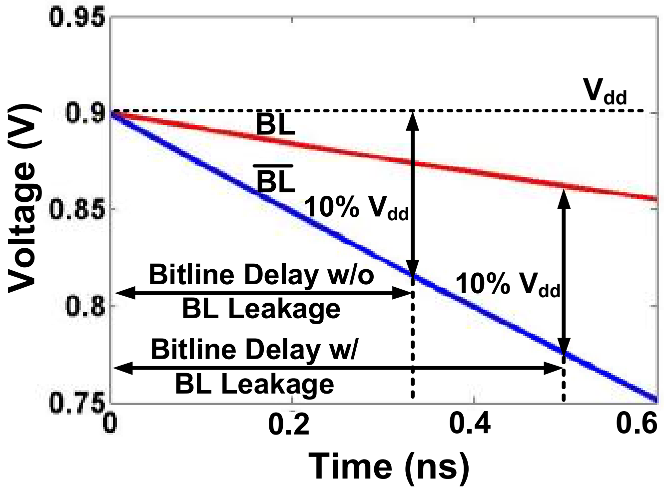


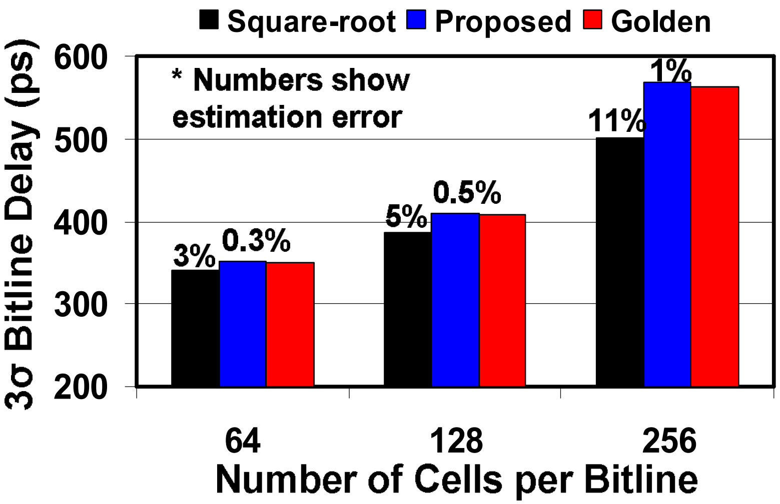


- Conventional approach underestimates the pull-down leakage misguiding the designer to use a smaller keeper
- This work shows that a 30% keeper size up is required to meet the target noise margin

Design Example: SRAM Bitline







- **Actual bitline delay is 3-12% longer than expected when using the conventional model due to underestimated bitline leakage**

Conclusions

- **“Microscopic” RDF leads to width-dependent V_T**
- **Conventional statistical V_T model is inaccurate**
  - Only capable of modeling on current (linear function of V_T)
  - Fails to model leakage current (exponential function of V_T)
  - Exhibits as much as 45% error in 3σ leakage value
- **Proposed width-dependent V_T model**
  - Simple closed-form expression with than 5% estimation error
  - Can be expanded to general sources of within-device variation
  - Handles both uncorrelated and correlated process variables
  - Useful for leakage-sensitive circuit designs such as dynamic circuits, SRAM bitlines, and subthreshold logic