ISPD 2006

Robust Extraction of Spatial Correlation

Jinjun Xiong, Vladimir Zolotov*, Lei He

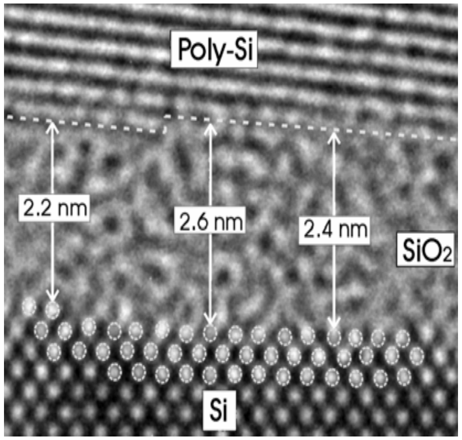
EE, University of California, Los Angeles
IBM T.J. Watson Research Center, Yorktown Heights*

Sponsors: NSF, UC MICRO, Actel

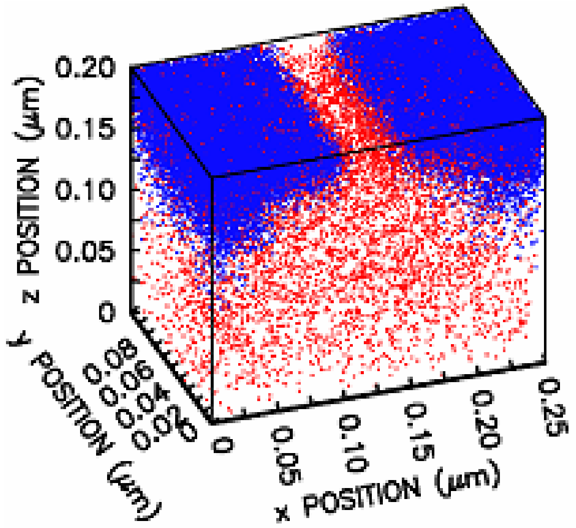
Acknowledgements to Dr. Chandu Visweswariah

Process Variations in Nanometer Manufacturing

- **Random fluctuations in process conditions ➔ changes physical properties of parameters on a chip**
  - What you design ≠ what you get
- **Huge impact on design optimization and signoff**
  - 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]



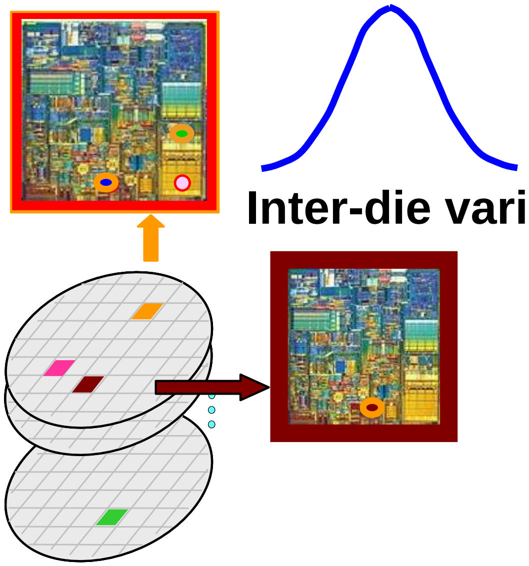
Oxide thickness



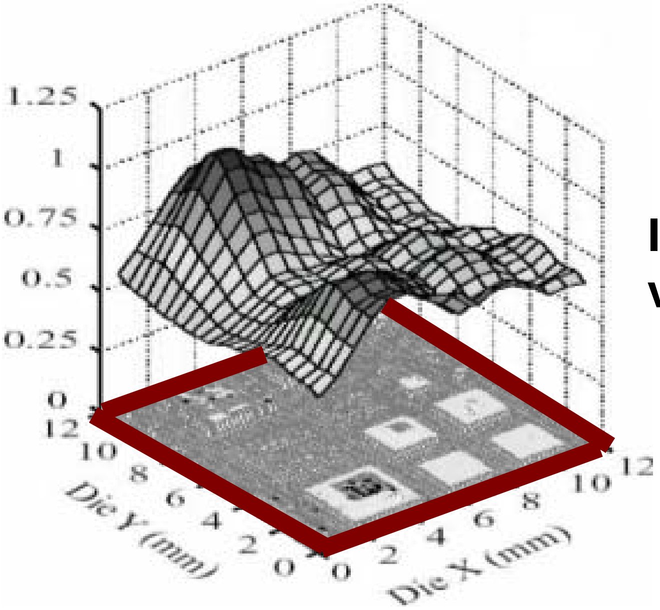
Random dopants

Process Variation Classification

- **Systematic vs random variation**
  - Systematic variation has a clear trend/pattern (deterministic variation [Nassif, ISQED00])
    - Possible to correct (e.g., OPC, dummy fill)
  - Random variation is a stochastic phenomenon without clear patterns
    - Statistical nature → statistical treatment of design
- **Inter-die vs intra-die variation**
  - Inter-die variation: same devices at different dies are manufactured differently
  - Intra-die (spatial) variation: same devices at different locations of the same die are manufactured differently



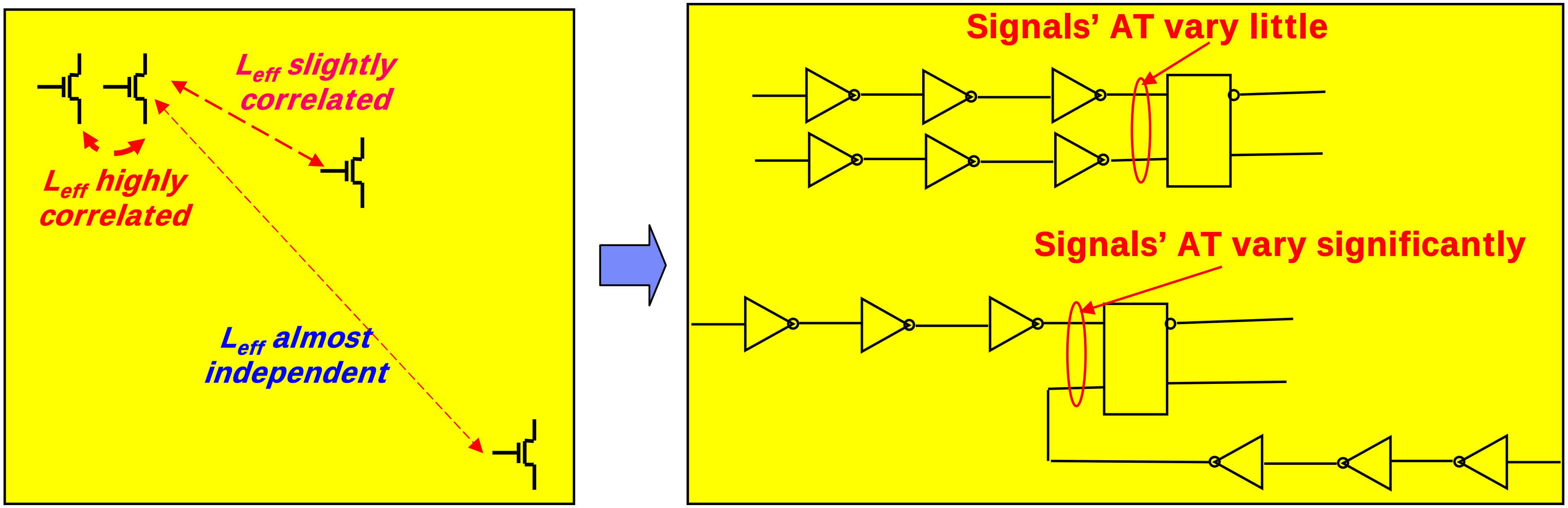
Inter-die variation



Intra-die variation

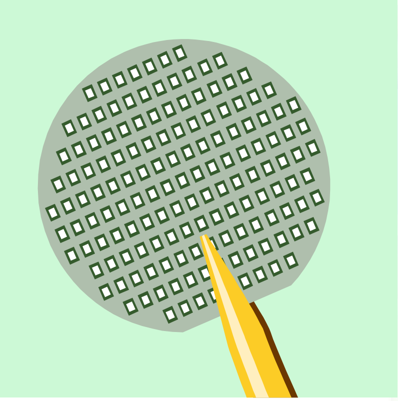
Spatial Variation Exhibits Spatial Correlation

- **Correlation of device parameters depends on spatial locations**
  - The closer devices → the higher probability they are similar
- **Impact of spatial correlation**
  - Considering vs not considering →30% difference in timing [Chang ICCAD03]
  - Spatial variation is very important: 40~65% of total variation [Nassif, ISQED00]

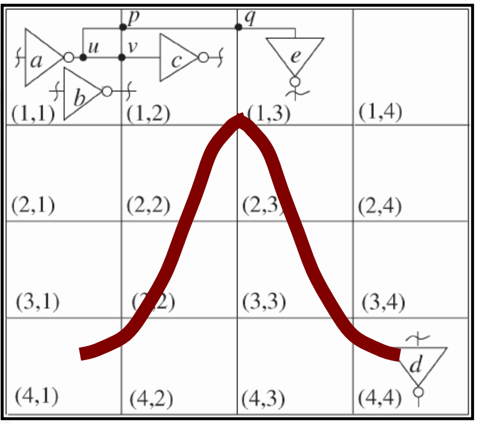


A Missing Link

- **Previous statistical analysis/optimization work modeled spatial correlation as a correlation matrix known a priori**
  - [Chang ICCAD 03, Su LPED 03, Rao DAC04, Choi DAC 04, Zhang DATE05, Mani DAC05, Guthaus ICCAD 05]
- **Process variation has to be characterized from silicon measurement**
  - Measurement has inevitable noises
  - Measured correlation matrix may not be valid (**positive semidefinite**)
- **Missing link: technique to extract a valid spatial correlation model**
  - Correlate with silicon measurement
  - Easy to use for both analysis and design optimization



Silicon Measurement



Statistical Design & Optimization

Agenda

- Motivations
- Process Variation Modeling
- Robust Extraction of Valid Spatial Correlation Function
- Robust Extraction of Valid Spatial Correlation Matrix
- Conclusion

Modeling of Process Variation

$$F = \underbrace{h_0 + h_1(Z_{D2D,sys}) + h_3(Z_{WID,sys})} + \underbrace{h_2(Z_{D2D,rnd}) + h_4(Z_{WID,rnd}) + X_r}$$

$$F = f_0 + F_r$$

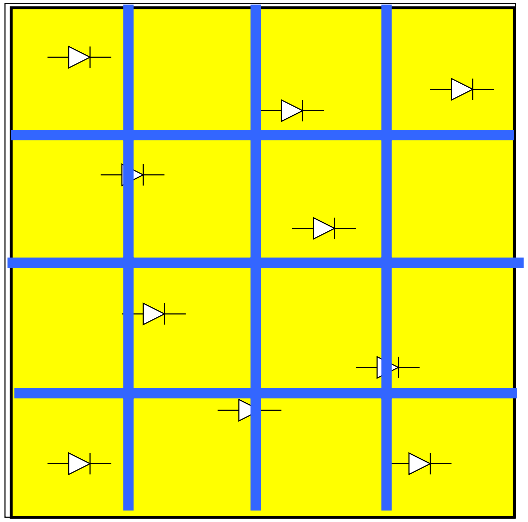
- **f_0 is the mean value with the systematic variation considered**
  - h_0: nominal value without process variation
  - $Z_{D2D,sys}$: die-to-die systematic variation (e.g., depend on locations at wafers)
  - $Z_{WID,sys}$: within-die systematic variation (e.g., depend on layout patterns at dies)
  - Extracted by averaging measurements across many chips
    - [Orshansky TCAD02, Cain SPIE03]
- **F_r models the random variation with zero mean**
  - $Z_{D2D,rnd}$: inter-chip random variation→ X_g
  - $Z_{WID,rnd}$: within-chip spatial variation→ X_s with spatial correlation ρ□
  - X_r: Residual uncorrelated random variation

$$F_r = X_g + X_s + X_r$$

- **How to extract F_r → focus of this work**
  - Simply averaging across dies will not work
  - Assume variation is Gaussian [Le DAC04]

Process Variation Characterization via Correlation Matrix

- **Characterized by variance of individual component + a positive semidefinite spatial correlation matrix for M points of interests**
  - In practice, superpose fixed grids on a chip and assume no spatial variation within a grid
- **Require a technique to extract a valid spatial correlation matrix**
  - Useful as most existing SSTA approaches assumed such a valid matrix
- **But correlation matrix based on grids may be still too complex**
  - Spatial resolution is limited→ points can't be too close (accuracy)
  - Measurement is expensive →can't afford measurement for all points



Overall variance $\sigma_F^2 = \sigma_G^2 + \sigma_S^2 + \sigma_R^2$

Global variance

Spatial variance

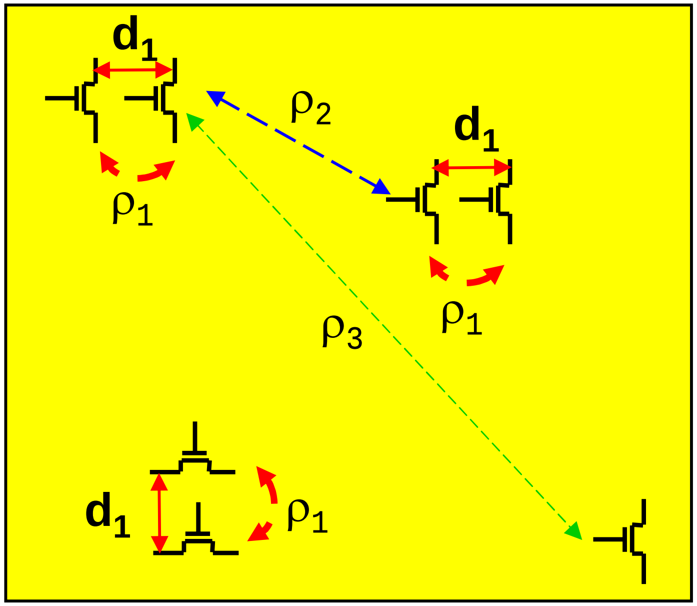
Random variance

Spatial correlation matrix

$$\Omega = \begin{pmatrix} 1 & \cdots & \rho_{1,M} \\ \vdots & \ddots & \vdots \\ \rho_{1,M} & \cdots & 1 \end{pmatrix}$$

Process Variation Characterization via Correlation Function

- **A more flexible model is through a correlation function**
  - If variation follows a homogeneous and isotropic random (HIR) field → spatial correlation described by a valid correlation function ρ(v)
    - Dependent on their distance only
    - Independent of directions and absolute locations
    - Correlation matrices generated from ρ(v) are always positive semidefinite
  - Suitable for a matured manufacturing process



Spatial covariance

$$\mathrm{cov}(F_i, F_j) = \sigma_G^2 + \rho(v)\sigma_S^2$$

Overall process correlation

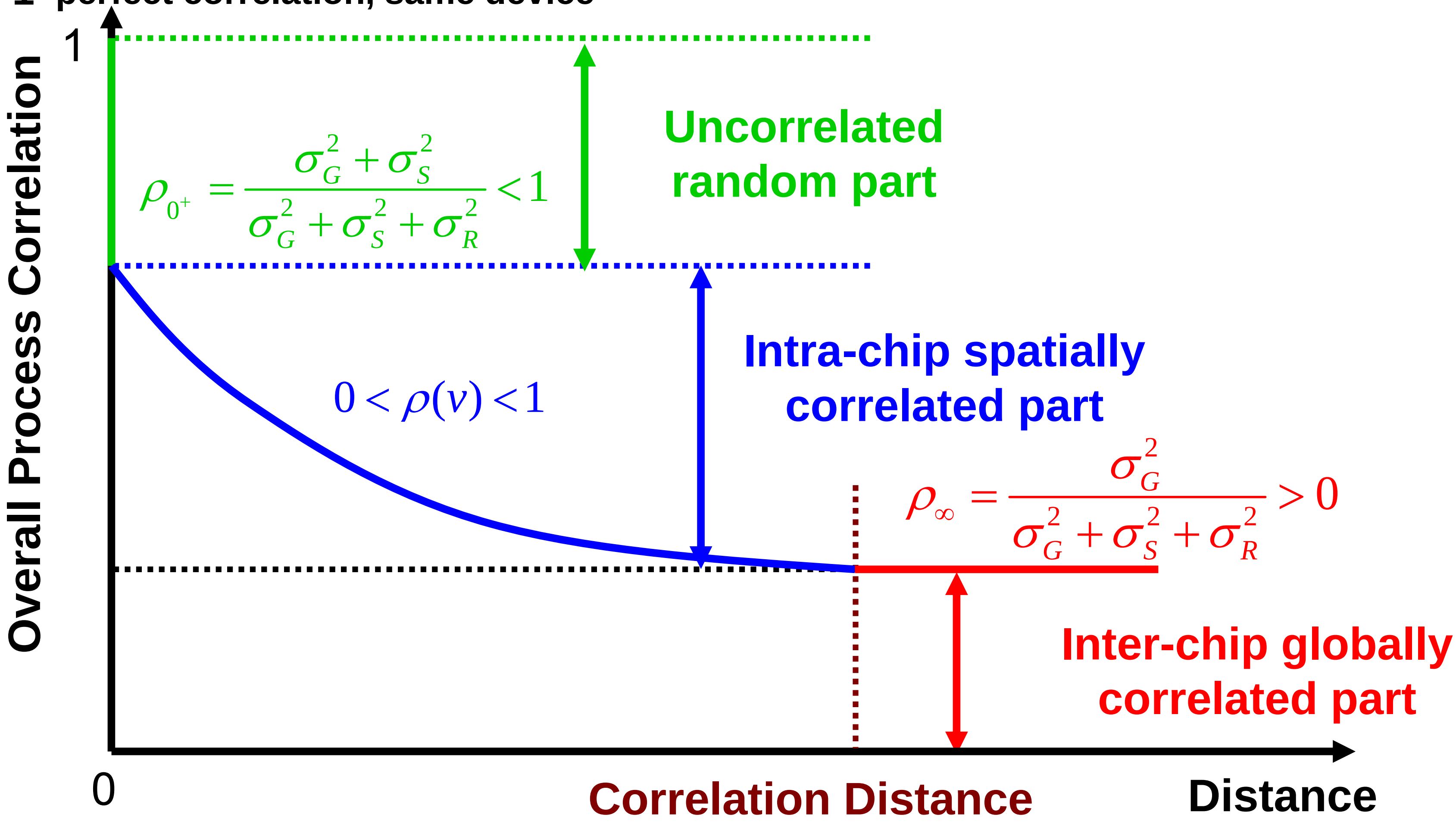
$$\rho_v = \frac{\mathrm{cov}(F_i, F_j)}{\sigma_i \sigma_j} = \frac{\sigma_G^2 + \rho(v)\sigma_S^2}{\sigma_G^2 + \sigma_S^2 + \sigma_R^2}$$

Overall Process Correlation without Measurement Noise

Overall process correlation

$$\rho_v = \frac{\mathrm{cov}(F_i, F_j)}{\sigma_i \sigma_j} = \frac{\sigma_G^2 + \rho(v)\sigma_S^2}{\sigma_G^2 + \sigma_S^2 + \sigma_R^2}$$

$\rho_v(0)=1$ perfect correlation, same device



$$\rho_{0^+} = \frac{\sigma_G^2 + \sigma_S^2}{\sigma_G^2 + \sigma_S^2 + \sigma_R^2} < 1$$

Uncorrelated random part

$$0 < \rho(v) < 1$$

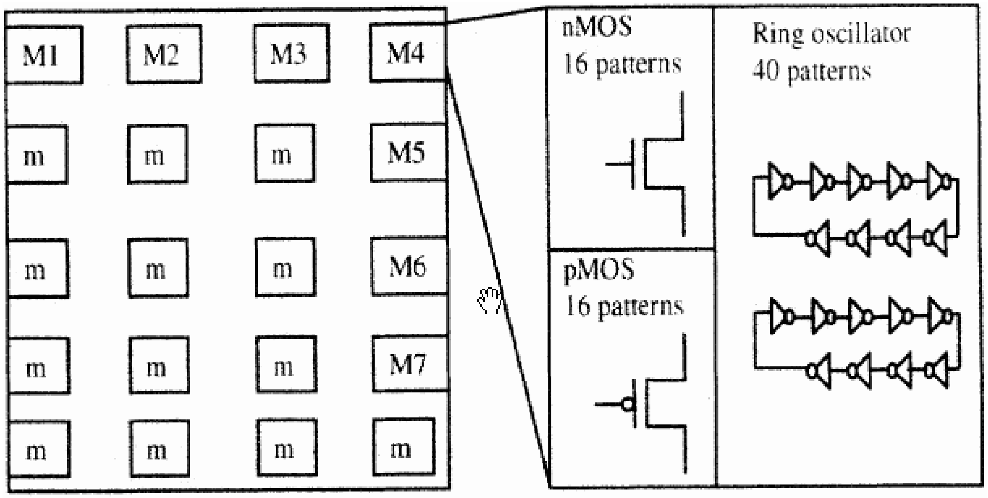
Intra-chip spatially correlated part

$$\rho_\infty = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_S^2 + \sigma_R^2} > 0$$

Inter-chip globally correlated part

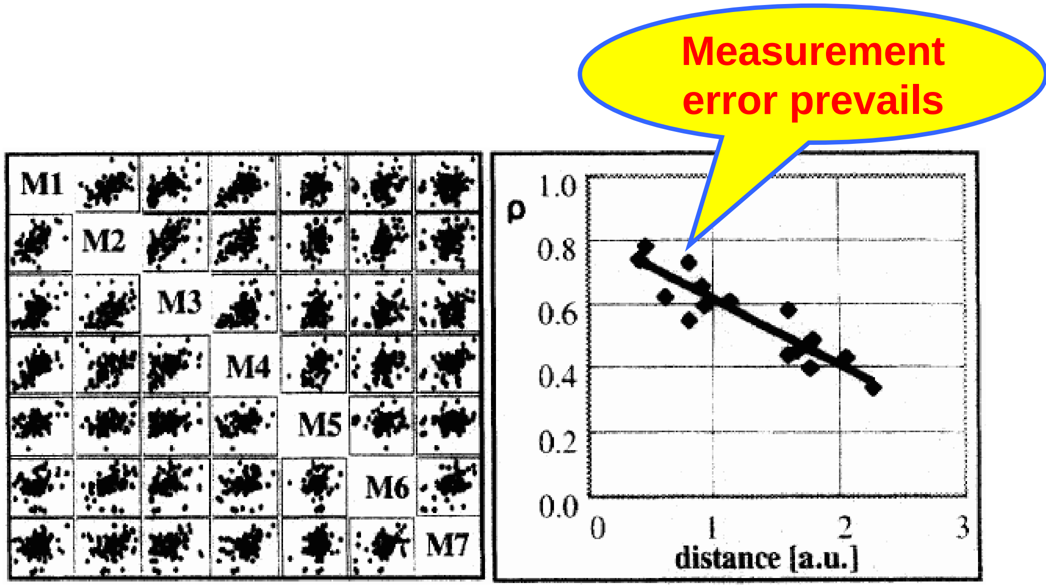
Die-scale Silicon Measurement [Doh et al., SISPAD 05]

- **Samsung 130nm CMOS technology**
- **4x5 test modules, with each module containing**
  - 40 patterns of ring oscillators
  - 16 patterns of NMOS/PMOS
- **Model spatial correlation as a first-order decreasing polynomial function**



Test chip (20mm x 20mm)

Test module (1200um x 600um)

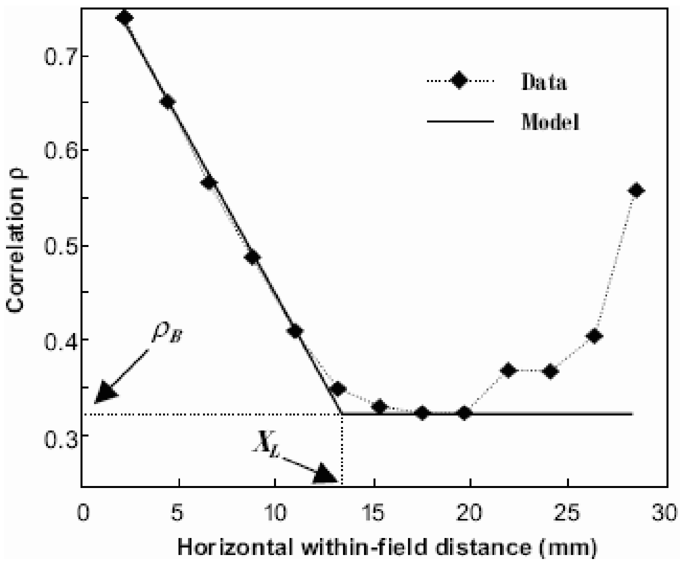


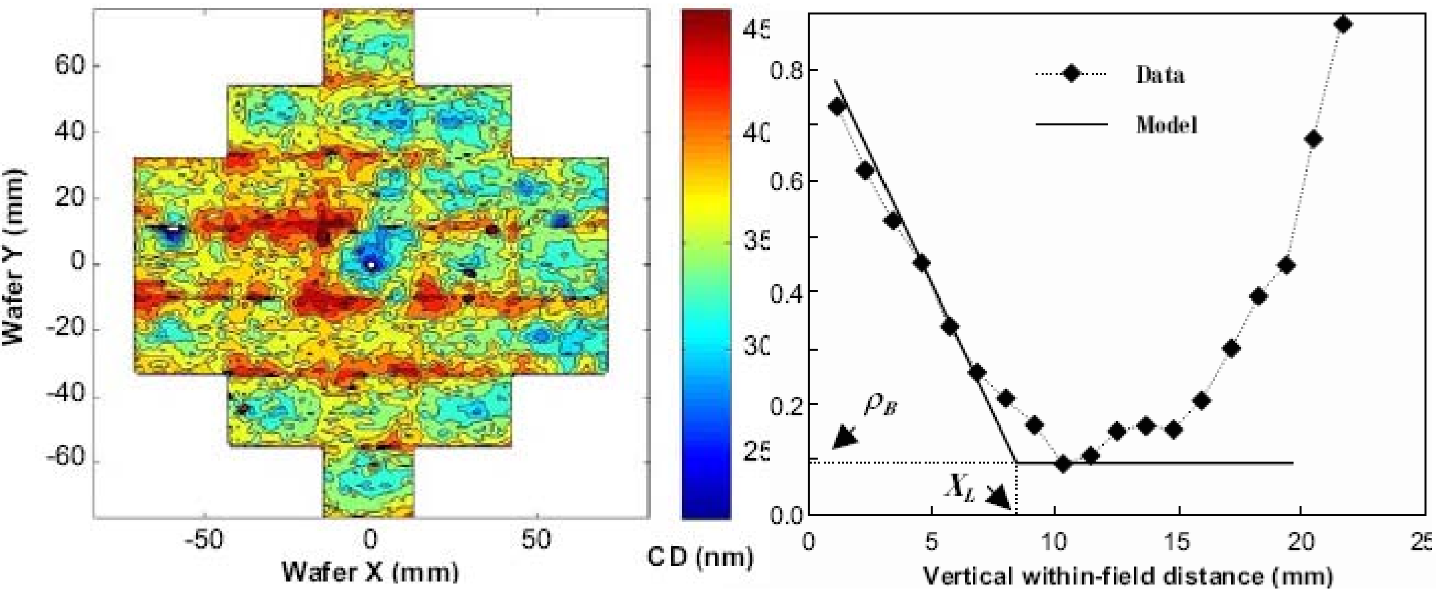
Measurement error prevails

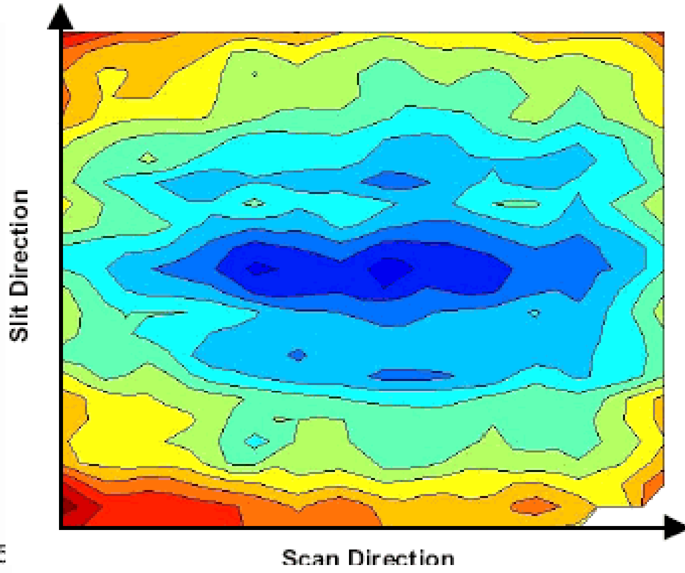
Correlation between measured NMOS saturation current

Wafer-scale Silicon Measurement [Friedberg et al., ISQED 05]

- **UC Berkeley Micro-fabrication Lab's 130nm technology**
- **23 die/wafer, 308 module/die, 3 patterns/module**
  - Die size: 28x22mm^2
- **Average measurements for critical dimension**
- **Model spatial correlation as a decreasing PWL function**

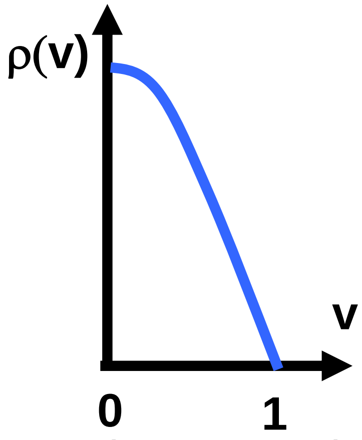


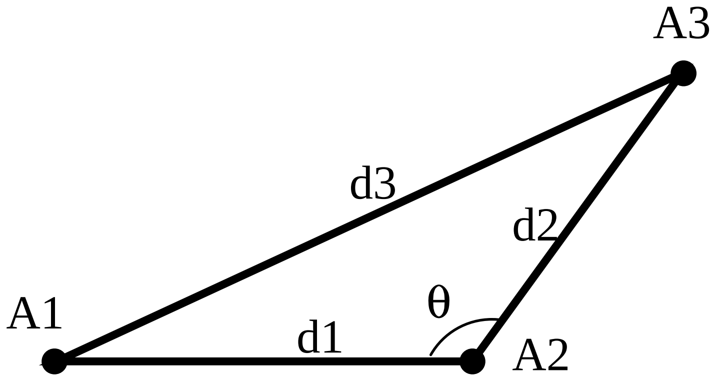




Limitations of Previous Work

- **Both modeled spatial correlation as monotonically decreasing functions (i.e., first-order polynomial or PWL)**
  - Devices close by are more likely correlated than those far away
- **But not all monotonically decreasing functions are valid**
  - For example, $\rho(v)=-v^2+1$ is monotonically decreasing on $[0,2^{1/2}]$





  - When d1=31/32, d2=1/2, d3=1/2, it results in a non-positive definite matrix

$$\Omega=\begin{pmatrix}1 & \rho(d_1) & \rho(d_3)\\ \rho(d_1) & 1 & \rho(d_2)\\ \rho(d_3) & \rho(d_2) & 1\end{pmatrix}=\begin{pmatrix}1 & 0.0615 & 0.75\\ 0.0615 & 1 & 0.75\\ 0.75 & 0.75 & 1\end{pmatrix}$$

Smallest eigen-value is -0.0303

Theoretic Foundation from Random Field Theory

- **Theorem: a necessary and sufficient condition for the function ρ(v) to be a valid spatial correlation function [Yaglom, 1957]**
  - For a HIR field, ρ(v) is valid iff it can be represented in the form of

$$\rho(v) = \int_0^{\infty} J_0(\omega v) d(\Phi(w))$$

    - where $J_0(t)$ is the Bessel function of order zero
    - $\Phi(\omega)$ is a real nondecreasing function such that for some non-negative p

$$\int_0^{\infty} \frac{d(\Phi(w))}{(1+w^2)^p} < \infty$$

  - For example:

$$\Phi(w) = 1 - (1 + w^2 / b^2)^{-0.5} \Longrightarrow \rho(v) = \exp(-bv)$$

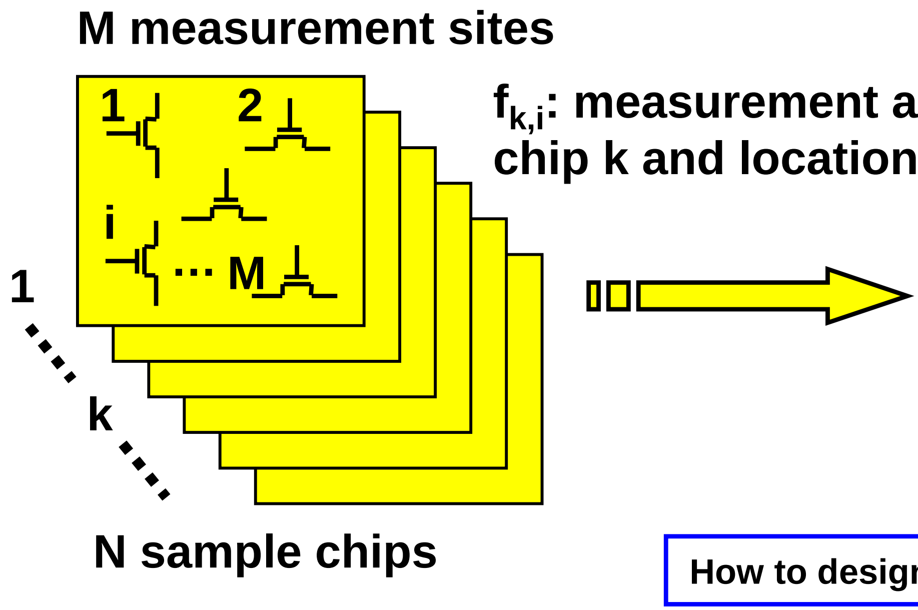
  - We cannot show whether decreasing polynomial or PWL functions belong to this valid function category → but there are many that we can

Agenda

- **Motivations**
- **Process Variation Modeling**
- **Robust Extraction of Valid Spatial Correlation Function**
  - Robust = immune to measurement noise
- **Robust Extraction of Valid Spatial Correlation Matrix**
- **Conclusion**

Robust Extraction of Spatial Correlation Function

- **Given**: noisy measurement data for the parameter of interest with possible inconsistency
- **Extract**: global variance σ_G^2, spatial variance σ_S^2, random variance σ_R^2, and spatial correlation function $\rho(v)$
- **Such that**: σ_G^2, σ_S^2, σ_R^2 capture the underlying variation model, and $\rho(v)$ is always valid



$f_{k,i}$: measurement at chip k and location i

$$\begin{cases} \sigma_F^2 = \sigma_G^2 + \sigma_S^2 + \sigma_R^2 \\ \rho(v) \end{cases}$$

Global variance

Spatial variance

Random variance

Valid spatial correlation function

How to design test circuits and place them are not addressed in this work

Extraction Individual Variation Components

- **Variance of the overall chip variation**

$$\sigma_F^2 \approx \sigma_f^2 = \frac{1}{MN-1}\left(\sum_i \sum_k f_{k,i}^2 - \frac{(\sum_i \sum_k f_{k,i})^2}{MN}\right)$$

Unbiased Sample Variance
[Hogg and Craig, 95]

- **Variance of the global variation**

$$\sigma_G^2 \approx \sigma_g^2 = \sigma_f^2 - \sigma_{fc}^2 = \sigma_f^2 - \frac{1}{M-1}\left(\sum_i f_{k,i}^2 - \frac{(\sum_i f_{k,i})^2}{M}\right)$$

- **Spatial covariance**

$$cov(F_i, F_j) = cov(v) \approx \frac{\sum_k f_{k,i} f_{k,j}}{N-1} - \frac{\sum_k f_{k,i} \sum_k f_{k,j}}{N(N-1)}$$

- **We obtain the product of spatial variance σ_S^2 and spatial correlation function $\rho(v)$**

$$\sigma_S^2 \cdot \rho(v) = cov(F_i, F_j) - \sigma_G^2 \approx cov(v) - \sigma_g^2$$

  - Need to separately extract σ_S^2 and $\rho(v)$
  - $\rho(v)$ has to be a valid spatial correlation function

Robust Extraction of Spatial Correlation

- **Solved by forming a constrained non-linear optimization problem**

$$\min_{\sigma_s^2,\rho(v)} : \left\| \sigma_s^2 \cdot \rho(v) - \mathrm{cov}(v) + \sigma_g^2 \right\|$$

  - Difficult to solve → impossible to enumerate all possible valid functions
- **In practice, we can narrow ρ(v) down to a subset of functions**
  - Versatile enough for the purpose of modeling
- **One such a function family is given by [Bras and Iturbe, 1985]**

$$\rho(v) = 2\left(\frac{b \cdot v}{2}\right)^{s-1} \cdot K_{s-1}(b \cdot v) \cdot \Gamma(s-1)^{-1}$$

  - K is the modified Bessel function of the second kind
  - Γ is the gamma function
  - Real numbers b and s are two parameters for the function family
- **More tractable → enumerate all possible values for b and s**

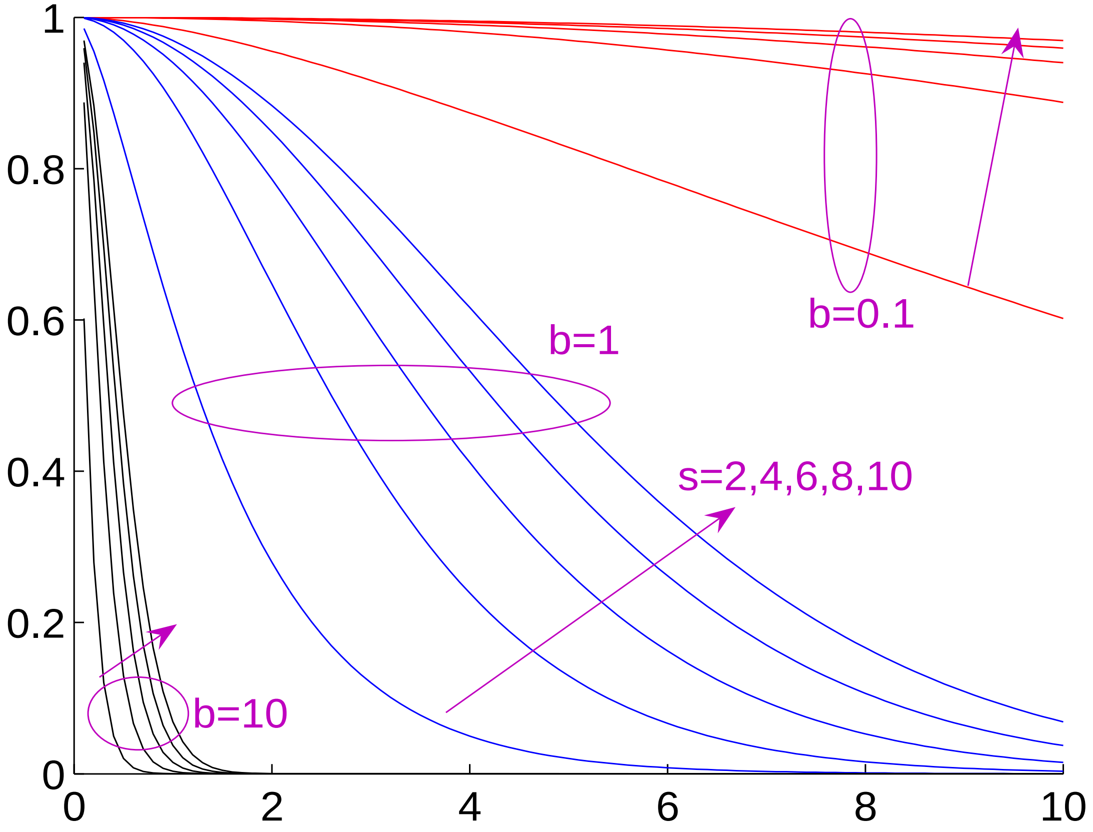
Robust Extraction of Spatial Correlation

- **Reformulate another constrained non-linear optimization problem**

$$\min_{\sigma_s^2, b, s} : \sum \left[2\sigma_s^2 \cdot \left(\frac{b \cdot v}{2} \right)^{s-1} \cdot K_{s-1}(b \cdot v) \cdot \Gamma(s-1)^{-1} - \mathrm{cov}(v) + \sigma_g^2 \right]^2$$

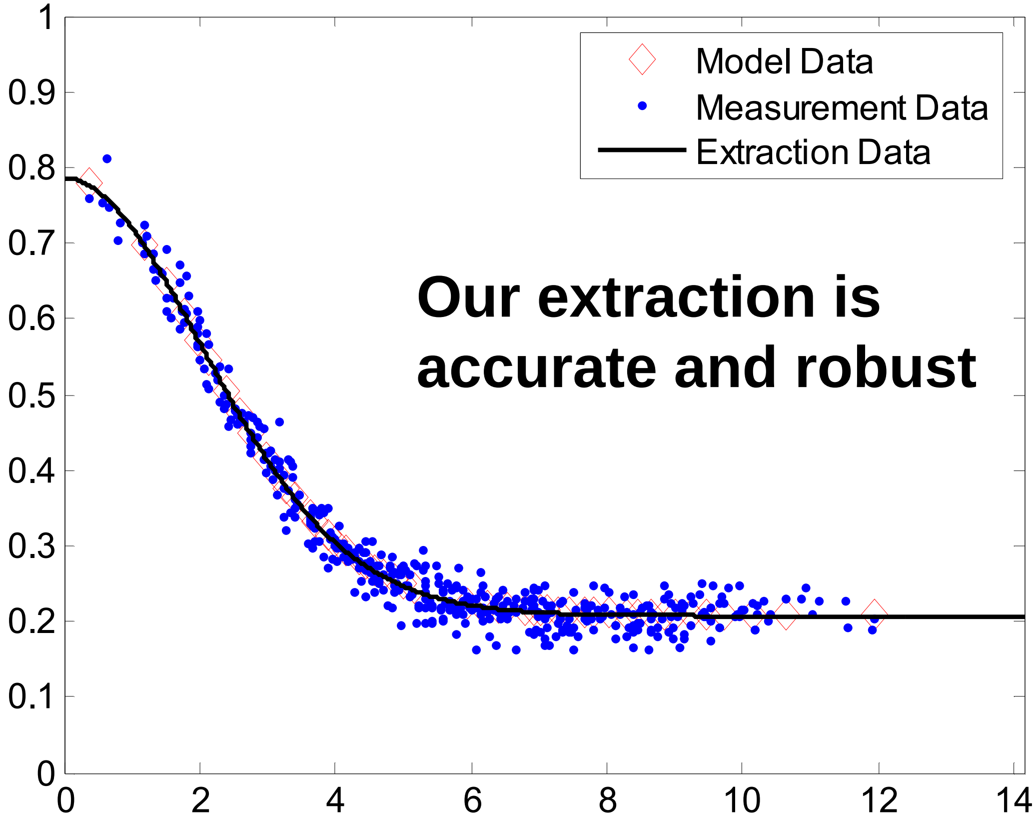
$$s.t.: \sigma_s^2 \leq \sigma_{fc}^2$$

Different choices of b and s ➔ different shapes of the function ➔each function is a valid spatial correlation function



Experimental Setup based on Monte Carlo Model

- **Monte Carlo model = different variation amount (inter-chip vs spatial vs random) + different measurement noise levels**
  - Easy to model various variation scenarios
  - Impossible to obtain from real measurement
- **Confidence in applying our technique to real wafer data**



Our extraction is accurate and robust

Results on Extraction Accuracy

Chip #	Site #	Noise level	Error(σ_g)	Error(σ_s)	Error(ρ(v))
2000	60	10%	0.40%	-1.90%	2.00%
		50%	0.30%	-2.80%	2.70%
		100%	0.30%	-2.60%	3.70%
1000	60	10%	7.50%	1.20%	1.00%
		50%	7.20%	1.00%	1.00%
		100%	6.90%	1.40%	1.00%
	50	10%	6.50%	0.80%	2.80%
		50%	5.70%	-0.40%	3.00%
		100%	5.10%	-3.00%	3.50%
	40	10%	8.60%	-4.10%	6.50%
		50%	8.70%	-3.90%	7.00%
		100%	8.90%	-2.30%	8.40%

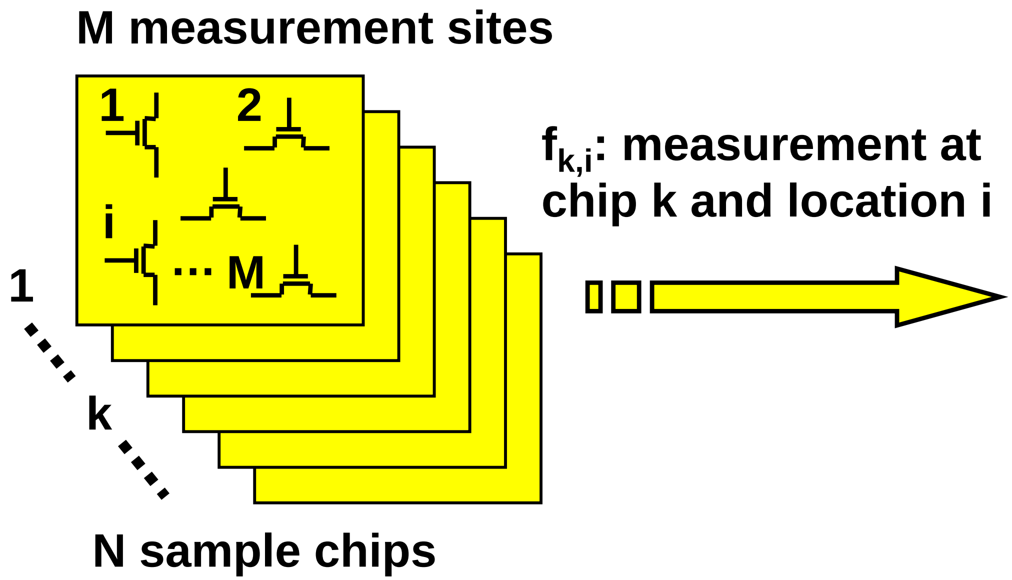
- **More measurement data (Chip# x site #) ➔ more accurate extraction**
  - More expensive
  - Guidance in choosing minimum measurements with desired confidence level

Agenda

- **Motivations**
- **Process Variation Modeling**
- **Robust Extraction of Valid Spatial Correlation Function**
- **Robust Extraction of Valid Spatial Correlation Matrix**
- **Conclusion**

Robust Extraction of Spatial Correlation Matrix

- **Given: noisy measurement data at M number of points on a chip**
- **Extract: the valid correlation matrix Ω that is always positive semidefinite**
- **Useful when spatial correlation cannot be modeled as a HIR field**
  - Spatial correlation function does not exist
  - SSTA based on PCA requires Ω to be valid for EVD



$f_{k,i}$: measurement at chip k and location i

$$\Omega = \begin{pmatrix} 1 & \cdots & \rho_{1M} \\ \vdots & \ddots & \vdots \\ \rho_{M1} & \cdots & 1 \end{pmatrix} \geq 0$$

Valid correlation matrix

Extract Correlation Matrix from Measurement

- **Spatial covariance between two locations**

$$cov(F_i, F_j) \approx \frac{\sum_k f_{k,i} f_{k,j}}{N-1} - \frac{\sum_k f_{k,i} \sum_k f_{k,j}}{N(N-1)}$$

- **Variance of measurement at each location**

$$\sigma_{F_i}^2 \approx \frac{1}{N-1}\left(\sum_k f_{k,i}^2 - \frac{(\sum_k f_{k,i})^2}{N}\right)$$

- **Measured spatial correlation**

$$A = \begin{pmatrix} \cdots & \cdots & \cdots \\ \cdots & \rho_{ij} & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

$$\boxed{\rho_{i,j} = \frac{cov(F_i, F_j)}{\sigma_{F_i} \sigma_{F_j}}}$$

- **Assemble all ρ_{ij} into one measured spatial correlation matrix A**
  - But A may not be a valid because of inevitable measurement noise

Robust Extraction of Correlation Matrix

$$\min_{\Omega} : \quad \| A - \Omega \|$$
$$s.t. : \quad \Omega \in correlation\ \ matrix.$$

- **Find a closest correlation matrix Ω to the measured matrix A**
- **Convex optimization problem [Higham 02, Boyd 05]**
- **Solved via an alternative projection algorithm [Higham 02]**
  - Details in the paper

```
$\Delta S_0 = 0,\ Y_0 = A$
for $k = 1,2,\ldots$
  $R_k = Y_{k-1} - \Delta S_{k-1}$
  $X_k = P_S(R_k)$
  $\Delta S_k = X_k - R_k$
  $Y_k = P_U(X_k)$
end
$\Omega = Y_k$
```

$$A = \begin{bmatrix} 1 & 0.0615 & 0.75 \\ 0.0615 & 1 & 0.75 \\ 0.75 & 0.75 & 1 \end{bmatrix}$$

$$\Omega = \begin{bmatrix} 1 & 0.0734 & 0.7326 \\ 0.0734 & 1 & 0.7326 \\ 0.7326 & 0.7326 & 1 \end{bmatrix}$$

$$\| A - \Omega \|_2 = 0.0313 \qquad \frac{\| A - \Omega \|_2}{\| A \|_2} = 0.0149$$

Results on Correlation Matrix Extraction

Sites	50	100	150	200
$\lambda(A)_{least}$	-0.83	-1.43	-1.84	-2.38
$\lambda(\Omega)_{least}$	0	0	0	0
$\|A-\Omega\|$	2.09	4.35	6.85	9.39
$\|A-\Omega\|/\|A\|$	5.2%	5.9%	6.6%	7.3%

- A is the measured spatial correlation matrix
- Ω is the extracted spatial correlation matrix
- λ is the smallest eigenvalue of the matrix
- Original matrix A is not positive, as λ is negative
- Extracted matrix Ω is always valid, as λ is always positive

Conclusion and Future Work

- **Robust extraction of statistical characteristics of process parameters is crucial**
  - In order to achieve the benefits provided by SSTA and robust circuit optimization
- **Developed two novel techniques to robustly extract process variation from noisy measurements**
  - Extraction of spatial correlation matrix + spatial correlation function
  - Validity is guaranteed with minimum error
- **Provided theoretical foundations to support the techniques**
- **Future work**
  - Apply this technique to real wafer data
  - Use the model for robust mixed signal circuit tuning with consideration of correlated process variations

Questions?

Thank You!