Non-Gaussian Parameter Modeling for SSTA with Confidential Interval Analysis

Lizheng Zhang, Jun Shao

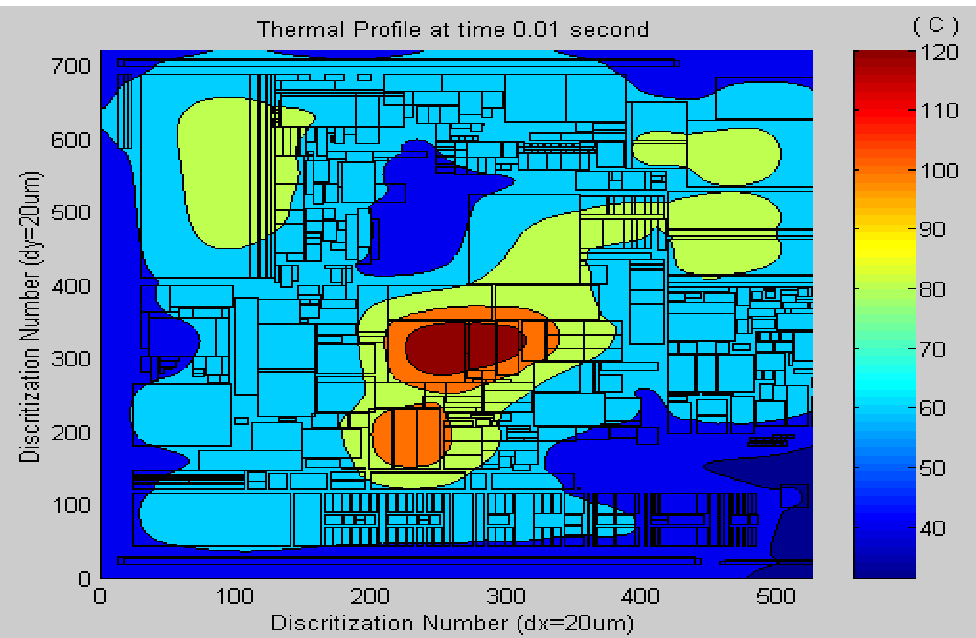
Charlie Chungping Chen

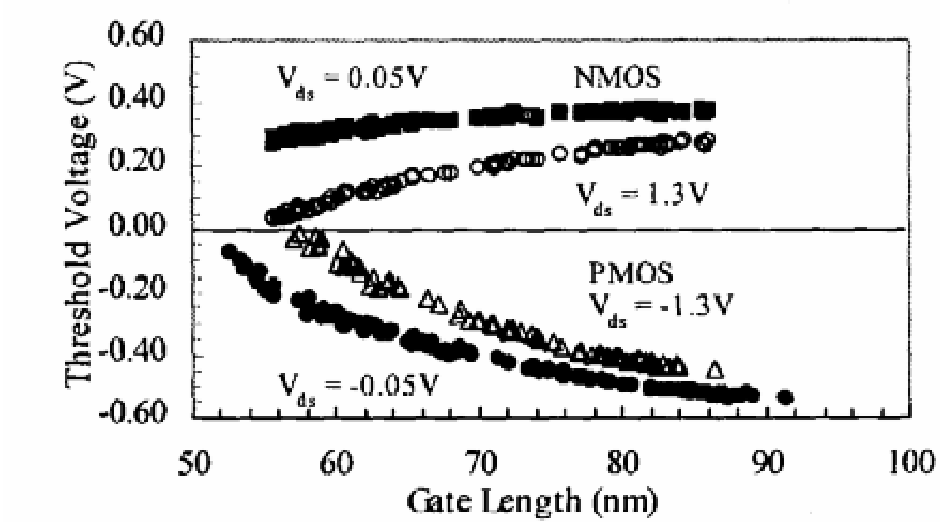
University of Wisconsin and

National Taiwan University

General Variation Sources

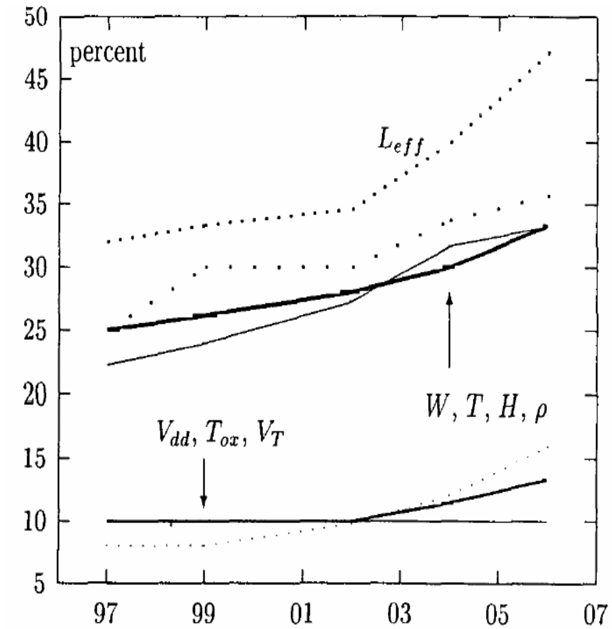
- Inter- and Intra-die Process Variation-W, L,…..
- PVT Variations:
  - Power Fluctuation: IR-drop,….
  - Non-uniform Temperature Distribution
  - Vt Variation
- Crosstalk: capacitive, inductive, and substrate coupling





Courtesy from Intel

Technology Parameter Variations



	1997	1999	2002	2005	2006
L_{eff} (nm)	250	180	130	100	70
T_{ox} (nm)	5	4.5	4	3.5	3
V_{dd} (V)	2.5	1.8	1.5	1.2	0.9
V_T (V)	0.5	0.45	0.4	0.35	0.3
W (μ)	0.8	0.65	0.5	0.4	0.3
H (μ)	1.2	1.0	0.9	0.8	0.7
ρ ($\frac{m\Omega}{\square}$)	45	50	55	60	75
L_{max} (μ)	*2123*	*1920*	*1670*	*1526*	*1303*

Table I: Technology parameters.

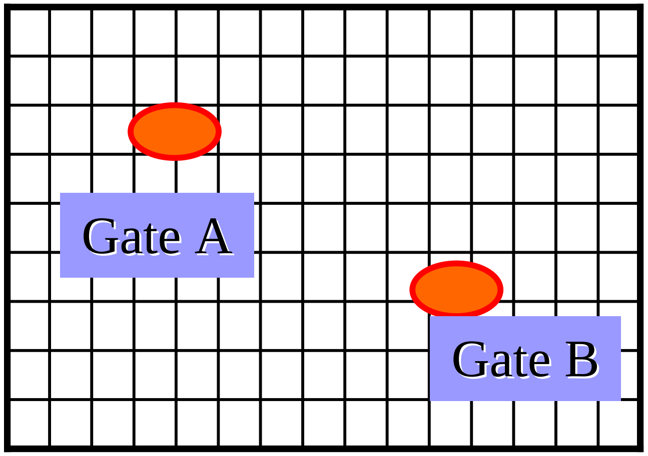
	1997	1999	2002	2005	2006
L_{eff} (nm)	80	60	45	40	33
T_{ox} (nm)	0.4	0.36	0.39	0.42	0.48
V_{dd} (V)	0.25	0.18	0.15	0.12	0.09
V_T (V)	50	45	40	40	40
W (μ)	0.2	0.17	0.14	0.12	0.1
H (μ)	0.3	0.3	0.27	0.27	0.25
ρ ($\frac{m\Omega}{\square}$)	10	12	15	19	25

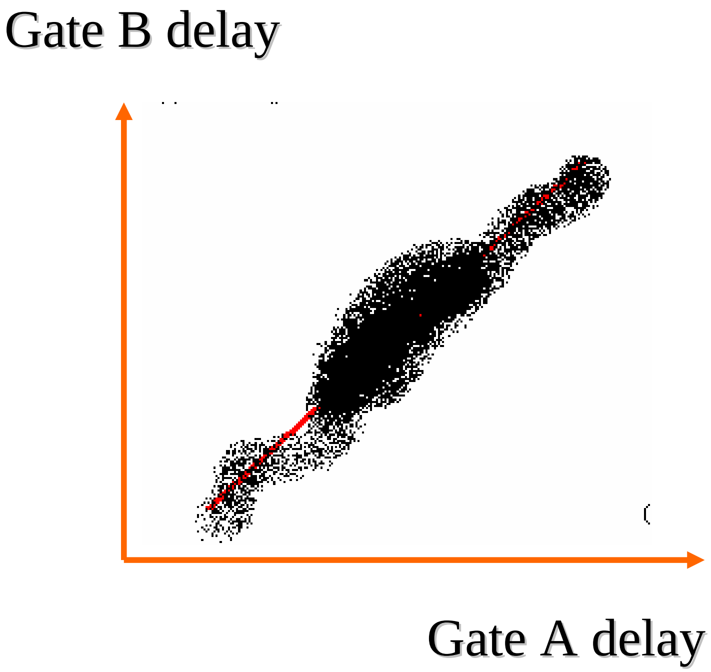
Table II: Technology parameter 3σ variations.

- Parameter variation as a percentage of its nominal value becomes larger and larger
  - Variations do not scale down or scale down slower than the nominal values
- Parameters may be correlated: there are three types of correlations
  - (1) Physical dependency; (2) Spatial correlation; (3) topological correlation
- Circuit performance is a function of technology parameters
  - Becomes a random variable and needs statistical method to compute

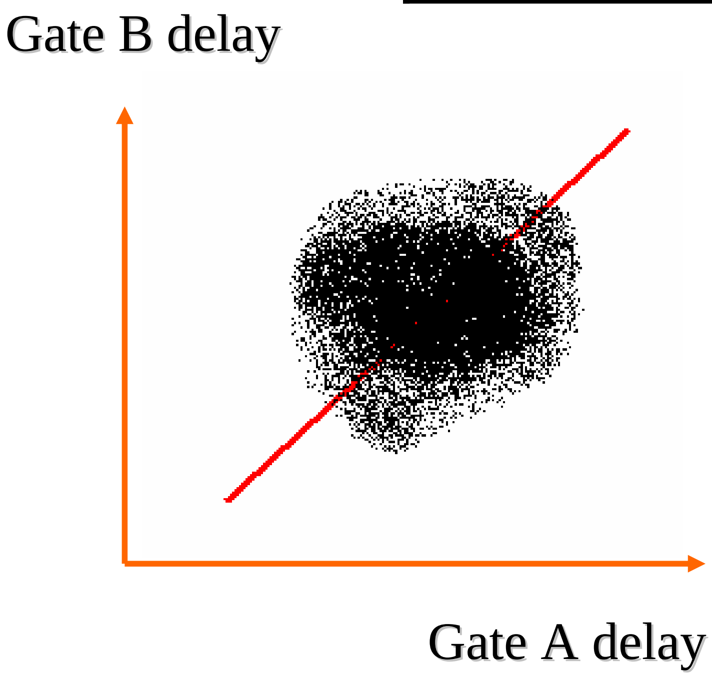
"Models of process variations in device and interconnect" Duane Boning, MIT & Sani Nassif, IBM ARL.

Correlated and Uncorrelated (independent) Variation



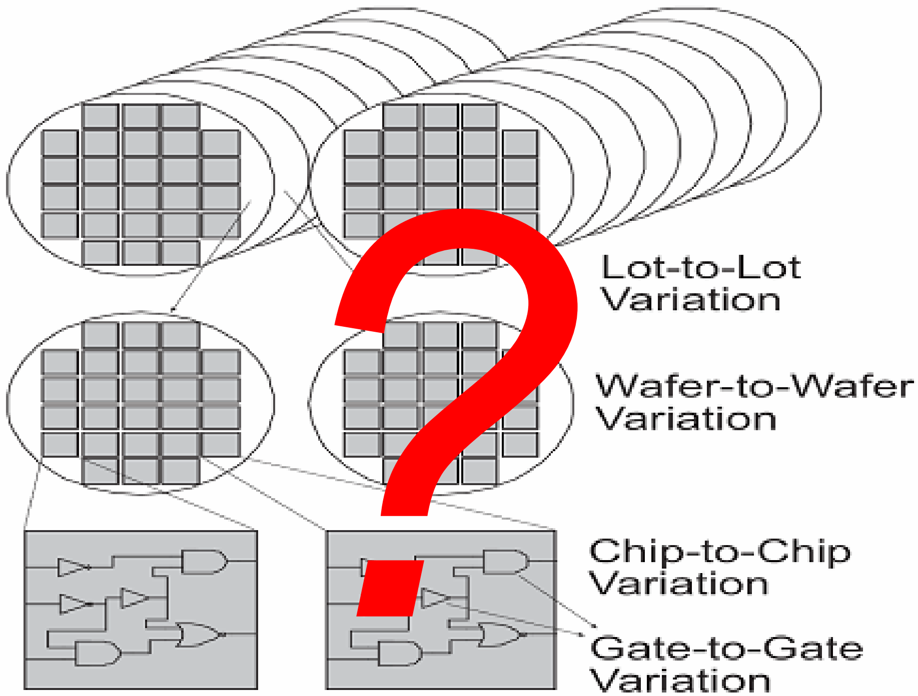


Correlated



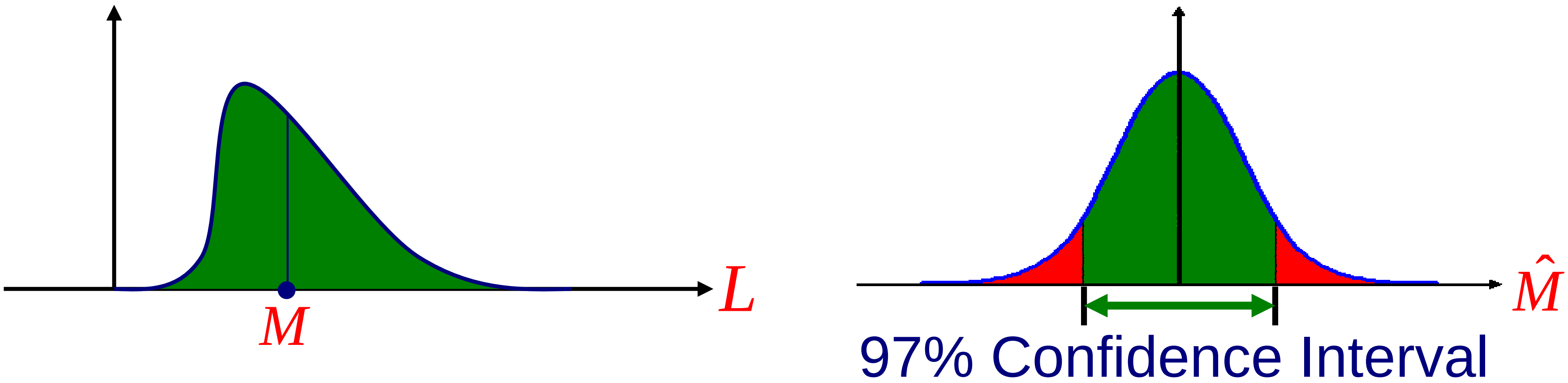
Uncorrelated

Process Parameter Variations



- Process parameter variations can be (*given by some poor guys*) represented by statistical distributions
  - Those distributions are characterized by several distribution parameters
- Distribution parameters can be estimated from physical measurements on test chips
  - Limited budget makes it infeasible to have infinite number of test chips
  - The distribution may be *non-Gaussian*
  - The interested variable (say Elmore Delay, RC~W_a/W_b) operation may be *non-linear*
- Distribution parameters have errors
  - Confidence interval analysis measures the fidelity of the given distribution parameters

Confidence Interval Analysis (QC for Data)

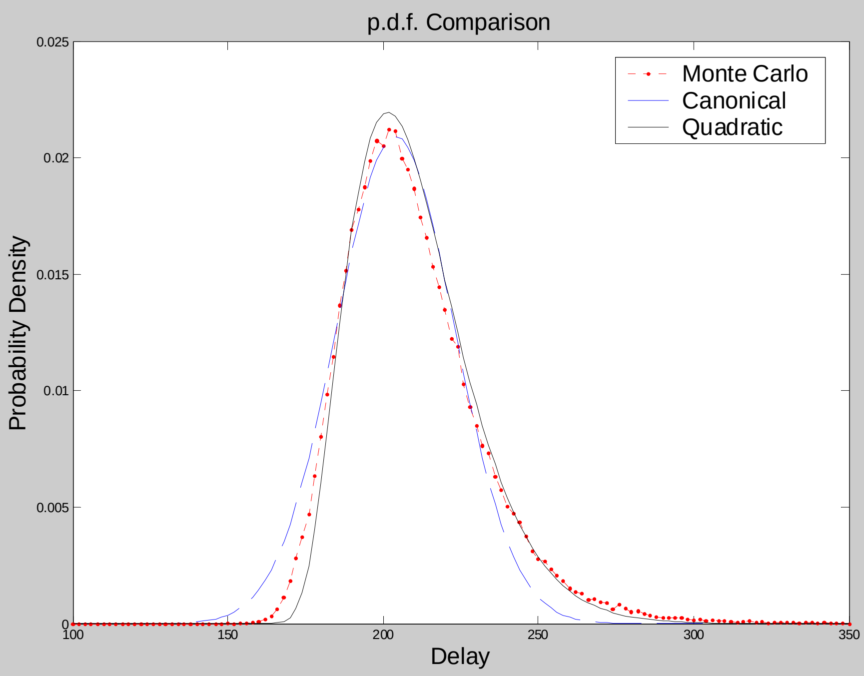


- Suppose gate length L follows an arbitrary distribution with true mean value of M. We get estimation of M from the channel length measurement on N gates: $L_1, L_2, \ldots, L_N$ as $\hat{M} = (L_1 + L_2 + \cdots + L_N)/N$
- The true mean M is deterministic, but the estimation $\hat{M}$ has limited accuracy due to limited test chips
- The gate length L may not be Gaussian, but the distribution of the estimation $\hat{M}$ can be close to Gaussian
- The 97% confidence interval for the gate length mean estimation is:
  - *chance to have the true mean value fall inside the interval is 97%*

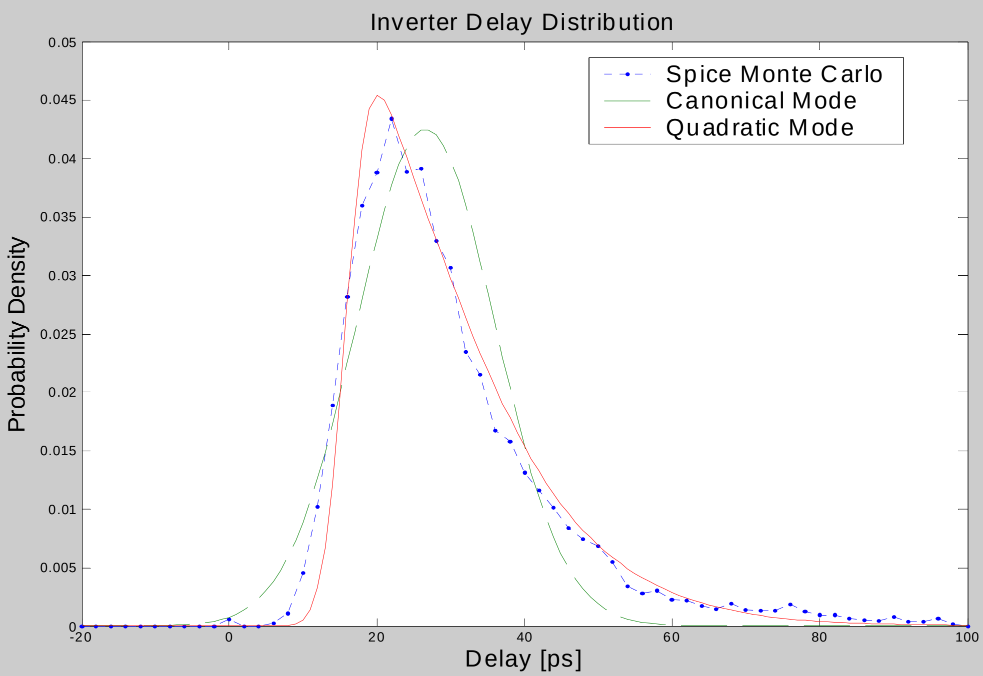
Non-Gaussian Gate and Interconnect Delay and Quadratic Gaussian Approximation

- Gate delay are functions of gate length, Vt, …
- Interconnect delay are functions of width, thickness,…

$$D_w = \sum_{i=1}^{N} \sum_{j=i}^{N} R_i C_j = \sum_{i=1}^{N} \sum_{j=i}^{N} \frac{r_s l^2 (c_s W_j + c_f T_j)}{W_i T_i}$$



Interconnect



Gate

Points to make

- Gate and Interconnect delay may be non-gaussian and also depend on width, length, vt, vdd, temperatures,….etc
- Delay calculation operation may be non-linear
- Due to finite test resources, confidence interval need to be carefully considered
- We propose to use quadratic gaussian model to match first three moments (mean, variation, and skewness) and correlations
- We provide systematic way to fit the measurement data
- We provide confidence interval analysis at the same time

Literature Review of SSTA

- In literature, there are two categories of SSTA approaches
  - Path-based SSTA: can consider false paths but with exponential complexity in the worst case
    - Michael Orshansky (DAC'02): statistical bounding
    - Kwang-Ting Cheng (DAC'02): delay testing oriented
  - Block-base SSTA: has linear complexity but is difficult to consider false paths
    - Chandu Visweswariah (DAC'04): parametric delay and direct linear MAX operation
    - Sachin S. Sapatnekar (ICCAD'03): parametric delay and projection-based linear MAX operation
    - Larry Pileggi (DAC'04): parametric delay and table-lookup based MAX operation
    - David Blauuw (TCAD'03): PDF-based approach
    - C. Chen (DAC' 05): Quadratic SSTA

Non-Gaussian Parameter

- Blindly using Gaussian variables to fit the process parameters may induce huge errors
- Correlation between non-Gaussian variables may be hard to manipulated directly, so direct non-Gaussian modeling may not be efficient
- We propose to use quadratic form of a standard Gaussian variable X to model the non-Gaussian parameter Y with distribution parameters of $a\ b\ c$

$$Y = aX^2 + bX + c$$

  - First three statistical moments (mean, variation, skewness) and correlations are exactly matched
  - Fully compatible with the our proposed quadratic SSTA

SSTA Delay Model

- Canonical Gaussian delay model $D = \mu + \alpha R + \sum_i \beta_i Y_i$
- Quadratic Gaussian delay model

$$D = \mu + \alpha R + \sum_i \beta_i Y_i + \sum_i \sum_j \Gamma_{ij} Y_i Y_j$$

$$\begin{aligned} D(L,V) &= m_g + \alpha R + \frac{\partial D}{\partial L} L + \frac{\partial D}{\partial V} V \\ &+ \frac{1}{2}\frac{\partial^2 D}{\partial L^2} L^2 + \frac{\partial^2 D}{\partial L \partial V} LV + \frac{1}{2}\frac{\partial^2 D}{\partial V^2} V^2 \\ &= m + \alpha R + \boldsymbol{\beta}^* \boldsymbol{\delta} + \boldsymbol{\delta}^* \boldsymbol{\Gamma} \boldsymbol{\delta} \end{aligned}$$

- Problems
  - Process parameters Y_1, Y_2, … are assumed to be Gaussian random variables
  - Process parameter distributions are known without any uncertainties

"***Correlation-Preserved Non-Gaussian Statistical Timing Analysis with Quadratic Timing Model***"
Lizheng Zhang, Weijen Chen, Yuhen Hu, John A. Gubner, Charlie Chung-Ping Chen, DAC'05

Moment Properties

- The mean, variation, and skewness of the measured data set ($Y_1, Y_2, \cdots, Y_N$) and modeling function ($Y = aX^2 + bX + c$) satisfy:

$$\mu_Y = a + c = \frac{1}{N} \sum_{i=1}^{N} Y_i$$

$$\sigma_Y^2 = 2a^2 + b^2 = \frac{1}{N-1} \sum_{j=1}^{N} \left(Y_j - \hat{Y}\right)^2$$

$$\kappa_Y^3 = 8a^3 + 6ab^2 = \frac{N}{(N-1)(N-2)} \sum_{j=1}^{N} \left(Y_j - \hat{Y}\right)^3$$

Moment Matching Process

- a will be one and the only one of the following three values whichever is real and within the range of $|a| < \sigma_Y / 1.414$:

$$a_1 = -\frac{2\sigma_Y^2 + \Delta^{2/3}}{2\Delta^{1/3}} \; ; \; a_2 = \frac{2(1+i\sqrt{3})\sigma_Y^2 + (1-i\sqrt{3})\Delta^{2/3}}{4\Delta^{1/3}}$$

$$a_3 = \frac{2(1-i\sqrt{3})\sigma_Y^2 + (1+i\sqrt{3})\Delta^{2/3}}{4\Delta^{1/3}} \; ; \; \Delta = \kappa_Y^3 + i\sqrt{8\sigma_Y^6 - \kappa_Y^6}$$

- Solutions for b and c can be found as:

$$b = \sqrt{\sigma_Y^2 - 2a^2} \text{ and } c = \mu_Y - a$$

$Y = aX^2 + bX + c$

Covariance Matched As Well

- Given data sets $(Y_{\alpha,1}, Y_{\alpha,2}, \cdots, Y_{\alpha,N})$ and $(Y_{\beta,1}, Y_{\beta,2}, \cdots, Y_{\beta,N})$ covariance between parameters Y_{α} and Y_{β}, $cov(Y_{\alpha}, Y_{\beta})$ can bet estimated as

$$cov(Y_\alpha, Y_\beta) = \frac{1}{N-1}\sum_{j=1}^{N}\left(Y_{\alpha,j}Y_{\beta,j} - \hat{Y}_\alpha \hat{Y}_\beta\right)$$

- Given quadratic models (Y_{α}¼$a_{\alpha} X_{\alpha}^{2}+b_{\alpha} X_{\alpha}+c_{\alpha}$) and ($Y_{\beta}$¼$a_{\beta} X_{\beta}^{2}+b_{\beta} X_{\beta}+c_{\beta}$), correlation between X_{α} and X_{β}, $\rho_X = cov(X_{\alpha}, X_{\beta})$ can be solved analytically as follows:

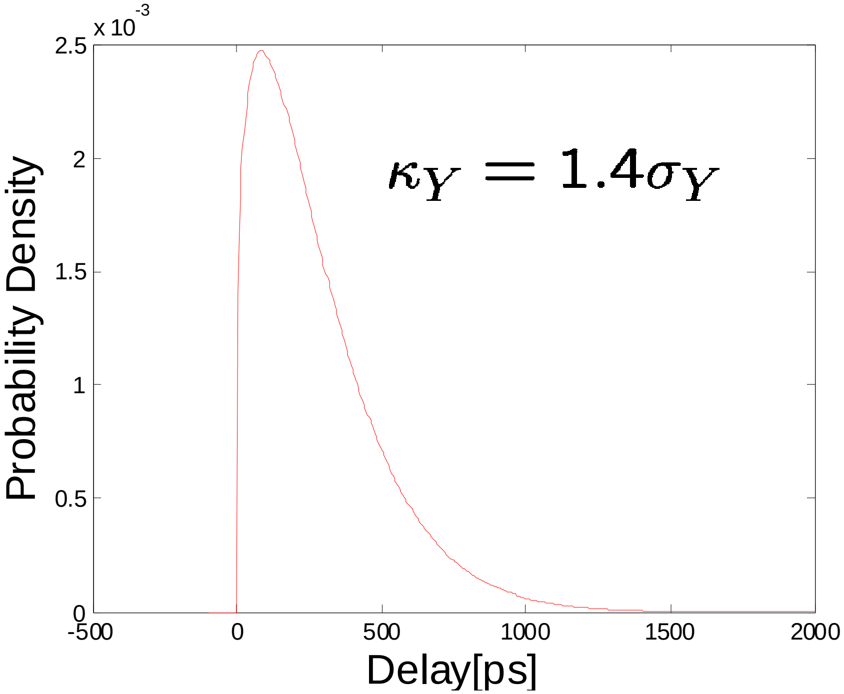
$$\rho_X = \frac{2cov(Y_\alpha, Y_\beta)}{b_\alpha b_\beta + \sqrt{b_\alpha^2 b_\beta^2 + 8a_\alpha a_\beta cov(Y_\alpha, Y_\beta)}}$$

Conditions for Real *a, b, c*

- Exact moment matching is possible only when Y's skewness

$$|\kappa_Y| \leq \sqrt{2}\sigma_Y$$

  - Most real parameter variations is within this constraint



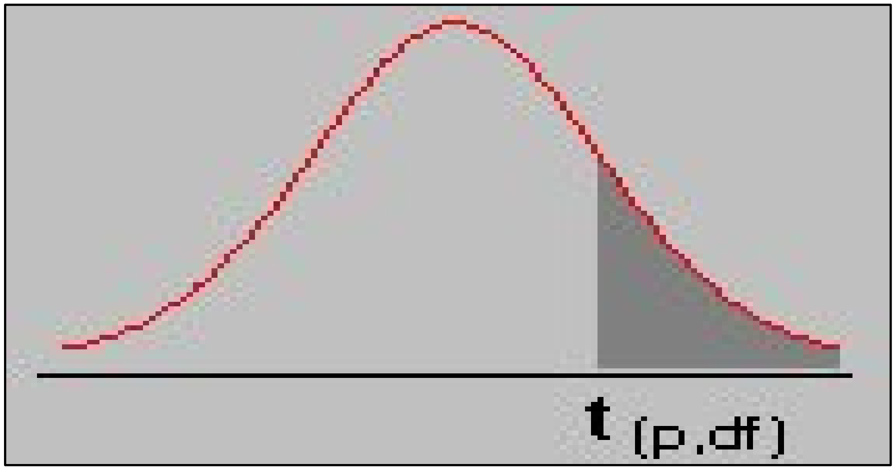
Confidence Interval

$$\text{Confidence Level}: p$$

$$\text{Degree of freedom}: df = N\text{-}1$$

$$\text{Student's t Distribution}: t(p, df)$$

$$(1\text{-}2p)\ \text{Confidence Interval}: \bar{X} - t(p, df)\frac{S}{\sqrt{N}} \le \mu \le \bar{X} + t(p, df)\frac{S}{\sqrt{N}}$$



df \ p	0.05(90%)	0.025 (95%)	0.01(98%)
1	6.314	12.71	31.82
2	2.920	4.303	6.965
3	2.353	3.182	4.541
4	2.132	2.776	3.747
5	2.015	2.571	3.365

Confidence Intervals Analysis

- *Jackknife* Method are used to estimate the confidence interval of coefficients *a, b, c*

Complete data set estimation: $Y_1, Y_2, \cdots, Y_N \;\Rightarrow\; \theta$

Reduced data set estimations:

$$\begin{cases} \text{Skip } Y_1\text{: } Y_2, Y_3 \cdots, Y_N & \Rightarrow \theta_1 \\ \text{Skip } Y_2\text{: } Y_1, Y_3 \cdots, Y_N & \Rightarrow \theta_2 \\ \vdots & \\ \text{Skip } Y_N\text{: } Y_1, Y_2 \cdots, Y_{N-1} & \Rightarrow \theta_N \end{cases}$$

θ = *a*, *b*, or *c*

Pseudo-value data set:

$$P_i = N\theta - (N-1)\theta_i \;\Rightarrow\; P_1, P_2, \cdots, P_N$$

Confidence interval: $\hat{P} \pm \frac{t_{N-1,\alpha/2}}{\sqrt{N}} \hat{S}_P$

Error reduces as $1/\sqrt{N}$

(One digit accuracy require 100 data points!)

Data Analysis Process for Non-Gaussian Mean Estimate

$$==$$

$$\text{Sectioning}: \underbrace{X_1 \cdots X_m}_{\overline{X}_{G1}} \quad \underbrace{X_{m+1} \cdots X_{2m}}_{\overline{X}_{G2}} \quad \underbrace{X_{2m+1} \cdots X_{3m}}_{\overline{X}_{G3}} \quad \cdots \quad \underbrace{\cdots X_N}_{\overline{X}_{Gk}}$$

$$-----------------------------------$$

$$---------N = mk \text{ and } k \text{ sections with size } m---------$$

$$-----------------------------------$$

$$\text{Normally Distributed}: \overline{X}_{G1}\ \overline{X}_{G2}\ \overline{X}_{G3} \cdots \overline{X}_{Gk} \quad \text{(Large Number Law)}$$

$$-----------------------------------$$

$$\text{Average and Stabdard Deviation}: \ \overline{X}_{XG} \text{ and } S_{XG}$$

$$-----------------------------------$$

$$\text{Confidence Interval}: \overline{X}_{XG} - t(p,df)\frac{S_{XG}}{\sqrt{k}} \le \mu \le \overline{X}_{XG} + t(p,df)\frac{S_{XG}}{\sqrt{k}}$$

$$==$$

Data Analysis Process for Non-Gaussian Variance Estimate

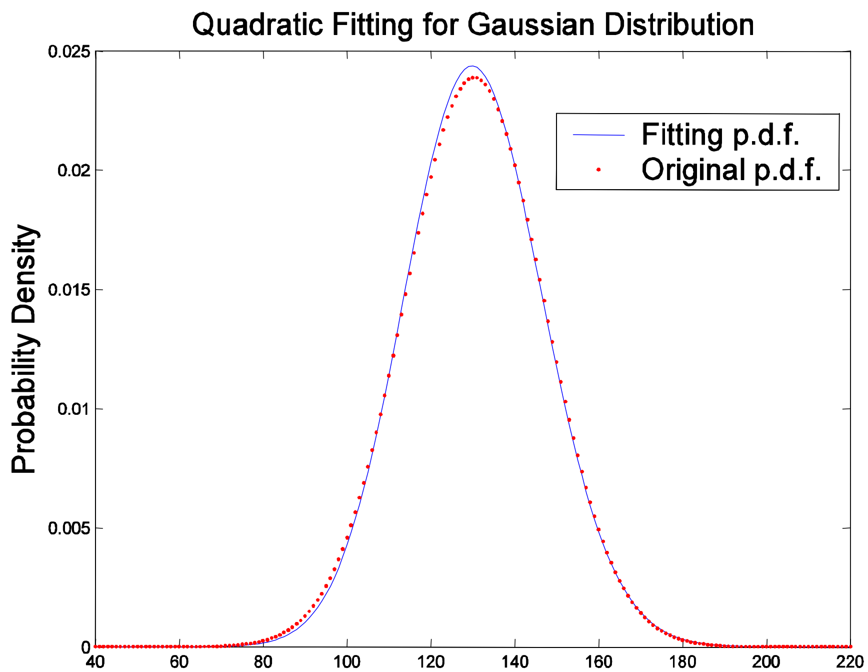
$$\text{Sectioning}: \underbrace{X_1 \cdots X_m}_{S_{G1}} \quad \underbrace{X_{m+1} \cdots X_{2m}}_{S_{G2}} \quad \underbrace{X_{2m+1} \cdots X_{3m}}_{S_{G3}} \quad \cdots \underbrace{\cdots X_N}_{S_{Gk}}$$

$$\text{Normally Distributed}: S_{G1}\ S_{G2}\ S_{G3} \cdots S_{Gk}$$

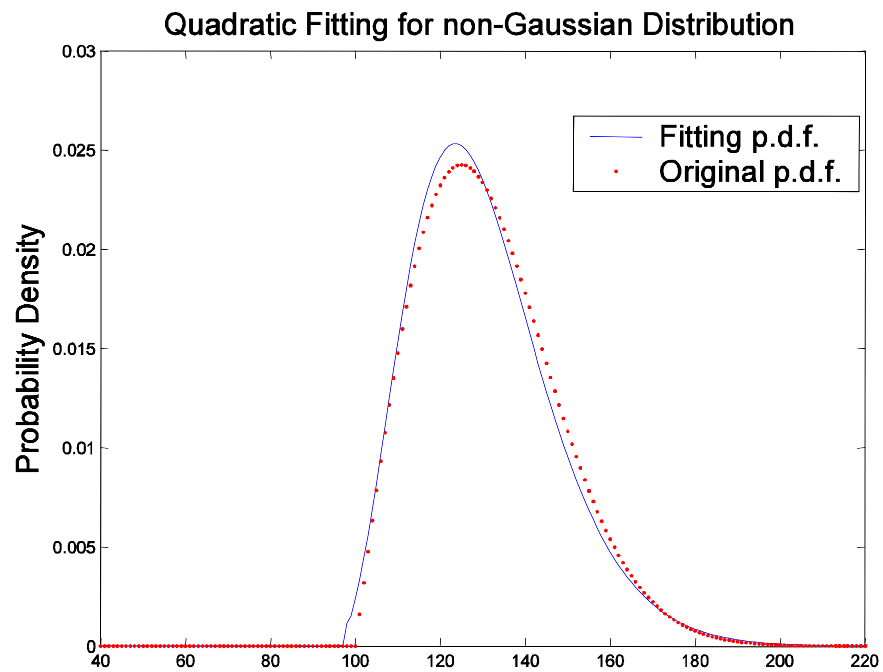
$$\text{Average and Stabdard Deviation}: \ \bar{S}_G \text{ and } S_{SG}$$

$$\text{Confidence Interval}: \bar{S}_G - t(p,df)\frac{S_{SG}}{\sqrt{k}} \le \sigma \le \bar{S}_G + t(p,df)\frac{S_{SG}}{\sqrt{k}}$$

Experimental Result: Moment Matching for Quadratic Parameter Modeling

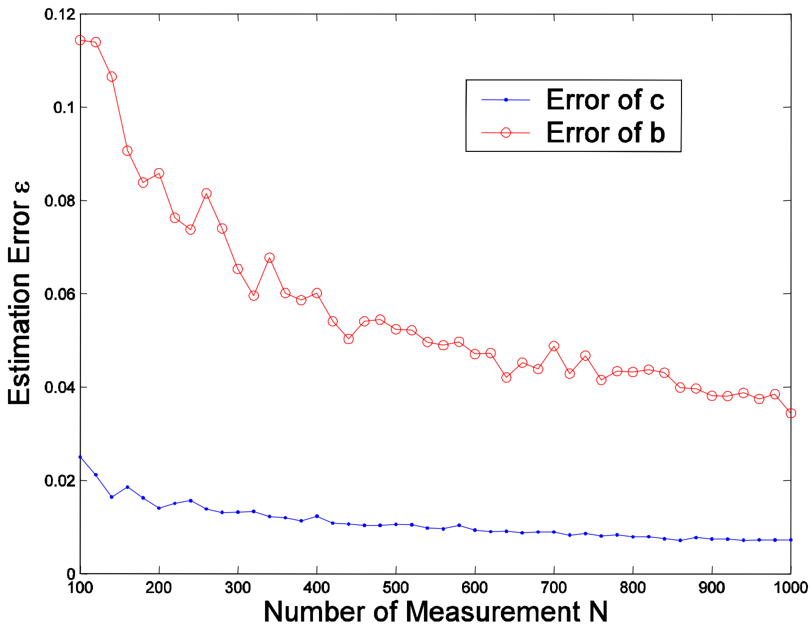


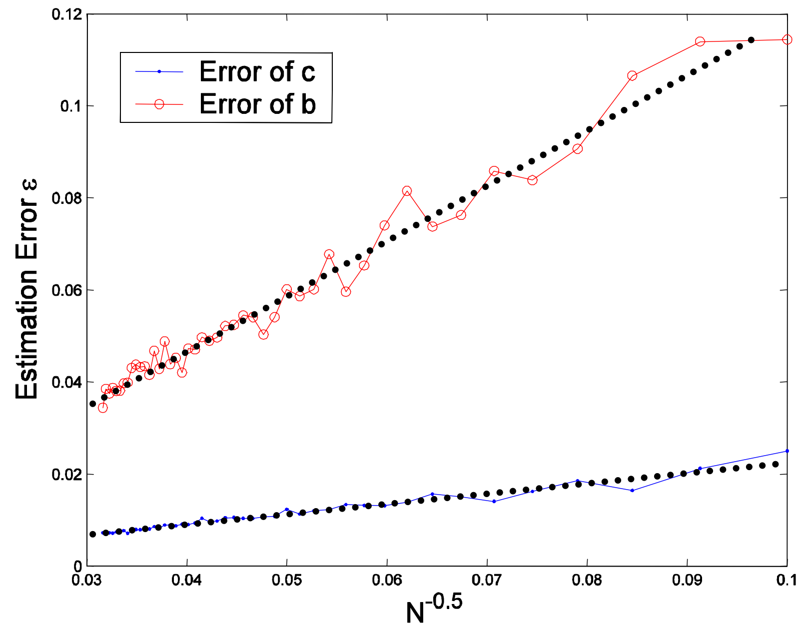
Skewness = 0



Skewness = 0.75

Experimental Result: Estimation Error v.s. Data Points





- Error is reduced in the rate of 1/sqrt(N)

Conclusions

- We propose to model Gaussian and non-Gaussian process parameters using quadratic format of Gaussian random variables
  - Three moments and covariance are matched
  - Fully compatible with the existing SSTA method
- We develop Jackknife-method-based confidence interval analysis to estimate the modeling error inherited from the limited resources
  - Generally speaking, to improve one digit of accuracy, 100X experiments will be required
  - Error reduces as 1/sqrt(N)….