

LVF²: A Statistical Timing Model based on Gaussian Mixture for Yield Estimation and Speed Binning

Junzhuo Zhou^{1,2}, Li Huang³, Haoxuan Xia³, Yihui Cai⁴, Leilei Jin⁴
Xiao Shi⁴, Wei Xing^{2*}, Ting-Jung Lin^{2*}, and Lei He^{2,1}

¹University of California, Los Angeles

²Ningbo Institute of Digital Twin, Eastern Institute of Technology, Ningbo, China

³University of Nottingham Ningbo China ⁴Southeast University, Nanjing, China

Abstract

As transistor size continues to scale down, process variation has become an essential factor determining semiconductor yield and economic return. The Liberty Variation Format (LVF) is the current industrial standard that expresses statistical timing behaviors based on single Gaussian model. However, it loses accuracy when the timing distribution is non-Gaussian due to growing process variations. This paper proposes a novel LVF² distribution model that combines two weighted skewed-normal (SN) distributions, which better captures the multi-Gaussian timing distribution while maintaining backward compatibility with LVF. Experiments using TSMC 22nm standard cells show that, compared to LVF, LVF² reduces binning error by 7.74× in delay and 9.56× in transition time, and reduces 3σ-yield error by 4.79× and 7.18× in delay and transition time, respectively. The error reduction for path delay is diminished due to Central Limit Theorem (CLT). But it is still 2× for a typical circuit path with 8 Fanout-of-4 (FO4) inverter delays.

Keywords: Speed binning, yield estimation, statistical timing modeling, process variation, LVF

1 Introduction

As transistor size continues to scale down, process variation has led to significant uncertainties in the performance evaluation of integrated circuits (ICs). Due to device parameter fluctuations, fabricated ICs vary from each other in performance. As significant spreads are observed in the delay due to process variation, speed binning has become essential during manufacturing tests for maximizing economic return [1]. The speed binning process sorts ICs into different bins according to the highest permissible operation frequency.

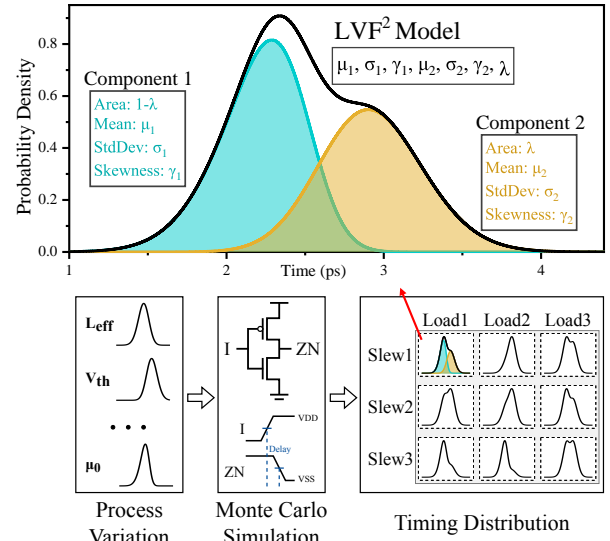


Figure 1. Timing distribution modeling by LVF².

The working ICs are then priced based on respective bins. Accurate statistical static timing analysis (SSTA) at the design stage provides a good insight into yield estimation and speed binning prediction, providing an early indicator for pricing strategy development.

Several works have focused on statistical models that quantify the probability density function (PDF) of delay distributions. The cell delay distribution was initially modeled by a Gaussian distribution, assuming a linear relationship between process variations and cell delays at normal supply range in [2]. Shown in bottom of Figure 1, as the technology node and supply voltage scale down, the non-linear process variation effect turns the delay distribution non-Gaussian. To model the asymmetry of the distribution, the SN model adds a parameter to shift the Gaussian density function to one side to match the actual delay distribution [3]. Eventually, the SN model is widely used in industry LVF [4], which considers skewness and uses a moment-based model to define timing distribution. To factor in the exponential effects of threshold voltage on the delay distribution, log-normal (LN) and log-skew-normal (LSN) models were proposed in [5] and [6]. In the near-threshold region, the delay distribution becomes asymmetric and has a long tail. By matching the kurtosis, the fourth moment of the cell delay distribution,

This research is partially supported by a research grant from AWS.

* Corresponding authors

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

DAC '24, June 23–27, 2024, San Francisco, CA, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0601-1/24/06.

<https://doi.org/10.1145/3649329.3655670>

[7] proposed the log-extended-skew-normal (LESN) model that enhanced the accuracy of the $\pm 3\sigma$ delay in either the sub-threshold or the near-threshold voltage region.

While previous works and the industrial standard mainly focus on fitting each statistical moment, there is mounting evidence of multiple Gaussian components in the timing distributions. One straightforward example is the multiple peaks in a PDF, which was shown in the experimental results of [8], but the work did not explore the phenomenon further. In other cases, although the multi-peak distribution is not significant, using multiple Gaussians can describe PDFs better than conventional single-peak-based models. Gaussian Mixture Model (GMM) is a probabilistic distribution model that represents normally distributed subpopulations within an overall population, which is widely used in natural science research [9]. A statistical timing model based on GMM was proposed as part of the SSTA algorithm in [10]. The work demonstrated that GMM effectively reduced the $\mu + 3\sigma$ error compared with the canonical Gaussian model. However, the model did not consider the skewness of the Gaussian components. On the other hand, without comprehensively verifying the timing distributions at the standard-cell level, the proposed SSTA method did not become a widely used industrial solution.

This paper proposes a novel statistical timing model based on GMM, denoted by LVF². As shown at the top of Figure 1, LVF² combines two weighted SN distributions, and can better address the multi-Gaussian and asymmetry in the timing distributions. At the same time, LVF² can reflect the kurtosis of PDFs. Regarding library definition, LVF² preserves backward compatibility with the conventional LVF library. Experiments using TSMC 22nm standard cells show that, compared to LVF, LVF² reduces binning error by 7.74 \times in delay and 9.56 \times in transition, and reduces 3 σ -yield error by 4.79 \times and 7.18 \times in delay and transition, respectively. The error reduction is reduced for path delay due to CLT. But it still improves 2 \times for a typical circuit path with 8-FO4 delay.

The rest of this paper is structured as follows. Section 2 introduces background. Section 3 describes LVF² modeling methodology. Section 4 presents experimental results. Finally, Section 5 concludes the paper.

2 Preliminary

2.1 Speed Binning

Figure 2 illustrates the concept of speed binning. Chips in the bin with a delay smaller than T_{min} suffer from an increased subthreshold leakage and are considered faulty. T_{max} is the target design delay that chips must satisfy. Only chips with a delay within T_{min} and T_{max} are considered usable. Each tested chip is classified into one of the bins based on its maximum operating frequency and will be priced accordingly. P_1 , P_2 , and P_3 in the figure represent the price for Bin_1 , Bin_2 , and Bin_3 , respectively. Faster chips will be sold higher, and profile decreases as the performance drops.

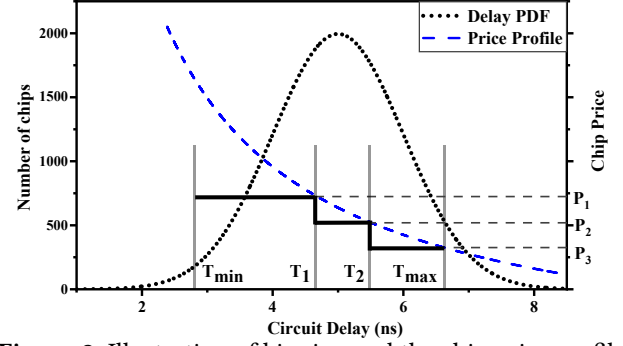


Figure 2. Illustration of binning and the chip price profile

Assume a binning process that has T_1, T_2, \dots, T_n as the boundaries of bins. The total number of bins in the PDF is $n + 1$. The probability of bin i is given by

$$P(Bin_i) = \begin{cases} P(t < T_1) & i = 1 \\ P(t < T_i) - P(t \leq T_{i-1}) & 2 \leq i \leq n \\ 1 - P(t \leq T_n) & i = n + 1 \end{cases} \quad (1)$$

where $P(t < T_i)$ represents the probability of the circuit delay t less than T_i .

2.2 Liberty Variation Format

Liberty Format is an industrial standard to define timing, power, and noise of standard cells. LVF [4] characterizes the statistical circuit behaviors considering on-chip variations. It is required for technology nodes of 22nm and below.

In LVF, each slew-load pair is associated with a delay distribution and a transition distribution. LVF uses four lookup tables (LUTs) to characterize the timing distributions, which include the nominal values and three statistical moments, mean μ , sigma σ , and skewness γ . There exist a bijection g between statistical moments vector $\theta = (\mu, \sigma, \gamma)$ and SN parameters vector $\Theta = (\xi, \omega, \alpha)$ [11]:

$$g : \theta \leftrightarrow \Theta \quad (2)$$

in most applications, the statistical moments vector θ of LVF defines a SN distribution whose PDF is

$$f_{LVF}(x|\theta) = f_{SN}(x|\Theta) = \frac{2}{\omega} \phi\left(\frac{x-\xi}{\omega}\right) \Phi\left(\alpha \frac{x-\xi}{\omega}\right), \quad (3)$$

where ϕ is the PDF of normal distribution and Φ is the cumulative distribution function (CDF) of normal distribution. Take `cell_rise` as an instance, the definition attributes are:

- `cell_rise`: the LUT of the nominal values.
- `ocv_mean_shift_cell_rise`: the LUT of the offset values from the mean to the nominal, and `mean = nominal + mean_shift`.
- `ocv_std_dev_cell_rise`: the LUT of the standard deviation.
- `ocv_skewness_cell_rise`: the LUT of the skewness.

3 LVF² Statistical Timing Model

This section presents the proposed statistical timing model, LVF², which inherits the conventional LVF model introduced in Section 2.2 and equips it with GMM to address multi-Gaussian and other non-Gaussian phenomena.

3.1 Statistical Definition of LVF²

As shown at the top of Figure 1, LVF² combines two SN distributions to address the multiple Gaussian components; (4) is its PDF expression, where $\theta_1 = (\mu_1, \sigma_1, \gamma_1)$ and $\theta_2 = (\mu_2, \sigma_2, \gamma_2)$ are the parameters of the two SN distributions, and λ is the weight coefficient.

$$f_{LVF^2}(x|\lambda, \theta_1, \theta_2) = (1 - \lambda)f_{LVF}(x|\theta_1) + \lambda f_{LVF}(x|\theta_2) \quad (4)$$

3.2 LVF² Model Fitting

To stay consistent with the industry standard, we use point estimation for the model parameters $\theta = (\lambda, \theta_1, \theta_2)$ (instead of the Bayesian approach that derives the posterior distribution of the parameters). However, even point estimation using maximum likelihood estimation is challenging due to the complexity of the mixture of two SN distributions. Specifically, the log-likelihood function for LVF² is

$$\log p(\theta; X) = \sum_{i=1}^n \log [(1 - \lambda)f_{LVF}(x_i|\theta_1) + \lambda f_{LVF}(x_i|\theta_2)], \quad (5)$$

where $X = \{x_1, x_2, \dots, x_n\}$ is the observed data. Due to the complex interaction between these parameters, the optimization is non-convex with multiple local maxima.

To simplify the optimization, the expectation-maximization (EM) algorithm [12] introduces latent variables $Z = \{z_1, z_2, \dots, z_n\}$, with each z_i indicating the likelihood of data point x_i belonging to a particular component of the mixture. The computation of z_i is based on the current estimates of the parameters $\theta^{(old)}$, and is given by the posterior probability of each component for a given data point:

$$z_i = \frac{(1 - \lambda)f_{LVF}(x_i|\theta_1^{(old)})}{(1 - \lambda)f_{LVF}(x_i|\theta_1^{(old)}) + \lambda f_{LVF}(x_i|\theta_2^{(old)})} \quad (6)$$

The complete-data log-likelihood is:

$$\log L_c(\theta; X, Z) = \sum_{i=1}^n [z_i \log((1 - \lambda)f_{LVF}(x_i|\theta_1)) + (1 - z_i) \log(\lambda f_{LVF}(x_i|\theta_2))] \quad (7)$$

The initial values θ are obtained by the combination of the K-means clustering algorithm [13] and the method of moments [14]. Initially, we partition the observed data X into two groups using the K-means algorithm. Subsequently, for each group, we derive the initial values for θ_1 and θ_2 utilizing moments estimators.

In the E-step, the expected value of the complete-data log-likelihood is computed with respect to the distribution of Z given the current parameter estimates $\theta^{(old)}$:

$$Q(\theta; \theta^{(old)}) = E_{Z|X, \theta^{(old)}} [\log L_c(\theta; X, Z)] \quad (8)$$

In the M-step, the parameters are updated by maximizing the expected log-likelihood:

$$\theta^{(new)} = \arg \max_{\theta} Q(\theta; \theta^{(old)}) \quad (9)$$

The EM algorithm iterates between the E-Step and M-Step, gradually updating the parameter estimates. This iterative process is suitable for finding a global maximum, avoiding the complexities of the original log-likelihood function.

3.3 Backward Compatibility

Compatibility with an existing industry standard is critical when introducing incremental features. We propose the Liberty Format of LVF² that ensures backward compatibility with LVF, i.e., SSTA tools that support LVF² can recognize LVF without additional effort, and library files can support LVF and LVF² simultaneously without conflicts.

Based on the parameters given in (4), seven new attributes are introduced to define LVF² in library. Take `cell_rise` as an instance, the definition is as follows:

- `ocv_mean_shfit1_cell_rise`: The LUT provides the mean of the first Gaussian (μ_1) - nominal. The **default** table values inherit from `ocv_mean_shift_cell_rise`.
- `ocv_std_dev1_cell_rise` (σ_1): The LUT specifies the standard deviation of the first Gaussian. The **default** values inherit from `ocv_std_dev_cell_rise`.
- `ocv_skewness1_cell_rise` (γ_1): The LUT specifies the skewness of the first Gaussian. The **default** table values inherit from `ocv_skewness_cell_rise`.
- `ocv_weight2_cell_rise` (λ): The LUT specifies the weight of the second Gaussian, with a range of $[0, 1]$. The **default** values are all zeros.
- `ocv_mean_shift2_cell_rise`: The LUT provides the mean of the second Gaussian (μ_2) - nominal.
- `ocv_std_dev2_cell_rise` (σ_2): The LUT provides the standard deviation of the second Gaussian.
- `ocv_skewness2_cell_rise` (γ_2): The LUT provides the skewness of the second Gaussian.

The above definition ensures that when an LVF²-capable SSTA tool processes a conventional LVF library, it automatically recognizes the SN distribution of LVF as the first component of LVF² in the approach of (10). Although LVF² assumes only two Gaussian components, one can easily extend the library to support more components by following similar attribute naming conventions.

$$\begin{aligned} f_{LVF^2}(x|\theta_{LVF}) &= f_{LVF^2}(x|\lambda = 0, \theta_1 = \theta_{LVF}, \theta_2 = \emptyset) \\ &= f_{LVF}(x|\theta_{LVF}) \end{aligned} \quad (10)$$

3.4 Under the Law of The Central Limit Theorem

It is worth investigating the benefit and limitation of LVF² in practice, particularly in SSTA, where the random timing delays are accumulated, i.e., the sum of random variables. The CLT states that the sum of a large number of independent random variables, each with finite mean and variance, will be approximately normally distributed [15].

To give a more quantitative analysis, let us consider the sum of n independent random variables X_1, X_2, \dots, X_n (not necessarily Gaussian) with finite mean μ_i and variance σ_i^2 to represent the timing delay of an n -stage circuit. We first normalize each distribution by subtracting the mean and dividing by the standard deviation, i.e., $Y_i = (X_i - \mu_i)/\sigma_i$.

Theorem 1 (Berry-Esseen Theorem [16]). *Let Y_1, Y_2, \dots, Y_n be i.i.d. with zero mean, variance $\sigma^2 = 1$, and finite third absolute moment $\rho = \mathbb{E}[|X_i|^3]$. Then, for all n , the CDF of*

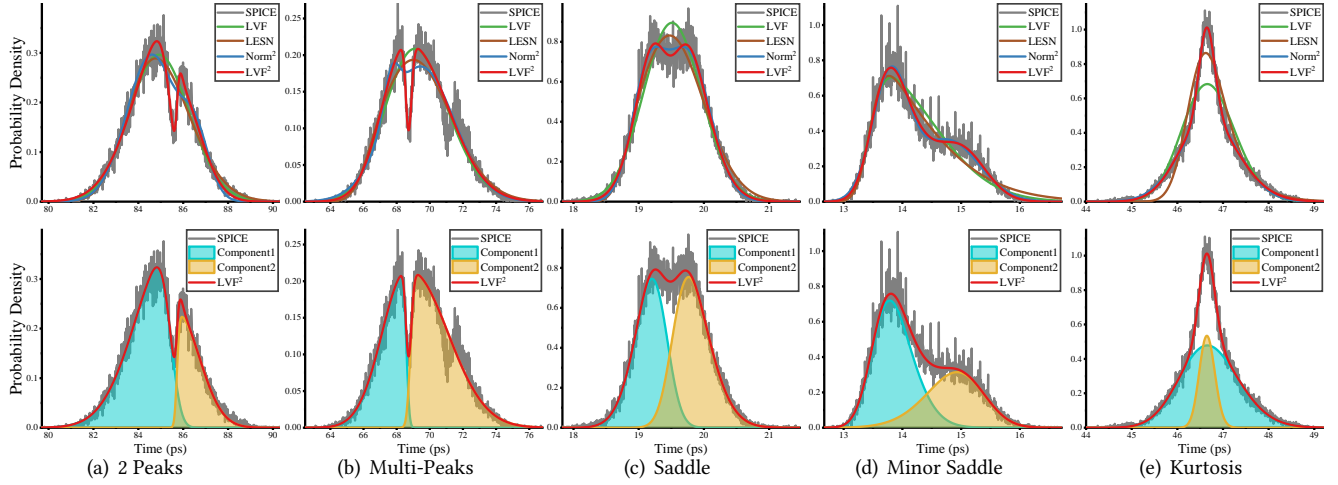


Figure 3. Fitting results of LVF, LESN, Norm², LVF² (above); and the decomposition of LVF² (bottom) for typical scenarios.

Y_n , denoted by $F_n(x)$, and the CDF of the standard normal distribution, $\Phi(x)$, satisfy

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sqrt{n}}, \quad (11)$$

where C is an absolute constant and ρ is the third moment.

Corollary 2. For the SSTA propagation, the convergence rate of the accumulated delay to be normal is $O(1/\sqrt{n})$.

Corollary 3. For the SSTA propagation, where X_i is close to Gaussian distribution, the third moment ρ will dominate the convergence rate.

What we can learn from the CLT and Berry-Esseen Theorem is that for SSTA problems, LVF²'s advantage will be less significant if the stage delay is close to Gaussian distribution or when the logic depth is deep. This gives us an insight of when to switch from LVF² to the compatible LVF in order to save storage space and computational time.

4 Experiments and Discussions

To assess LVF², we performed experiments using the TSMC 22nm standard cells at the TTGlobal_LocalMC corner. The supply voltage is 0.8V and the temperature is 25°C. For each timing distribution, 50k process variation samples were generated by Latin Hypercube Sampling (LHS) SPICE Monte Carlo (MC) simulation with all variations turned on.

The experiments compared four timing models: LVF², Norm² [10], LESN [7], and LVF [4]. Norm² is the combination of two weighted normal distributions that can be defined by five parameters $(\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2)$. The difference between Norm² and LVF² is that Norm² does not consider the skewness of the Gaussian components. LESN is a state-of-art statistical moments-based model. LVF is the current industry standard based on SN, which serves as the baseline.

In this section, we assess LVF² with cells in Section 4.1, 4.2, and 4.3. Section 4.4 evaluates the improvement of LVF² with circuit paths. To evaluate the accuracy of binning prediction, the binning boundaries are set to $\mu \pm 3\sigma$, $\mu \pm 2\sigma$, $\mu \pm \sigma$, and μ , which lead to eight speed bins. To facilitate a thorough discussion, we use three evaluation metrics in the following experiments: bin probability, 3σ yield, and root mean square

Table 1. Scenarios Assessment among Models.

Scenarios	Binning Error Reduction (×)			
	LVF ²	Norm ²	LESN	LVF
2 Peaks	12.65	1.01	1.02	1
Multi-Peaks	29.65	7.67	10.68	1
Saddle	9.62	5.06	1.88	1
Minor Saddle	16.27	10.58	0.84	1
Kurtosis	8.63	8.16	3.43	1

error (RMSE) of CDF. Error reduction is used to normalize each metric in a unified way (12), where *baseline* is the result of LVF and *golden* is from SPICE MC simulations.

$$\text{Error Reduction} = \frac{|\text{baseline} - \text{golden}|}{|\text{result} - \text{golden}|} \quad (12)$$

4.1 PDF Fitting Assessment For Cells

From the distributions generated from cells, we select five representative non-Gaussian scenarios. Although not all the non-Gaussian distributions can be categorized into one of these scenarios, it is commonly found that a real distribution is the mixture of them. As shown in Figure 3, LVF² fits these scenarios better than the other models. Table 1 summary the binning error reduction. Below, we discuss these scenarios case by case. All the quantitative synthesis is assumed for cells, but the conclusion is valid for circuit paths as well.

- **2 Peaks:** In Figure 3(a), two prominent peaks exist due to the considerable distance between their locations and the minor standard deviations. LVF² perfectly fits this case with two peaks. A sharp edge indicates a significant skewness. As a result, Table 1 suggests that skewness is an indispensable parameter as LVF² outperforms Norm² by 12.5× binning error reduction.
- **Multi-Peaks:** The case in 3(b) is similar to 3(a), in which both peaks have significant skewness. LVF² successfully identifies the two dominant peaks and shows the highest binning error reduction, 29.65×.
- **Saddle:** The case in Figure 3(c) is characterized by two similar peaks with slight skewness and comparable standard deviations. A small standard deviation usually indicates a higher peak. LVF² can perfectly fit the distribution with an error reduction of 9.62×.

Table 2. Standard Cell Library Assessment among Models.

Std. Cell Type	Test Arcs Number [*]	Delay Binning Error Reduction [†]				Transition Binning Error Reduction [†]				Delay 3 σ -Yield Error Reduction [†]				Transition 3 σ -Yield Error Reduction [†]			
		LVF ²	Norm ²	LESN	LVF	LVF ²	Norm ²	LESN	LVF	LVF ²	Norm ²	LESN	LVF	LVF ²	Norm ²	LESN	LVF
INV	24	11.87	4.95	3.76	1	16.98	7.30	4.02	1	12.32	4.28	3.60	1	8.68	8.09	4.38	1
BUFF	21	7.62	4.64	3.61	1	5.25	4.14	5.49	1	8.25	4.15	3.99	1	5.72	3.70	5.76	1
NAND2	57	11.87	4.06	5.19	1	14.43	5.91	5.73	1	4.50	4.45	4.52	1	10.19	8.64	6.19	1
NAND3	39	6.01	3.33	6.38	1	16.77	5.54	4.57	1	4.18	4.43	4.56	1	18.17	8.30	5.23	1
NAND4	28	5.67	3.36	3.79	1	12.72	4.62	5.88	1	6.13	7.41	3.30	1	11.43	7.07	6.00	1
AND2	20	10.06	3.14	4.28	1	6.33	3.06	5.56	1	11.35	3.39	5.10	1	3.36	2.73	3.21	1
AND3	22	4.08	3.53	4.37	1	6.58	2.72	5.86	1	3.81	5.32	3.59	1	4.56	5.70	5.75	1
AND4	11	10.23	2.99	3.38	1	6.39	2.41	5.34	1	3.94	3.18	2.70	1	5.94	3.57	5.87	1
NOR2	14	5.79	2.86	3.11	1	12.61	5.47	5.82	1	3.06	2.91	2.83	1	9.26	6.29	6.48	1
NOR3	13	9.82	3.11	5.43	1	10.43	5.14	4.93	1	6.85	4.15	4.82	1	8.03	5.84	4.93	1
NOR4	25	5.50	3.53	4.25	1	10.85	5.53	5.21	1	2.78	3.75	3.12	1	6.79	3.95	4.29	1
OR2	17	5.54	3.08	7.31	1	8.16	2.94	4.94	1	3.42	3.54	3.11	1	7.72	6.56	10.37	1
OR3	12	13.56	5.05	5.72	1	8.10	4.35	6.01	1	1.77	8.44	3.73	1	4.01	4.65	8.00	1
OR4	23	8.10	4.06	5.04	1	8.98	3.91	5.34	1	4.85	7.71	2.36	1	14.06	7.35	5.07	1
XOR2	32	6.41	3.61	4.21	1	8.88	2.48	5.01	1	3.75	3.10	7.87	1	3.57	5.11	7.04	1
XOR3	49	5.80	4.14	4.28	1	7.64	3.30	5.32	1	3.62	3.02	4.31	1	4.74	5.30	5.40	1
XOR4	74	6.04	4.24	4.88	1	7.57	2.40	5.80	1	3.32	6.07	3.79	1	3.51	3.49	6.09	1
XNOR2	30	7.68	3.40	4.74	1	11.38	3.88	5.22	1	4.05	3.54	3.16	1	5.21	6.09	7.54	1
XNOR3	48	5.99	7.42	3.69	1	10.60	2.92	5.69	1	5.83	2.50	4.64	1	11.30	3.25	7.38	1
XNOR4	45	12.06	5.42	4.59	1	8.46	2.63	6.03	1	3.28	3.46	4.98	1	7.18	7.72	7.50	1
MUX2	31	6.39	4.91	6.07	1	7.72	4.18	5.04	1	2.70	3.34	3.02	1	5.96	5.52	4.63	1
MUX3	40	5.19	2.90	3.75	1	10.39	3.54	5.93	1	3.40	3.63	3.90	1	6.22	3.41	6.46	1
MUX4	40	8.58	3.27	3.88	1	9.23	3.40	5.68	1	4.06	3.60	4.88	1	6.44	3.11	6.44	1
FA	25	6.01	4.27	3.50	1	6.30	3.38	7.70	1	5.91	2.80	4.02	1	3.62	3.51	7.56	1
HA	7	7.63	2.96	4.77	1	6.34	2.13	6.63	1	2.73	2.67	5.44	1	3.80	7.16	11.02	1
Overall	747	7.74	3.93	4.56	1	9.56	3.89	5.55	1	4.79	4.19	4.05	1	7.18	5.44	6.34	1

^{*} Each timing arc contains 8×8 delay and 8×8 transition distributions, and each distribution contains 50k samples; [†] In the unit of improvement multiples (×).

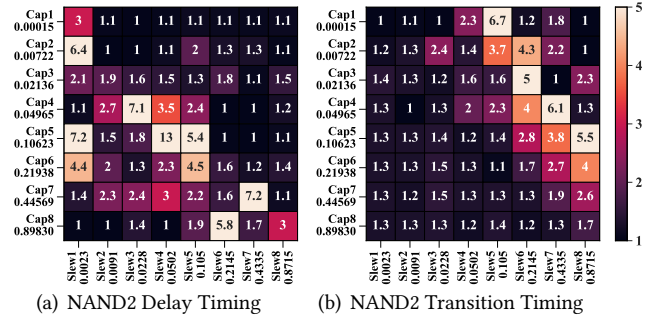
- **Minor Saddle:** The case in Figure 3(d) features one Gaussian dominating another, and the two Gaussians having deviated standard deviations. In this case, LVF² also significantly improves the fitting result, reducing the binning error by 16.27×.
- **Kurtosis:** This case occurs when the distribution exhibits a high level of kurtosis. In 3(e), LVF² considers two peaks with similar centers but different weights and deviations. This leads to a high kurtosis. We can see that both LVF² and Norm² significantly improve the binning error. It is worth noting that even without considering the skewness, the Norm² model can still accurately fit distributions with large kurtosis.

4.2 Standard Cell Characterization

Our benchmark encompasses 25 distinct standard combinational cells, exploring different driving strengths and different timing arcs among each cell type. Considering 8×8 slew-load pairs, each timing arc has 64 conditions to simulate and is characterized into 64 transition time and 64 delay distributions. We employed LVF², Norm², LESN, and LVF models to fit the distributions and evaluated the results with binning and 3 σ -yield error reduction.

Table 2 summarizes the experimental results. LVF², Norm², and LESN demonstrate great improvements in binning and 3 σ -yield prediction, compared to LVF. Since LESN focuses more on the tail fitting, the results show its advantages in estimating the 3 σ -yield. Norm² address non-Gaussian distributions with two Gaussians, which is more observed in the PDF of transition times. Hence, we see a more significant improvement by Norm² in transition than delay.

In most cases, LVF² performs the best. To summarize, LVF² demonstrates its strengths and robustness across all ranks, with improvements of 7.74× and 9.56× in delay and transition binning predictions, and 4.79× and 7.18× in the 3 σ -yield estimations of delay and transition, respectively.



(a) NAND2 Delay Timing

(b) NAND2 Transition Timing

Figure 4. Accuracy Pattern: LVF² CDF RMSE Reduction in 8×8 Timing Tables.

4.3 Accuracy Patterns For LVF²

We explored the patterns of multi-Gaussian distributions occurring in the transition and the delay. Each table is indexed with the input slew (ns) and output load (pf), which increase non-linearly and combine into 8×8 table entries. We use LVF²'s CDF RMSE reduction as an indicator. The score reflects the merit of LVF² fitting, which can quantify the degree of multi-Gaussian phenomenon of a timing distribution.

Figure 4 illustrates a well-characterized accuracy pattern, which was found to be a general case in our experiments. We observed that the multi-Gaussian phenomenon appears in a diagonal pattern. We assume that the confrontation of different variations causes the regularity, and the variation is influenced by the slew-load combination. With the assumption, we see that when the multi-Gaussian is apparent at position (i, j) of the table, which means there are two (or more) variations, denoted as \mathcal{A} and \mathcal{B} , evenly matched against each other; the multi-Gaussian is weaker at $(i \pm 1, j)$ and $(i, j \pm 1)$, which means one of \mathcal{A} and \mathcal{B} dominates when only one of slew and load changes and the other is fixed; and the multi-Gaussian phenomenon appears again at $(i \pm 1, j \pm 1)$,

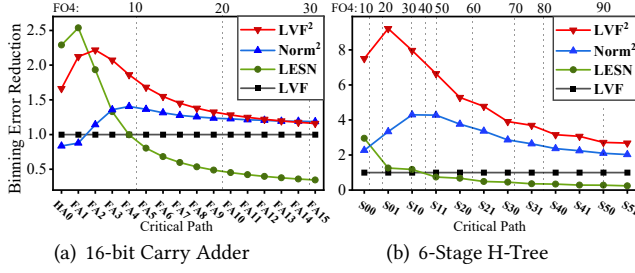


Figure 5. Comparison of Binning Error Reduction along Two Circuit Critical Path.

which reflects \mathcal{A} and \mathcal{B} are back to the confrontation state when slew and load increase or decrease together.

The regularity of the diagonal patterns leads to the conclusion that the multi-Gaussian phenomenon is not coincidental. In many cases, it is imperative to upgrade LVF to LVF^2 . Furthermore, LVF^2 can serve as a helpful tool to quantify the multi-Gaussian phenomenon, and then assist us in exploring the reasons behind multi-Gaussian in timing distributions in future research. Meanwhile, the accuracy pattern serves as an excellent prediction of PDF shapes, which can be useful for developing efficient methods to characterize statistical timing behaviors across multiple slew-load pairs.

4.4 Validation For Paths

According to Section 3.4, timing distribution becomes Gaussian along propagation, decaying the benefits of LVF^2 and other non-Gaussian models. However, modern VLSI design usually adopts shallow logic depths to optimize energy efficiency, and an optimal path with delay can be as short as 6 to 8 FO4 [17]. We validated LVF^2 for SSTA flow. The first benchmark is a 16-bit carry adder with a typical structure and a critical path delay of 30-FO4. The second is a 6-stage H-tree with a delay of 95-FO4, in which each stage consists of 2 buffer cells and metal wires described with the Π -model. Although there are some formularized non-Gaussian SSTA approaches [18] [19], to avoid losing generality, the block-based SSTA [20] is used to calculate the propagation timing distribution of each model along the critical path. The golden is obtained based on MC simulation of timing critical paths.

The SSTA propagation result shown in Figure 5 agrees with the CLT and $O(1/\sqrt{n})$ convergence rate discussed in Section 3.4. LVF^2 and Norm^2 show their strengths in the first few stages, but the improvements gradually converge towards one as the stage increases. In the 16-bit carry adder benchmark, LVF^2 improves by 2 \times compared to LVF at 8-FO4, and 1.15 \times at the last cell of the path. For the H-tree benchmark, LVF^2 improves 8 \times at 8-FO4, and 2.68 \times at the end. The H-tree is deeper in path but slower in convergence. We believe it is because the accumulation of random variables slows down due to the simple structure of buffer cells compared to the former. The results of LFSN did not meet our expectations. A possible reason is an error introduced during moment matching, which accumulates during propagation and eventually results in an inaccurate timing distribution.

5 Conclusions

This paper proposes a novel statistical timing model, LVF^2 , to accurately address the multiple Gaussian components and asymmetry in timing distributions. LVF^2 combines two weighted SN distributions and can be solved by the EM algorithm. Experimental results based on TSMC 22nm standard cells prove LVF^2 's capability to significantly improve accuracy. Compared to the industry-standard LVF, LVF^2 achieves error reductions of 7.74 \times and 9.56 \times in delay and transition binning, and 4.79 \times and 7.18 \times in the 3σ -yield of delay and transition, respectively. With a real circuit, SSTA experiments show a 2 \times binning error reduction with a typical path of 8-FO4 delay. Finally, the paper discusses the accuracy pattern of multi-Gaussian characteristics occurring in the standard cell library. We anticipate that the accuracy pattern will become an essential indicator in future research exploring non-Gaussian timing distributions. Meanwhile, assuming such an accuracy pattern can provide significant insight to speed up the statistical characterization that includes MC simulations across multiple slew-load pairs.

References

- [1] B. D. Cory et al. Speed Binning with path Delay Test in 150-nm Technology. *IEEE Design & Test of Computers*, 20(5):41–45, 2003.
- [2] H. Mahmoodi et al. Estimation of Delay Variations due to Random-dopant Fluctuations in Nanoscale CMOS Circuits. *JSSC*, 40(9), 2005.
- [3] K. Chopra et al. A New Statistical Max Operation for Propagating Skewness in Statistical Timing Analysis. In *IEEE/ACM ICCAD*, 2006.
- [4] Synopsys. Liberty Release Notes Version 2017.06, 2017.
- [5] S. Keller et al. A Compact Transregional Model for Digital CMOS circuits operating Near Threshold. *IEEE TVLSI*, 22(10), 2014.
- [6] H. A. Balef et al. All-Region Statistical model for delay variation Based on Log-Skew-Normal Distribution. *IEEE TCAD*, 35(9):1503–1508, 2016.
- [7] L. Jin et al. A Statistical Cell Delay Model for Estimating the 3σ Delay by Matching Kurtosis. *IEEE TCAS II*, 69(6):2932–2936, 2022.
- [8] L. Jin et al. A Novel Delay Calibration Method Considering Interaction between Cells and Wires. In *DATE*, pages 1–6, 2023.
- [9] J. Hossain et al. Multi-peak Gaussian Fit Applicability to Wind Speed Distribution. *Renew. Sustain. Energy Rev.*, 34:483–490, 2014.
- [10] H. Takahashi et al. A Gaussian Mixture Model for Statistical Timing Analysis. In *46th ACM/IEEE DAC*, pages 110–115, 2009.
- [11] A. Azzalini et al. Statistical applications of the multivariate skew normal distribution. *JRSS. Series B (Methodological)*, 61(3), 1999.
- [12] A. P. Dempster et al. Maximum Likelihood from Incomplete Data via the EM Algorithm. *JRSS. Series B (Methodological)*, 39(1):1–38, 1977.
- [13] J. A. Hartigan and M. A. Wong. A k-means clustering algorithm. *JRSS. Series C (Applied Statistics)*, 28(1):100–108, 1979.
- [14] B. C. Arnold et al. The nontruncated marginal of a truncated bivariate normal distribution. *Psychometrika*, 58:471–488, 09 1993.
- [15] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, January 1968.
- [16] A. C. Berry. The accuracy of the gaussian approximation to the sum of independent variates. *Trans. of the AMS*, 49(1):122–136, 1941.
- [17] M. S. Hrishikesh et al. The Optimal Logic Depth per Pipeline Stage is 6 to 8 FO4 Inverter delays. In *29th ISCA*, pages 14–24, 2002.
- [18] L. Cheng et al. Non-linear statistical static timing analysis for non-gaussian variation sources. In *ACM/IEEE DAC*, pages 250–255, 2007.
- [19] Cheng. L. et al. Non-Gaussian Statistical Timing Analysis using Second-order Polynomial Fitting. In *ASP-DAC*, pages 298–303, 2008.
- [20] A. Devgan and C. Kashyap. Block-based Static Timing Analysis with Uncertainty. In *2003 IEEE/ACM ICCAD*, pages 607–614, 2003.