Yield-driven Clock Skew Scheduling Based on Generalized Extreme Value Distribution

30th Asia and South Pacific Design Automation Conference

Kaixiang Zhu, Wai-Shing Luk* and Lingli Wang
State Key Laboratory of Integrated Chips and Systems
Fudan University, China
kxzhu23@m.fudan.edu.cn, {luk, llwang}@fudan.edu.cn

Jan. 21st, 2025

1 Useful Skew Design: Why and Why not?

- Adjust clock skew to enhance IC performance or robustness.

Good👍:

If you do it right,

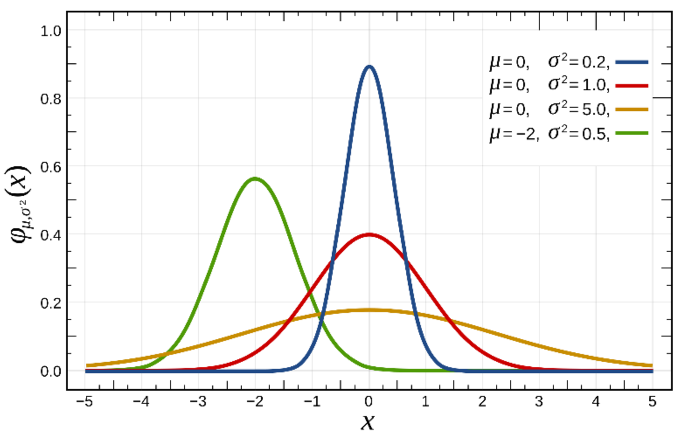
- spend less time struggling about timing, or
- get better chip performance or timing yield.

Bad 👎:

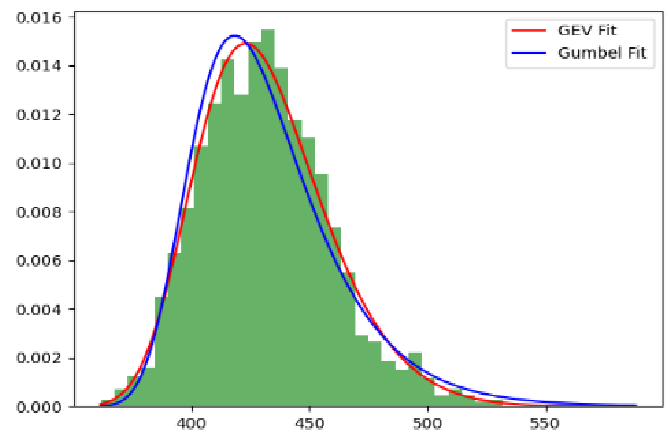
- Need more engineer training.
- Balanced clock-trees are harder to build.
- Don't know how to handle **process variation**, multi-corner multi-mode, ..., etc.

1 Useful Skew Design: Why and Why not?

- 65nm and below, non-Gaussian distribution path delay.
  - Skewed and long-tailed
- Traditional Statistical Clock Skew Scheduling (CSS) relies on Gaussian distribution.
- Real-world path delays: asymmetric, making Gaussian-based methods insufficient.



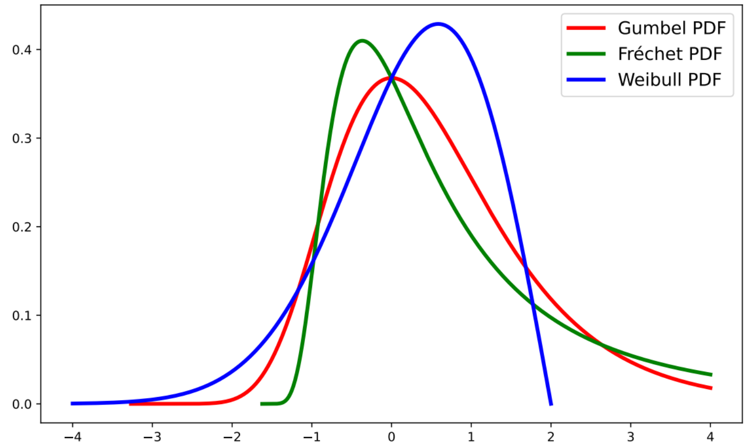
Gaussian distribution.



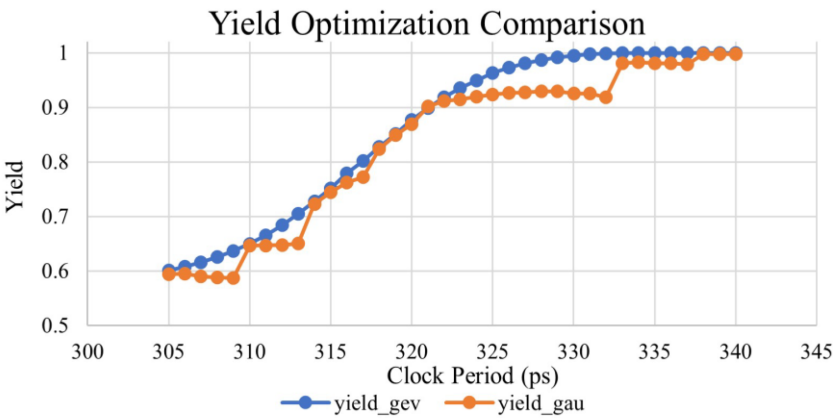
Real path delay distribution.

2 Proposed Solution:

- Generalized Extreme Value (GEV) Distribution:
  - Flexible, asymmetric distribution from extreme value theory.
  - Capture skewness and asymmetry in path delay distributions effectively.
- Advantages Over Gaussian:
  - High accuracy in approximating delay distributions.
  - Better modeling for timing yield-driven optimization.



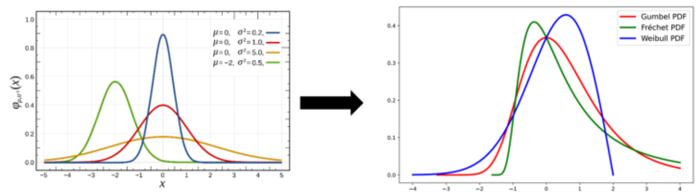
GEV distribution.



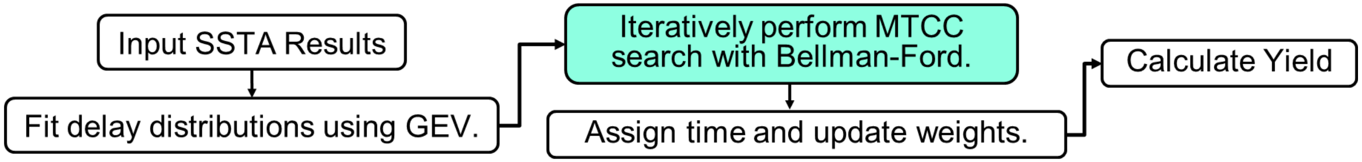
Better timing yield.

2 Contributions of Our Work:

- **Introduce GEV distribution**
  - Replace Gaussian distribution with GEV for clock skew scheduling.



- Propose a framework for **timing yield-driven CSS**

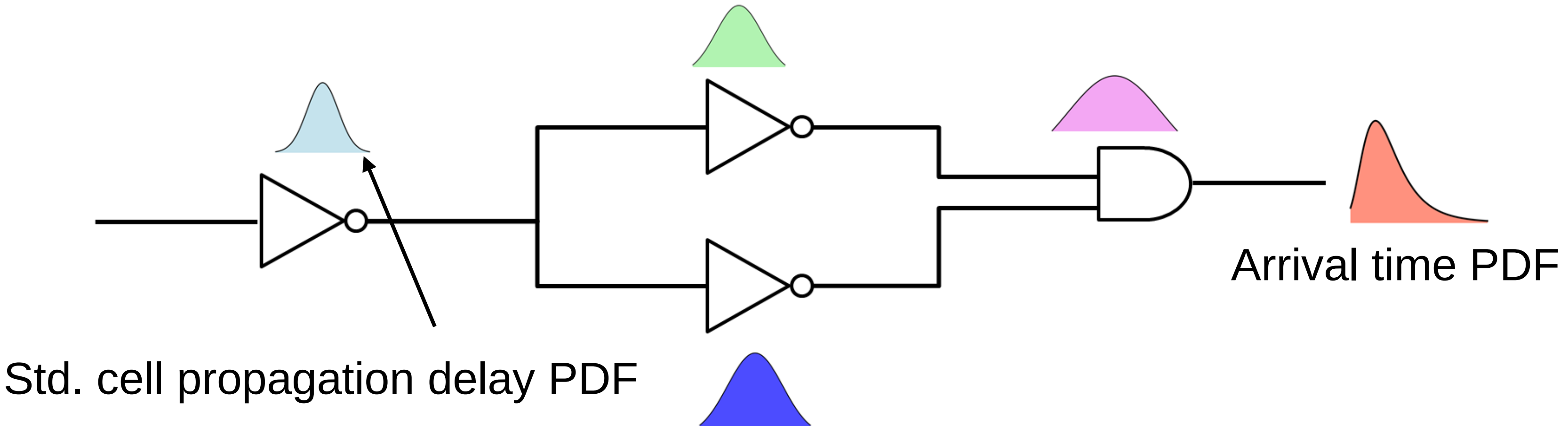


- Evaluate Parameter Estimation Methods
  - Compare MLE, MoM, and L-Moments for fitting accuracy.
- Achieves **superior timing yield**
  - Improve timing yield by 8% on benchmark circuits.

3 Preliminaries

Statistical Static Timing Analysis (SSTA)

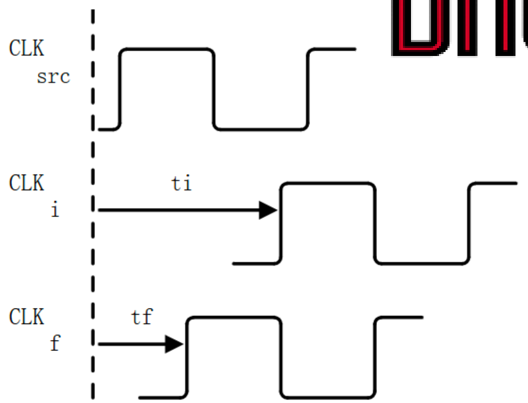
- SSTA: an extension of traditional STA, accounts for process variations.
- Modern STA tools based on Gaussian distribution
  - Provide 3-sigma statistics for slacks/path delays (POCV).
- However, the full probability density function (PDF) information are not available.



SSTA example.

3 Preliminaries

- $T_{skew}(i,j) = t_i - t_j$, where
  - t_i: clock signal delay at the initial register
  - t_j: clock signal delay at the final register



- Setup time constraint:

$$T_{skew}(i,j) \leq T_{cp} - D_{ij} - T_{setup}$$

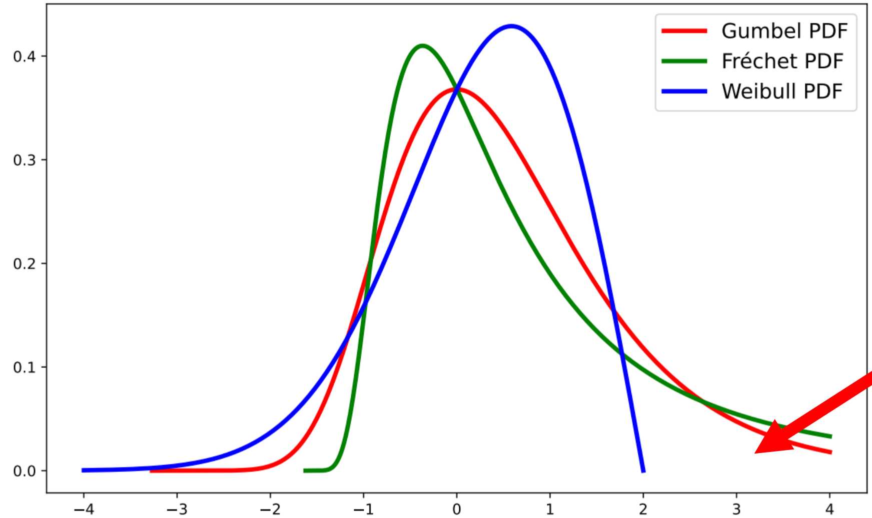
  While this constraint destroyed, cycle time violation occurs.
- Hold time constraint

$$T_{skew}(i,j) \geq T_{hold} - d_{ij}$$

  While this constraint destroyed, race condition occurs.
- Primary goal of Yield-driven Clock Skew Scheduling: minimize the timing yield loss.
  - Timing Yield = (functional correct times) / sample number * 100%

4 Generalized Extreme Value

- A family of continuous probability distributions developed in extreme value theory.
- Model the maximum or minimum of a set of random variables.
- Defined by three parameters:
  - Location parameter (μ)
  - Scale parameter ($\sigma > 0$)
  - Shape parameter (ξ)
- Three types:
  - Gumbel (ξ=0)
  - Fréchet (ξ>0)
  - Weibull (ξ<0)
- Gaussian has only two parameters:
  - Location parameter (μ)
  - Scale parameter (σ)

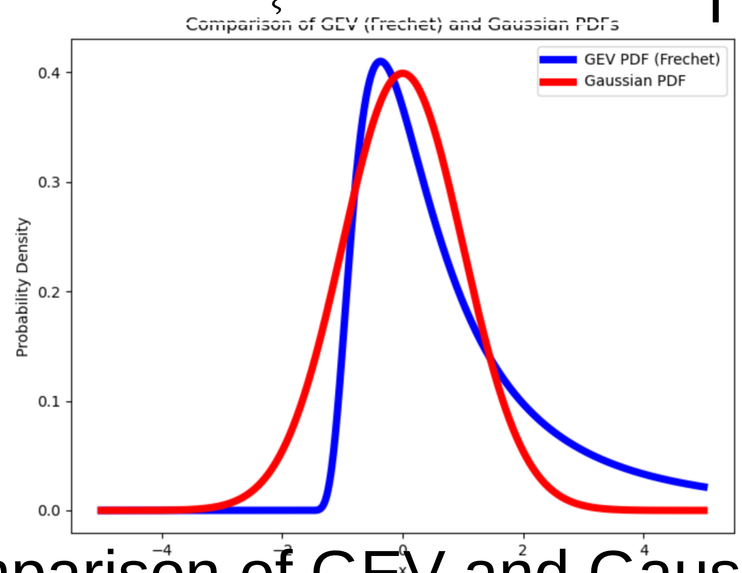


GEV PDF

4 Generalized Extreme Value

- **PDF, CDF Comparison**

	GEV	Gaussian
Probability Density Function (PDF)	$f(x) = \frac{1}{\sigma}[1+\xi\left(\frac{x-\mu}{\sigma}\right)]^{-1/\xi^{\xi+1}} e^{-[1+\xi\left(\frac{x-\mu}{\sigma}\right)]^{-1/\xi}}$ Or $f(x) = \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right)e^{-\exp\left(-\frac{x-\mu}{\sigma}\right)}$, when $\xi = 0$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Cumulative Distribution Function (CDF)	$F(x) = e^{-t(x)}$, $x \in [\frac{\mu-\sigma}{\xi}, +\infty)$ when $\xi > 0$, $x \in (-\infty, +\infty)$ when $\xi = 0$, $x \in (-\infty, \frac{\mu-\sigma}{\xi}]$ when $\xi < 0$.	$F(x) = \frac{1}{2}[1+\mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)]$



Comparison of GEV and Gaussian PDFs

4 Generalized Extreme Value

Parameter Estimation:

- Three estimation techniques:
- Scipy:
  - **MLE:** Maximum likelihood estimation
    - Accurate for distributions with well-defined likelihoods; suitable for small datasets.
  - **Mom**: Method of moments
    - Simple and computationally efficient; faster in large-scale scenarios.
- scikit-extreme
  - **L-Moments**: Linear combinations of order statistics
    - Robust for extreme values; better for capturing tail behavior in skewed distributions.

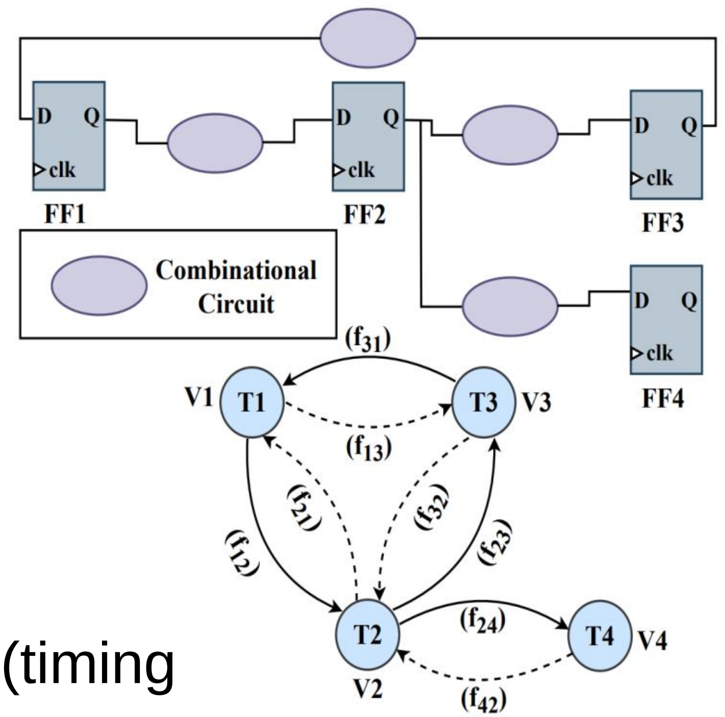
5 Yield-driven Optimization

Timing Constraint Graph

- Create a graph *G* by:
  - h-*edge:* $-(T_{hold} - d_{ij})$ from $\mathrm{FF_j}$ to $\mathrm{FF_i}$
  - s-edge: $T_{cp} - D_{ij} - T_{setup}$ from $\mathrm{FF_i}$ to $\mathrm{FF_j}$.
- Sum of clock skews of any cycle C in *G* is greater than 0.

$$\sum_{e_{ij} \in C} T_{skew}(i,j) \geq 0$$

- If a negative cycle exits, timing violation (timing failure) occurs.



Timing const. graph.

5 Yield-driven Optimization

- Max-Min Formulation:
  - $\min\{\max\{Pr\{t_i - t_j \le \widetilde{W_{ij}}\}\}\}$
  - Equivalent to $\max\{\min\{Pr\{t_i - t_j \le \widetilde{W_{ij}}\}\}\}$
  - No need for correlation information between $\widetilde{W_{ij}}$
- Equivalent to:

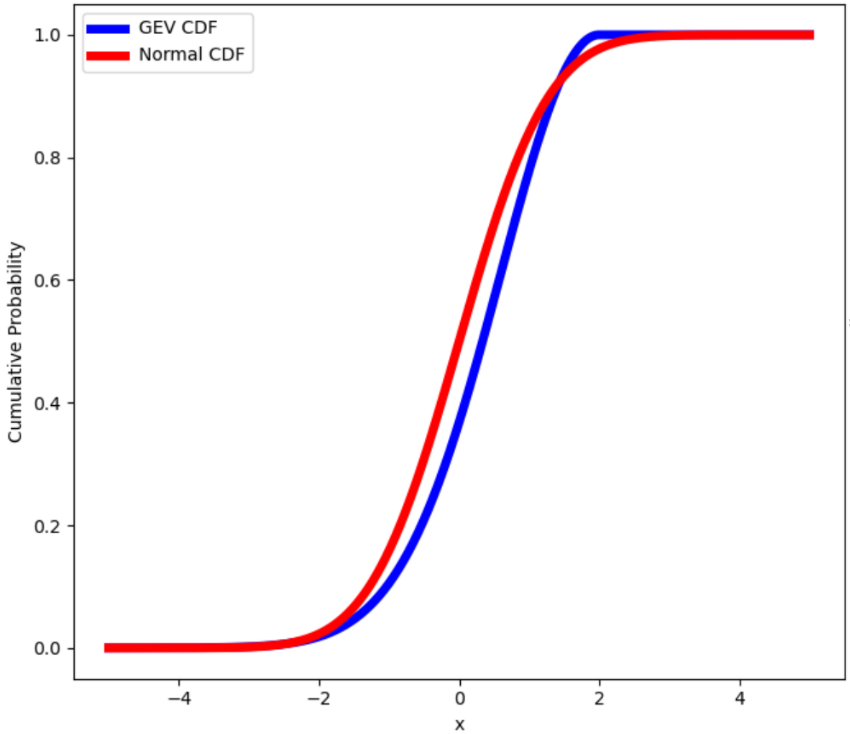
$$\begin{array}{ll} \text{maximum} & \beta \\ \text{subject to} & Pr\{t_i - t_j \le T_{CP} - \widetilde{D_{ij}}\} \ge \beta \\ & Pr\{t_j - t_i \le \widetilde{H_{ij}}\} \ge \beta \end{array}$$

  - β: timing satisfaction probability
  - $\widetilde{H_{ij}}$and $\widetilde{D_{ij}}$ represent the minimum and maximum path delays.

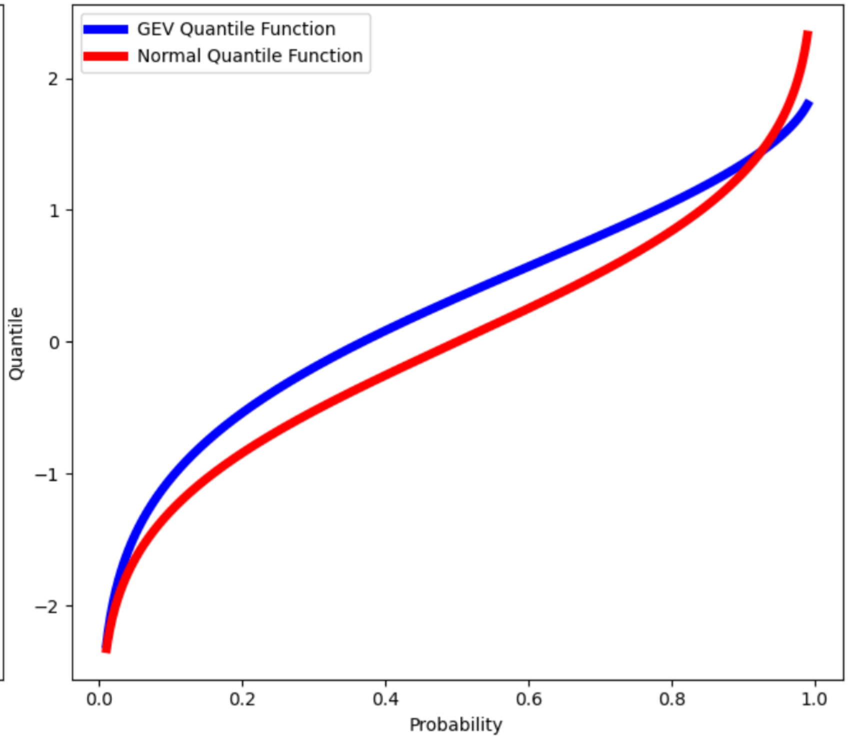
5 Yield-driven Optimization

- Use quantile function (Inverse of the CDF function)

$$
\begin{aligned}
\text{maximum} \quad & \beta \\
\text{subject to} \quad & t_i - t_j \le T_{CP} - \Phi^{-1}_{D_{ij}}(\beta) \\
& t_j - t_i \le \Phi^{-1}_{H_{ij}}(1-\beta)
\end{aligned}
$$



CDF function



Quantile function

5 Yield-driven Optimization

		Gaussian
Quantile Function		$\Phi^{-1}(\beta) = \mu + \sigma\sqrt{2}erf^{-1}(2\beta - 1))$
Reduce to	maximum	β
	subject to	$t_i - t_j \le T_{CP} - (\mu_{D_{ij}} + \sigma_{D_{ij}}\sqrt{2}erf^{-1}(2\beta - 1))$ $t_j - t_i \le \mu_{H_{ij}} + \sigma_{H_{ij}}\sqrt{2}erf^{-1}(2(1-\beta) - 1)$
Linearization	maximum	$\beta' = erf^{-1}(2\beta - 1))$
	subject to	$t_i - t_j \le T_{CP} - \mu_{D_{ij}} - \sigma_{D_{ij}}\beta'$ $t_j - t_i \le \mu_{H_{ij}} - \sigma_{H_{ij}}\beta'$

- Equivalent to the minimum cost-to-time ratio cycle (linear).
- However, actual path delay distributions are non-Gaussian .

5 Yield-driven Optimization

	Gumbel ($\xi = 0$) in GEV
Quantile Function	$\Phi^{-1}(\beta) = \mu - \sigma \cdot \ln(-\ln\beta)$
Reduce to	maximum β subject to $t_i - t_j \leq T_{CP} - (\mu_{D_{ij}} - \sigma_{D_{ij}} \ln(-\ln\beta))$ $t_j - t_i \leq \mu_{H_{ij}} - \sigma_{H_{ij}} \ln(-\ln\beta))$
Linearization	maximum $\beta' = -\ln(-\ln\beta)$. subject to $t_i - t_j \leq T_{CP} - \mu_{D_{ij}} - \sigma_{D_{ij}} \beta'$ $t_j - t_i \leq \mu_{H_{ij}} + \sigma_{H_{ij}} \beta'$

- Also equivalent to the minimum cost-to-time ratio cycle (linear).

5 Yield-driven Optimization

	GEV ($\xi \neq 0$)
Quantile Function	$\Phi^{-1}(\beta) = \mu + \frac{\sigma}{\xi}((-\ln\beta)^{-\xi} - 1)$
Reduce to	maximum β subject to $t_i - t_j \leq T_{CP} - (\mu_{D_{ij}} + \frac{\sigma_{D_{ij}}}{\xi_{D_{ij}}}((-\ln\beta)^{-\xi_{D_{ij}}} - 1)$ $t_j - t_i \leq \mu_{H_{ij}} + \frac{\sigma_{H_{ij}}}{\xi_{H_{ij}}}((-\ln\beta)^{-\xi_{H_{ij}}} - 1)$
Linearization	**Non-linear**

- Non-linear, can be solved using numerical or binary search method.

5 Yield-driven Optimization

- Consider $\sum_{e_{ij}\in C} T_{skew}(i,j) = 0$

Gaussian

$$\text{maximum} \quad \beta'$$

$$\text{subject to} \quad \sum_{e_{ij}\in C} t_i - t_i = 0 \le \sum_{e_{ij}\in C} T_{CP} - \mu_{D_{ij}} - \sigma_{D_{ij}}\beta'$$

$$\sum_{e_{ij}\in C} t_j - t_i = 0 \le \sum_{e_{ij}\in C} \mu_{H_{ij}} - \sigma_{H_{ij}}\beta'$$

Gumbel ($\xi = 0$)

$$\text{maximum} \quad \beta'$$

$$\text{subject to} \quad \sum_{e_{ij}\in C} t_i - t_i = 0 \le \sum_{e_{ij}\in C} T_{CP} - \mu_{D_{ij}} - \sigma_{D_{ij}}\beta'$$

$$\sum_{e_{ij}\in C} t_j - t_i = 0 \le \sum_{e_{ij}\in C} \mu_{H_{ij}} + \sigma_{H_{ij}}\beta'$$

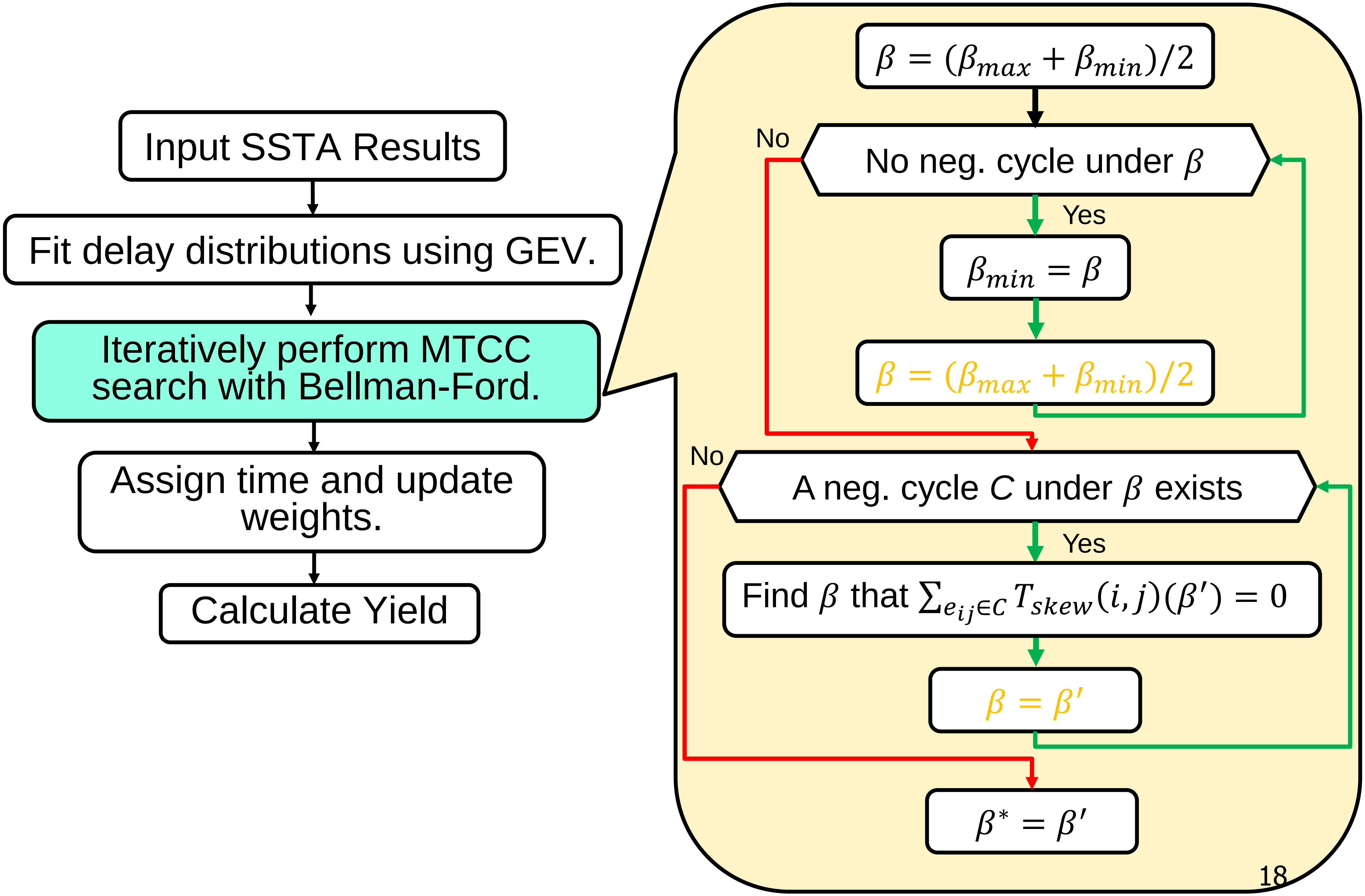
GEV ($\xi \neq 0$)

$$\text{maximum} \quad \beta$$

$$\text{subject to} \quad \sum_{e_{ij}\in C} t_i - t_i = 0 \le \sum_{e_{ij}\in C} T_{CP} - (\mu_{D_{ij}} + \frac{\sigma_{D_{ij}}}{\xi_{D_{ij}}}((-\ln\beta)^{-\xi_{D_{ij}}} - 1)$$

$$\sum_{e_{ij}\in C} t_j - t_i = 0 \le \sum_{e_{ij}\in C} \mu_{H_{ij}} + \frac{\sigma_{H_{ij}}}{\xi_{H_{ij}}}((-\ln\beta)^{-\xi_{H_{ij}}} - 1)$$

6 Yield-Driven CSS Algorithm



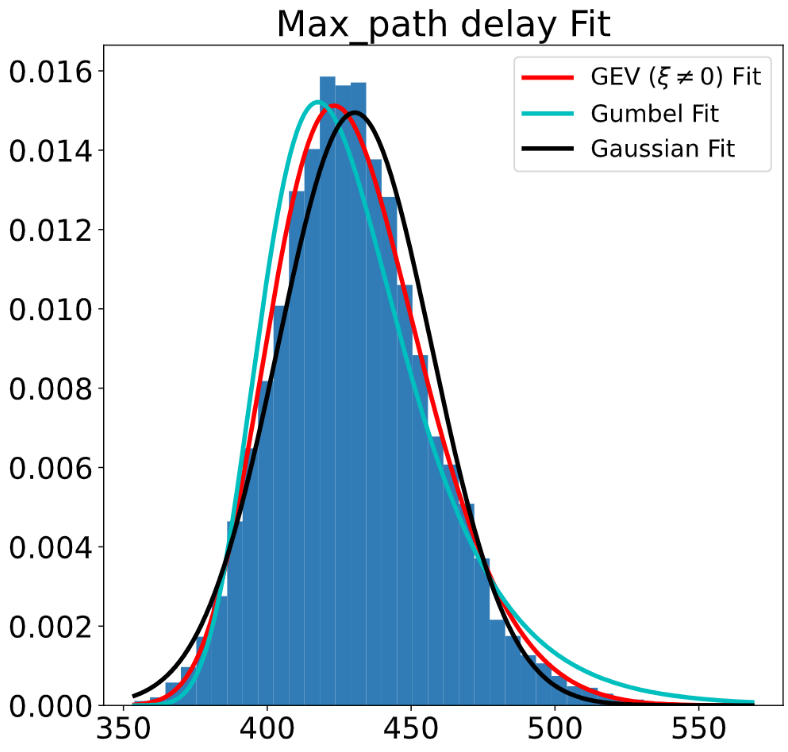
7 Experimental setup

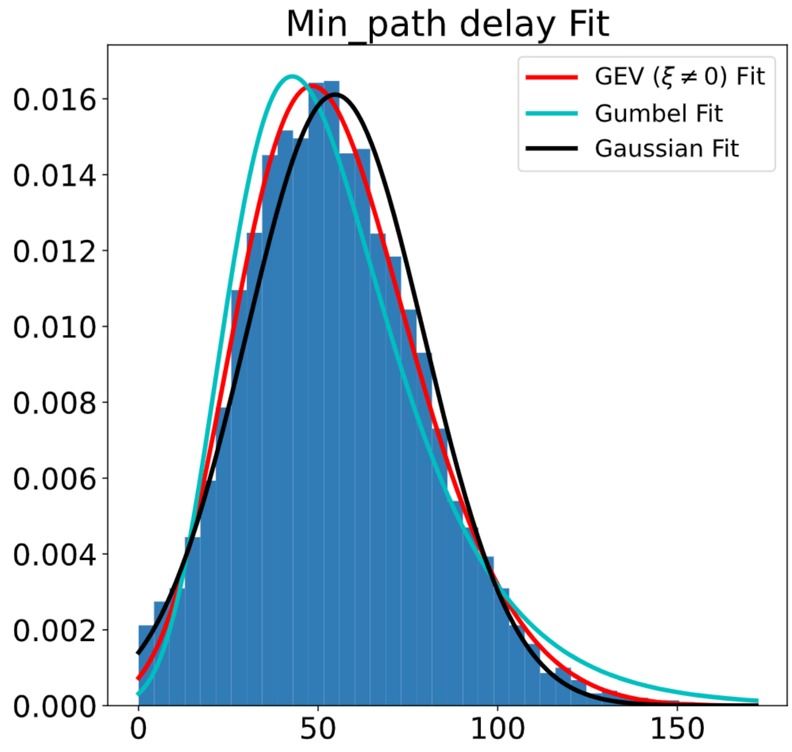
- Benchmark: OpenTimer
- Process library: 45nm
- Experimental platform: Customized OpenSTA
  - Each cell is given a delay that follows a Gaussian distribution
- Fitting SSTA path delay results with different probability distributions
  - Kolmogorov-Smirnov test is used to measure the results
- Yield-driven CSS is performed with different probability distributions.

8 Experimental Result

Fit comparison with the Gaussian Distribution

- GEV ($\xi \neq 0$) (red) best approximate
  - Horizontal axis represents delay (*ps*)
  - Vertical axis shows the probability distribution

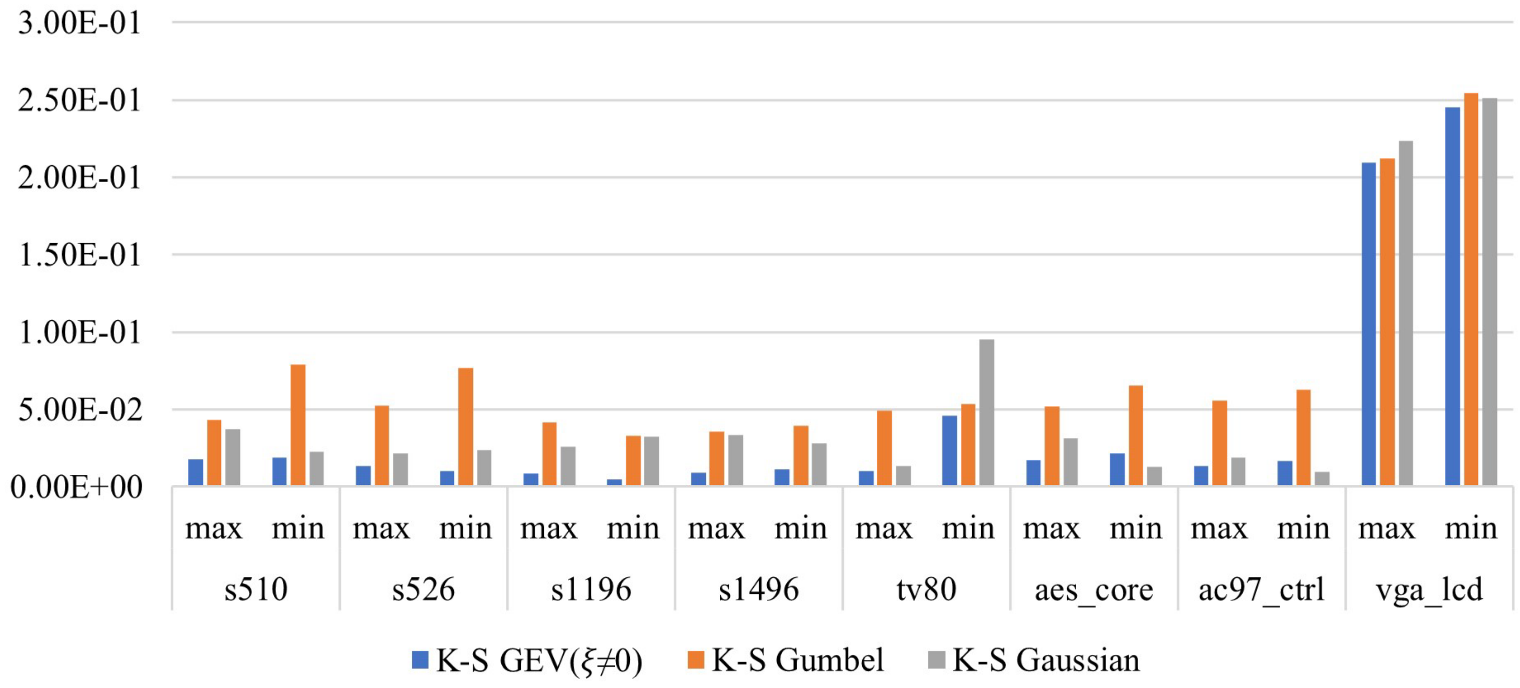




8 Experimental Result

KS Statistic Comparison with the Gau. Distribution

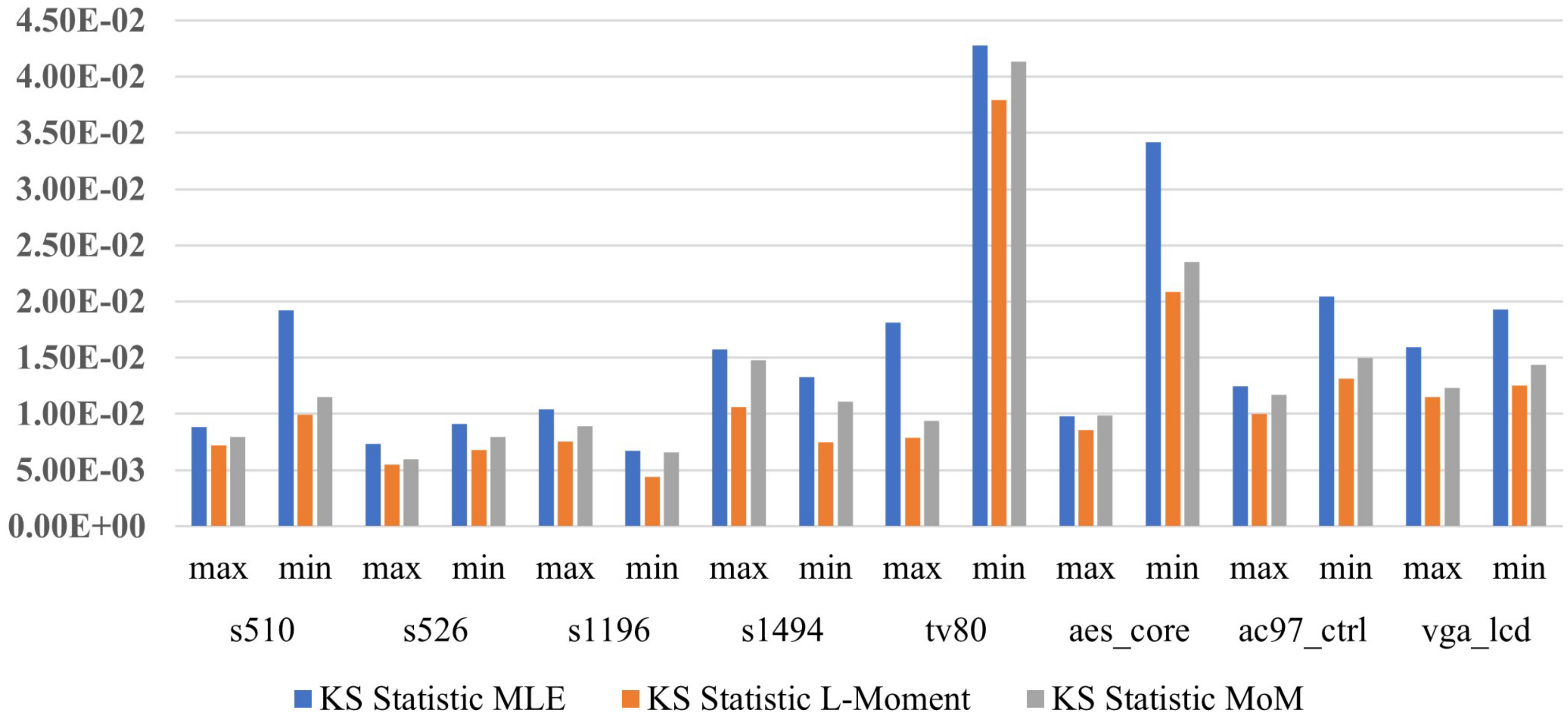
- Lower KS statistic indicates a better result.
- GEV ($\xi \neq 0$) is the best👍, 40% lower than Gaussian.
- Gaussian is the worst👎.



8 Experimental Result

KS Statistic Comparison for Different Fit Methods

- No significant difference between three methods.



8 Experimental Result

Yield-driven CSS Experiment

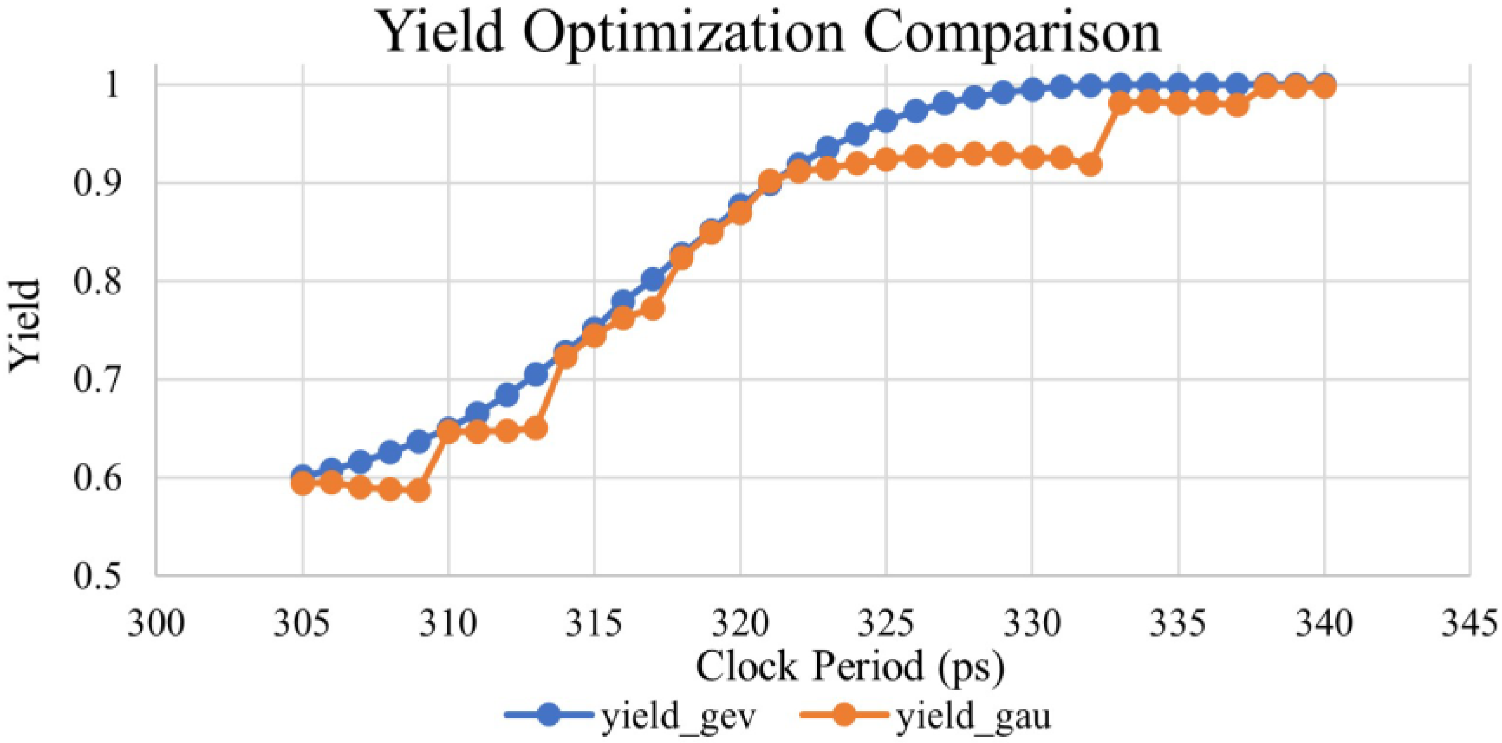
- Timing yield: calculated by 10,000 Monte Carlo simulations.
- Improve 8% on average.

Benchmark	#Constr.	Tcp(ps)	Timing Yield(%) GEV	Gaussian	Improv.
s510	34	291	89.0	55.7	33.3
s526	118	270	96.3	95.4	0.9
s1196	110	299	94.2	93.2	1.0
s1494	30	313	99.7	92.2	7.5
gcd	708	370	85.6	82.2	3.4
tv_80	6200	540	91.2	88.7	2.5

8 Experimental Result

Yield-driven CSS Experiment

- Experiment on s1494.
  - Horizontal axis represents clock period (*ps*)
  - Vertical axis shows the timing yield.
- Gaussian based method shows discontinuities.



9 Conclusion

- Use the **GEV** distribution to more accurately fit the results of SSTA.
- **Timing yield-driven CSS** based on GEV is proposed.
- Experimental results show that the GEV distribution can more accurately approximate the CDF of path delays.
- CSS based on GEV produces **superior timing yield** compared to that based on Gaussian.

Thanks!

Q & A 🙋‍♀️