A Time Efficient Comprehensive Model of Approximate Multipliers for Design Space Exploration

ARITH 2024

Ziying Cui, Ke Chen, Bi Wu, Chenggang Yan, Yu Gong and Weiqiang Liu

College of Integrated Circuits

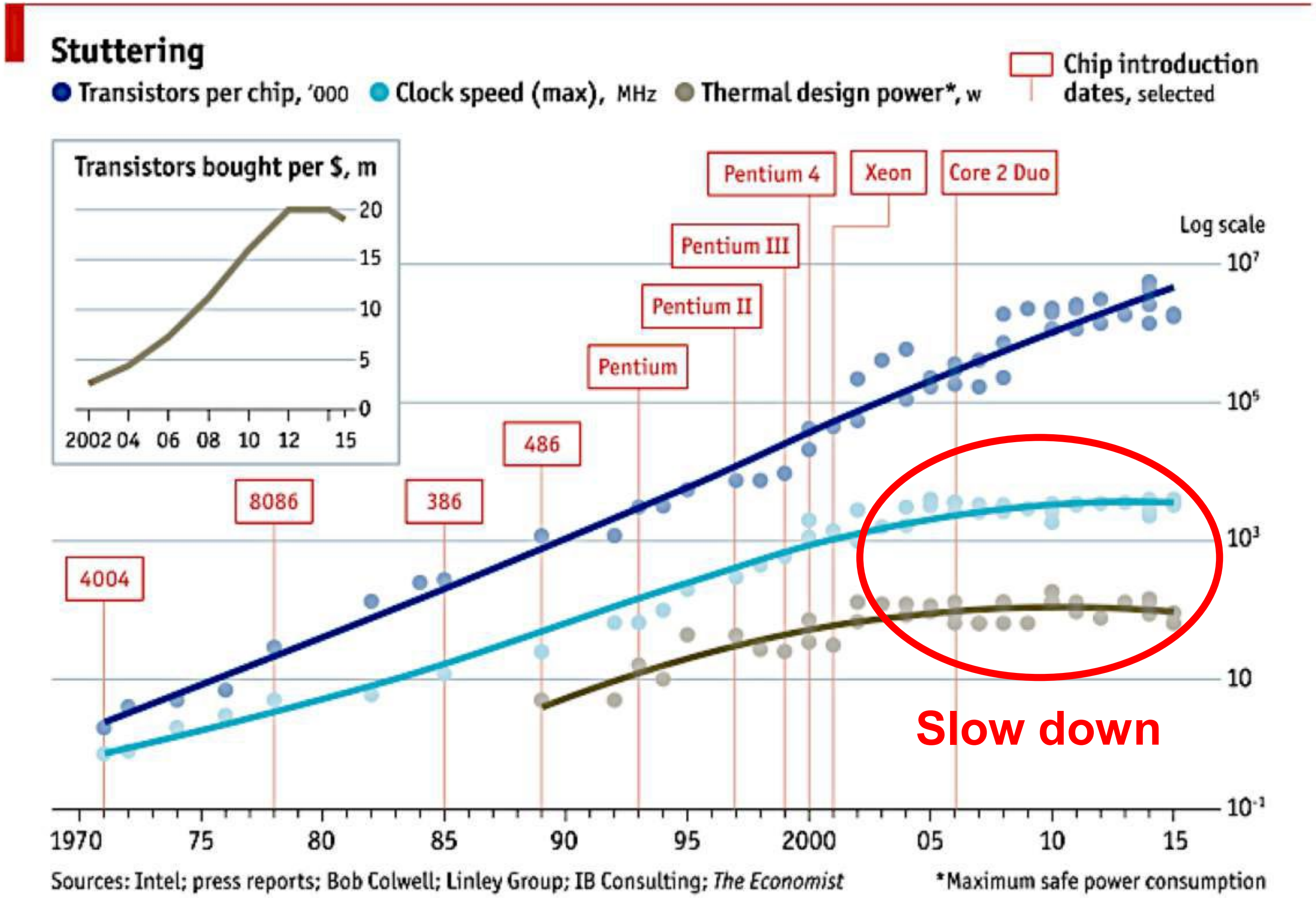
Nanjing University of Aeronautics and Astronautics

Málaga, Spain, June 10-12, 2024

Content

- Introduction and Motivation
- Proposed Comprehensive Model
- Evaluation and Analysis
- Verification
- Conclusion

Introduction



Demise of Moore's Law and Dennard scaling Law

- **Performance improvement**
- **Power reduction**

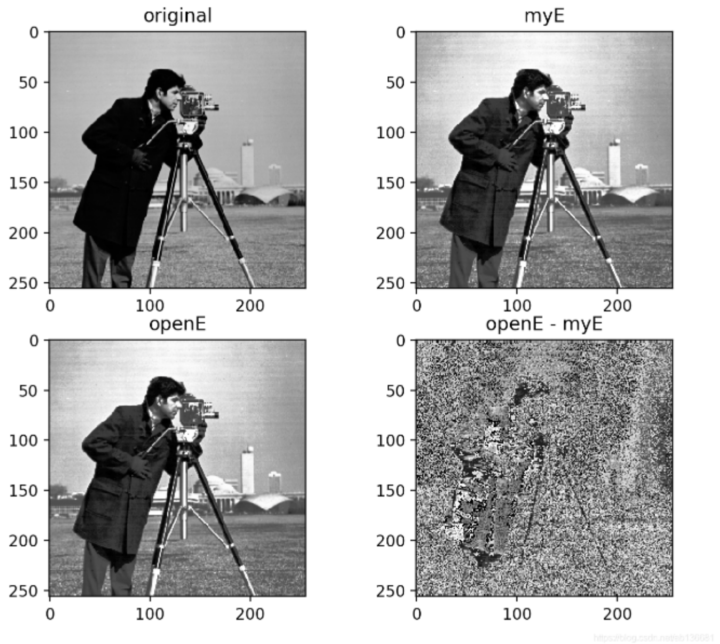
Introduction

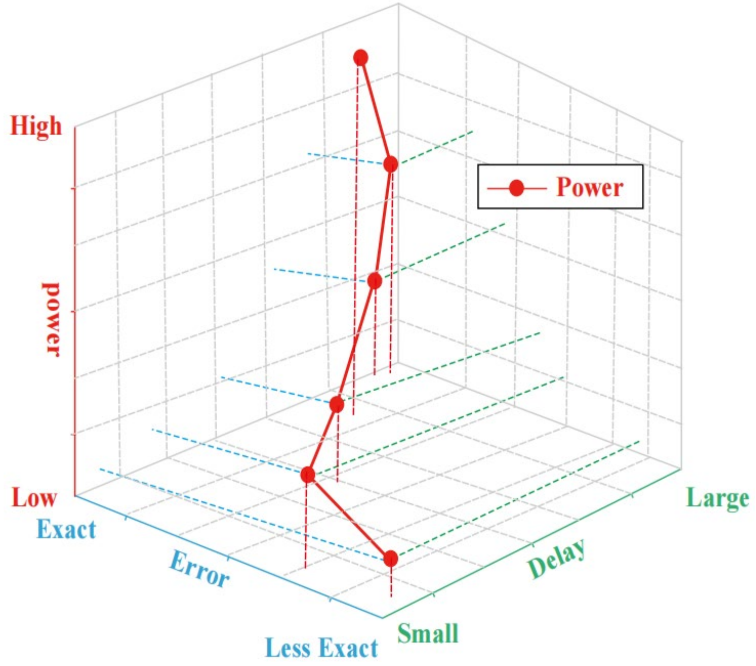
Intrinsic fault tolerance → Approximate Computing → Approximate Multipliers

Adders, MAC, ……

human perception / noise floor

Trade-off performance, power, and error





Fast and accurate model for Design Space Exploration

Motivation

- Runtime experiencing exponential growth with input bit-width

Exhaustive Approach

- Unknown iteration times
- Regardless of mechanics

Monte Carlo

- Non-universal
- Computationally intensive

Analytical model

- Specialized input representations
- Incomprehensive metrics
- ……

Other methods

Proposed model

extensible | universal | Input-aware | Comprehensive | Configuration inputs

Content

- Introduction and Motivation
- Proposed Comprehensive Model
- Evaluation and Analysis
- Verification
- Conclusion

Proposed Comprehensive Model

➢ The model is mainly applied to the deliberately designed approximate multiplier.

Approximate Approaches

- Input Truncation and Compensation
- Approximate Booth Encoding and Decoding
- Partial Product Truncation and Compensation
- Approximate Compressors

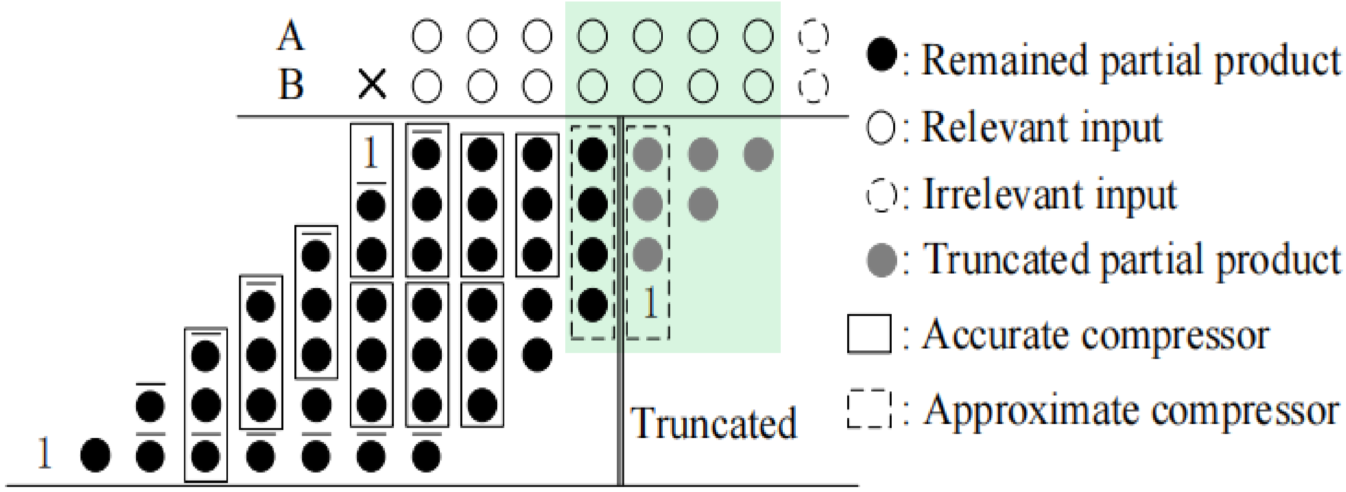


Fig. 1: Demo

➢ For other types of multipliers, only one traversal is needed.

Proposed Error Model

- The model can generate error metrics including

$$MED,\ MRED,\ MAED,\ RMS_{ed} \text{ and } Var_{ed}.$$

Common metrics (MED, MRED, MAED, RMS)

Crucial for NN (RMS, Var)

- An explicit model for uniform input distributions is provided, offering enhanced computational speed.

Tip1: When a and b is mutual independence, we have

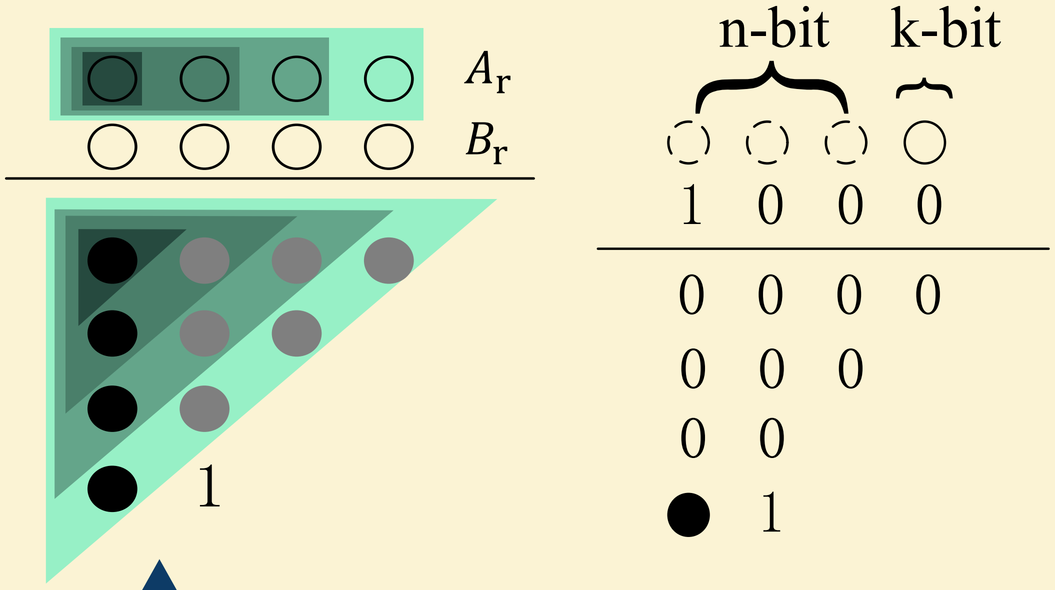
$$\sum_{i=0}^{m}\sum_{j=0}^{n} a_i b_j = \sum_{i=0}^{m} a_i \sum_{j=0}^{n} b_j$$

Proposed Error Model

- A new paradigm is needed for approximate methods concerning two inputs.

Tip2: Utilizing the one-to-many relationships between relevant partial products and input combinations.

- Only the high n bits of A_r are relevant to the first n rows of partial products.
- When the n LSBs of B_r are all zero, it cuts the connection to the high n bits of A_r.
- The ED is pre-stored in a Look-up Table.



★ Combinations: 256→170

✶ extensible

Proposed Error Model

- Take MED of an N-bit multiplier using it-bit input-truncation for example.

$$A/B \xrightarrow{\text{separate}} A_{21}2^{it} + A_0$$

$$ED = E - A_{21}2^{it} \times B_{21}2^{it} = 2^{it} \times (A_{21}B_0 + B_{21}A_0) + A_0B_0$$

$$MED' = \sum_{a_{21}=-Ed_{it}}^{Ed_{it}-1} \sum_{b_{21}=-Ed_{it}}^{Ed_{it}-1} \sum_{a_0=0}^{2^{it}-1} \sum_{b_0=0}^{2^{it}-1} a_0 b_0 \times p_a p_b$$

$$MED' = \sum_{a_0=0}^{2^{it}-1} a_0 \times p_{a_0} \sum_{b_0=0}^{2^{it}-1} b_0 \times p_{b_0}$$

$$ED_{it} = 2^{N-it-1} \qquad p_{a_0} = P(A_0 = a_0) = \sum_{a_{21}=-Ed_{it}}^{Ed_{it}-1} P\left(A = a_{21}2^{it} + a_0\right)$$

Uniform distribution

$$p_{a_0} = \sum_{a_{21}=-Ed_{it}}^{Ed_{it}-1} \frac{1}{2^N} = 2^{-it}$$

Proposed Hardware Model

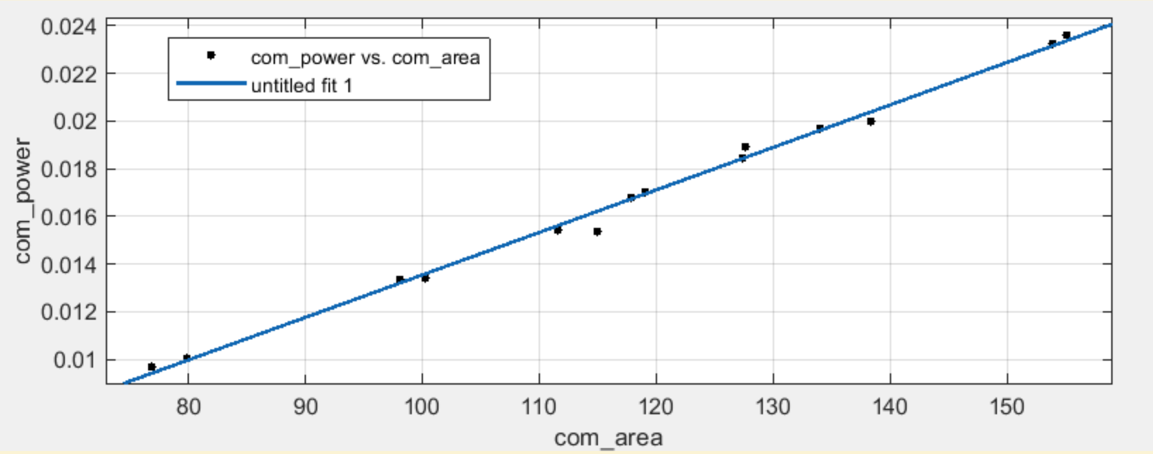
Partition Counting

- It can be expanded to different **technology**, approximate **Booth encoding algorithm**, and approximate **compressor**.

Area/Delay : multiply total/maximum number by corresponding area/delay.

$$A_{com}/T_{com} = C_{ppg}A_{ppg}/T_{ppg} + C_{ppc}A_c/T_{c\_s} + C_aA_a/T_{a\_c}$$

Power: curve fitting based on area due to strong correlation.



Taking array multiplier as an example:

$$P_{com} = 0.0001768A_{com} - 0.004179$$

Content

- Introduction and Motivation
- Proposed Comprehensive Model
- Evaluation and Analysis
- Verification
- Conclusion

Evaluation and Analysis: Error Model Precision

Table Ⅰ: The Accuracy and Runtime Compared with Exhaustive Simulation of MED, $MRED$, $MAED$ and RMS_{ed}.

Input distribution	MED			$MRED$			$MAED$			RMS_{ed}		
	Diff/ 10^{-13}	ex_time/ m_time(s)	runtime ratio	Diff/ 10^{-15}	ex_time/ m_time(s)	runtime ratio	Diff/ 10^{-13}	ex_time/ m_time(s)	runtime ratio	Diff/ 10^{-13}	ex_time/ m_time(s)	runtime ratio
8-bit array multiplier with 2-bit input-truncation, 3-bit partial-product-truncation and compensation												
uniform	0	1.40 / 0.0012	1166.67	0	1.48 / 0.0031	477.42	0	1.47 / 0.088	16.7	0	1.45 / 0.0016	906.25
$\mu = 0$ $\sigma = 100$	4.09	1.51 / 0.0047	321.28	-0.059	1.57 / 0.0075	209.33	-3.41	1.61 / 0.34	4.74	-6.82	1.56 / 0.016	97.5
$\mu = 50$ $\sigma = 200$	8.6	1.50 / 0.0047	319.16	-0.247	1.60 / 0.0075	213.33	7.67	1.56 / 0.32	4.88	0.568	1.55 / 0.017	91.18
10-bit radix_4 booth multiplier with 3-bit input-truncation and compensation, 2-bit partial-product-truncation and compensation												
uniform	0	45.16 / 0.0021	21504.76	0	46.14 / 0.0077	5992.21	0	45.21 / 0.66	68.5	0	44.89 / 0.0029	15479.31
$\mu = 0$ $\sigma = 100$	3.77	27.34 / 0.008	3417.5	26	47.84 / 0.025	1913.6	23.3	46.88 / 3.77	12.44	3.41	47.54 / 0.038	1251.05
$\mu = 50$ $\sigma = 200$	2.82	46.64 / 0.0087	5360.92	1.11	47.72 / 0.025	1908.8	-235	47.92 / 3.83	12.51	-5.68	49.69 / 0.036	1380.28
12-bit radix_8 booth multiplier with 1-bit input-truncation and compensation, 4-bit partial-product-truncation												
uniform	0	377.84 / 0.018	20991.11	-141	388.10 / 0.087	4460.92	0	402.45 / 1.33	302.59	0	390.2 / 0.021	18580.95
$\mu = 0$ $\sigma = 100$	24.1	411.9 / 0.13	3168.46	-241	421.95 / 0.35	1205.57	1550	414.22 / 17.15	24.15	0.426	414.63 / 0.58	714.88
$\mu = 50$ $\sigma = 200$	75	407.36 / 0.13	3133.54	-17.3	419.75 / 0.33	1354.03	12000	426.53 / 17.28	24.68	-1.71	421.83 / 0.58	727.29

Gaussian Distribution

Multiplier description

Evaluation and Analysis: Hardware Model Precision

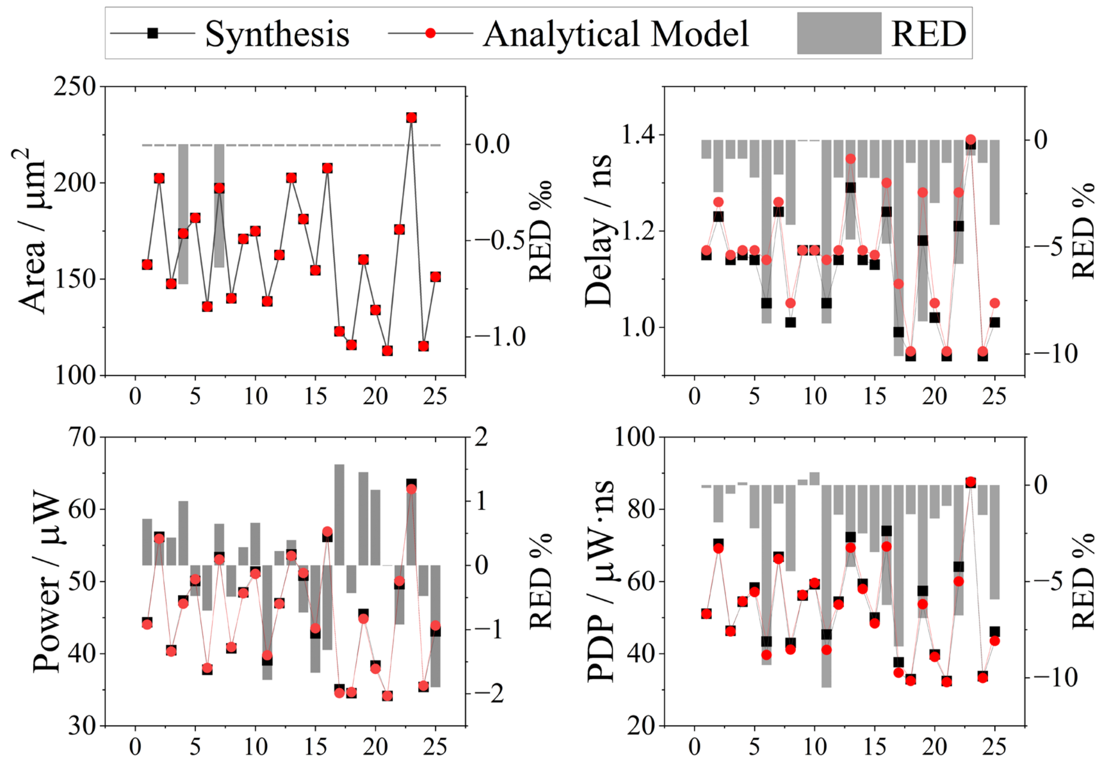


Fig.2: Hardware evaluation using 25 representative multipliers for area, power consumption and PDP.

	MRED	WC_RED
Area	-0.055‰	-0.726‰
Power	-0.28‰	-1.90%
Delay	-3.18%	-10%
PDP	3.2%	-10.5%

- Area model has a high precision owe to accurate partition counting and almost constant cell area.
- Power model is slightly worse than that of the area.
- Delay model is affected by varying delays among the same elements.

Evaluation and Analysis: Error Model Performance

Exhaustive simulation: runtime exploration

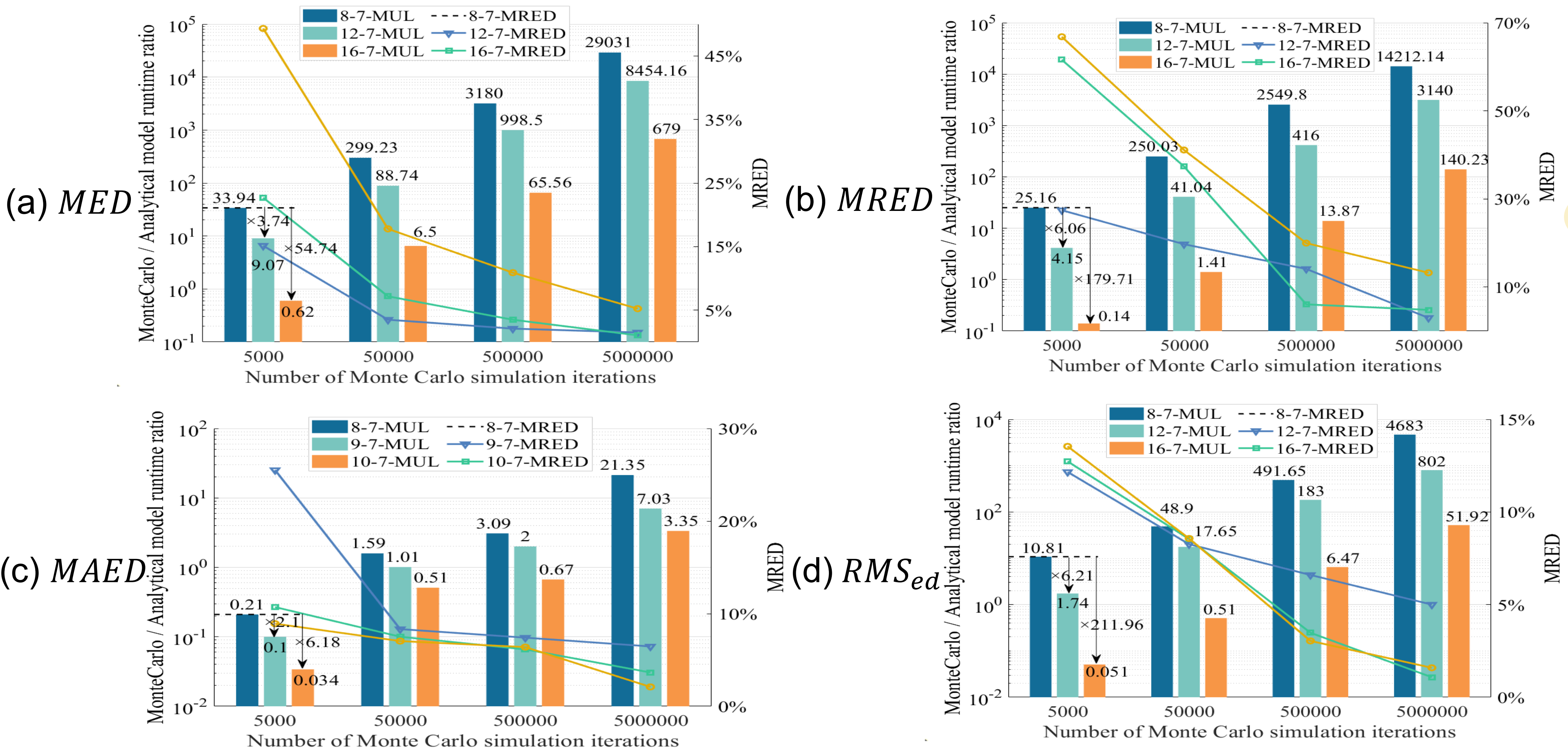
Compare the proposed model with Monte Carlo simulation

- The number of multipliers are 14, 28, and 35 for different bit-widths separately.
- The approximate bit-width is 7.

Bit-width	Multiplier configuration[1]
8	[0,0,7,0,0],[0,1,5,0,0],[0,1,5,1,0],[0,2,3,0,0],[0,2,3,1,0],[0, 3,1,0,0],[0,3,1,1,0][1,0,7,0,0],[1,1,5,0,0],[1,1,5,1,0], [1,2,3,0,0],[1,2,3,1,0],[1,3,1,0,0],[1,3,1,1,0]
9, 10	8.append([0,0,8,0,1],[0,1,6,0,1],[0,1,6,1,1],[0,2,4,0,1],[0, 2,4,1,1],[0,3,2,0,1],[0,3,2,1,1],[1,0,8,0,1],[1,1,6,0,1], [1,1,6,1,1],[1,2,4,0,1],[1,2,4,1,1],[1,3,2,0,1],[1,3,2,1,1])
12, 16	9.append([2,0,7,0,0],[2,1,5,0,0],[2,1,5,1,0],[2,2,3,0,0], [2,2,3,1,0],[2,3,1,0,0],[2,3,1,1,0])

[1]Encoding scheme, input-truncation bit-width, partial-product-truncation bit-width, compensation for input-truncation and compensation for partial-product-truncation

Evaluation and Analysis: Error Model Performance



2549.8× for 8-bit

120.85× for 12-bit

Fig.3: The time ratio of the Monte Carlo simulation method to the proposed method and the errors for four metrics

Content

- Introduction and Motivation
- Proposed Comprehensive Model
- Evaluation and Analysis
- Verification
- Conclusion

Verification: Convolution

➢ 20 representable multipliers form a 794 billion design space.

APP
- 3*3 convolution circuit
- Gaussian Blur

Set
- Quality: sum of the MSEs [9]
- Hardware: [10]
- Search technique: MCT [11]

Verify
- Quality: 5 million sets
- Hardware: synthesis using DC

Comparable Pareto-optimal sets with 10 times faster

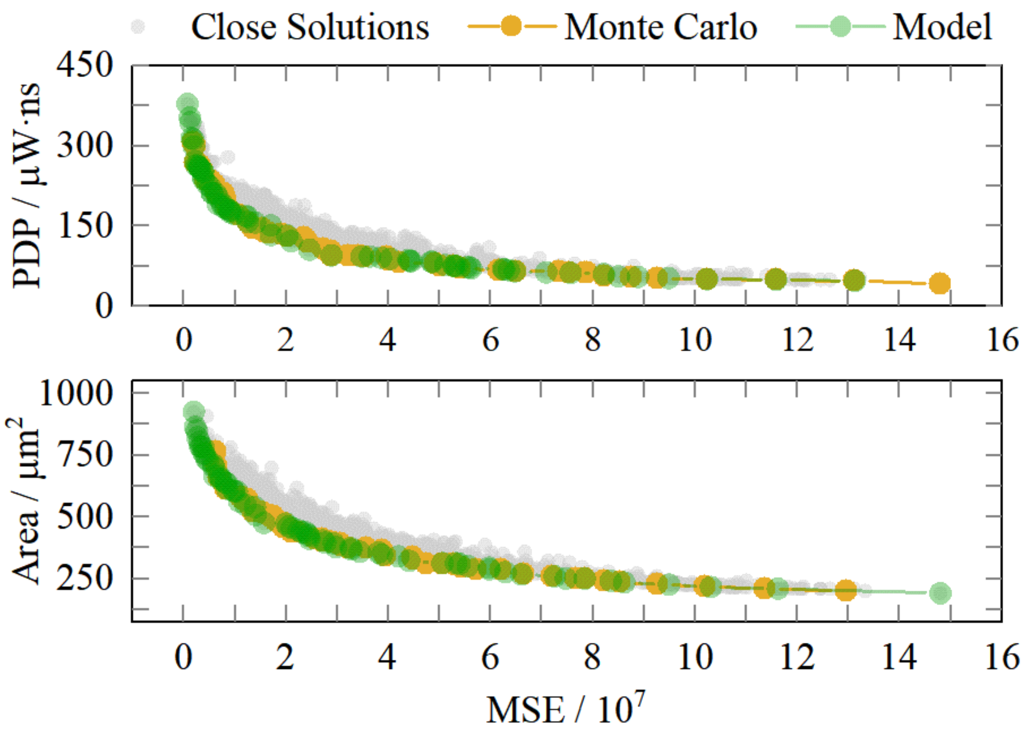


Fig.4: The Pareto-optimal values using the proposed model and Monte Carlo method.

[9] D. Sengupta, F. S. Snigdha, J. Hu, and S. S. Sapatnekar, "SABER: Selection of Approximate Bits for the Design of Error Tolerant Circuits," in 54th Annual Design Automation Conference, 2017, pp. 1–6.
[10] J. Castro-Godnez, J. Mateus-Vargas, M. Shafique, and J. Henkel, "AxHLS: Design space exploration and high-level synthesis of approximate accelerators using approximate functional units and analytical models," in Proceedings of the 39th International Conference on Computer-Aided Design, 2020, pp. 1–9.
[11] M. Awais, H. G. Mohammadi and M. Platzner, "An MCTS-based Framework for Synthesis of Approximate Circuits," IFIP/IEEE International Conference (VLSI-SoC), Verona, Italy, 2018, pp. 219-224.

Verification: Convolution

- Select the optimal values under four accuracy constraints for two objects.

MSE_set $/10^7$	MSE_ex $/10^7$	PSNR_ex /dB	Synthesis	Analytical Model
PDP-optimal			PDP/$\mu w \cdot ns$	
0	0	$+\infty$	856.96	/
1	0.98	33.45	176.02	173.05
2	2.01	30.34	131.06	131.00
6	5.98	25.61	74.28	73.02
10	9.98	23.38	50.50	52.94
Area-optimal			Area/μm^2	
0	0	$+\infty$	2158.76	/
1	0.98	33.45	604.93	604.95
2	2.00	30.36	475.27	474.61
6	5.98	25.61	287.28	288.46
10	9.96	23.39	225.41	226.20

MRED of model error is **2.3‰** and **2.1%** for PDP and Area.

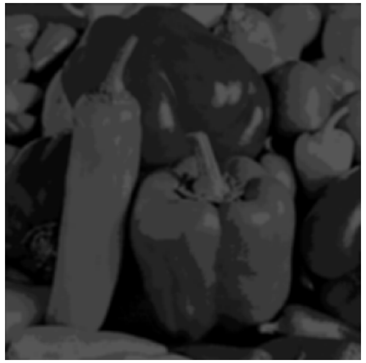
79.46% reduction of PDP is achieved when the MSE is constrained to 10^7.

Area can be reduced by **71.98%** when the optimal area is required

Verification: Gaussian Blur

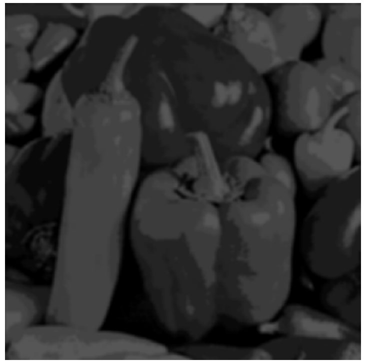
- 25 8-bit grayscale maps for input distribution
- 20dB and 30dB restriction
- **68.59%** less PDP for 30dB restriction
- **56.21%** less area for 30dB restriction

$Area = 951.05\mu\text{m}^2$
$PDP = 294.98\mu W \cdot \text{ns}$



$PDP, 20.37dB$
$PDP = 34.32\mu W \cdot \text{ns}$

$PDP, 30.00dB$
$PDP = 92.66\mu W \cdot \text{ns}$



$Area, 20.37dB$
$Area = 221.89\mu\text{m}^2$

$Area, 30.04dB$
$Area = 416.43\mu\text{m}^2$

Fig.5: Resulting images of Gaussian blur for different targets.

Content

- Introduction and Motivation
- Proposed Comprehensive Model
- Evaluation and Analysis
- Verification
- Conclusion

Conclusions

- In proposed comprehensive model, five distinct quality metrics in addition to four hardware metrics are derived.
- In comparison to the Monte Carlo simulation method, the proposed model demonstrates a remarkable reduction in runtime, with an average decrease of 120.85, and in specific instances, as low as 2,500.
- The 3*3 convolution circuit and Gaussian Blur application is employed to verify the proposed model with 79.46% reduction in PDP and a 71.98% reduction in area when compared to the accurate counterpart .

Thanks!