Department of Electrical and Computer Engineering
UNIVERSITY OF WISCONSIN–MADISON

In-Stream Stochastic Division and Square Root via Correlation

Di Wu and Joshua San Miguel

Stochastic Computing

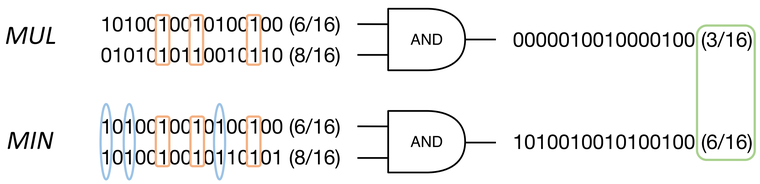
SC data representation

- *Bernoulli Sequence* as unipolar *Bit Stream (BS).*
  - Value equal to the ratio of 1s
  - Value irrelevant to order of 0s and 1s
  - Value within (0, 1)

A=0.5	1	1	0	1	0	1	1	1	0	0	1	0	0	1	0	0
B=0.5	0	0	1	0	0	1	1	1	0	1	0	1	0	1	0	1
C=0.75	1	0	1	1	0	1	1	1	1	1	0	1	0	1	1	1

SC circuit and correlation

- SC multiplication with a single AND gate
  - *Extremely simple logic*
  - *Ultra-low power*
  - *Resistant to noise*
- Level of paired 1s indicate SC correlation
- Varying correlation leads to function varying (from *MUL* to *MIN*)



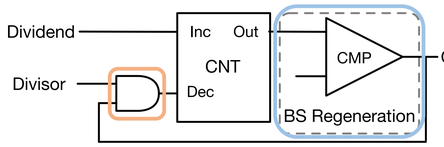
Nonlinear functions in neural networks

NN	Operation
CNN	Conv, FC, Pooling, ReLu
LSTM	*, Div, Exp, Tanh
Graph CNN	*, Div, Exp, Log, Sqrt
CapsNet	*, Div, Exp, Log, Sqrt

- More nonlinear functions in emerging NNs
- Stochastic computing for better hardware efficiency

Existing SC DIV and SQRT

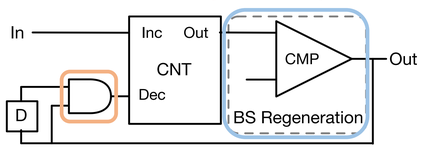
Gaines DIV (GDIV)



System reaches equilibrium:

$$P_{Dividend} = P_{Quotient} * P_{Divisor}$$

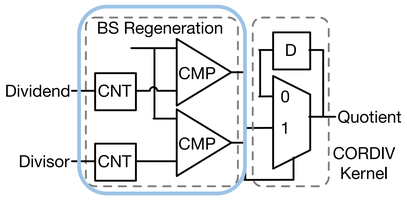
Gaines SQRT (GSQRT)



System reaches equilibrium:

$$P_{In} = P_{Out} * P_{Out}$$

Correlated DIV (CORDIV)



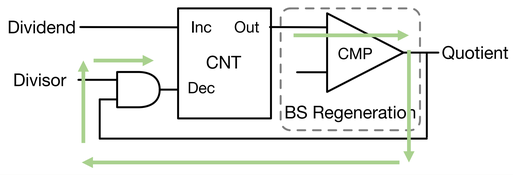
BS inputs to CORDIV Kernel are positively correlated

$$P_{Quotient} = P_{Dividend}/P_{Divisor} = \frac{\#(1)_{Dividend}}{\#(1)_{Divisor}}$$

Hardware Problem

- Highly relying on near zero correlation, restricting the RNG quality
- Expensive Bit Stream Regeneration, increasing cost

Statistical Problem

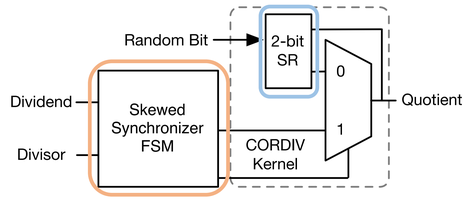


Feedback loops introduce fluctuation in resulting BS, leading to output deviation.

Proposed SC DIV and SQRT

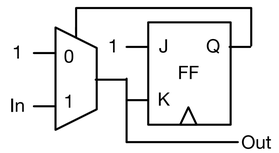
In-Stream Correlation Based Division (ISCBDIV)

- Leveraging CORDIV Kernel with correlation



- Skewed Synchronizer pairs up 1s for maximal correlation.
- 2-bit Shift Register improves randomness.

JKDIV based BISQRT

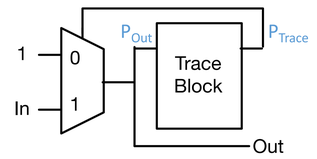


Connecting port J to 1 leads to

$$P_Q = P_J/(P_J + P_K) = \frac{1}{1 + P_K}$$

Bit Inserting Square Root (BISQRT)

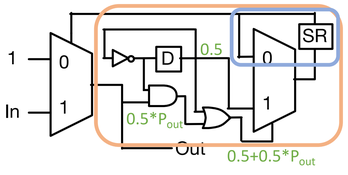
- SC SQRT output is less than input. $\sqrt{0.81} = 0.9 > 0.81$
- Insert 1s to input leads to correct SQRT in SC.



Functionality is guaranteed if

$$P_{Trace} = 1/(1 + P_{Out})$$

ISCBDIV based BISQRT



Simplified ISCBDIV

Output of CORDIV Kernel comes from the SR for isolation

$$P_{Quotient} = 0.5/(0.5 + 0.5 * P_{Out}) = \frac{1}{1 + P_{Out}}$$

Evaluation and Implementation

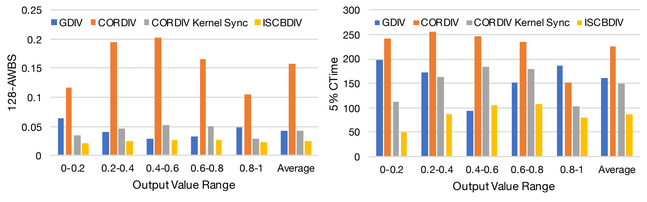
Evaluating metrics

New metrics to deal with feedback loops

Metric	Definition
SCC	Stochastic Cross Correlation for 2 BSs.
SAC	Stochastic Auto Correlation for 1 BS.
WBS	N-Window Bias: accumulative error for the most recent N bits; ranging from -1 to 1; expecting 0 for no error.
AWBS	Average N-Window Bias: statistical root-mean-square of multiple N-WBS values.
p% CTime	p% Convergence Time: cycle count required for SCU to achieve a stable N-WBS of less than p%; smaller is better.

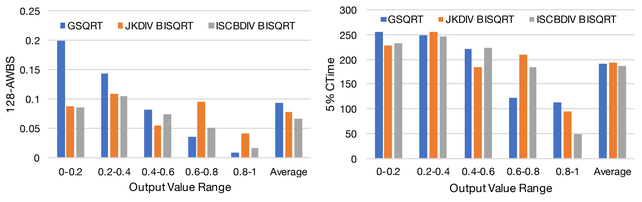
ISCBDIV performance

- 43.3% lower in error
- 46.3% faster in convergence



BISQRT performance

- JKDIV-BISQRT reduces error by 16.8%.
- ISCBDIV-BISQRT reduces error by 29.0%.



Meeting correlation requirements for different operators improves accuracy.

Module	Desired	Actual
MUX	0	-
JKDIV	-	+
ISCBDIV	-	+

SC correlation table

Hardware implementation

- For DIV, energy reduction is 67.6%.
- For JKDIV/ISCBDIV SQRT, reduction is 72.8%/47.8% in energy/sqrt.

Design	*Area* (μm^2)	*Power* (μW)	*Latency* ($cycles$)	*TPA* ($1/(\mu m^2 \cdot s)$)
GDIV(Depth-5)	74.3	21.0	158	34,073
CORDIV	211.2	60.9	226	8,384
Proposed ISCBDIV	**40.4**	**12.5**	**86**	**115,128**
GSQRT(Depth-5)	78.3	23.5	192	26,607
Proposed JKDIV BISQRT	**11.3**	**6.3**	195	**181,529**
Proposed ISCBDIV BISQRT	25.4	12.6	**187**	84,214