高性能計算基盤

– High Performance Computing Platforms–

#7

Stochastic Computing

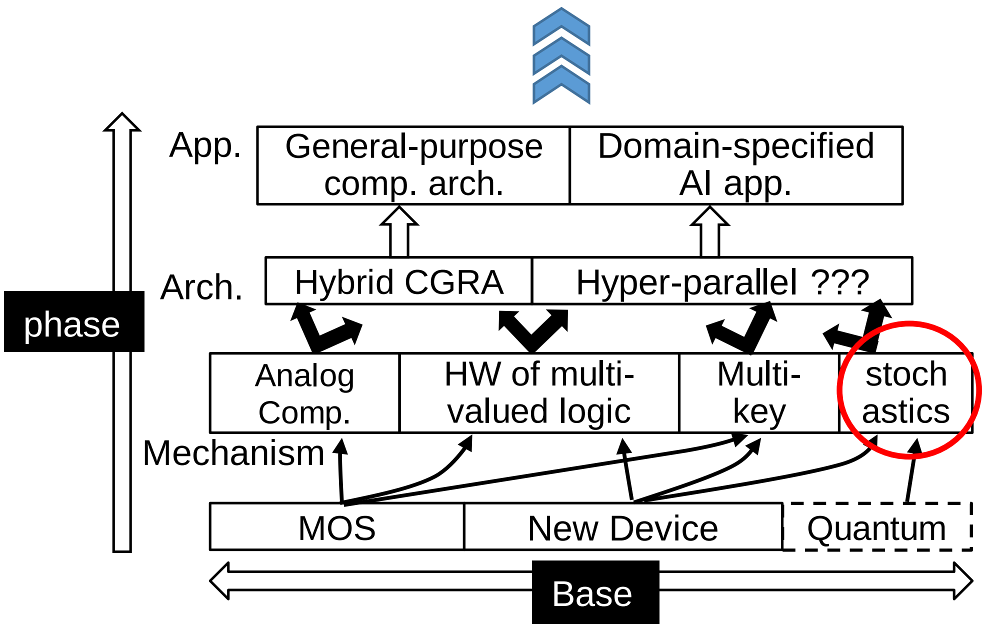
Renyuan Zhang

2020/06/18

NAIST

Road-Map of emerging tech.s for computing

Inventing new topologies and algorithms



Challenges:

- Almost no Mathematical fundamentals
- How to make good use of Appr. Comp.

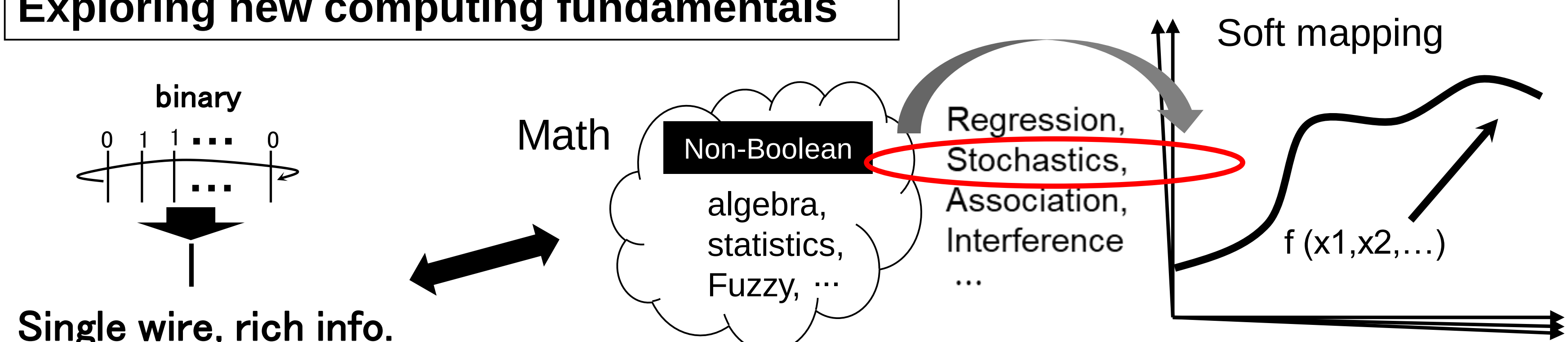
Key Candidates:

- Programmable Analog Computing
- Multi-Domain Stochastic Computing
- New-device + Neuromorphic

Platform:

Hybrid Coarse-Grained Reconfigurable Array (CGRA)

Exploring new computing fundamentals



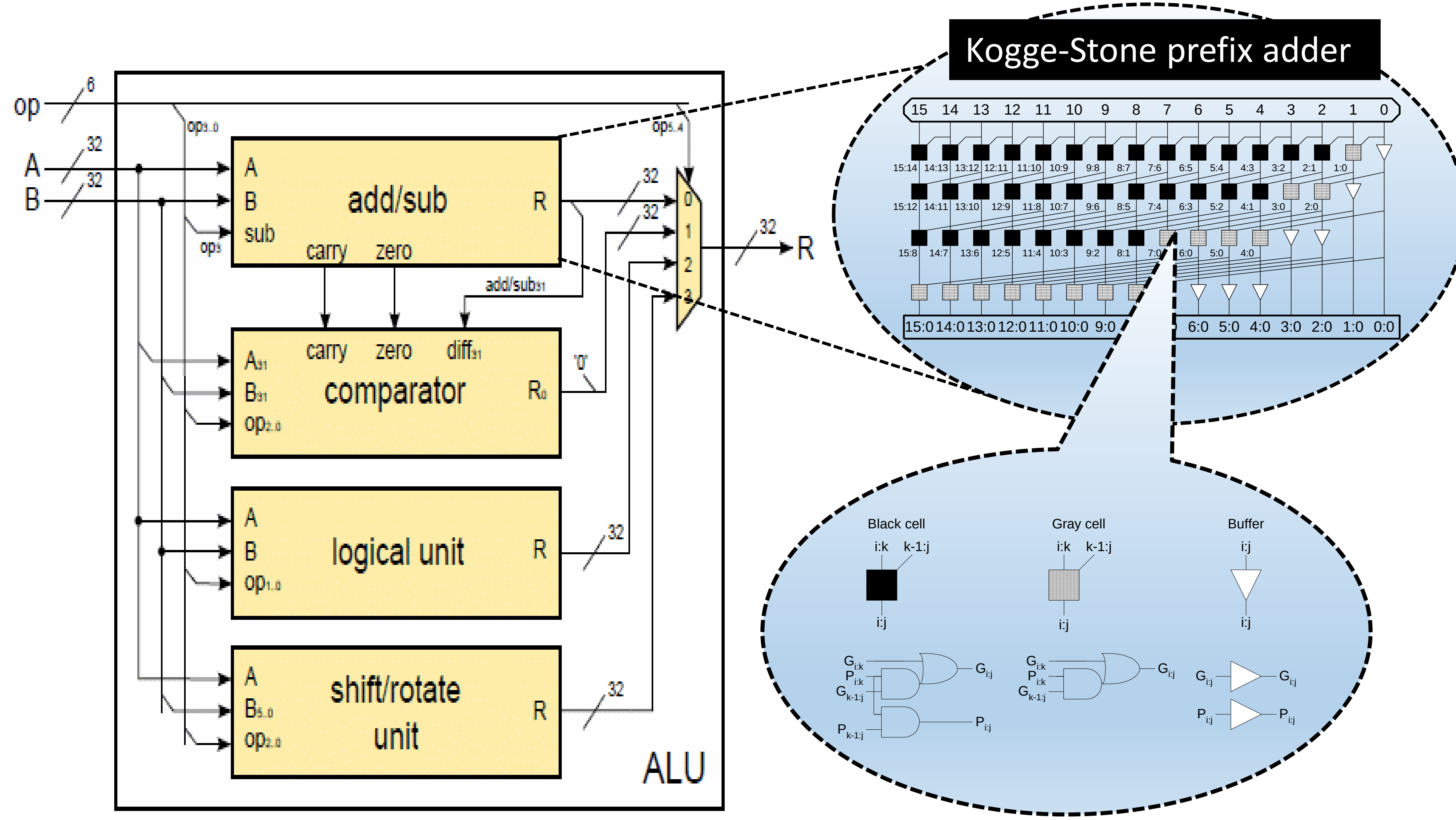
Outline

- What is stochastic computing (SC)
  - Mechanism of SC
  - Elements of SC
  - Implementation of SC
- Time based stochastic computing (TBSC)
  - Mechanism of TBSC
  - Hybrid TBSC
  - Analysis

What is stochastic computing

Complexity of computational unit

A simple review of digital type ALU (arithmetic logic unit) ➔ core part = adder

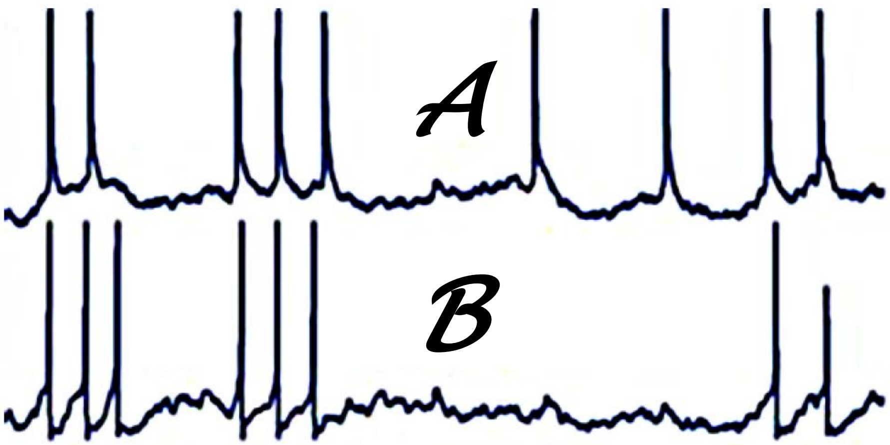


What is stochastic computing

To shrink calculator's size

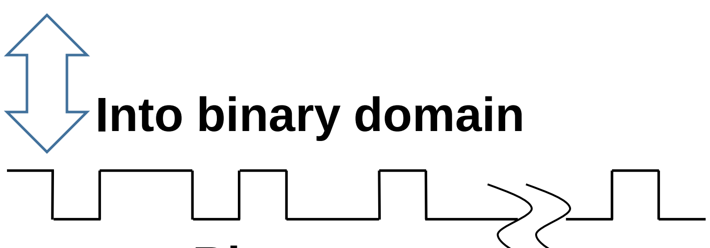
Again, reconsider the data representation: try to use something like ``probability"

11000111000010010011



11101110000000000011

Time sequence



Into binary domain

Bit stream

Considering the probability of pulse-appearance:

$$P_A = \frac{9}{20} = 0.45$$

$$P_B = \frac{8}{20} = 0.4$$

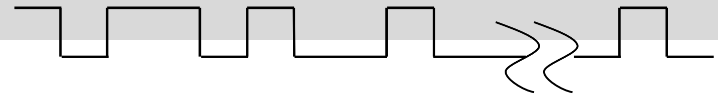
Similar but different from bio-signal: irrelevant to timing, positioning, and strength…

For ``B", it is incorrect somehow; But it loos like no impact to the representation. Is it true?

What is stochastic computing

Definition

Given a bit stream ``X" with length of N: ``1" appearance counting = N1; ``0" appearance counting = N0. $P_X = N_1/N$ or shortly, $X = P_X$
This bit stream is called stochastic number (SN)

X


Property 2-1

1. The SN representation is NOT unique;
2. Only total counting indicates info., → position&pattern = meaningless;
3. The SN with N bit only represent the number in set {0/N, 1/N, 2/N,…, (N-1)/N, N/N} in total of N+1 numbers (= resolution);
4. Range = [0,1] (but extendable by following)

Format to address the real number domain

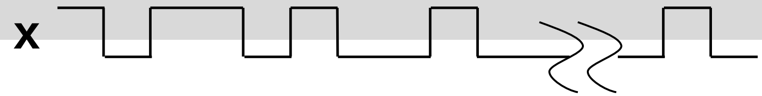
Format	Number value	Number range	Relation to unipolar value p_X
Unipolar (UP)	N_1/N	$[0, 1]$	p_X
Bipolar (BP)	$(N_1 - N_0)/N$	$[-1, +1]$	$2p_X - 1$
Inverted bipolar (IBP)	$(N_0 - N_1)/N$	$[-1, +1]$	$1 - 2p_X$
Ratio of 1's to 0's	N_1/N_0	$[0, +\infty]$	$p_X/(1 - p_X)$

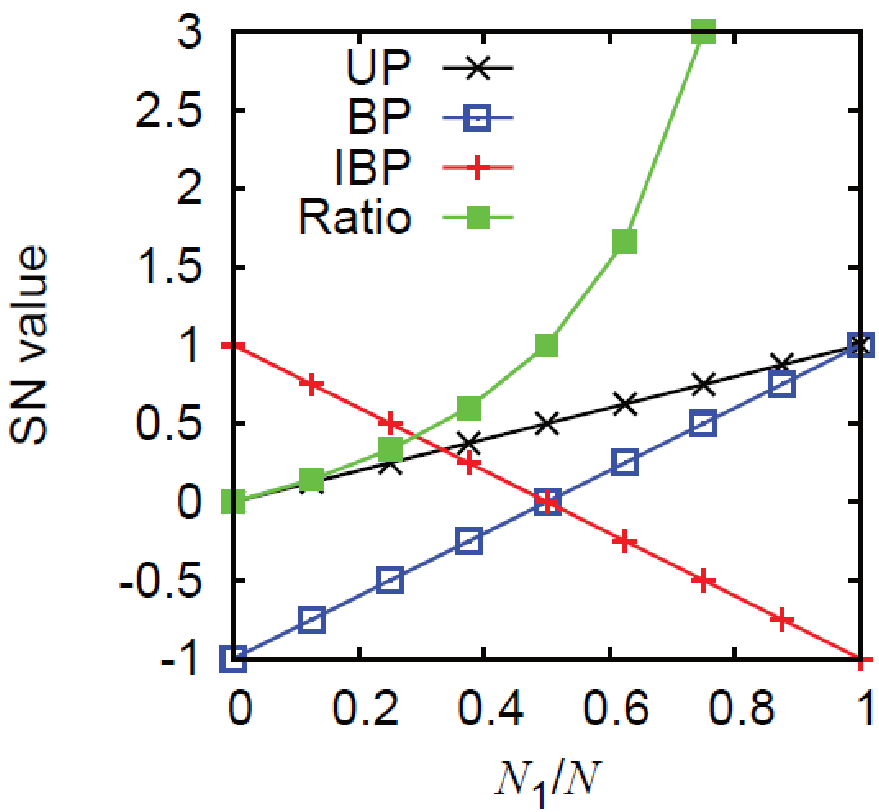
Here, the inaccuracy is observed over data-representation itself → resolution

What is stochastic computing

Definition

Given a bit stream ``X” with length of N: ``1” appearance counting = N1; ``0” appearance counting = N0. $P_X = N_1/N$ or shortly, $X = P_X$
This bit stream is called stochastic number (SN)





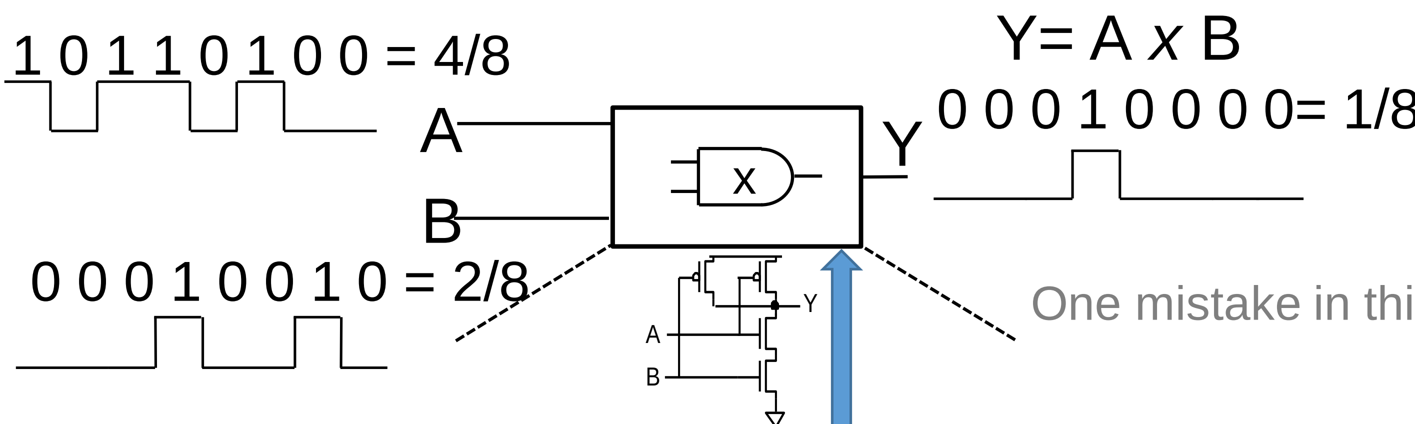
Quiz 2-1 Complete it into full range = 5 min.s

Bit-stream	UP	BP	IBP	Ratio
00000000	0	-1	+1	0
00000001	1/8	-3/4	+3/4	1/7
00000011	2/8	-2/4	+2/4	1/3
00000111				
00001111				
00011111				
00111111				
01111111				
11111111				

What is stochastic computing

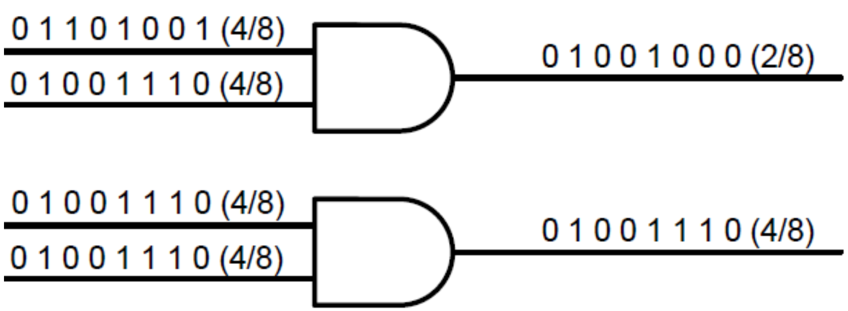
To calculate the ``probability”

Given two stream of A and B, the logic gate ``AND” performs multiplication of Y=AxB



One mistake in this figure, where?

See other examples:



Here, the inaccuracy is observed over calculation→ so far, we suffer from two types of inaccuracy

Why? = position&pattern means something
How? = make them ``random” and uncorrelated

STOCHASTIC

Property 2-2

$$P_{X \cdot Y} = P_X \cdot P_Y$$

If and only if X and Y are random (Bernoulli) and independent (uncorrelated)

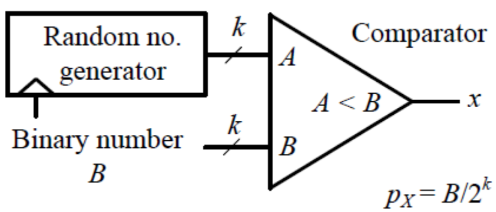
Outline

- What is stochastic computing (SC)
  - Mechanism of SC
  - Elements of SC
  - Implementation of SC
- Time based stochastic computing (TBSC)
  - Mechanism of TBSC
  - Hybrid TBSC
  - Analysis

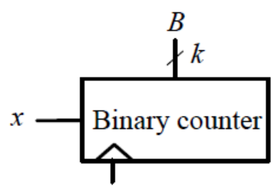
Elements of stochastic computing HW

Generating SN

It is different from generating random number. More than that.



Convert binary to SN

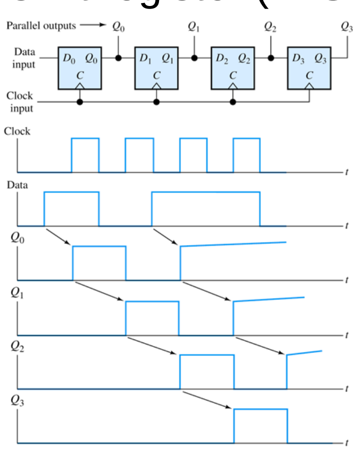


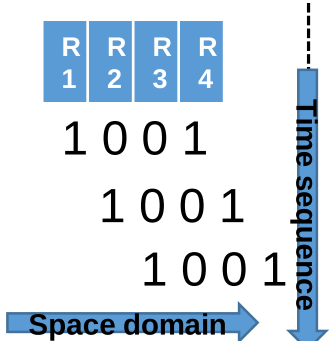
Convert SN to binary

Inaccuracy in SC has several distinct sources: random fluctuations in SN representation, similarities (correlations) among the numbers that are being combined, and physical errors that alter the numbers.

Generally, use long stream…

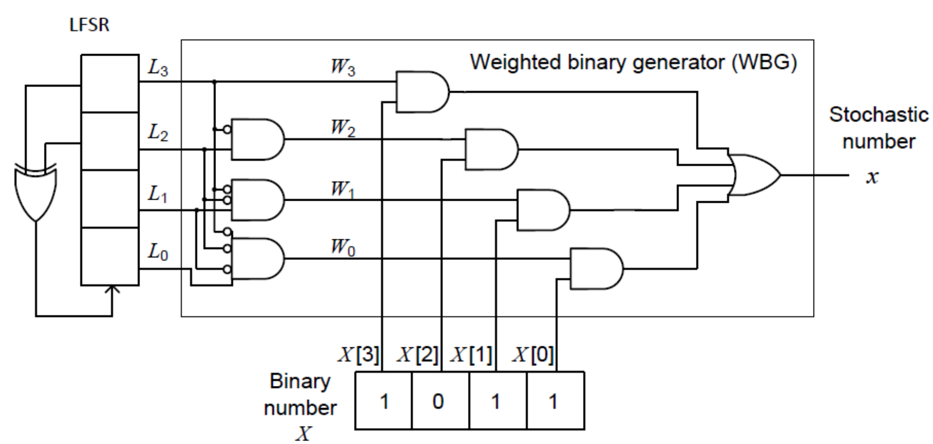
Preliminary, linear feedback shift register (LFSR)





Move to right

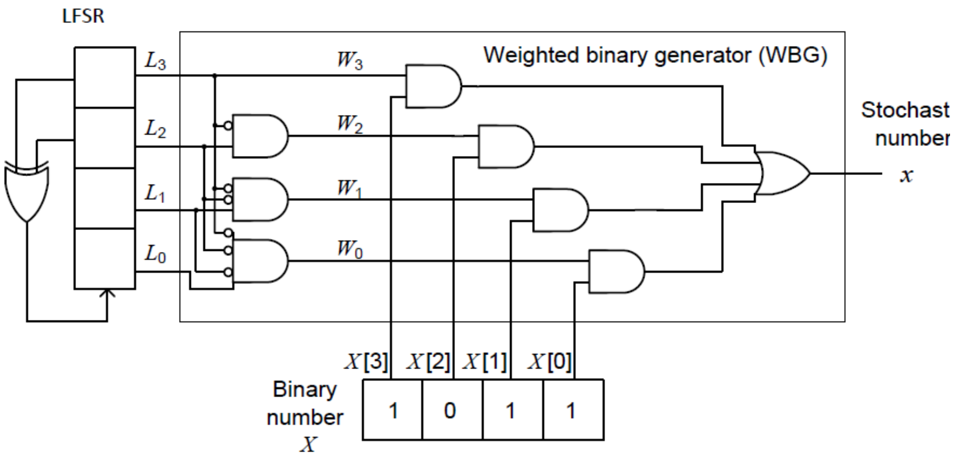
Another (smaller SN generator)



Elements of stochastic computing HW

Generating SN

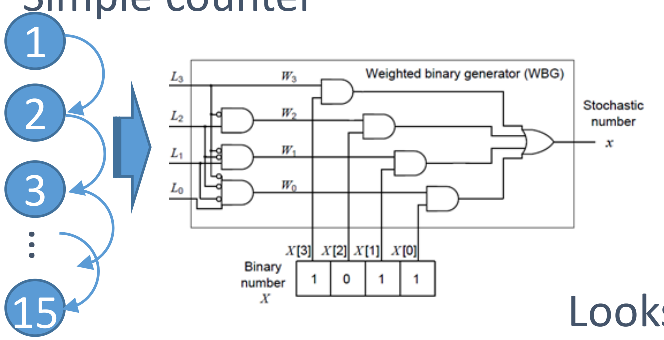
``weight” is attached to final bit stream from the binary radix. Quiz 2-2: explain why



Signal	Bit-stream	Value
L_3	0 0 1 0 1 0 1 1 1 1 0 0 0 0 1 1	8/16
L_2	0 1 0 1 0 1 1 1 1 0 0 0 0 1 1 0	8/16
L_1	1 0 1 0 1 1 1 1 0 0 0 0 1 1 0 0	8/16
L_0	0 1 0 1 1 1 1 0 0 0 0 1 1 0 0 1	8/16
W_3	0 0 1 0 1 0 1 1 1 1 0 0 0 0 1 1	8/16
W_2	0 1 0 1 0 1 0 0 0 0 0 0 0 1 0 0	4/16
W_1	1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0	2/16
W_0	0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0	1/16
x	1 0 1 0 1 0 1 1 1 1 0 1 1 0 1 1	11/16

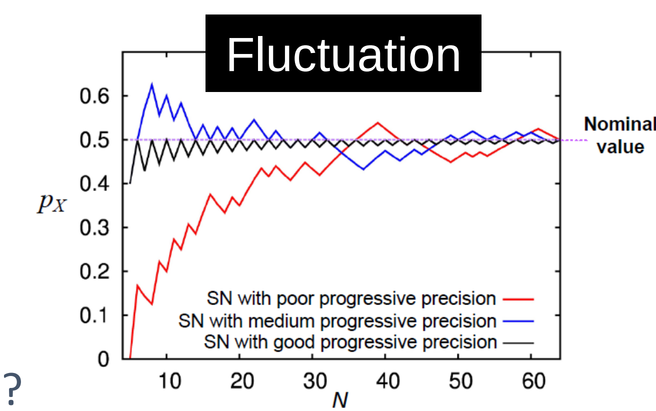
Time sequence

Simple counter



Signal	Bit-stream	Value
L_3	0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1	8/16
L_2	0 0 0 0 1 1 1 1 0 0 0 0 1 1 1 1	8/16
L_1	0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1	8/16
L_0	0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1	8/16
W_3	0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1	8/16
W_2	0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0	4/16
W_1	0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0	2/16
W_0	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1/16
x	0 1 1 1 0 0 0 0 1 1 1 1 1 1 1 1	11/16

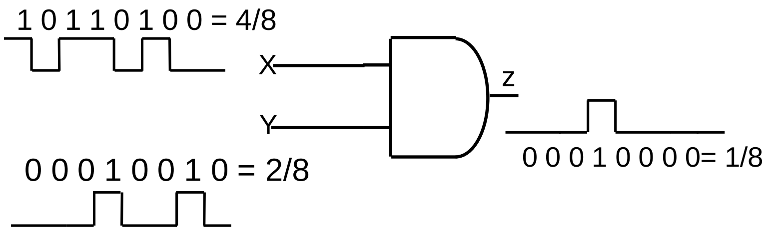
Looks the same, but is that true?

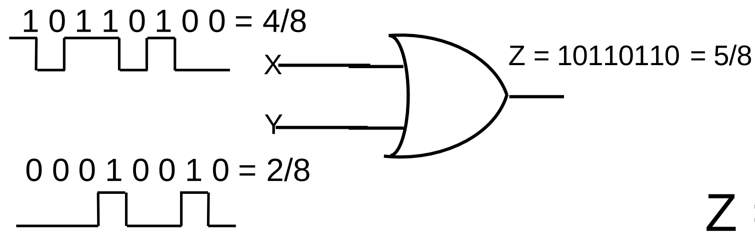


Elements of stochastic computing HW

Calculate SNs

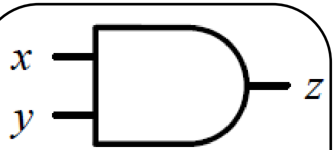
ANG gate is used for multiplication; summation would be = ?? OR gate??





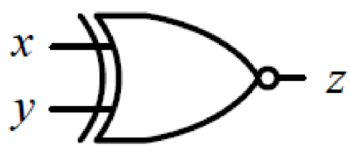
Quiz2-3
Prove it

Z = X+Y-XY



$Z_{AND} = XY$

Multiplier for UP



$Z_{NXOR} =$
$(1-X)(1-Y) + XY =$
$1 - X - Y + 2XY$

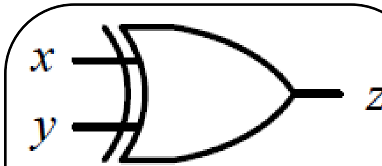
For BP, $X = \frac{X'-1}{2}$

$\frac{Z'_{NXOR} - 1}{2} =$

$1 - \frac{X'-1}{2} - \frac{Y'-1}{2} + 2(\frac{X'-1}{2})(\frac{Y'-1}{2})$

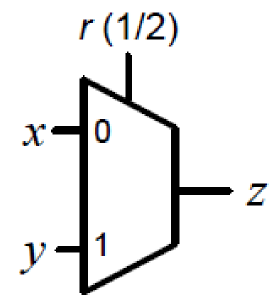
$Z'_{NXOR} = X'Y'$

Multiplier for BP



$Z''_{AND} = X''Y''$

Multiplier for IBP



$Z_{MUX} = Xr' + Yr$
$= 0.5(X+Y)$

Summation for UP, BP, IBP

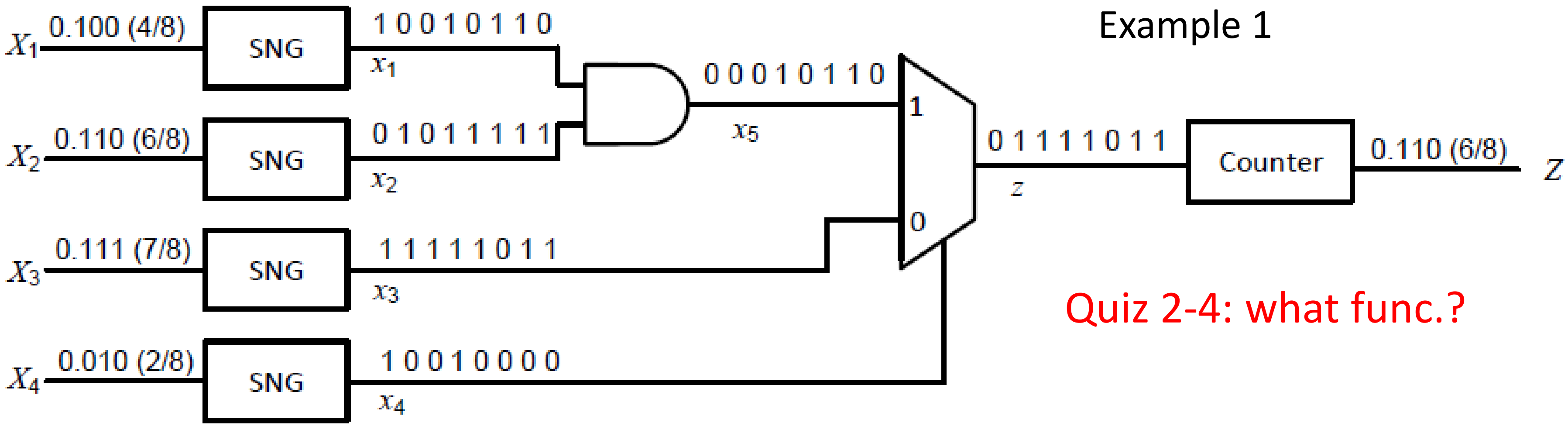
Outline

- What is stochastic computing (SC)
  - Mechanism of SC
  - Elements of SC
  - Implementation of SC
- Time based stochastic computing (TBSC)
  - Mechanism of TBSC
  - Hybrid TBSC
  - Analysis

Implementation of stochastic computing HW

Complex func.

Linear combination is achieved by multiplier and adders



Example 1

Quiz 2-4: what func.?

But, for the complex NON-linear functions, the simple implementations are insufficient. The ONLY option is to ``approximate” them by simple items.

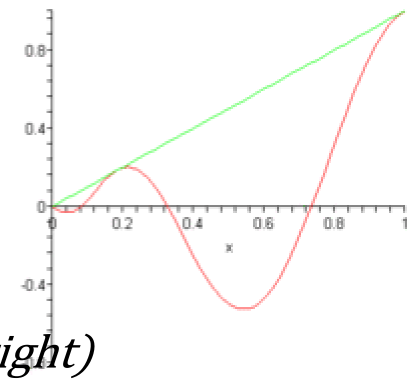
1. Item expansion technologies: arbitrary non-linear function is approximated by, for instance, Taylor Expansion, Bernstein Polynomial etc.
2. Machine learning regression: refer to previous lecture
3. Special tech.s for stochastic

Implementation of stochastic computing HW

Complex func.

Non-linear functions are approximated by Bernstein Polynomial

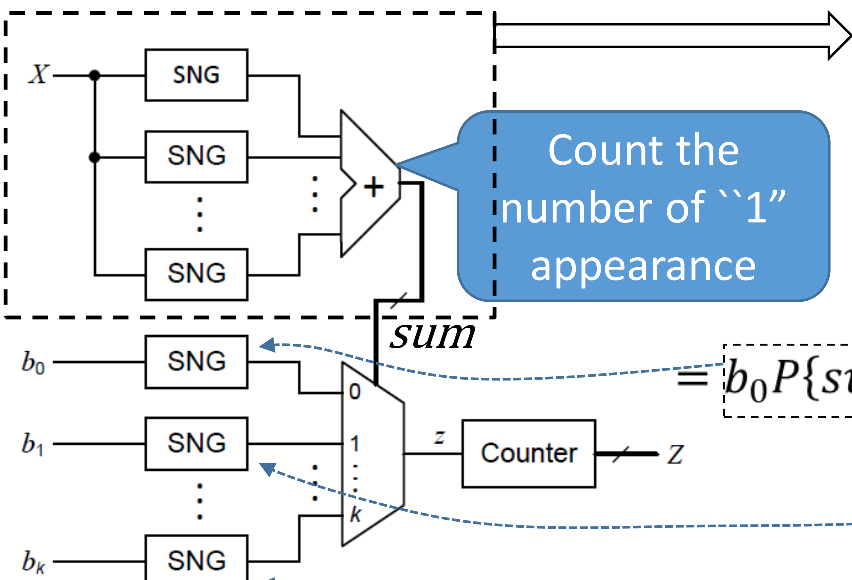
In the mathematical field of numerical analysis, a Bernstein polynomial, named after Sergei Natanovich Bernstein, is a polynomial in the Bernstein form, that is a linear combination of Bernstein basis polynomials. [wikipedia]



$$f(x) \approx \sum_{i=0}^{k} b_k B_{i,k}(x) \text{ where } B_{i,k}(x) = \binom{k}{i} x^i (1-x)^{k-i}$$

$$for\ SC, Z = \sum_{i=0}^{k} b_k B_{i,k}(X) \text{ where } B_{i,k}(X) = \binom{k}{i} X^i (1-X)^{k-i}$$

Where degree k controls the accuracy of approximation. (see right)



k times Bernoulli test: how many ``1" appear?

$$P\{sum = i\} = \binom{k}{i} X^i (1-X)^{k-i}$$

⇩ Addressing proper bi by MUX

$$P\{1\ appears@Z\}$$

$$= b_0 P\{sum = 0\} + b_1 P\{sum = 1\} + \cdots + b_k P\{sum = k\}$$

$$= \sum_{i=0}^{k} b_k B_{i,k}(X)$$

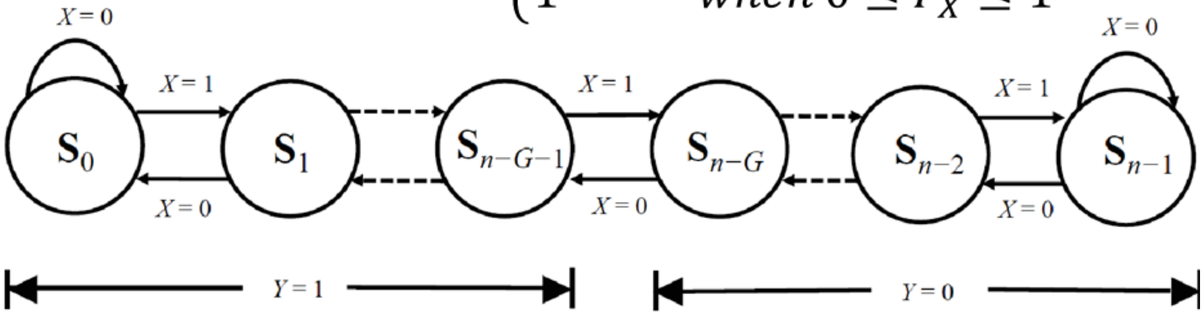
Implementation of stochastic computing HW

Complex func.

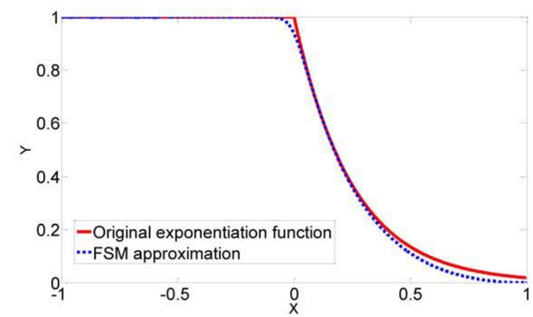
Special technology by finite state machine (FSM) → ONLY very few func.s are available

Example 1

$$P_Y \approx \begin{cases} e^{-2GP_X} & when\ 0 \le P_X \le 1 \\ 1 & when\ 0 \le P_X \le 1 \end{cases}$$

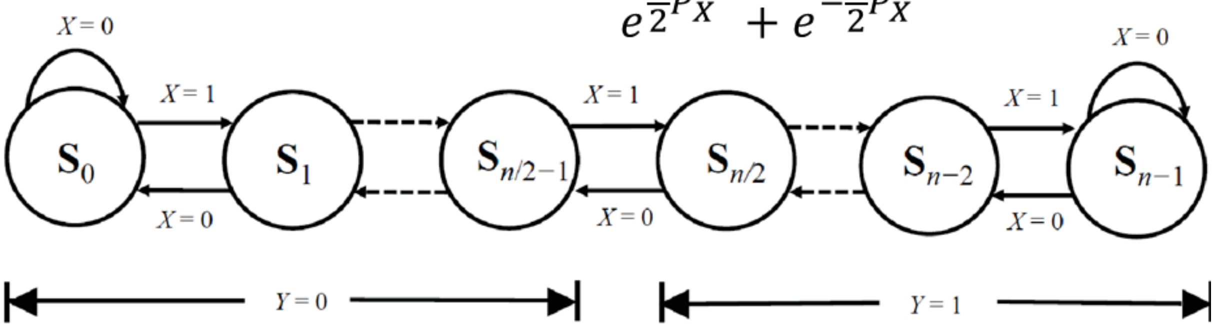


X = BP; Y = UP

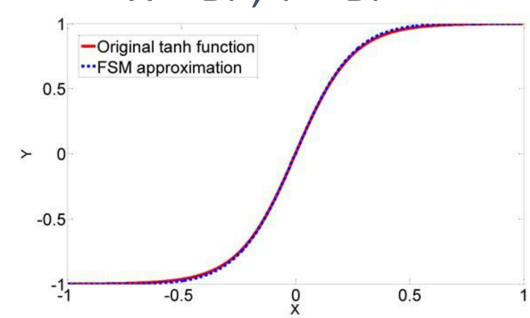


Example 2

$$P_Y \approx \frac{e^{\frac{n}{2}P_X} - e^{-\frac{n}{2}P_X}}{e^{\frac{n}{2}P_X} + e^{-\frac{n}{2}P_X}}$$



X = BP; Y = BP



Implementation of stochastic computing HW

Summary

	Good	Bad
Circuit size and power	Tiny arithmetic components	Many random number sources and stochastic-binary conversion circuits
Operating speed	Short clock periods Massive parallelism	Very long bit-streams
Result quality	High error tolerance Progressive precision	Low precision Random number fluctuations Correlation-induced inaccuracies
Design issues	Rich set of arithmetic components	Theory not fully understood Little CAD tool support at present

Error (inaccuracy)

Error type	Why	How
Approximation, Quantization	Non-linear target functions, Low-degree polynomial approximation, Low-precision constant number generation	Increase polynomial degree, Increase number of bits in constant number generation
Random fluctuation	Inherent randomness, Short bit-stream length	Increase bit-streams length, Use deterministic or low-discrepancy sequences
Insufficient randomness	High error tolerance Progressive precision	Increase random sources, De-correlate correlated signals, Use better number sources (larger LFSRs)
Soft errors	Environmental noise, Component variability,	Use circuit-level error-resilience techniques,

Outline

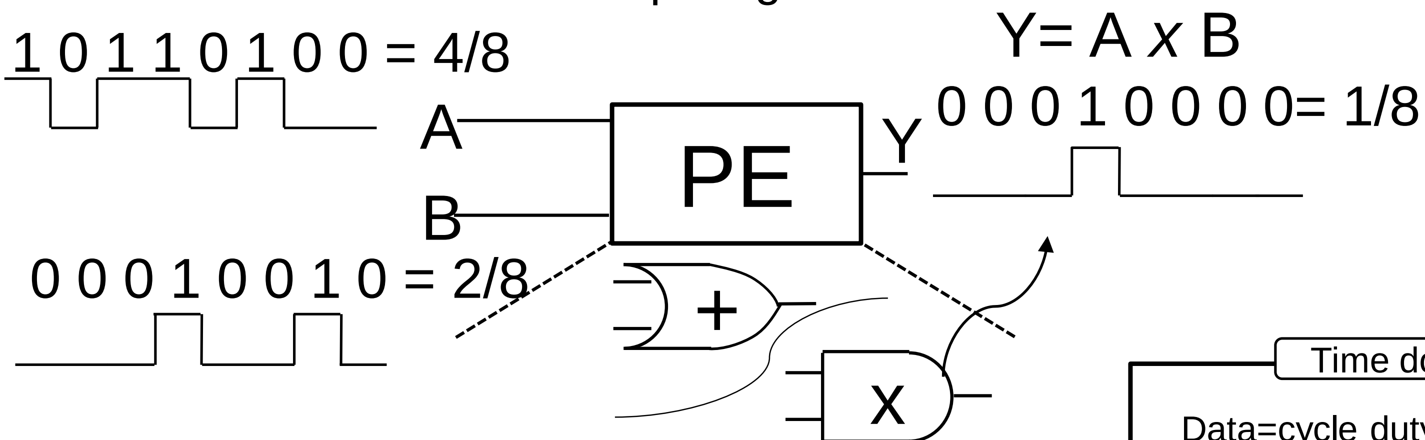
- What is stochastic computing (SC)
  - Mechanism of SC
  - Elements of SC
  - Implementation of SC
- Time based stochastic computing (TBSC)
  - Mechanism of TBSC
  - Hybrid TBSC
  - Analysis

Time-Based Stochastic Computing

Original stochastic computing

Counting probability of ``0” or ``1”

Conventional stochastic computing



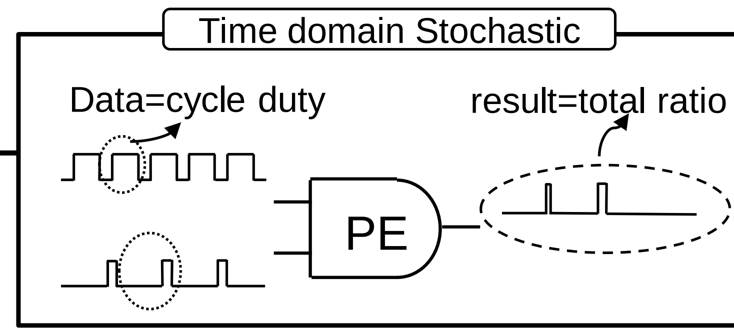
With benefits, But inefficient

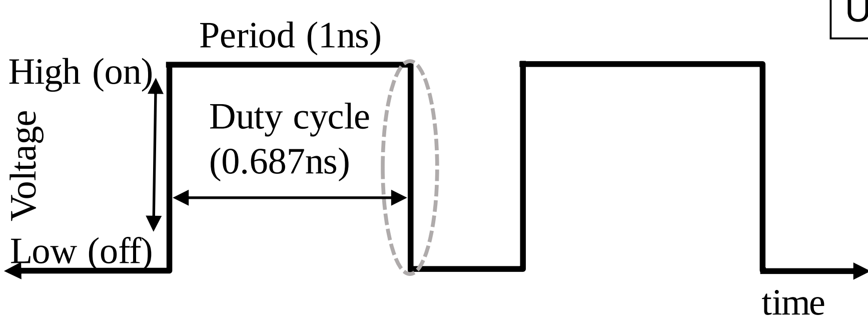
Existing solution

Make it faster?

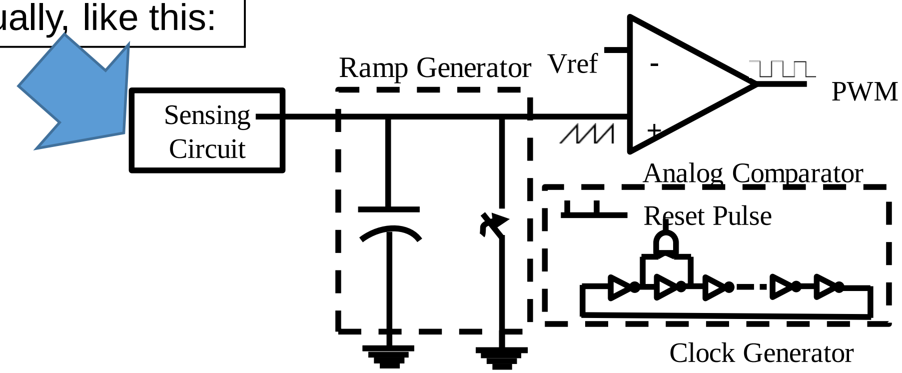
Time-domain SC

Only one problem: how to tune the duty cycle, efficiently



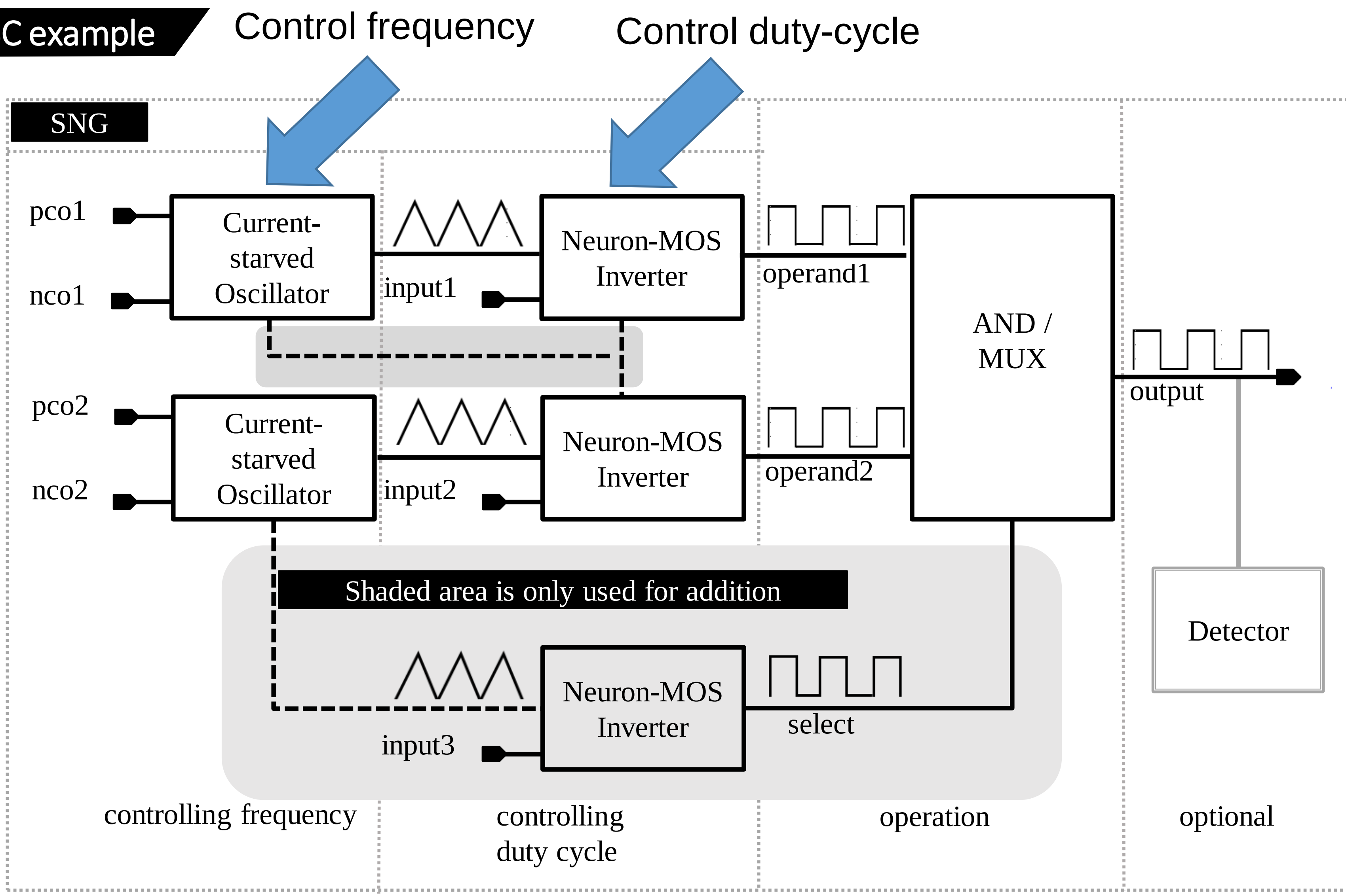


Usually, like this:



Time-Based Stochastic Computing

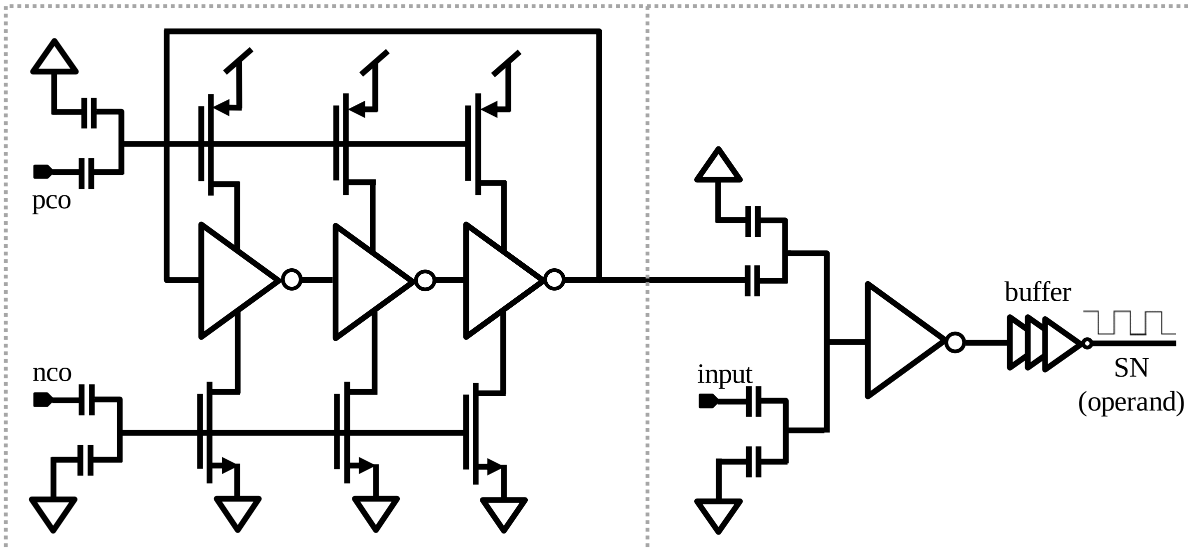
TBSC example

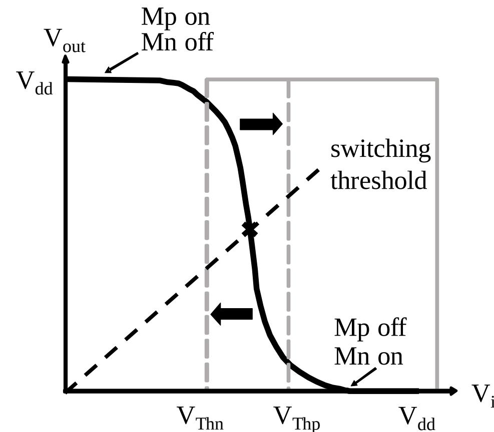


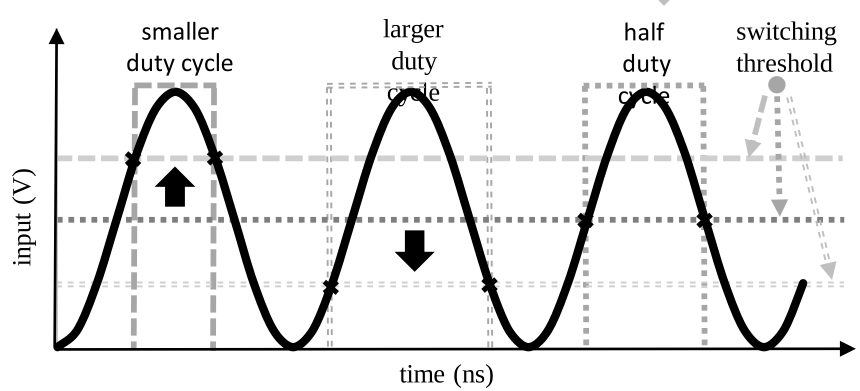
Time-Based Stochastic Computing

implementation

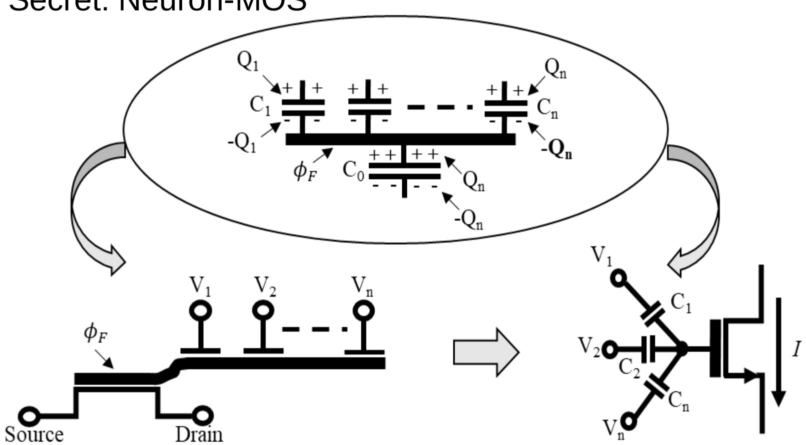
Generate stochastic number in time-domain ➔ duty cycle







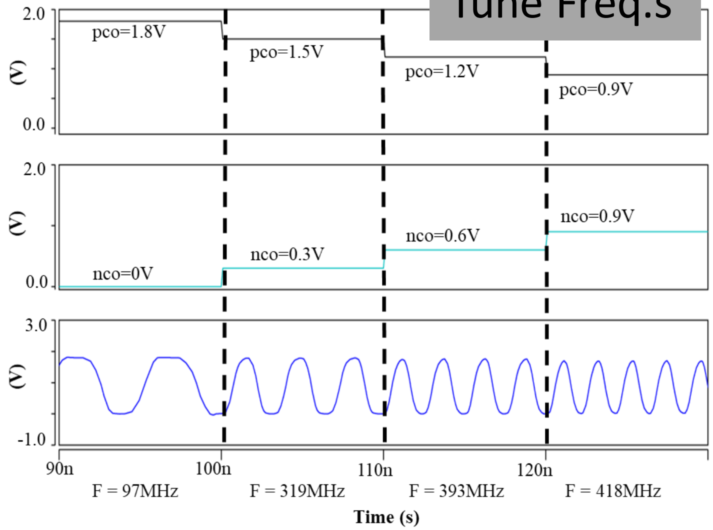
Secret: Neuron-MOS



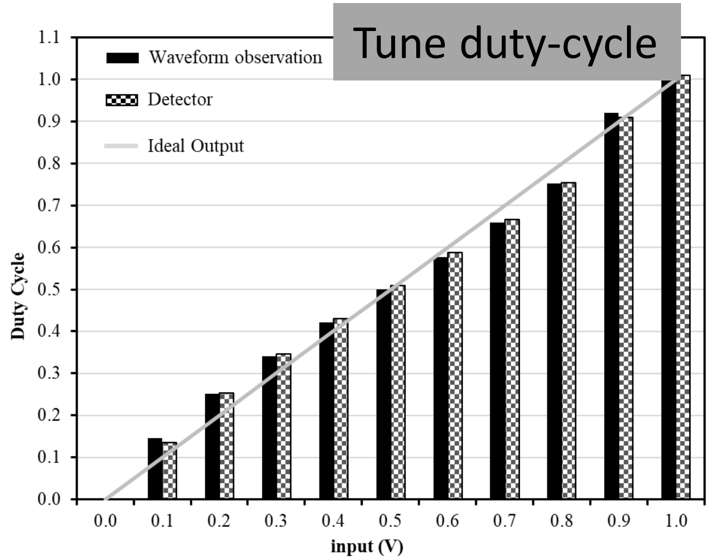
Time-Based Stochastic Computing

Result: example

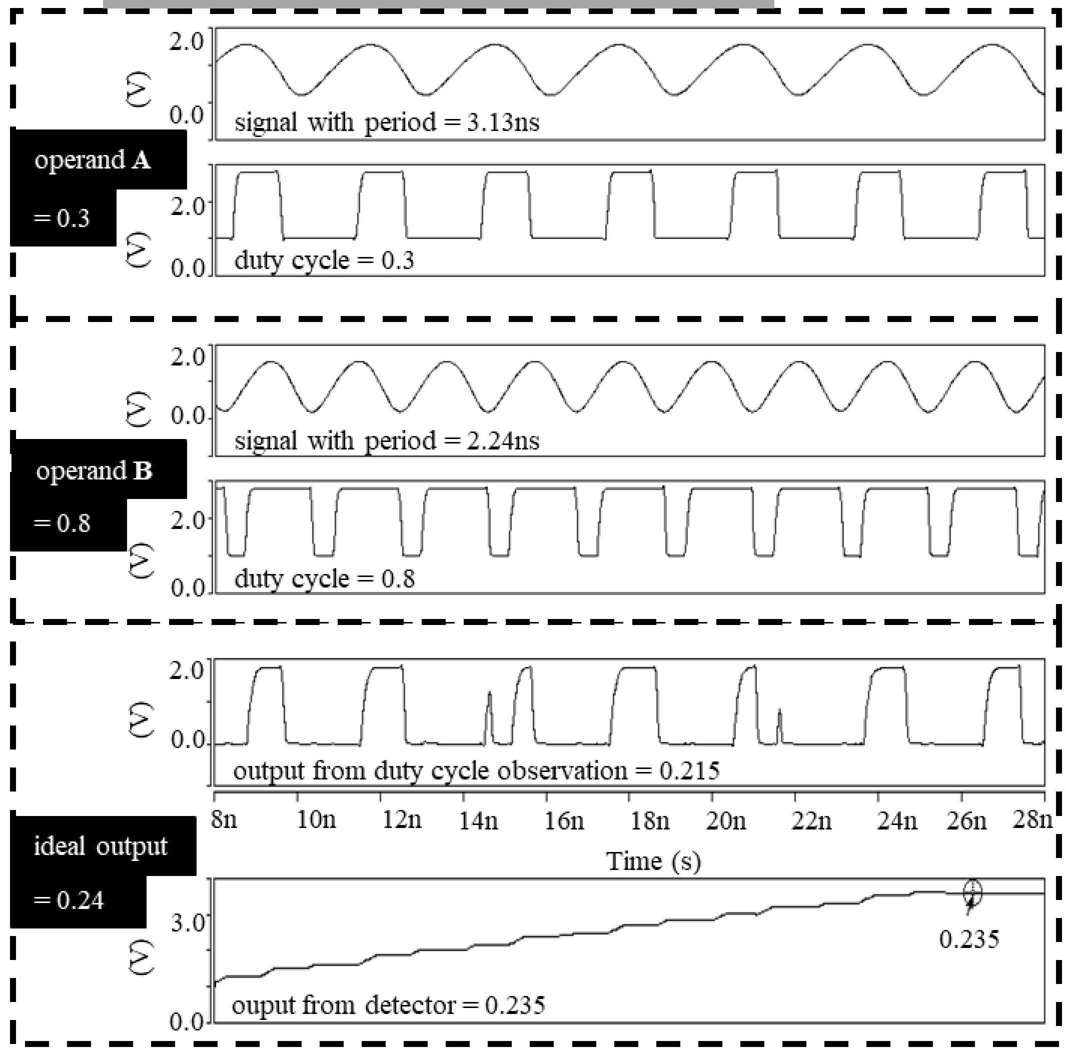
Tune Freq.s



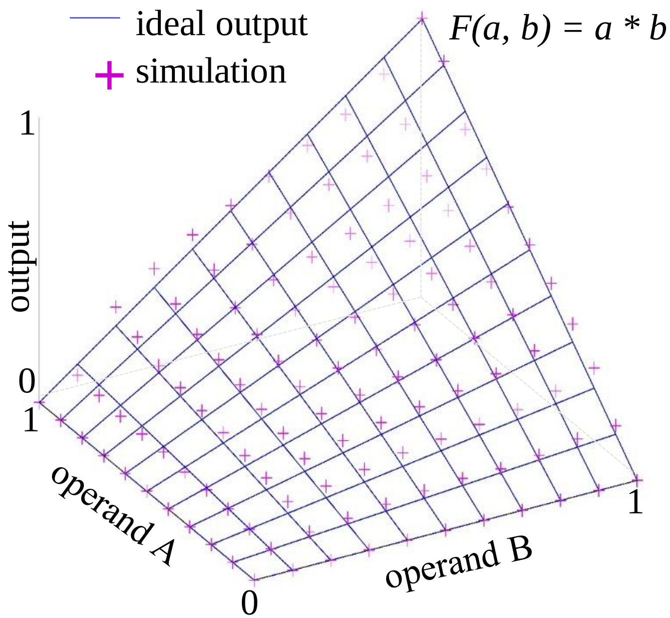
Tune duty-cycle

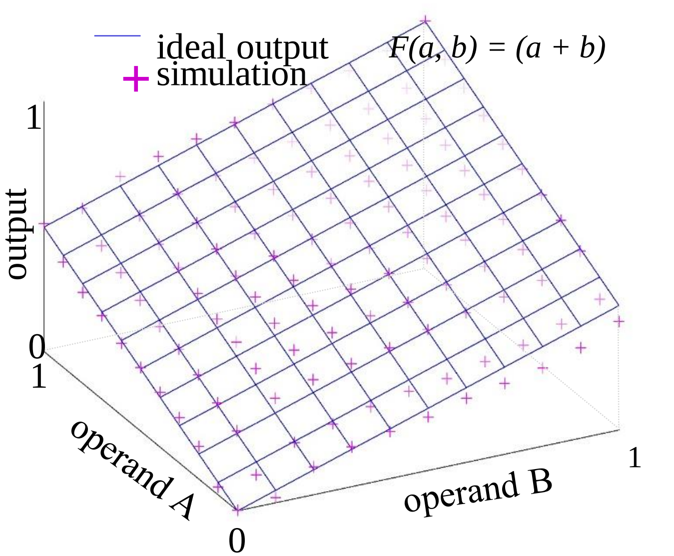


Function ➔ 0.3*0.8 = 0.24



Time-Based Stochastic Computing





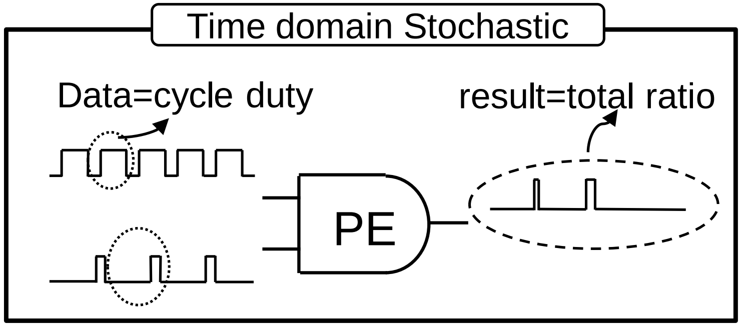
	*	Proposed Circuit
Technology	45nm	180nm
Strategies	time-based values	time-based values
Components (of SNG)	Comparator Ramp Generator Clock Generator	Current-starved Oscillator Neuron-MOS Inverter PWM Detector (optional)
Input	Analog Current	Analog Voltage
Speed (ns)	7 (mul.) 7 (add.)	7 (mul.) 7 (add.)
Accuracy (%)	98.6 (mul.) 98.6 (add.)	96.6 (mul.) 96.7 (add.)
# of trans.	967 (mul.) 1512 (add.)	140 (mul.) 210 (add.)
Energy (pJ)	4.5 (mul.) 6.8 (add.)	1.8 (mul.) 2.5 (add.)

[*] H. Najafi et al., Time-encoded values for highly efficient stochastic circuits, IEEE Trans. VLSI Systems 2017

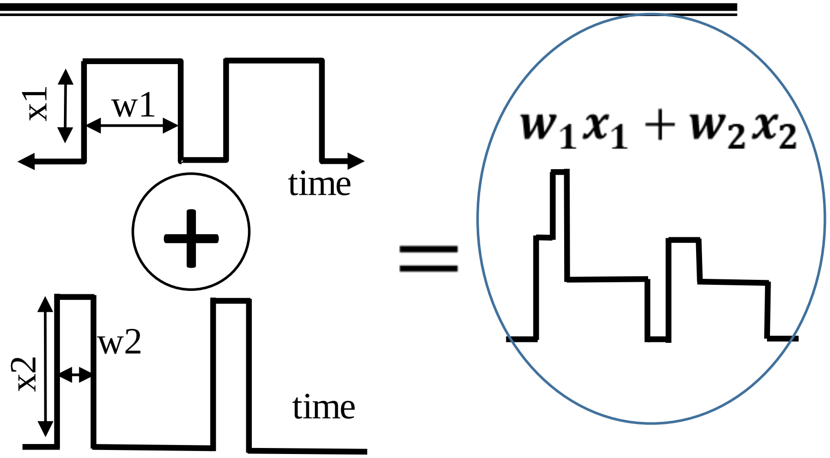
Outline

- What is stochastic computing (SC)
  - Mechanism of SC
  - Elements of SC
  - Implementation of SC
- Time based stochastic computing (TBSC)
  - Mechanism of TBSC
  - Hybrid TBSC
  - Analysis

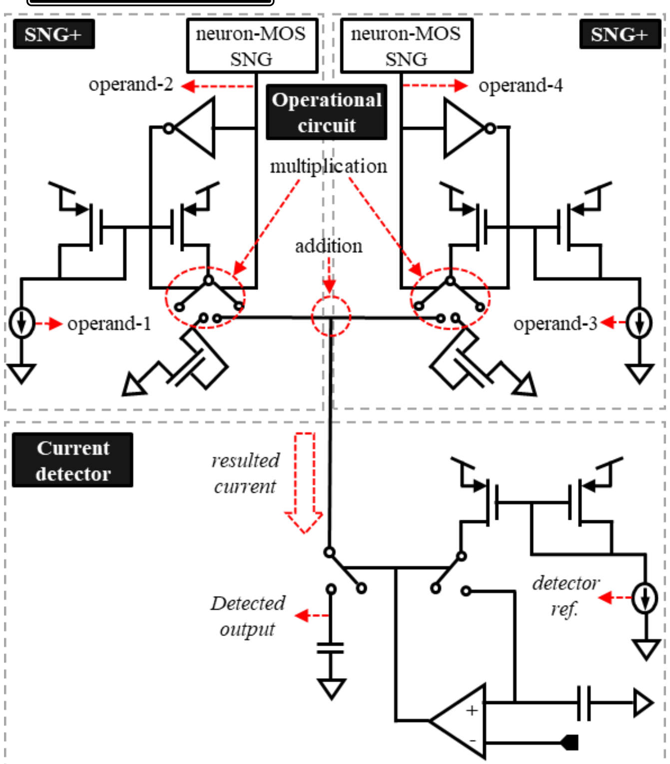
Hybrid: stochastic + analog



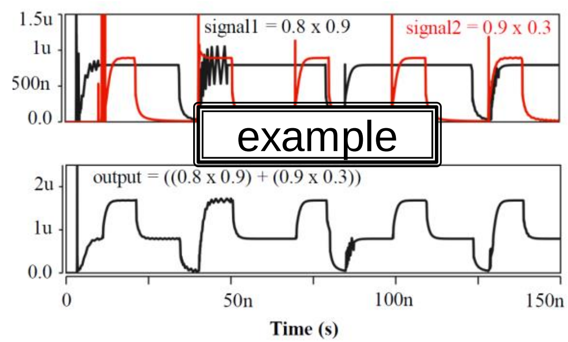
Not only period
But also strength

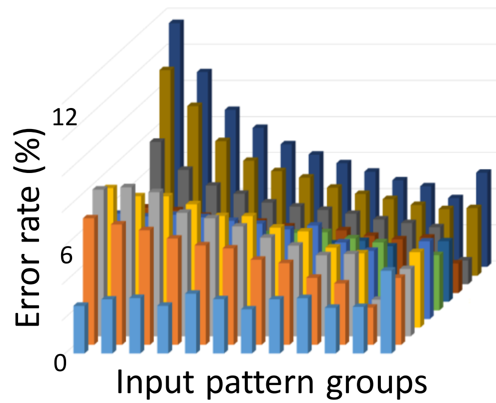


Circuits

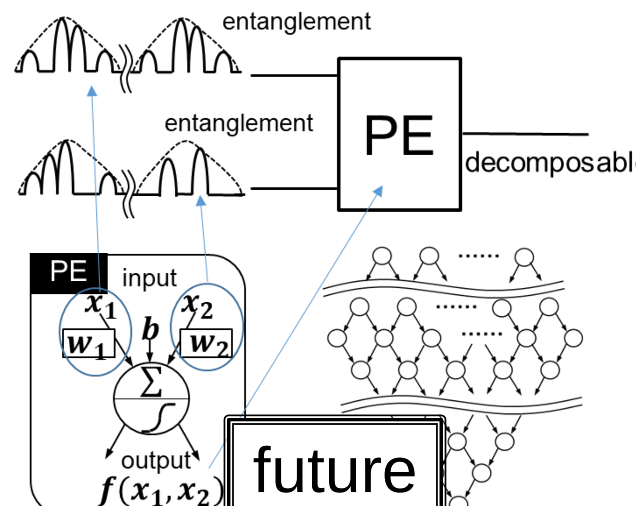


example



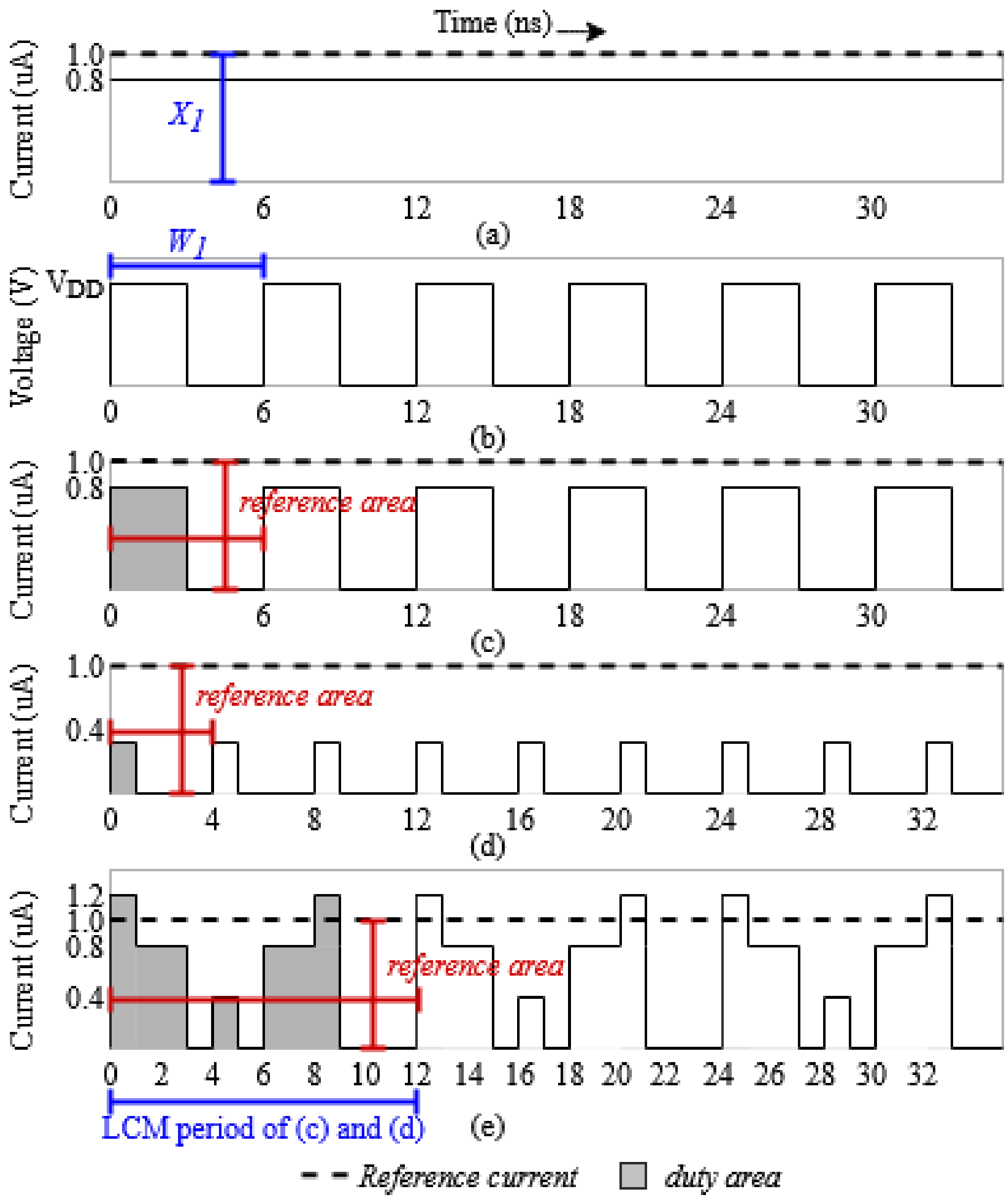


Very possible to
implement for DiaNet



future

Hybrid: stochastic + analog



Key idea:
In stochastic computing, we concern ``probability”. Then, why always use discrete fashion?

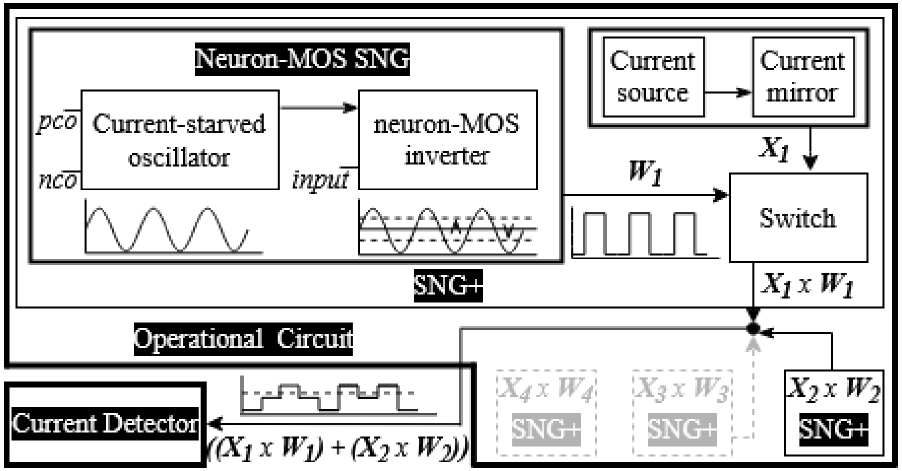
Hybrid: use continuous probability distribution instead of discrete Bernoulli test; integral instead of bit counting

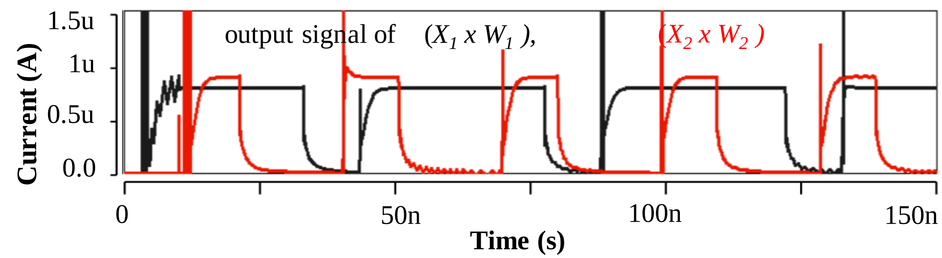
Merits: short time; infinite range; easy for summation; light SNG

Demerit: almost no theory; circuit design expertise

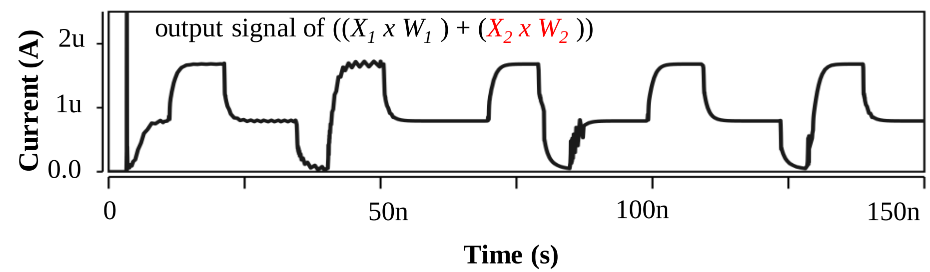
Hybrid: stochastic + analog

implementation

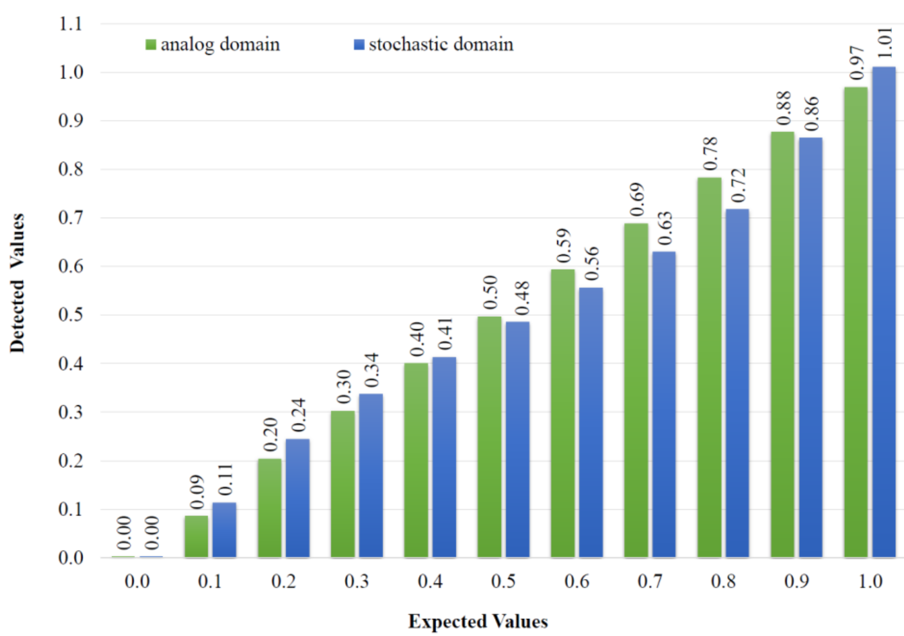


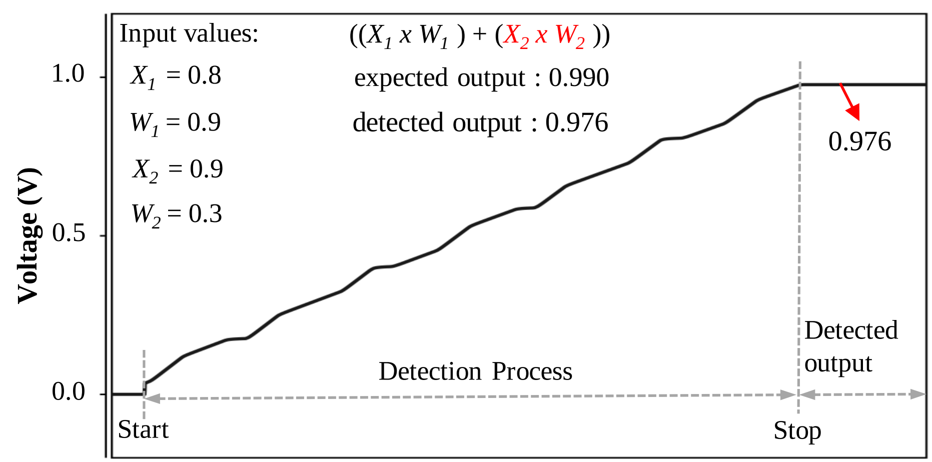


(a)



(b)





(c)

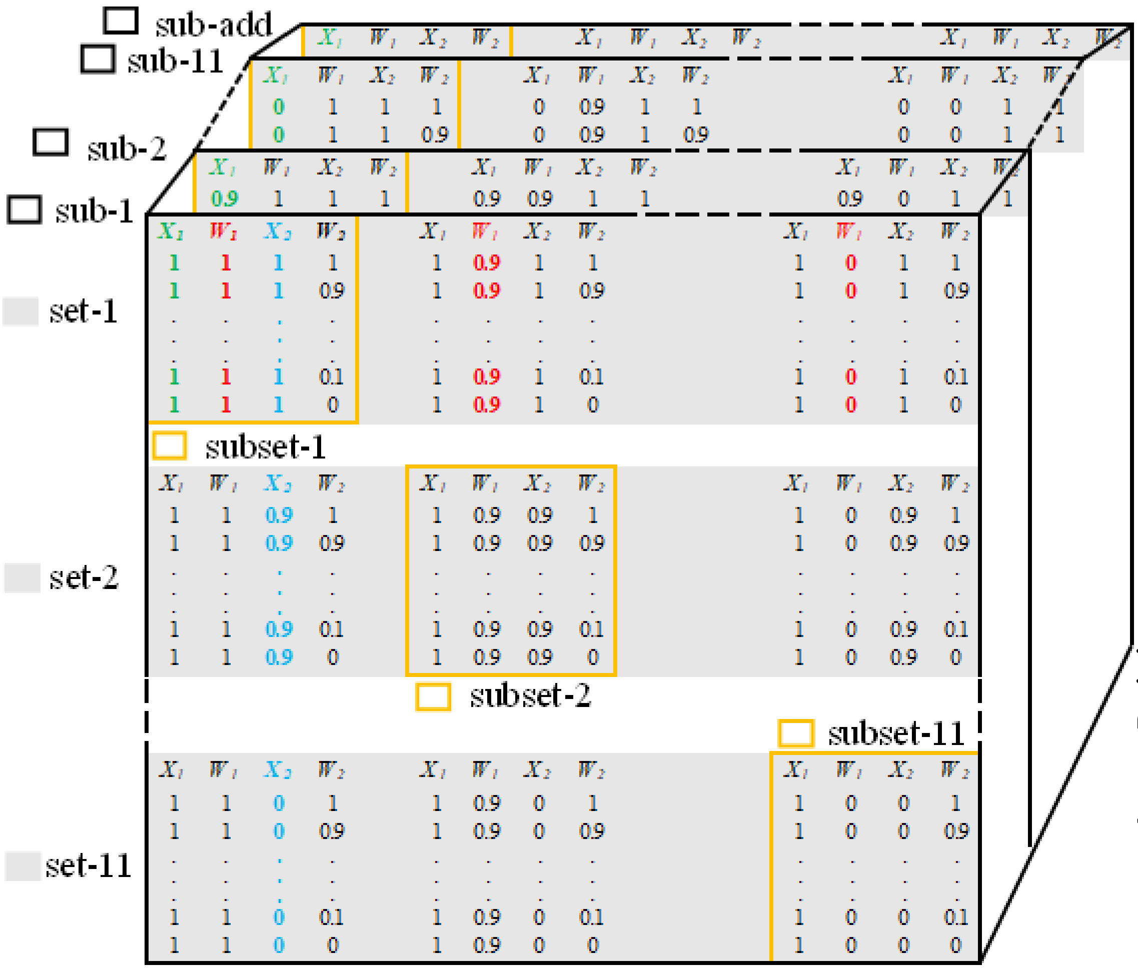
Outline

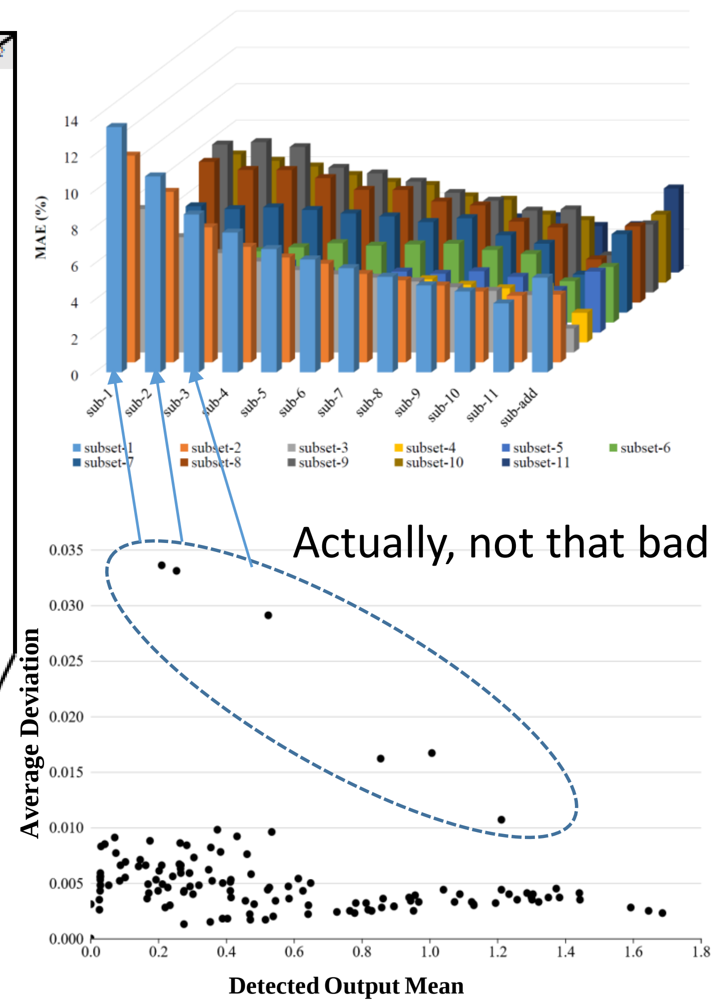
- What is stochastic computing (SC)
  - Mechanism of SC
  - Elements of SC
  - Implementation of SC
- Time based stochastic computing (TBSC)
  - Mechanism of TBSC
  - Hybrid TBSC
  - Analysis

Hybrid: stochastic + analog

Error analysis

X1*w1+X2*w2 contains four operands, full pattern test is impossible. Thus, sampling.

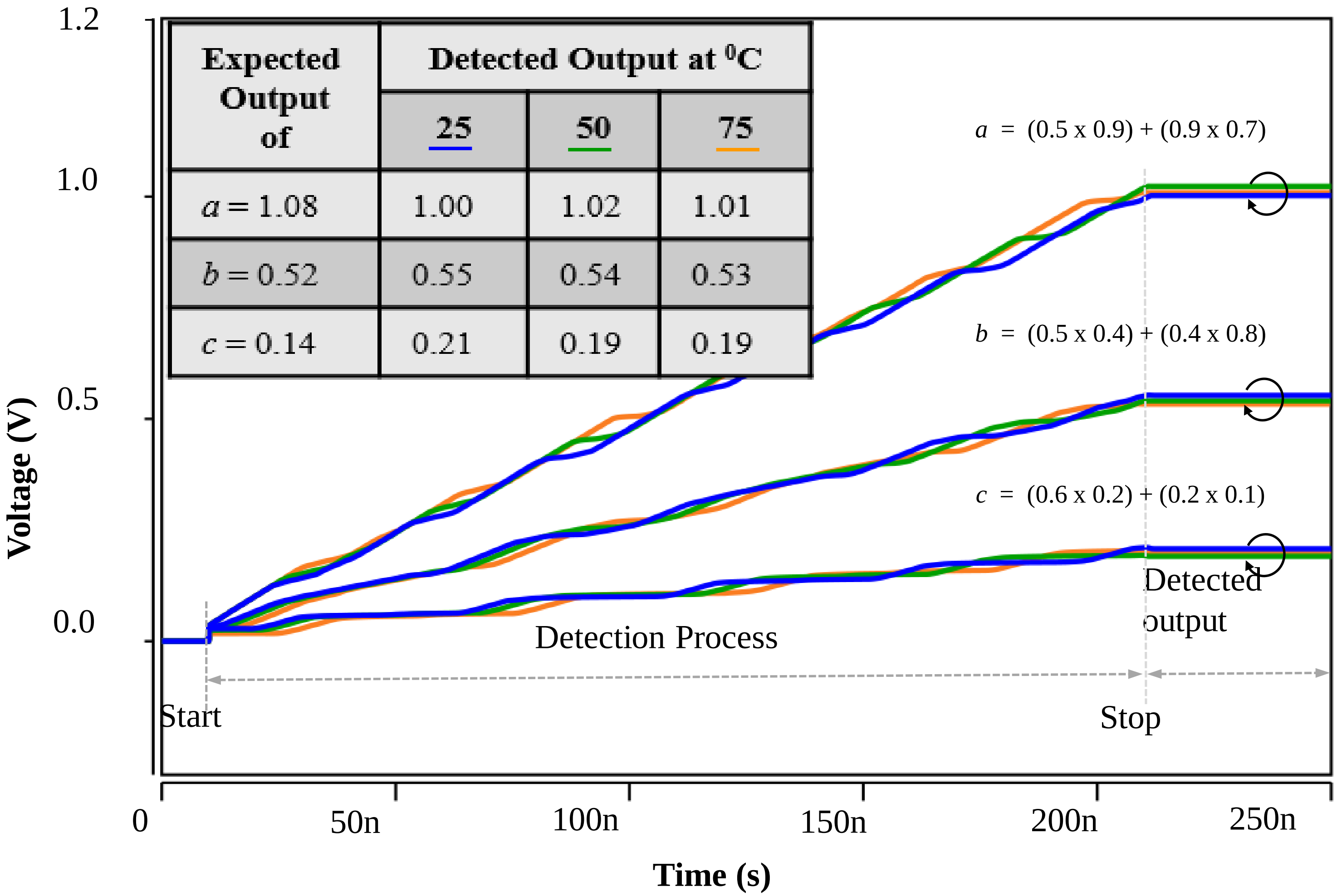




Hybrid: stochastic + analog

Error analysis

Errors also come from the analog side: process and temperature variations etc.

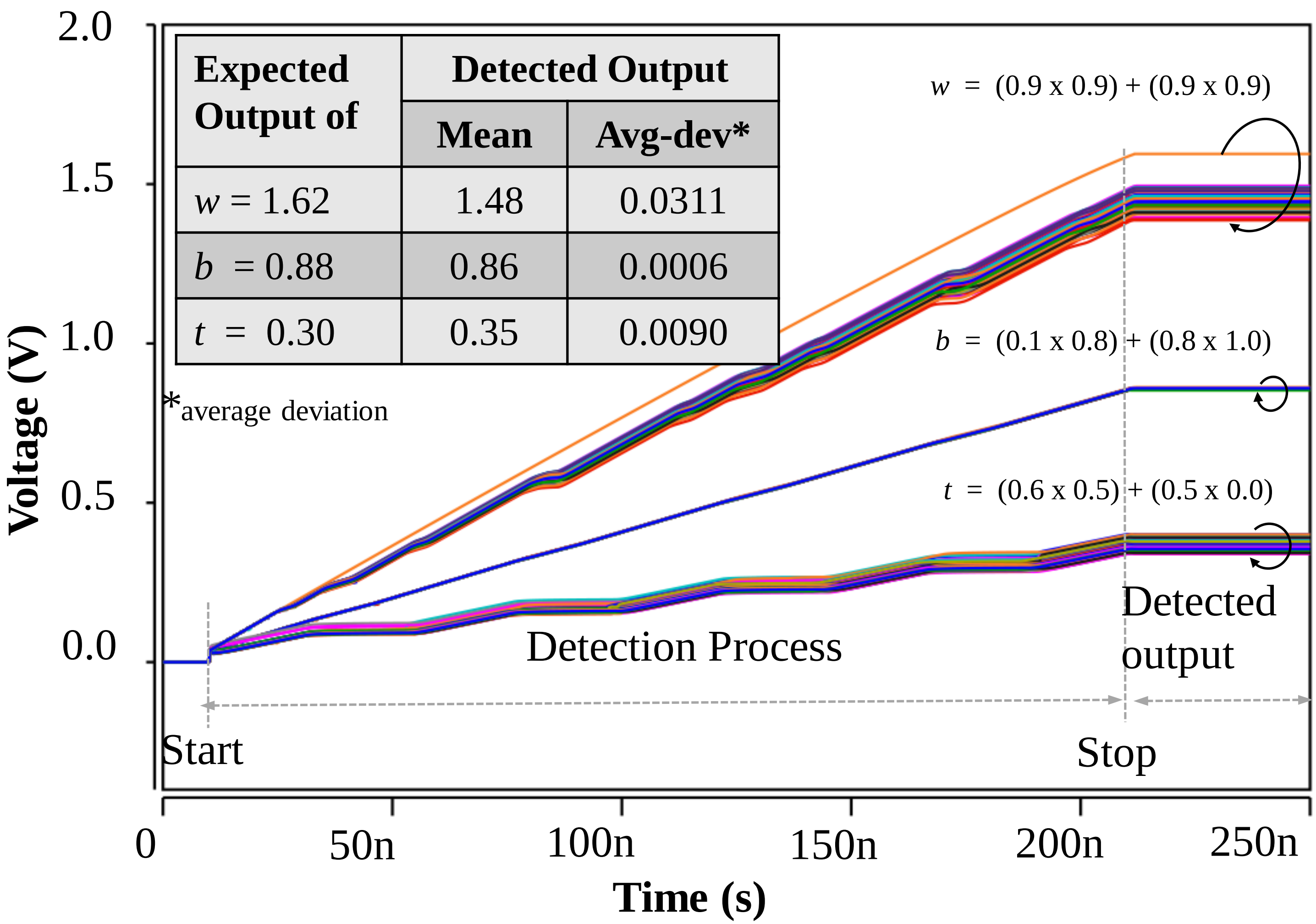


Expected Output of	Detected Output at ⁰C		
	25	50	75
$a = 1.08$	1.00	1.02	1.01
$b = 0.52$	0.55	0.54	0.53
$c = 0.14$	0.21	0.19	0.19

Hybrid: stochastic + analog

Error analysis

Errors also come from the analog side: process and temperature variations etc.



Expected Output of	Detected Output	
	Mean	Avg-dev*
w = 1.62	1.48	0.0311
b = 0.88	0.86	0.0006
t = 0.30	0.35	0.0090

Thousands of trials for each by Monte Carlo

Much better than pure analog

inaccuracy is eaten up by SN error-tolerance somehow

End

Thank you very much.