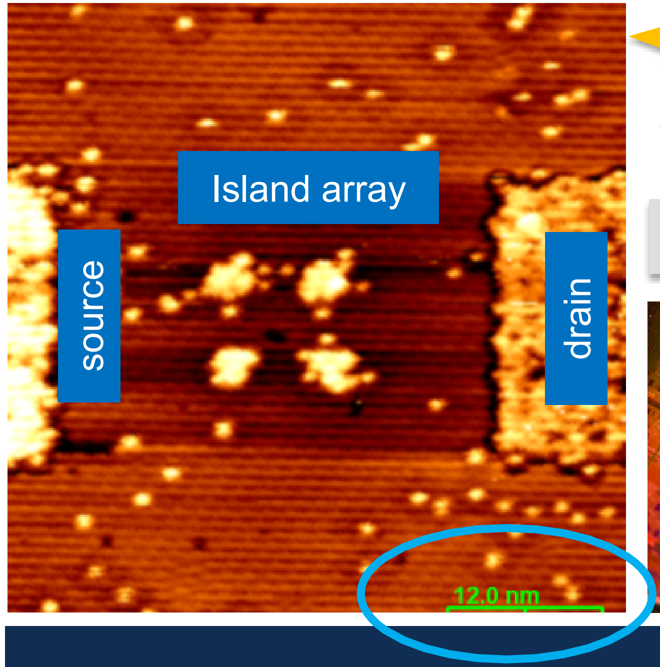
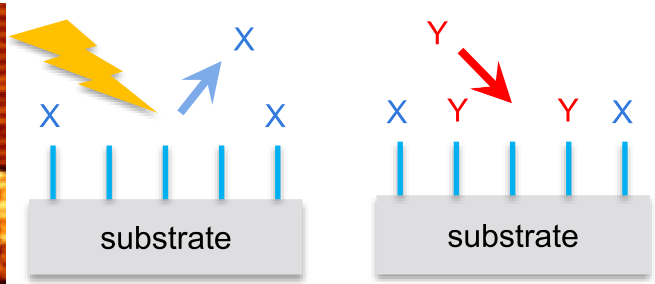
SAND2021-8477PE

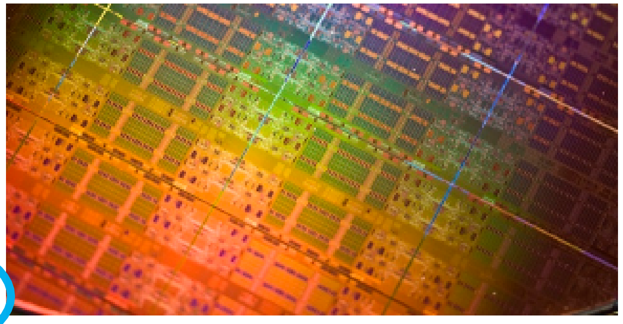
CO-designed Improved Neural Foundations Leveraging Inherent Physics Stochasticity (COINFLIPS)

Exceptional service in the national interest

Sandia National Laboratories







Shashank Misra

Fourth EAB Meeting

July 21, 2021

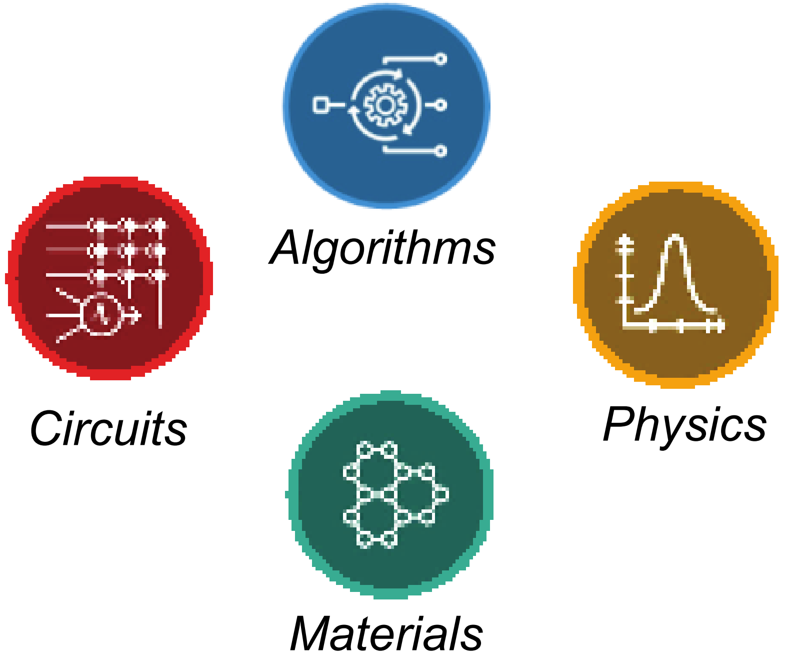
U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Is probabilistic computing a useful paradigm?

DOE-SC Call: Multiscale Codesign

BES, ASCR, HEP, FES, NP

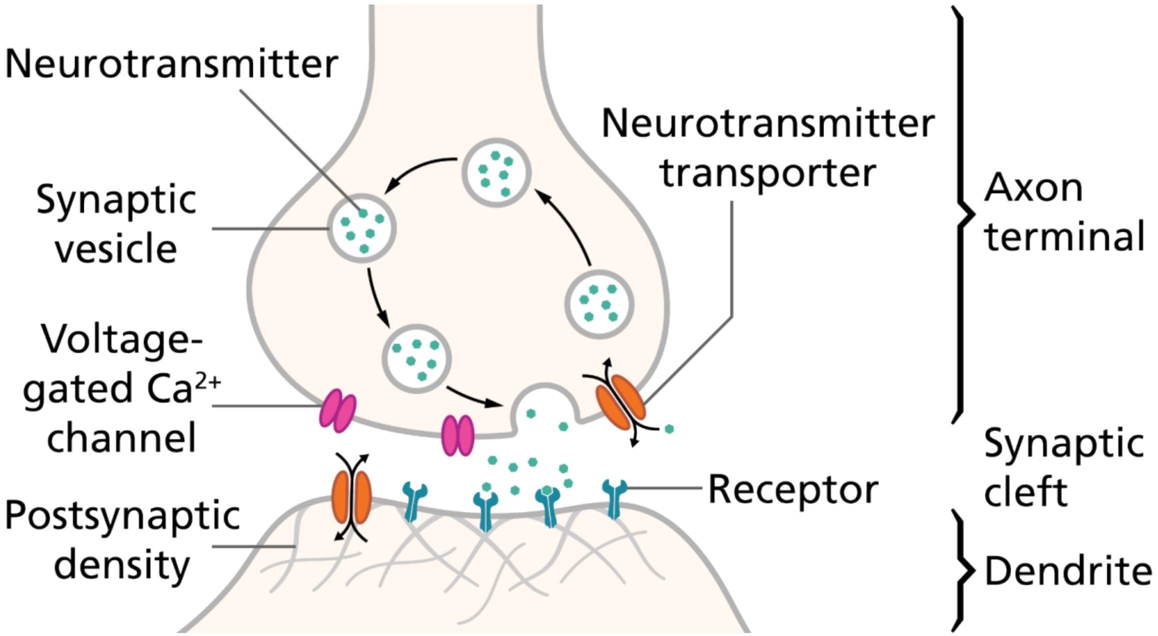


New computing paradigms and architectures

New materials, chemistry, synthesis, and fabrication

Inspiration from Biology:

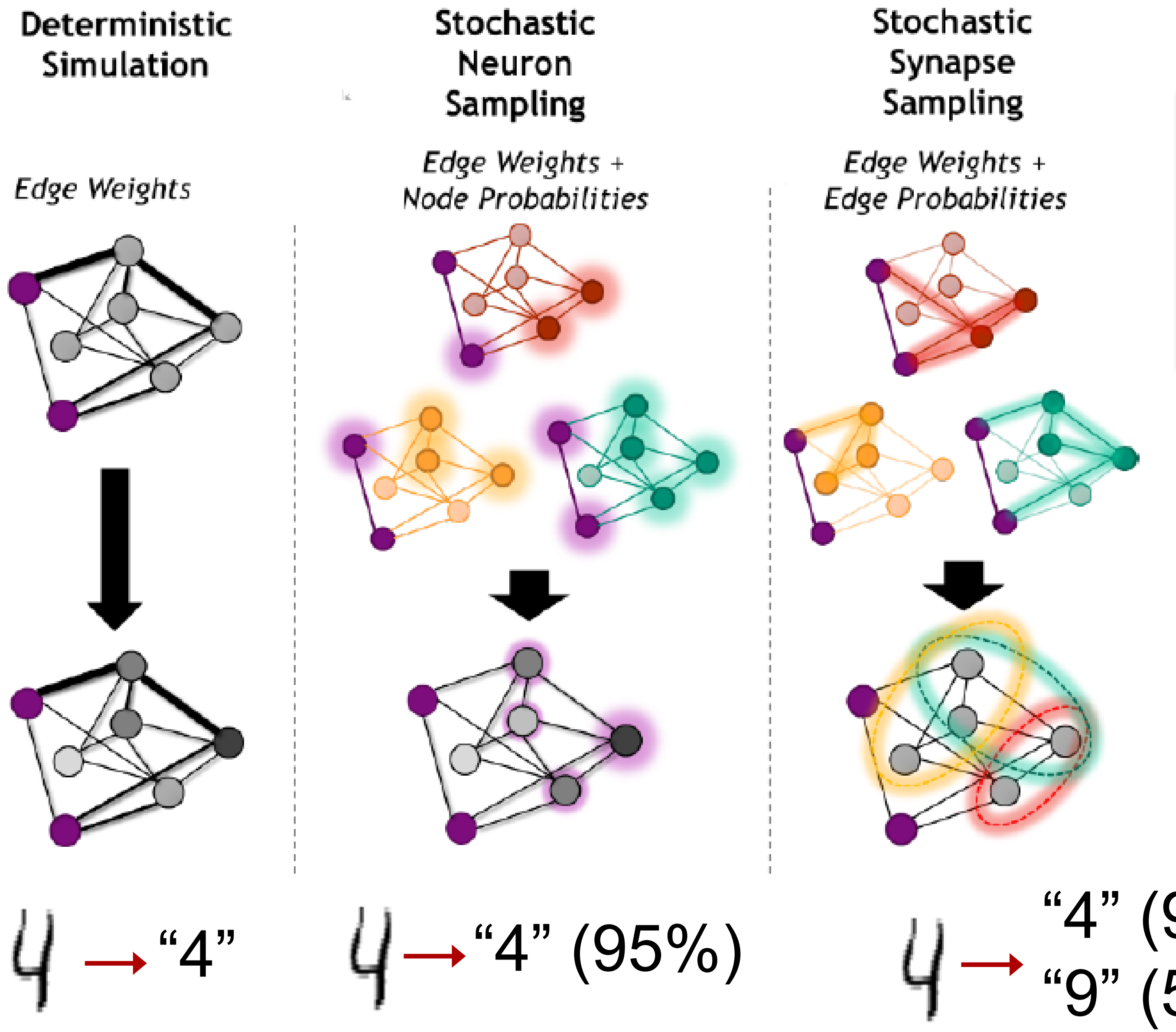
Brain has 10^{15} probabilistic synapses, firing at a rate > 1 Hz

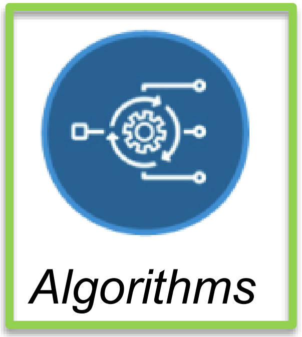


Idea development – HAANA grand challenge, SNL works

How can randomness be incorporated into artificial neural networks (ANNs)?

Stochasticity in ANNs will provide new functionality



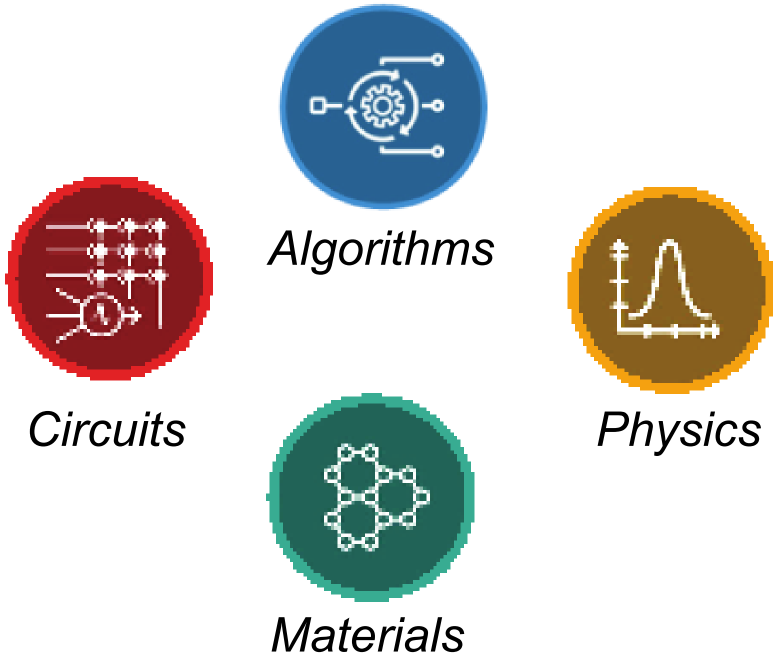


COINFLIPS Thrusts

Probabilistic computing

Formalizing the paradigm of probabilistic computing

Applications of probabilistic computing



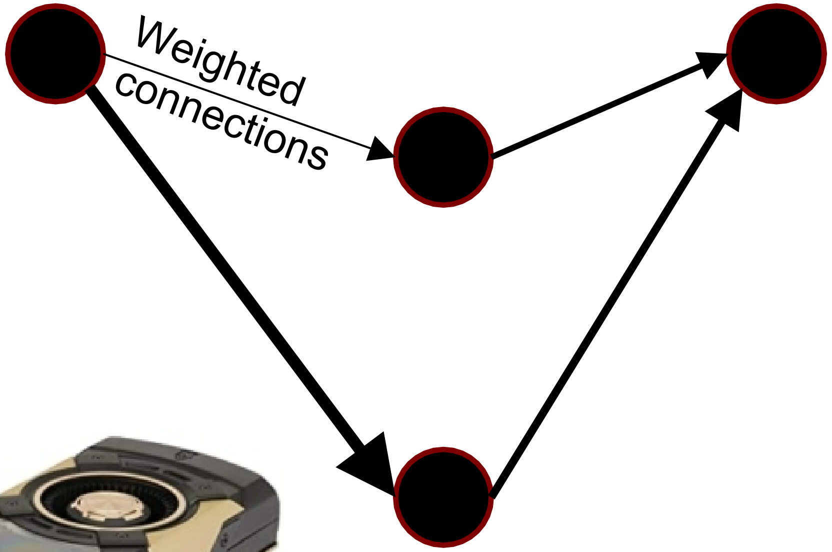
Linking probabilistic circuits with neural architectures

Achieving high quality, ubiquitous probabilistic devices

Multiscale codesign

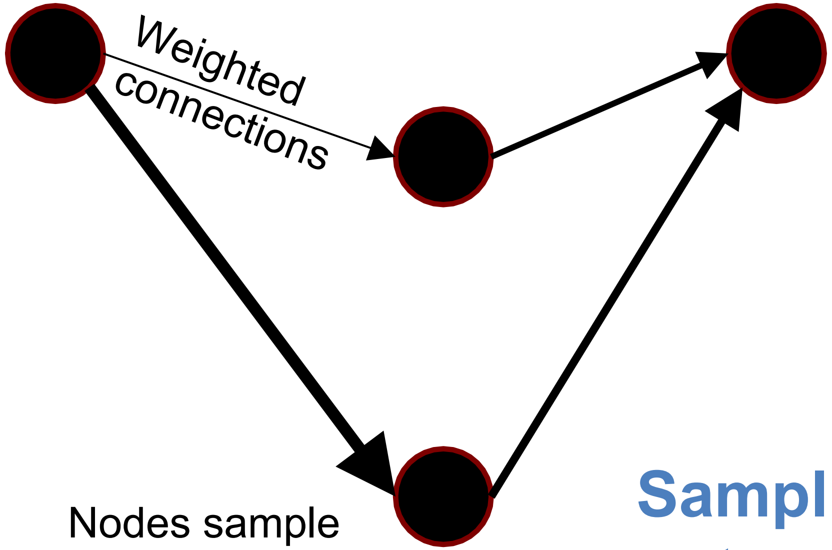
What is the state of the art?

Artificial neural networks developed to utilize linear algebra, which GPUs are well-matched to run.

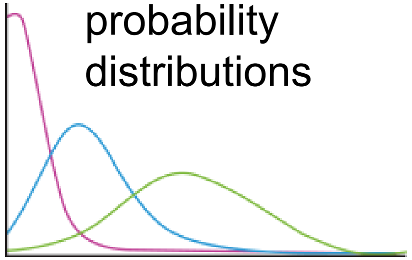


GPU: fast linear algebra

Probabilistic computing – no algorithm, no hardware



Nodes sample probability distributions



Sampling is untenable in software:
> 200 kW for 1 Gb/μs

How do generate samples of a distribution function?

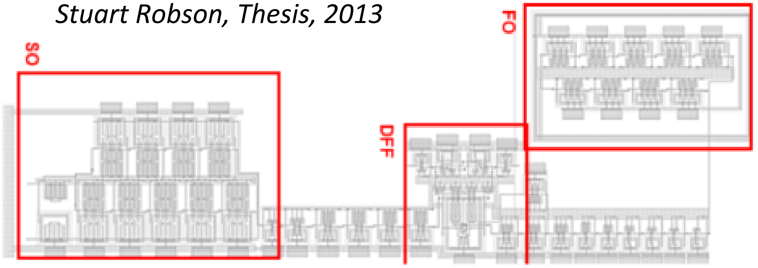
Random number generator (RNG) → Sampling application-relevant distribution

Device faithfully produces 50:50 weighted coinflip

Poisson distribution

Ring oscillator RNG

Stuart Robson, Thesis, 2013



1 Gb/μs = all of a 6 inch wafer

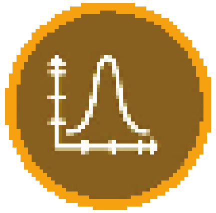
Perfect coinflips are expensive

Expensive bit of math

Requires a few thousand DSPs!

Transforming them is expensive

Algorithms



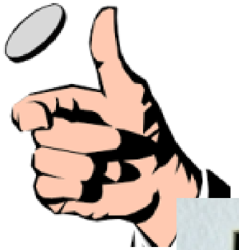
Physics

Circuits

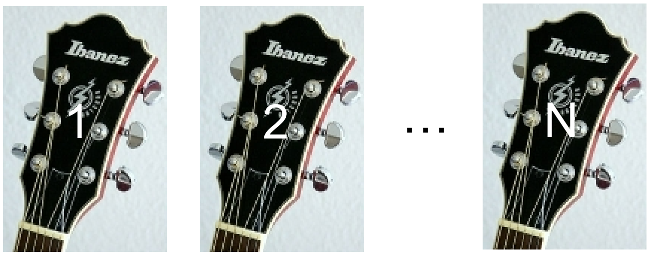
Staggering inefficiency for algorithmic simplicity

Codesign an efficient path linking RNG and probability distributions

Tunable RNG

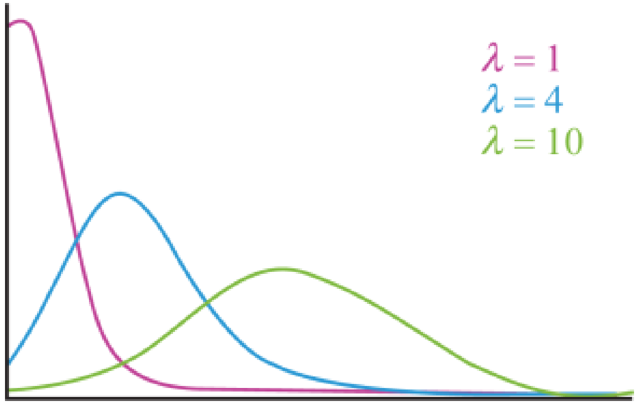


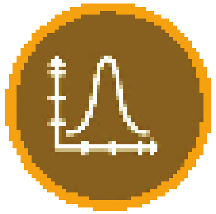
Game theory approach



Repeat draws tune internal degrees of freedom

Probability distributions





Physics

Circuits

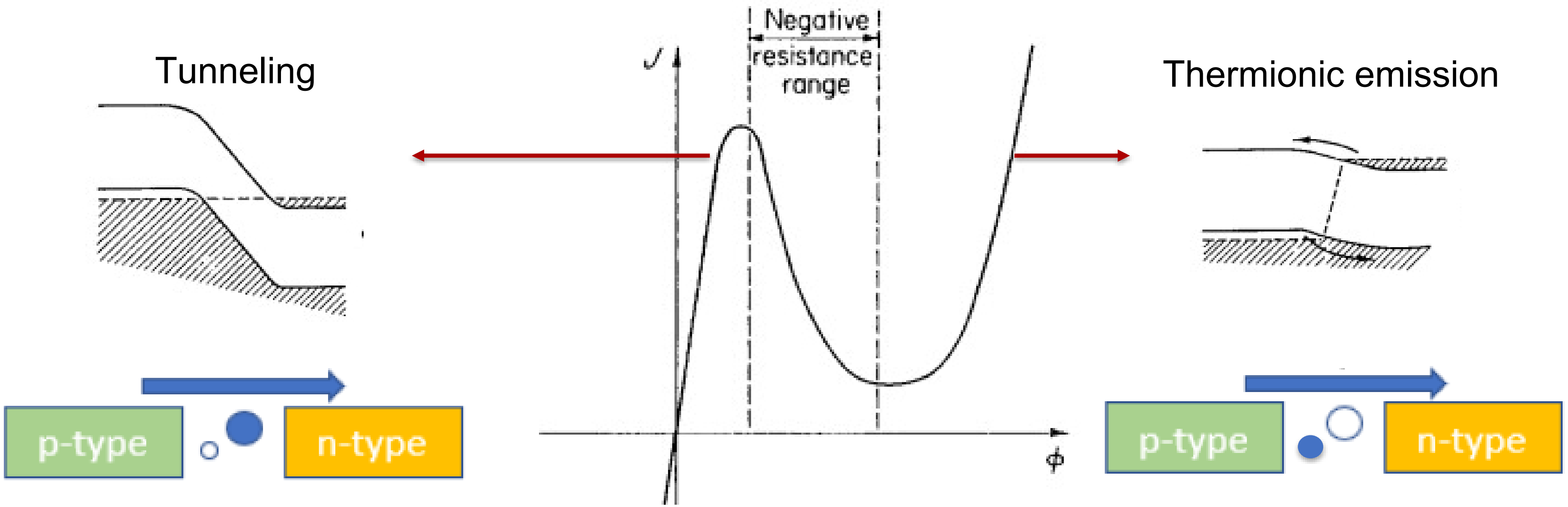
Discover reinforcement learning method to map between tunable RNG and probability distribution

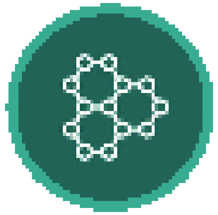
How do I make an efficient and tunable RNG? Leverage physics of device.

Efficient, tunable RNG #1: APAM Tunnel diode

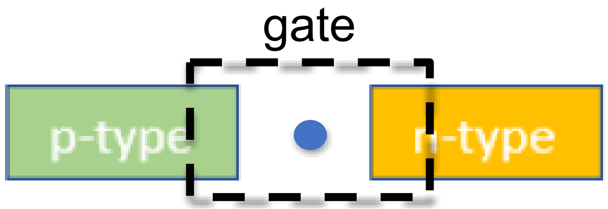
Sandia National Laboratories

RNG from randomness of defects. Diode is an efficient detector.



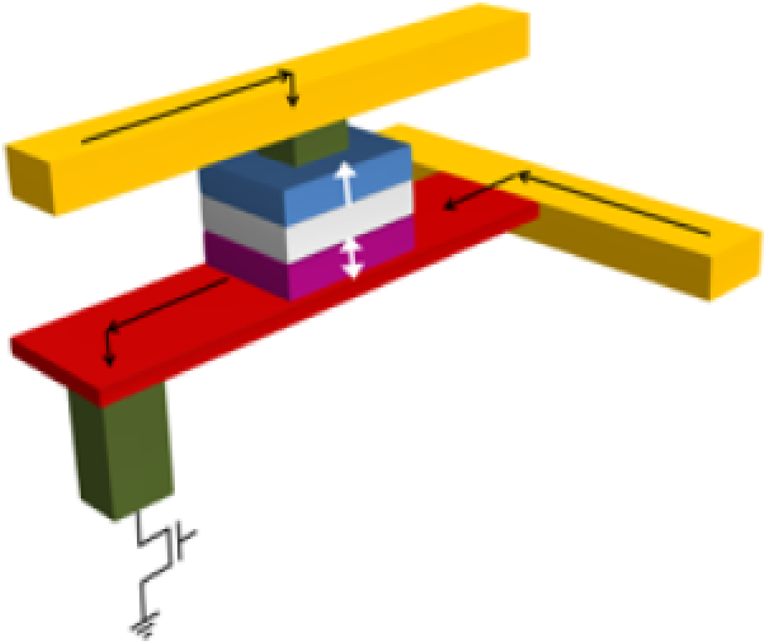


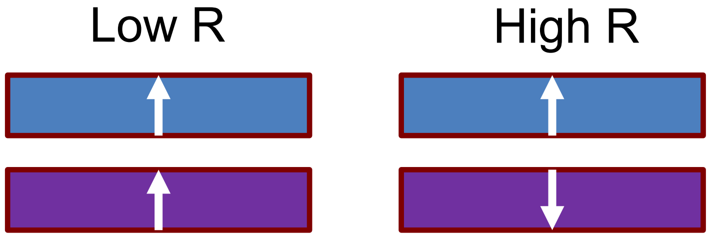
APAM: put a single defect in barrier, use gate to change its state

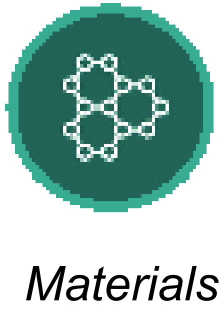


What parameters control RNG speed and tunability?

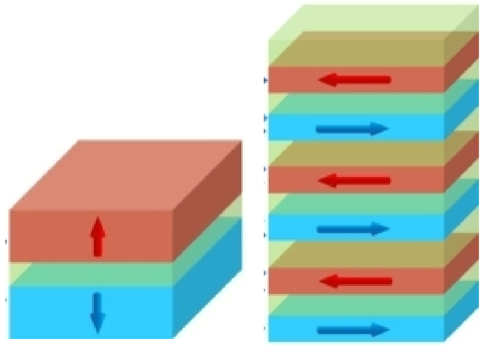
Efficient, tunable RNG #2: Magnetic tunnel junction



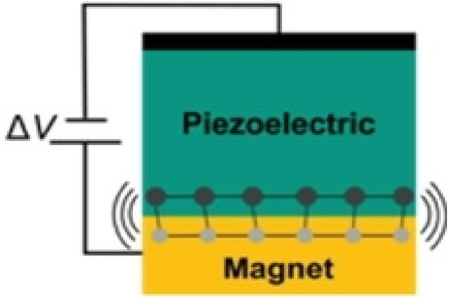




Materials exploration – speed



Piezoelectric – tuning

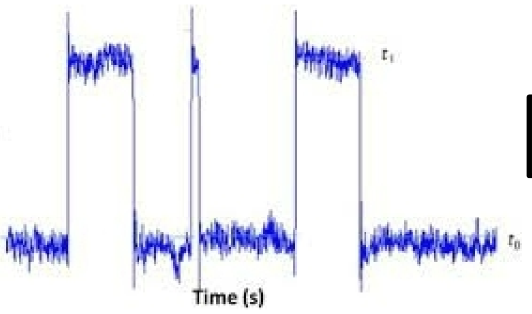


What are the fitness criteria for devices?

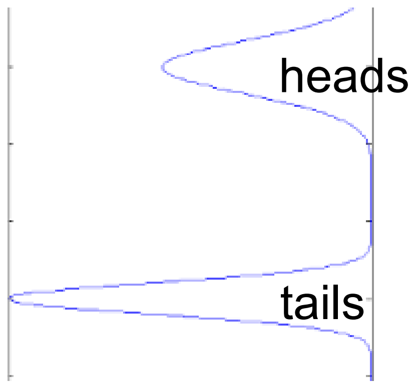
p-type

n-type

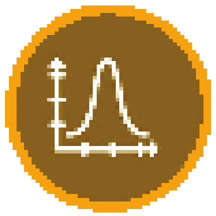
Efficient, nanoscale current-biased tunnel diode



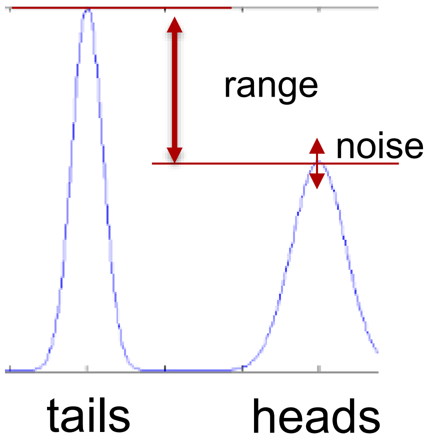
Telegraph noise



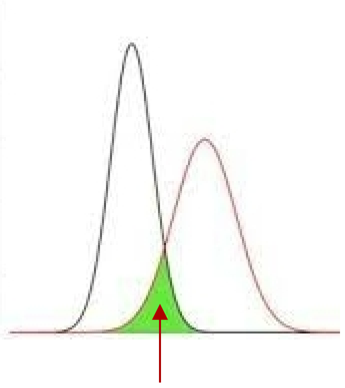
Distribution of outcomes



Physics



Vs.



Fundamental question: what aspects of fidelity, noise, etc. are important for sampling?

COINFLIPS Proposal

National lab call

$2 M/yr for 3 yr
$500 k/yr for APAM

Algorithms

Circuits

Physics

Materials

Larger DOE call anticipated in future