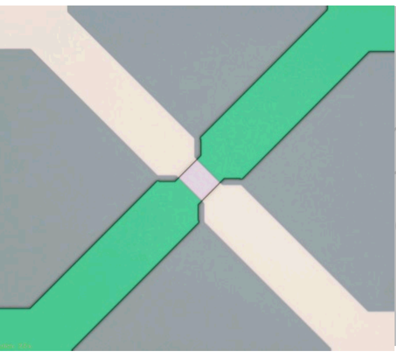
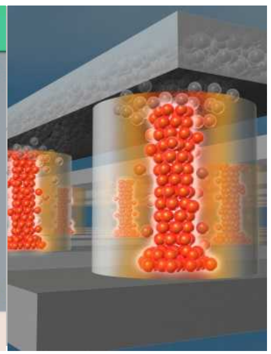
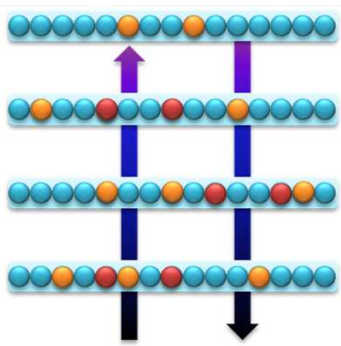
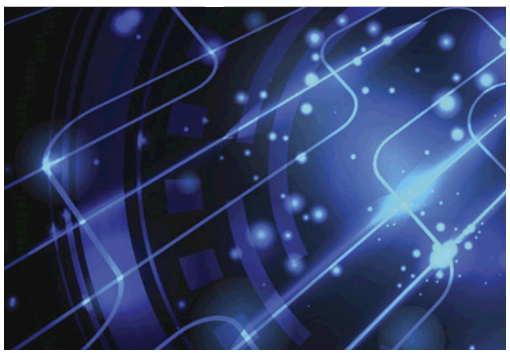
SAND2016-10652C

Exceptional service in the national interest

Sandia National Laboratories

Implementation of a Liquid State Machine with Temporal Dynamics on a Novel Spiking Neuromorphic Architecture

Michael R. Smith[1], Aaron Hill[1], Kristofer D. Carlson[1], Craig M. Vineyard[1], Jonathon Donaldson[1], David R. Follett[2], Pamela L. Follett[2,3], John H. Naegle[1], Conrad D. James[1], James B. Aimone[1]
[1]Sandia National Laboratories, [2]Lewis Rhodes Labs, [3]Tufts University

U.S. DEPARTMENT OF ENERGY   NNSA National Nuclear Security Administration   LDRD LABORATORY DIRECTED RESEARCH & DEVELOPMENT

Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. XXXXXXXXXX

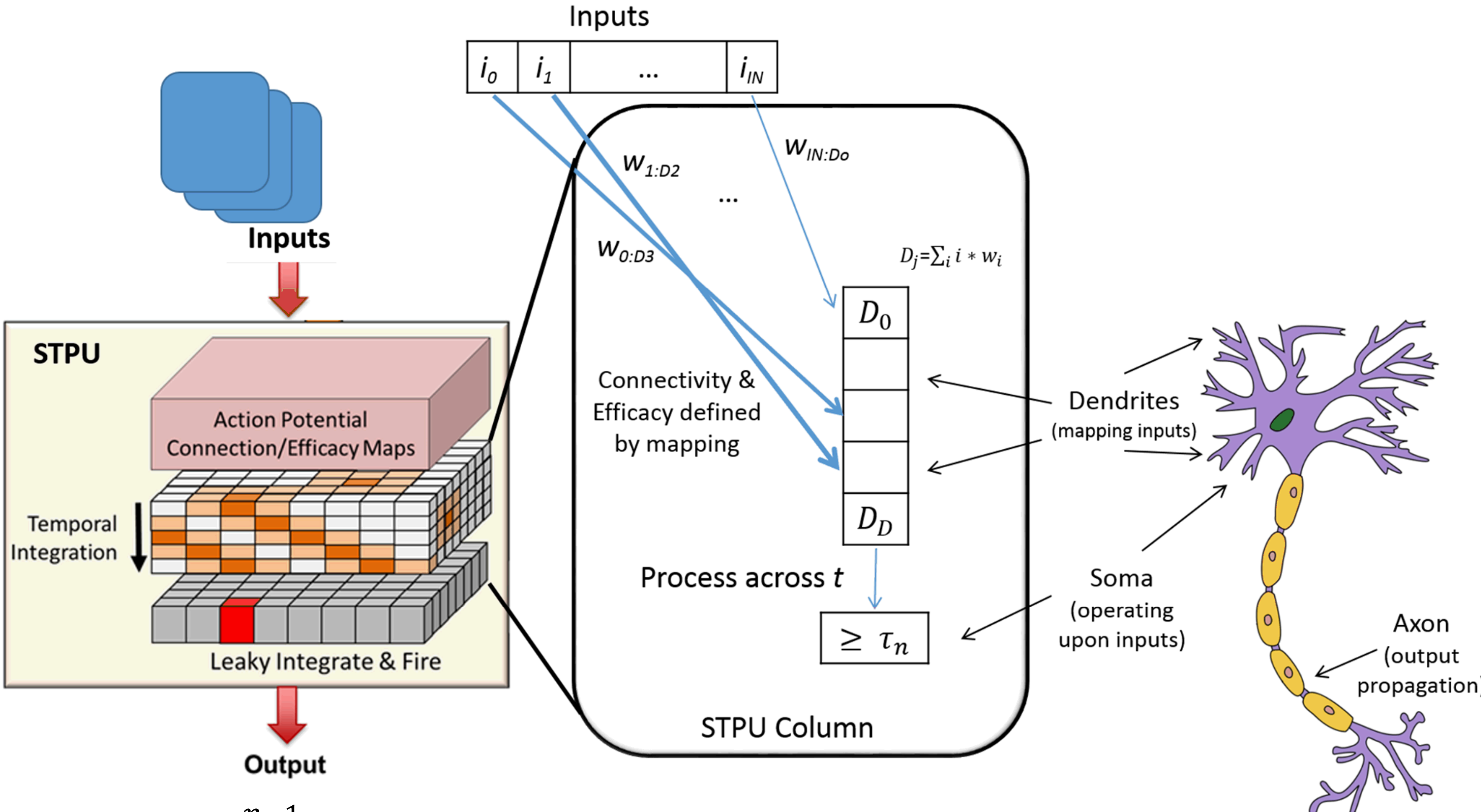
Outline

- **A novel neuromorphic architecture: the Spiking Temporal Processing Unit**
- A neuro-inspired algorithm for the neuromorphic computing: the Liquid State Machine
- Mapping the LSM onto the STPU
- Lessons learned and future development

Spiking Temporal Processing Unit

- Neuro-inspired
  - Brain innately parallel
  - Simple computational units (neurons)
  - Functionality is encoded via:
    - Sparse connectivity
    - Unique efficacies
    - Temporal latencies
    - Binary spikes
- Fast
- Lower power consumption
- High band in and out
- Spiking Temporal Processing Unit (STPU)
  - Composed of leaky-integrate and fire (LIF) spiking neurons
  - Each LIF has a associated temporal stack

Spiking Temporal Processing Unit

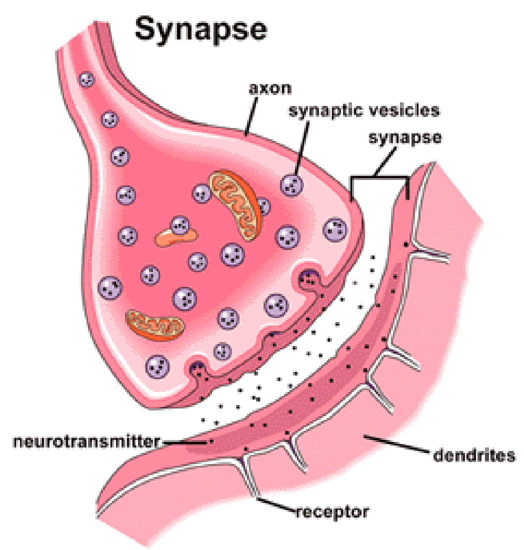


$$v_m^n = v_m^{n-1} - \frac{v_m^{n-1}}{\tau_m} + \sum_i \sum_j w_{mi} \cdot s(t - t_{ij} - d_i)$$

Spiking Temporal Processing Unit

- Synapses encoded in efficacy mapping
- Signals along axon are digital (binary spikes)
- Synapses are analog and have exponential behavior
- Generally ignored in other neuromorphic architectures
  - Expensive to implement in hardware
- Can be implemented via the temporal stack

Input	Neuron	Δt	Efficacy	
1	4	3	3	
1	7	1	5	
1	7	2	7	
2	6	1	5	
…	…	…	…	



Exponential Synaptic Response Functions in Digital Hardware

$$v_m^n = v_m^{n-1} - \frac{v_m^{n-1}}{\tau_m} + \sum_i \sum_j w_{mi} \cdot s(t - t_{ij} - d_i)$$

- Static

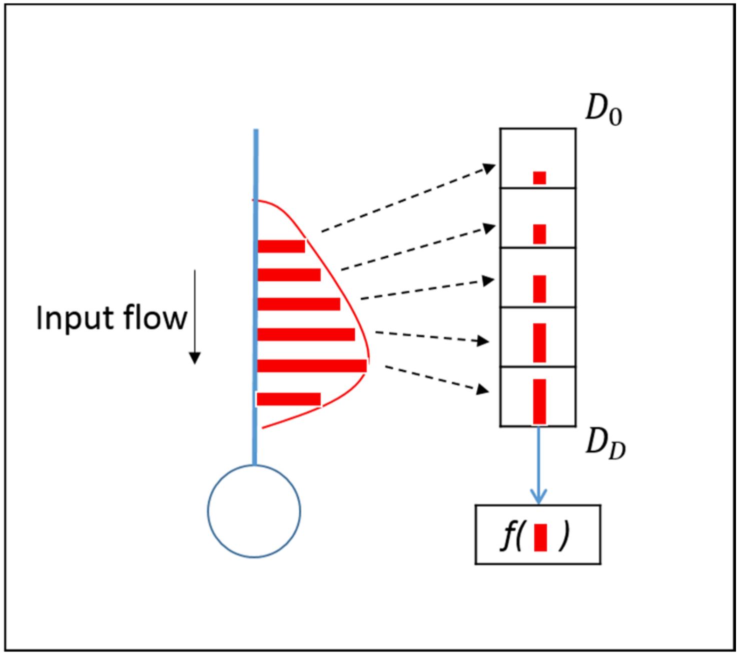
$$\delta(t - t_{ij} - d_j)$$

- First-order response

$$\frac{1}{\tau_m} e^{-\frac{t - t_{ij} - d_j}{\tau^s}} \cdot H(t - t_{ij} - d_j)$$

- Second-order response

$$\frac{1}{\tau_1^s - \tau_2^s} \left(e^{-\frac{t - t_{ij} - d_j}{\tau_1^s}} - e^{-\frac{t - t_{ij} - d_j}{\tau_2^s}}\right) \cdot H(t - t_{ij} - d_j)$$



Neuromorphic Architectures

True North

- Specific neuron implementations
- Only communicate with a single neuron (1 to 1 mapping)
- Low power

SpiNNaker

- Neuron implementation is more flexible (basic instructions)
- Routing table (1 to many mapping)
- Fast

Neuromorphic Architectures

True North

- Specific neuron implementations
- Only communicate with a single neuron (1 to 1 mapping)
- Low power

SpiNNaker

- Neuron implementation is more flexible (basic instructions)
- Routing table (1 to many mapping)
- Fast

Spiking Temporal Processing Unit

- Hardware implementation of neurons (specific)
- Routing table enabling 1 to many neuron mapping
- Each neuron has a temporal buffer

Outline

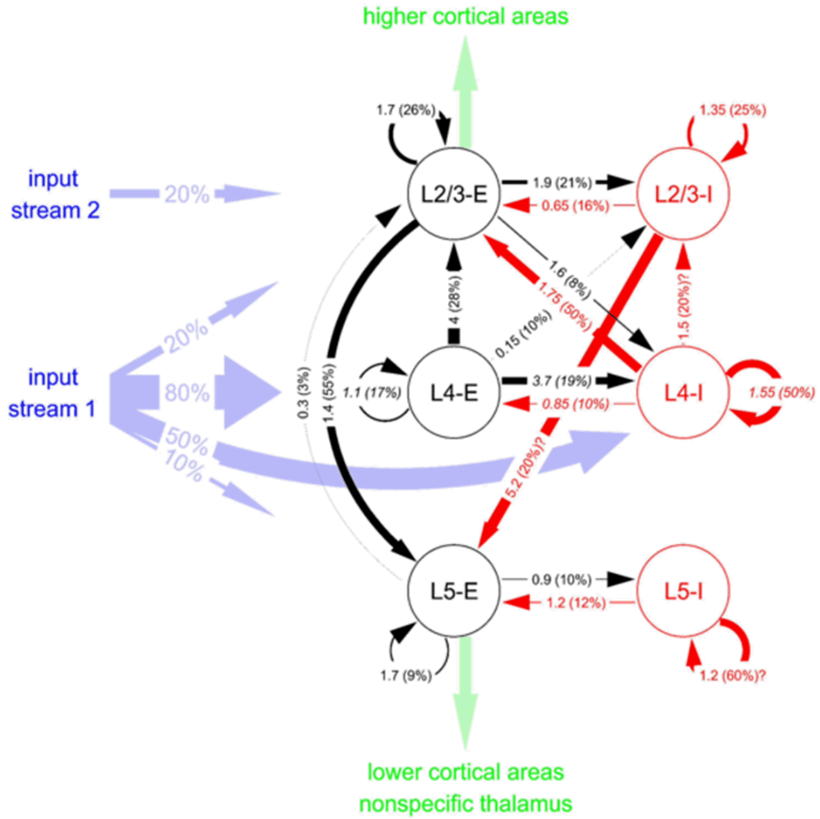
- A novel neuromorphic architecture: the Spiking Temporal Processing Unit
- **A neuro-inspired algorithm for the neuromorphic computing: the Liquid State Machine**
- Mapping the LSM onto the STPU
- Lessons learned and future development

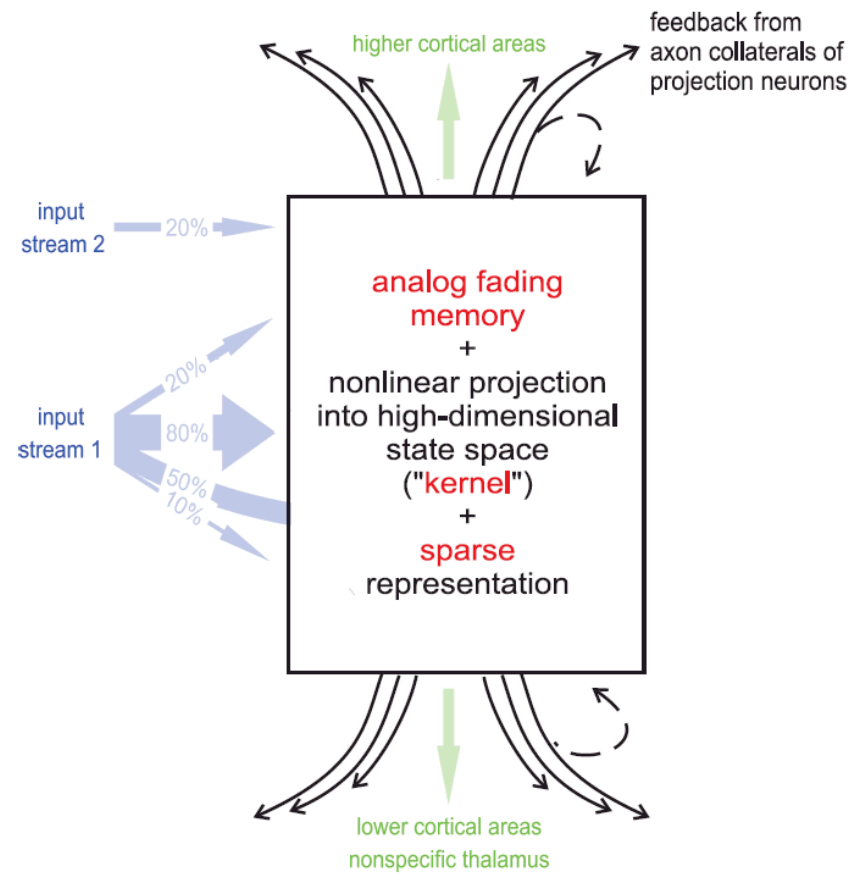
Liquid State Machine

- Developed by Wolfgang Maass
- Reservoir computing
  - Echo State Machines
  - Liquid State Machines
- Different items at different locations at different times
- Differences between the patterns are amplified by the liquid
- Mimics brain functionality

Maass, W., Markram, H., *On the Computational Power of Recurrent Circuits of Spiking Neurons*, Journal of Computer and System Sciences 69(4): 593-616, 2004.

Cortical Microcircuit

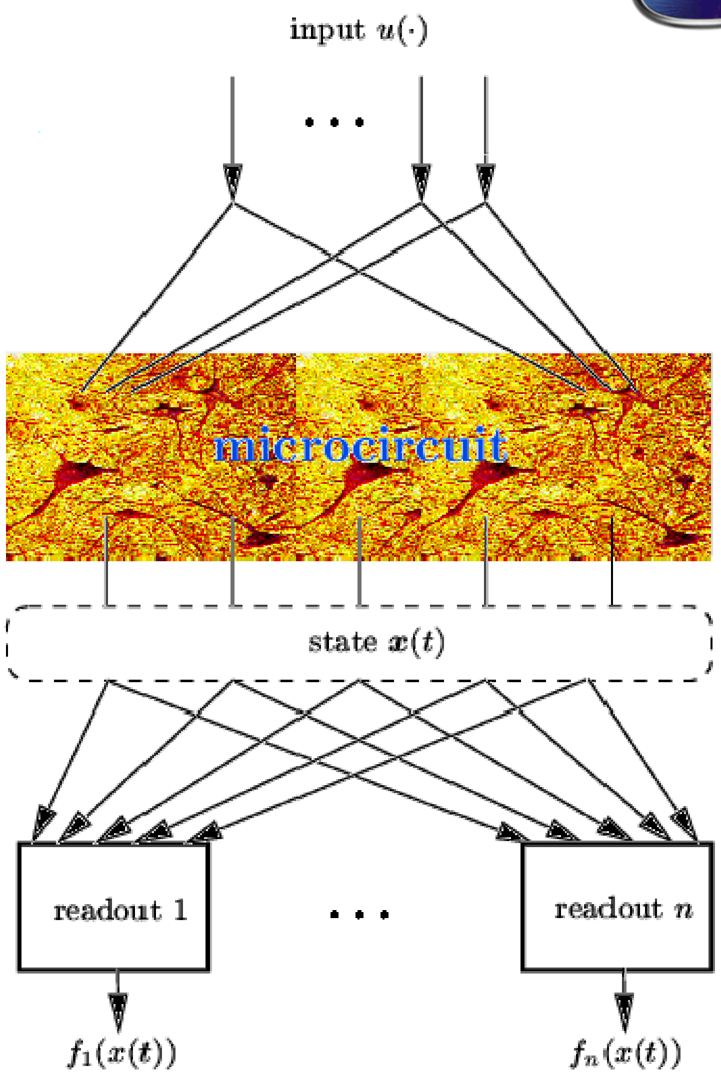




Maass, W., Liquid state machines: motivation, theory, and applications. In: Computability in context: computation and logic in the real world (Cooper B, Sorbi A, eds), pp 275–296 (2010)

Liquid State Machine

- Input (spike trains)
  - Maps input streams to output streams
- Liquid (or microcircuit)
  - A recurrent neural network of spiking neurons (leaky integrate and fire)
  - Acts a preprocessor (temporal)
- State
  - Measure the state of the liquid at any given time t
- Readout neurons
  - Plastic synapses
  - By assumption, has no temporal integration capability of its own

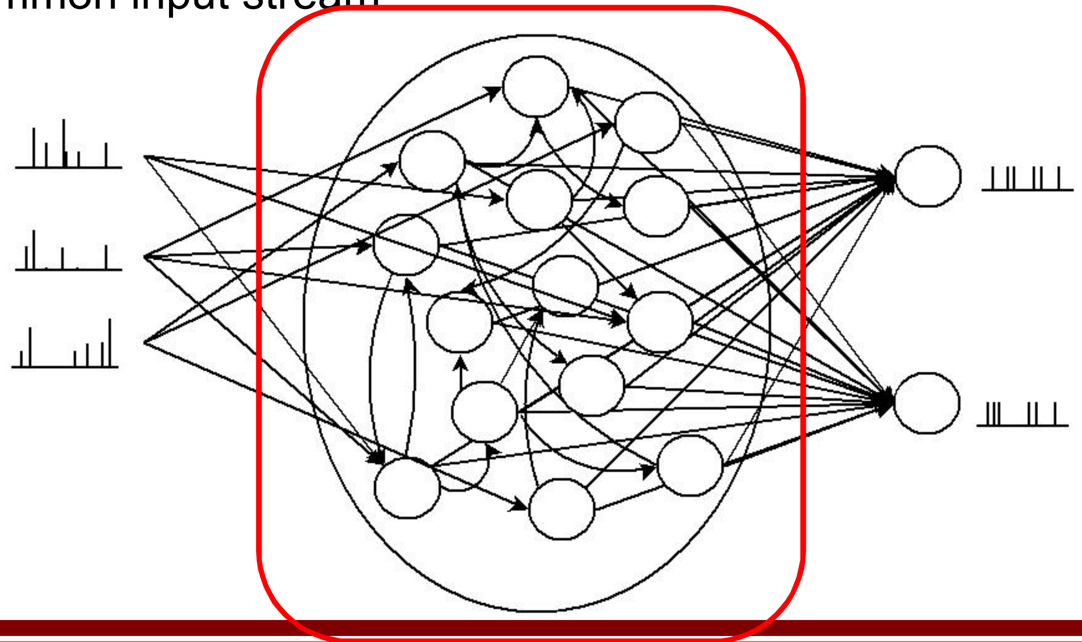


Natschläger, T., "The Liquid State Machine Framework." *Neural Micro Circuits,* http://www.lsm.tugraz.at/learning/framework.html. Accessed 26 September 2016

The Liquid

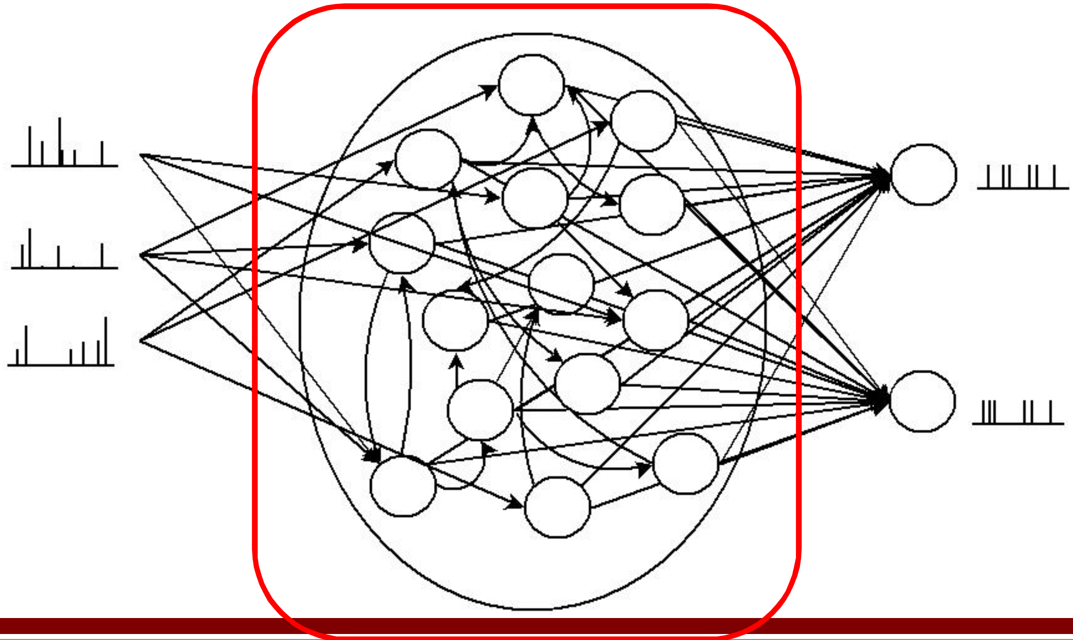
- Neurons are randomly (heuristically) connected
  - Follow properties observed in cortical microcircuits
  - Some attempts to “learn” the liquid (Spike-timing-dependency plasticity)
  - Multiplexing diverse computations on a common input stream

- Serves a temporal preprocessor
  - Provides all temporal integration of information for readout neurons
  - Fading memory
- Maps input streams to output streams



The Liquid

- Heuristic properties
  - 80% excitatory neurons, 20% inhibitory neurons
  - 30% connectivity
  - Stochastic connectivity based on distance between neurons
- Accumulates information overtime

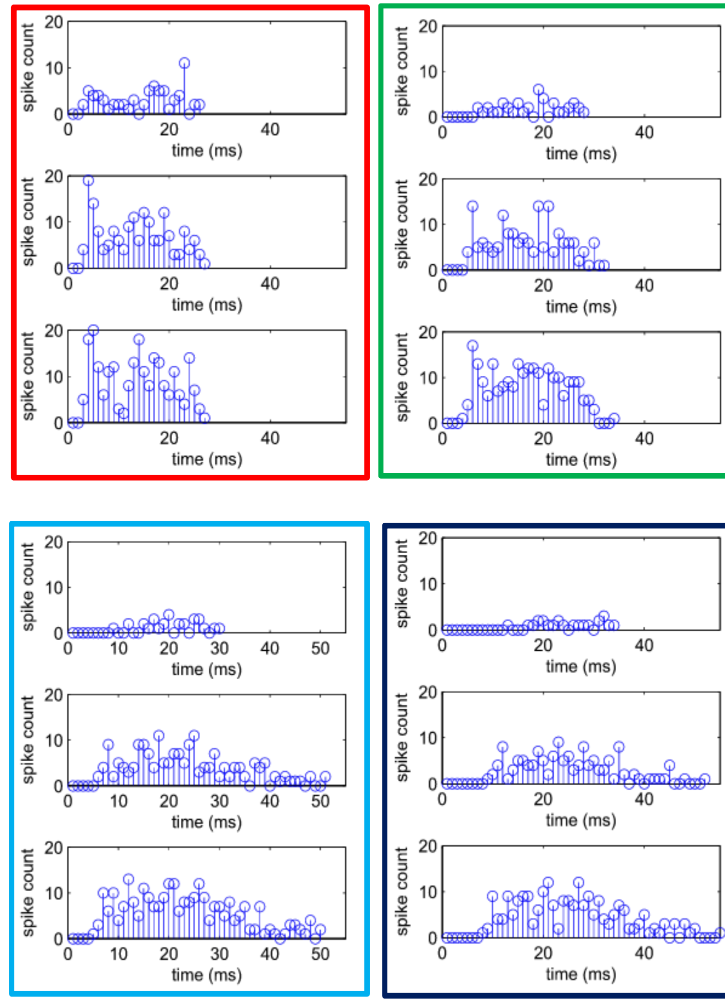


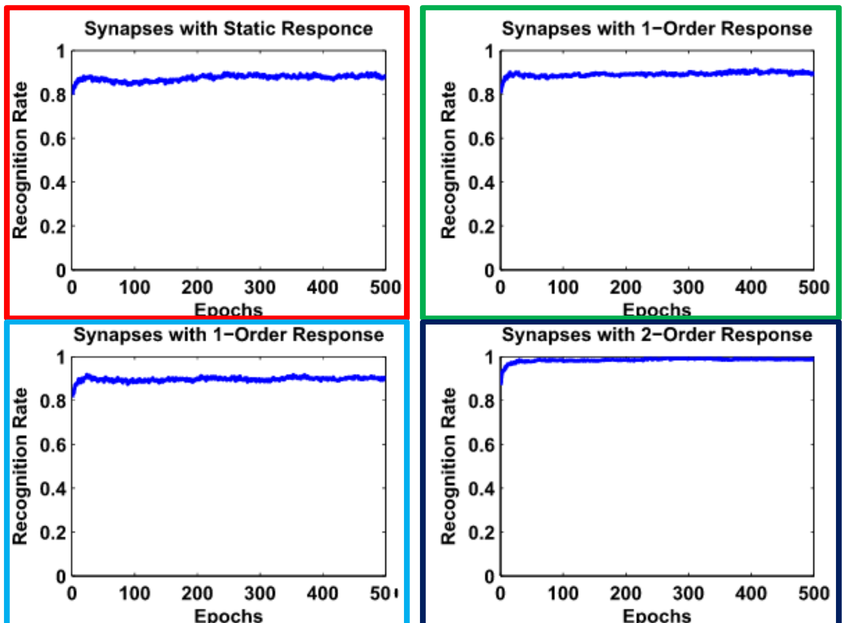
The Liquid

- Two properties must be fulfilled for the system to work well
  - **Separation property** for a class of basis filters β: if we have two input functions $u(\cdot), v(\cdot)$ with $u(s) \neq v(s)$ for some $s \leq t$ a basis filter $B \in \beta$ with $(Bu)(t) \neq (Bv)(t)$
    - The amount of distance between the trajectories of different input streams into the liquid
  - **Approximation property**: The ability of readout units to distinguish trajectories from an input stream and connect them to the target outputs

Living on the Edge of Chaos

- Fading memory
  - Feedback loops and synaptic properties
  - Do not want to evolve to a steady state



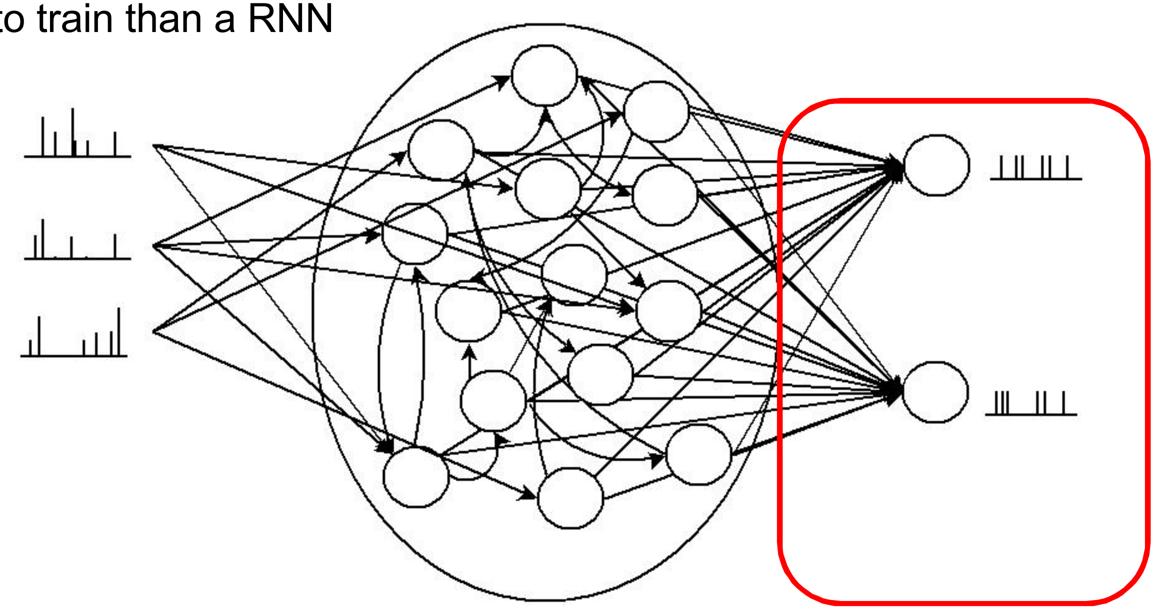


Zhang, Y., Li, P., Jin, Y. & Choe, Y. (2015). A Digital Liquid State Machine With Biologically Inspired Learning and Its Application to Speech Recognition.. *IEEE Trans. Neural Netw. Learning Syst.*, 26, 2635-2649.

Readout Neurons

- Can be any function
- Theoretically, should be memory-less
- Generally, use a linear readout function (SVM…)
  - Fast
  - Cannot get stuck in local minima
  - Entails superior generalization since its VC-dimension is equal to the dimensionality of its input plus 1
  - Easier to train than a RNN

- Multiple readouts for a given liquid and input stream
  - This means that the liquid only needs to be computed once, giving the LSM an inherent parallel processing capability
- Can make a prediction at any time
  - Before all of the inputs have arrived
- Proven to be a universal function approximator



Applications

- Speech and audio recognition
- Image Pattern Recognition
- Music Classification
- Robot Path Planning
- Fingerprint Scanners
- Facial emotion recognition

Difficulties

- Applications lacking state awareness
  - LSM research focuses on modeling dynamical and representational phenomena in biological neural networks, rather more at engineering applications
  - How to bridge the gap between research and practicality
- Getting data into spike trains
  - Looking at using raw input of the data
- Varying time between inputs
  - Could vary the feedback time for different neurons

Outline

- A novel neuromorphic architecture: the Spiking Temporal Processing Unit
- A neuro-inspired algorithm for the neuromorphic computing: the Liquid State Machine
- **Mapping the LSM onto the STPU**
- Lessons learned and future development

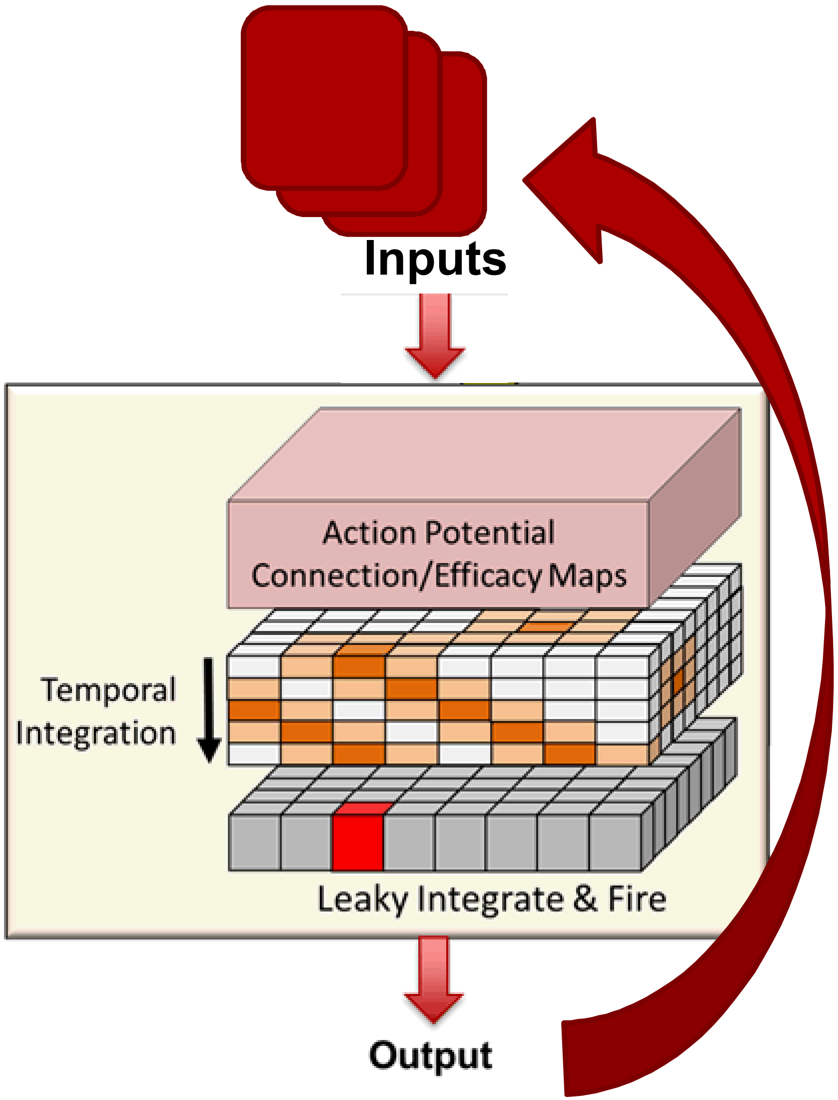
LSM Hardware implementations

- Several implementations in academia
  - 2006 (NIPS): Edge of chaos computation in mixed-mode VLSI-“A hard liquid”
  - 2008 (IJCNN): Compact hardware liquid state machines on FPGA for real-time speech recognition
  - 2015 (TNNLS): A digital liquid state machine with biologically inspired learning and its application to speech recognition
    - Cross-bar architecture with digital processing units
- One patent
  - 2008 Physical neural network liquid state machine utilizing nanotechnology

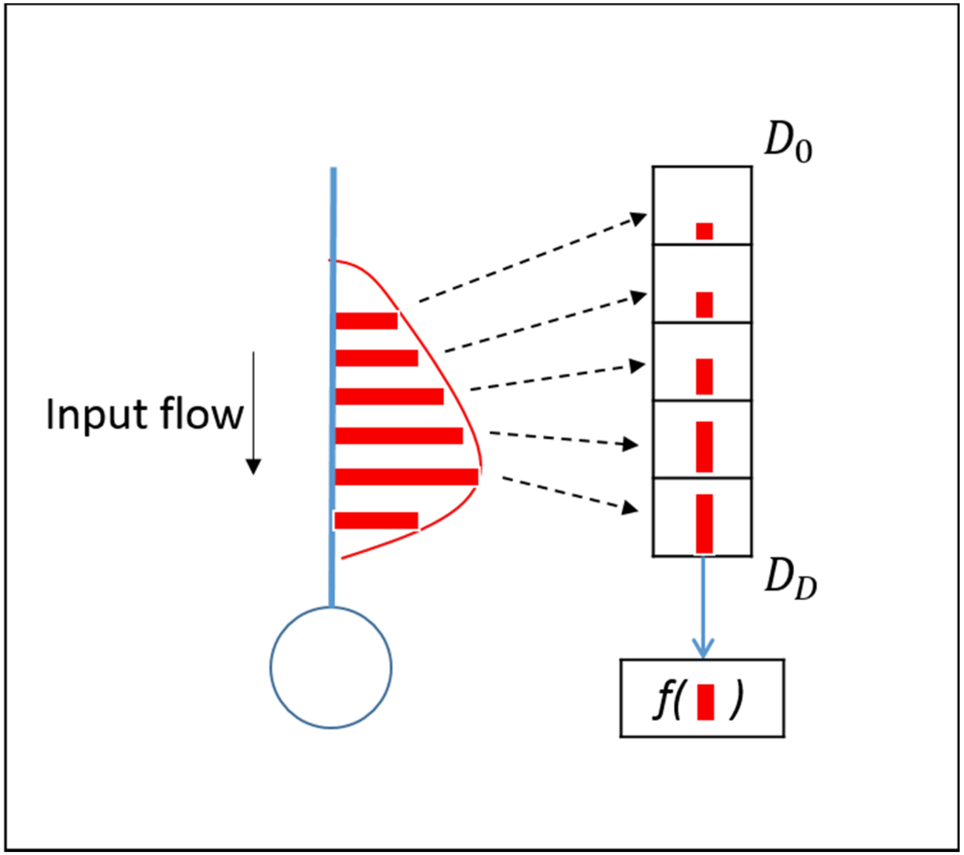
Mapping LSMs map onto the STPU

- Goals:
  - Implement the rich dynamics of a LSM efficiently
  - Recurrence
  - Exponential synaptic response functions
- Currently implemented LSM in simulator matlab code
- Can do speech recognition with minimal parameter tuning using ridge regression
- Comparison with Zhang et al. 2015:
  - Use state variables to keep track of synaptic responses. Time constants are binary (division becomes bit shifting)
  - STPU uses weights to put values into the temporal stack

Recurrence


Inputs
Action Potential
Connection/Efficacy Maps
Temporal
Integration
Leaky Integrate & Fire
Output

Second-Order Synaptic Response Functions



Outline

- A novel neuromorphic architecture: the Spiking Temporal Processing Unit
- A neuro-inspired algorithm for the neuromorphic computing: the Liquid State Machine
- Mapping the LSM onto the STPU
- **Future development and lessons learned**

Research Directions

- Algorithmic drives improvements (STPU V2)
  - Internal recurrence
  - Improved routing
  - On-line learning (on chip)
  - Linear discriminator using spiking neurons

Lessons Learned

- Several neuromorphic architectures
  - Some are used to better understand neuroscience
  - Many lack applications
- Algorithmic instantiation provided insights into short comings
  - LSM is a complex algorithm that exposed many un-foreseen shortcomings of the STPU
- Shortcomings
  - Converting to temporal domain

Exceptional service in the national interest

Sandia National Laboratories

Photos placed in horizontal position with even amount of white space between photos and header

FIN