Energy-Efficient Recurrent Spiking Neural Processor with Unsupervised and Supervised Spike-Timing-Dependent-Plasticity

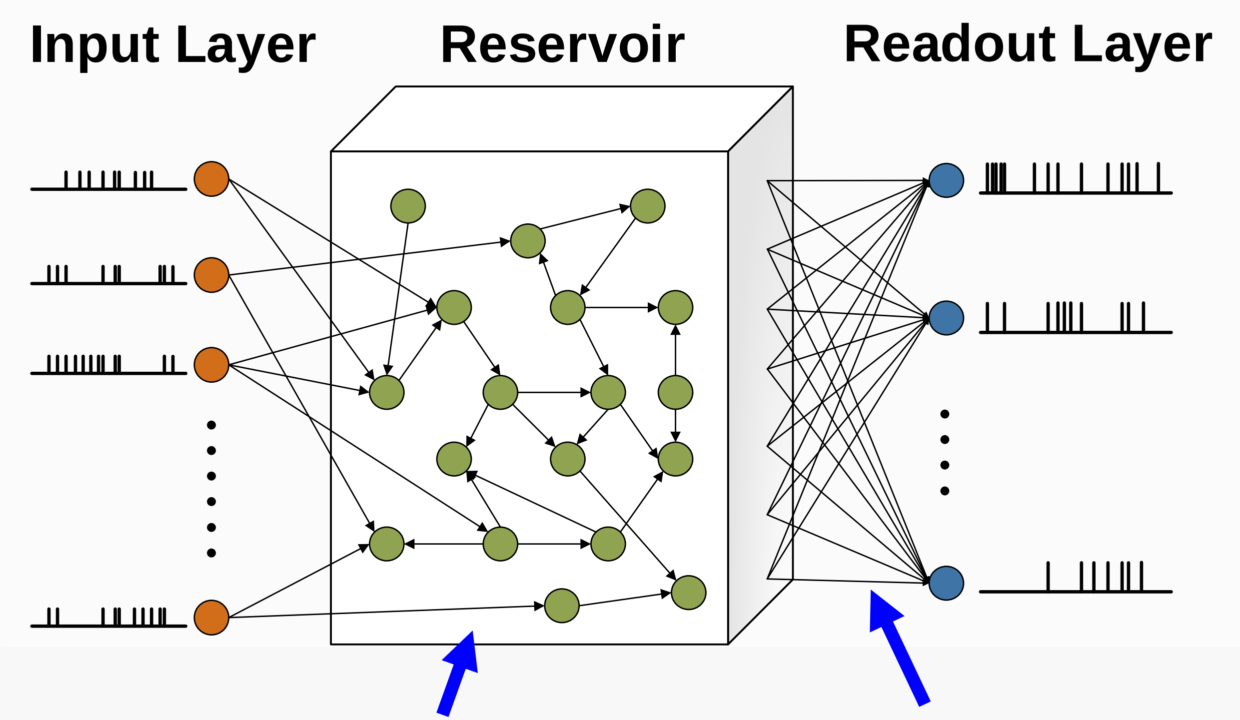
Yu Liu, Peng Li
Dept. of Electrical & Computer Engineering
Texas A&M University
{yliu129, pli}@tamu.edu

Neuromorphic Computing based on Spiking Neural Nets

- **Spiking Neural Networks (SNNs)**
  - Biologically realistic
  - Rate and temporal codes
  - Ultra-low energy, event-driven processing
- **Present Challenges**
  - Cognitive Principles:
    - Rich inspiring ideas, limited successfully demonstration in real-world tasks
  - Network Architecture:
    - Mostly simple networks such as feedforward
  - Training
    - Locality constraints: algorithms for ANNs does not satisfy
    - Lack of powerful spike-based training methods

(Spiking) Liquid State Machine (LSM)

- **Tradeoffs between biological plausibility, design complexity and performance.**
- **Recurrent reservoir structure**



Liu and Li. Energy-Efficient Recurrent Spiking Neural Processor with Unsupervised and Supervised Spike-Timing-Dependent-Plasticity

In This Work:

- **Improve learning performance of LSM neural accelerators with power efficiency with proposed unsupervised and supervised STDP training algorithms.**

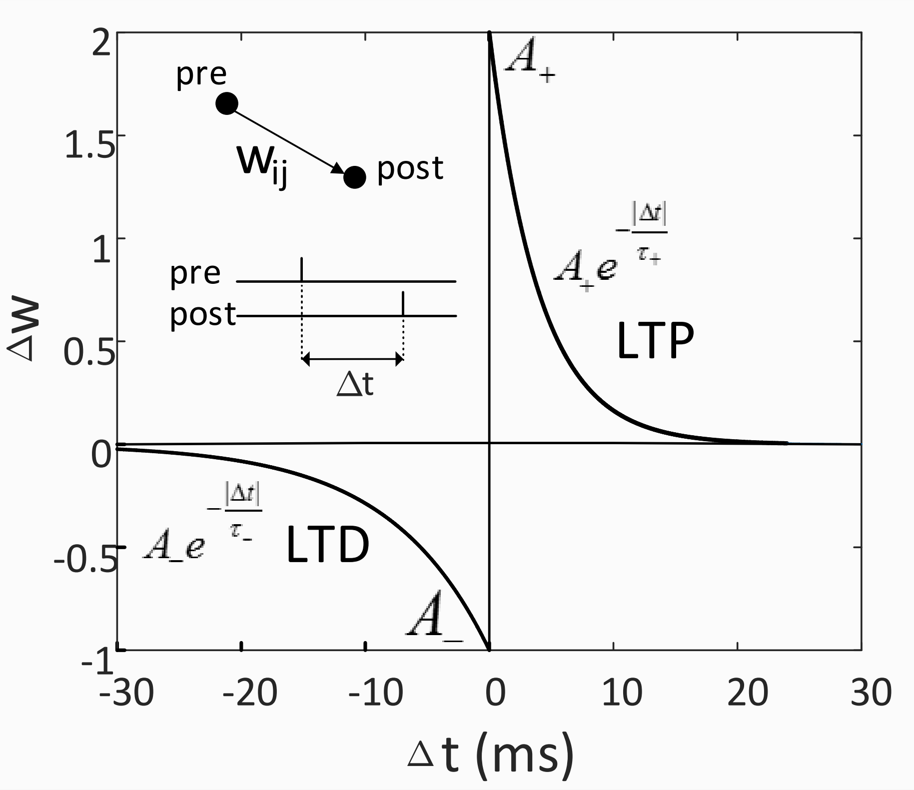
Unsupervised STDP	Supervised STDP
❖ Reservoir training	❖ Readout training
❖ **Supplement** to classification training on readout	❖ **Maximize the distance of firing frequency** between desired and undesired neurons
❖ Sparse synaptic connectivity from **self-organizing reservoir tuning**	❖ Sparse synaptic connectivity **without degrading performance**

Jin, Yingyezhe, and Peng Li. "Calcium-modulated supervised spike-timing-dependent plasticity for readout training and sparsification of the liquid state machine." Neural Networks (IJCNN), 2017 International Joint Conference on. IEEE, 2017.

Spike-Timing-Dependent Plasticity (STDP) Reservoir Training

- **Adjust the connection strengths based on the relative timing of spike pairs** **[Bi & Poo, Ann. review of neurosci.'01]**

$$\Delta w^{+} = A_{+}(w) \cdot e^{-\frac{|\Delta t|}{\tau_{+}}} \quad if\ \Delta t > 0$$
$$\Delta w^{-} = A_{-}(w) \cdot e^{-\frac{|\Delta t|}{\tau_{-}}} \quad if\ \Delta t < 0$$



- **Locally tune the synaptic weights**
- **Naturally lead to sparse**

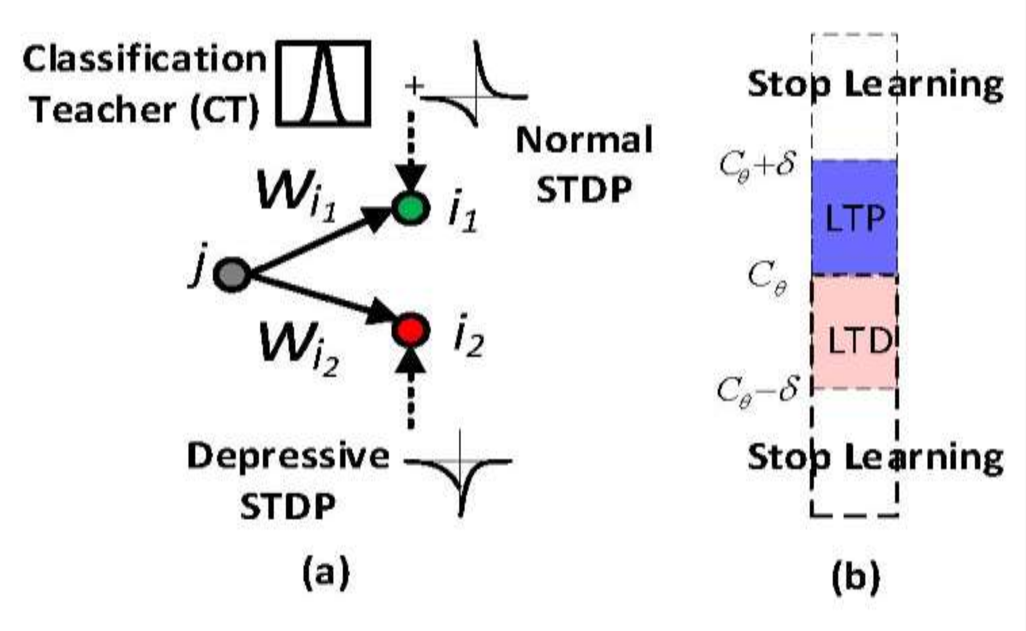
Jin, Yingyezhe, Yu Liu, and Peng Li. "SSO-LSM: A sparse and self-organizing architecture for liquid state machine based neural processors." Nanoscale Architectures (NANOARCH), 2016 IEEE/ACM International Symposium on. IEEE, 2016.

Supervised STDP Readout Training

- ***CAL-S^{2}TDP:* Calcium-modulated Learning Algorithm Based on STDP**
  - **Supervisory** signal (CT) combined with **depressive STDP**
  - Improving **memory retention:**

    Probabilistic weight update
  - Preventing **weight saturation:**

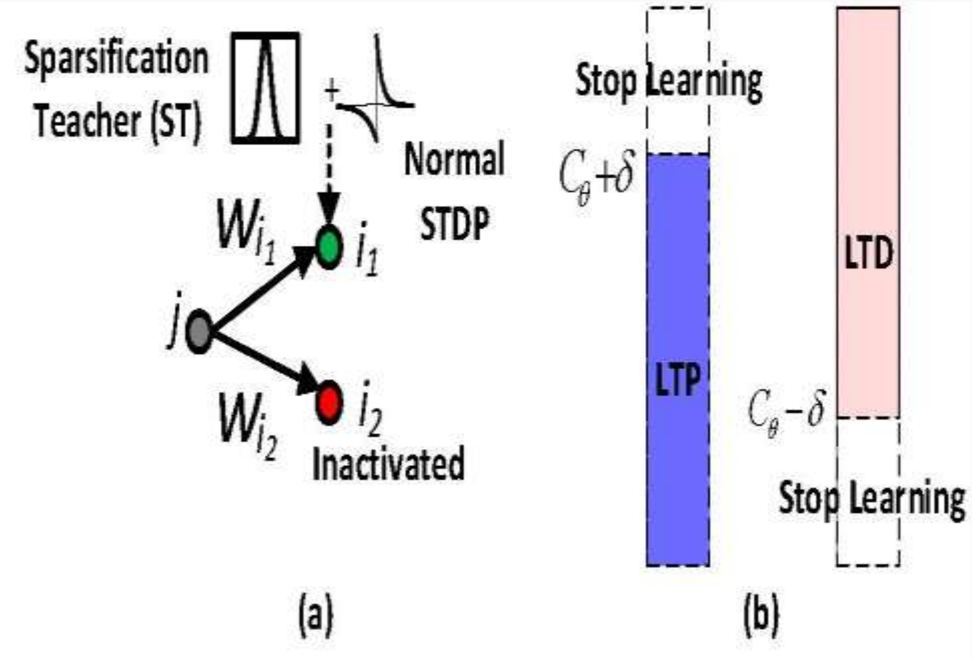
    Calcium-modulated stop learning



$$w \leftarrow w + d \;\; w/prob. \propto |\Delta w^+|, \;\; if\ \Delta t > 0 \ \&\& \ c_\theta < c < c_\theta + \delta$$
$$w \leftarrow w - d \;\; w/prob. \propto |\Delta w^-|, \;\; if\ \Delta t < 0 \ \&\& \ c_\theta - \delta < c < c_\theta$$

Supervised STDP Readout Training

- ***CAS-S²TDP:*** **Calcium-modulated Sparsification Algorithm Based on STDP**
  - Fully connected readout synapses
    - **Overfitting**
    - Large **hardware overhead**
  - Random dropouts lead to **significant performance drop**.
  - Embed class information into to maximize the sparsity and secure learning performance.



$$w \leftarrow w + d\,w/\,prob. \propto |\Delta w^+|,\ \ if\ \Delta t > 0\ \&\&\ \ c < c_\theta + \delta$$
$$w \leftarrow w - d\,w/\,prob. \propto |\Delta w^-|,\ \ if\ \Delta t < 0\ \&\&\ \ c_\theta - \delta < c$$

Results

- **Adopted Benchmark:**
  - TI46 speech of English letters (single speaker, 260 samples)
- **Training Settings**
  - 5-fold cross-validation, 500 training iterations on readout layer
  - Baseline is a competitive spike-dependent non-STDP supervised training algorithm*.

		Baseline	**Proposed**
Inference Accuracy	135 Reservoir Neurons	$92.3 \pm 0.4\%$	$93.8 \pm 0.5\%$
	90 Reservoir Neurons	$89.6 \pm 0.5\%$	$92.3 \pm 0.4\%$

** Yong Zhang, Peng Li, Yingyezhe Jin, and Yoonsuck Choe, "A digital liquid state machine with biologically inspired learning and its application to speech recognition," IEEE Trans. on Neural Networks and Learning Systems, Nov. 2015.*

Acknowledgement

- We thank High Performance Research Computing (HPRC) at Texas A&M University for providing computing support. **Resource Utilization:**
  - **Cluster**: Terra
  - **Software**: CUDA
  - **Core & Memory**: 1 GPU, 2GB
  - **Typical Runtime**: 0.5 ~ 2 days

- This material is based upon work supported by the National Science Foundation under Grant No. 1639995 and the Semiconductor Research Corporation(SRC) under task # 2692.001.