Neural Generative Models with Stochastic Synapses Capture Richer Representations

Maruan Al-Shedivat[1], Emre Neftci[2] and Gert Cauwenberghs[2]

[1]King Abdullah University of Science and Technology (KAUST)
Thuwal 23955-6900, KSA

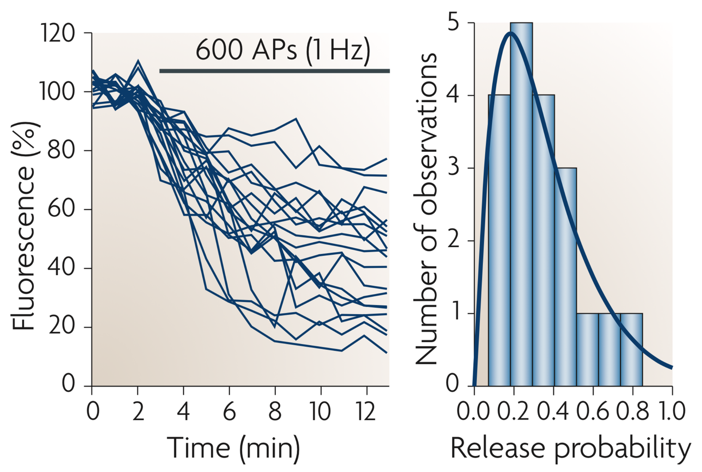
[2]Institute for Neural Computation, University of California, San Diego
La Jolla, CA 92037, USA

Abstract

Stochasticity in synaptic vesicle release is one of the major sources of noise in the brain. While the concept of cellular neural noise gave rise to computational models of biological learning such as deep belief networks and algorithms such as spike-sampling, the functional implications of synaptic stochasticity on learning remain unascertained and are often limited to filtering, decorrelation, or regularization. This work approaches synaptic stochasticity from the perspective of learning representations in the context of Boltzmann machines.

Motivation

Stochasticity of the synaptic release has an extensive experimental grounding [1]. However, the purpose and function of such behavior is poorly understood, especially on the level of neural circuits.



CNS area	P_{release}
Cat and rat L2/3	0.5 ± 0.05
	0.46 ± 0.26
	0.65 ± 0.18
Rat L5/6	$0.16 - 0.9$
	0.53 ± 0.22

Figure 1 and Table 1 : Measured release probabilities for pyramidal cells [1].

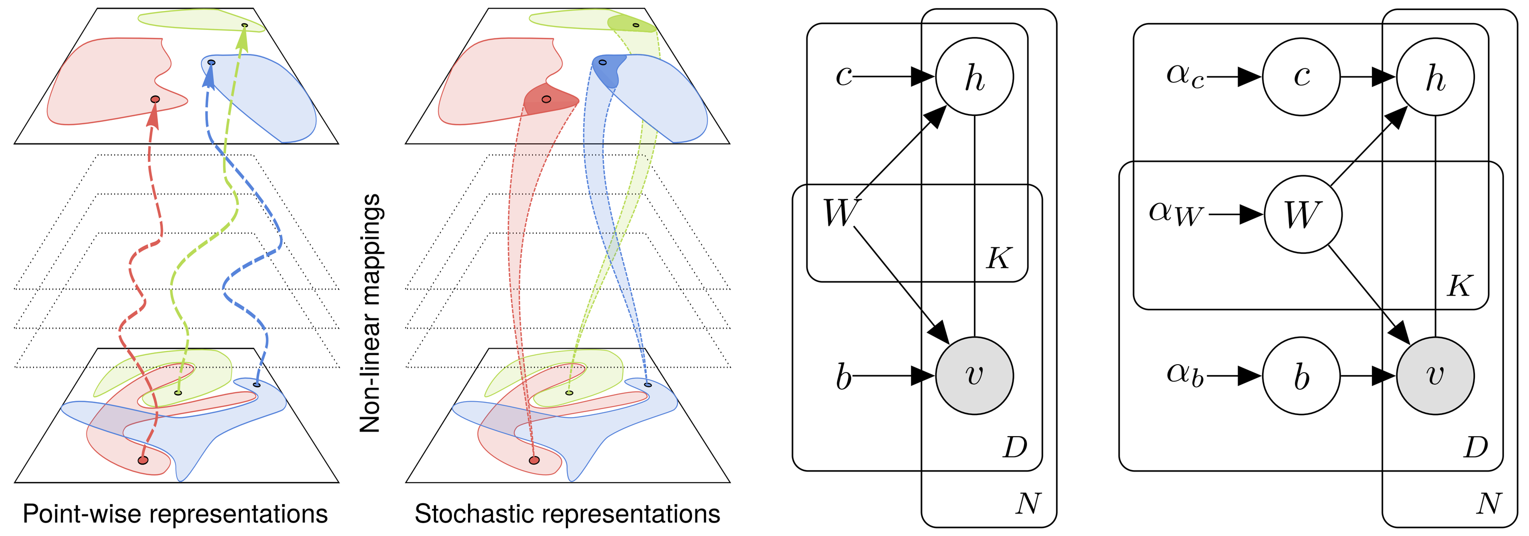
Synapses can individually adjust their neurotransmitter release probabilities dynamically through local field regulation [2]. **What is the information learned and stored in the synaptic probability distributions?**

Possible functional implications of the unreliable synapses:

- Stochastic filtering of the synaptic spike-trains.
- Stochastic facilitation mechanism (similar to stochastic resonance).
- Generation of the Poisson-like spiking variability in cortical microcircuits.

Open Questions

- What are the implications of the synaptic stochasticity on learning?
- How does the adjustable probabilistic synaptic release affect the representations learned by a neural network? Can we quantify this effect?
- Which computational models are the most suitable for studying learning with stochastic plastic synapses?
- How to connect synaptic stochasticity with the cognitive function of the brain?



(a) Point-wise vs. stochastic mapping (b) RBM (c) RBSE

Figure 2 : The concept of stochastic representations and RBM and RBSE models.

Restricted Boltzmann Stochastic Ensemble (RBSE)

The RBSE model is defined as follows:

$$P(\boldsymbol{v}, \mathbf{h}, \boldsymbol{W}; \alpha) = P(\boldsymbol{v}, \mathbf{h} \mid \boldsymbol{W}) P(\boldsymbol{W}; \alpha) = \frac{e^{-E(\boldsymbol{v}, \mathbf{h}, \boldsymbol{W})}}{Z(\boldsymbol{W})} P(\boldsymbol{W}; \alpha),$$

where $\boldsymbol{v}$, $\mathbf{h}$ are the visible and hidden neurons, respectively, $E(\boldsymbol{v}, \mathbf{h}, \boldsymbol{W})$ is the Boltzmann energy that depends on synaptic strengths $\boldsymbol{W}$, and $P(\boldsymbol{W}; \alpha)$ is the synaptic reliability distribution.

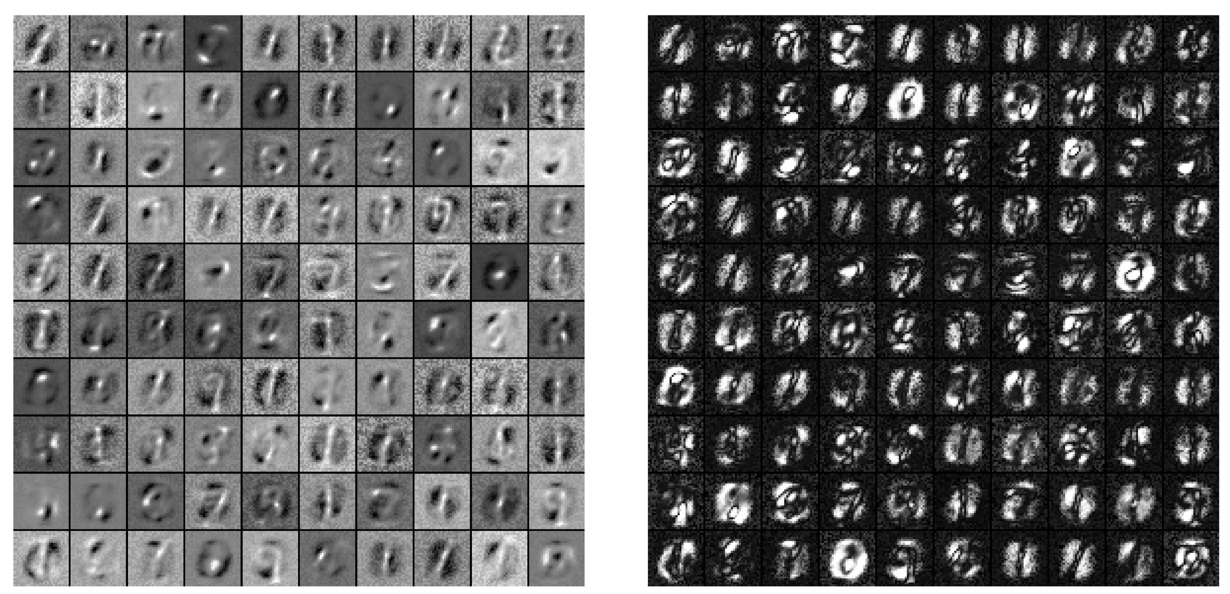
The gradient of the log-likelihood for RBSE has the following form:

$$\frac{\partial \log P(\boldsymbol{v}; \alpha)}{\partial \alpha} = \mathbb{E}\left[\frac{\partial \phi(\theta; \alpha)}{\partial \alpha}\right]_{P(\theta; \alpha)} - \mathbb{E}\left[\frac{\partial \phi(\theta; \alpha)}{\partial \alpha}\right]_{P(\theta \mid \boldsymbol{v}; \alpha)},$$

which resembles the standard contrastive divergence [3]. The expectations over the model space are estimated through Monte Carlo sampling:

$$\mathbb{E}[\,\cdot\,]_{P(\theta;\alpha)} \approx \int d\theta\; P(\theta \mid \tilde{\boldsymbol{v}}, \tilde{\mathbf{h}})[\,\cdot\,], \quad \mathbb{E}[\,\cdot\,]_{P(\theta \mid \boldsymbol{v};\alpha)} \approx \int d\theta\; P(\theta \mid \boldsymbol{v}, \hat{\mathbf{h}})[\,\cdot\,],$$

where $\tilde{\boldsymbol{v}}, \tilde{\mathbf{h}} \sim P(\boldsymbol{v}, \mathbf{h})$ and $\hat{\mathbf{h}} \sim P(\mathbf{h} \mid \boldsymbol{v})$. For more details on the training see [4].

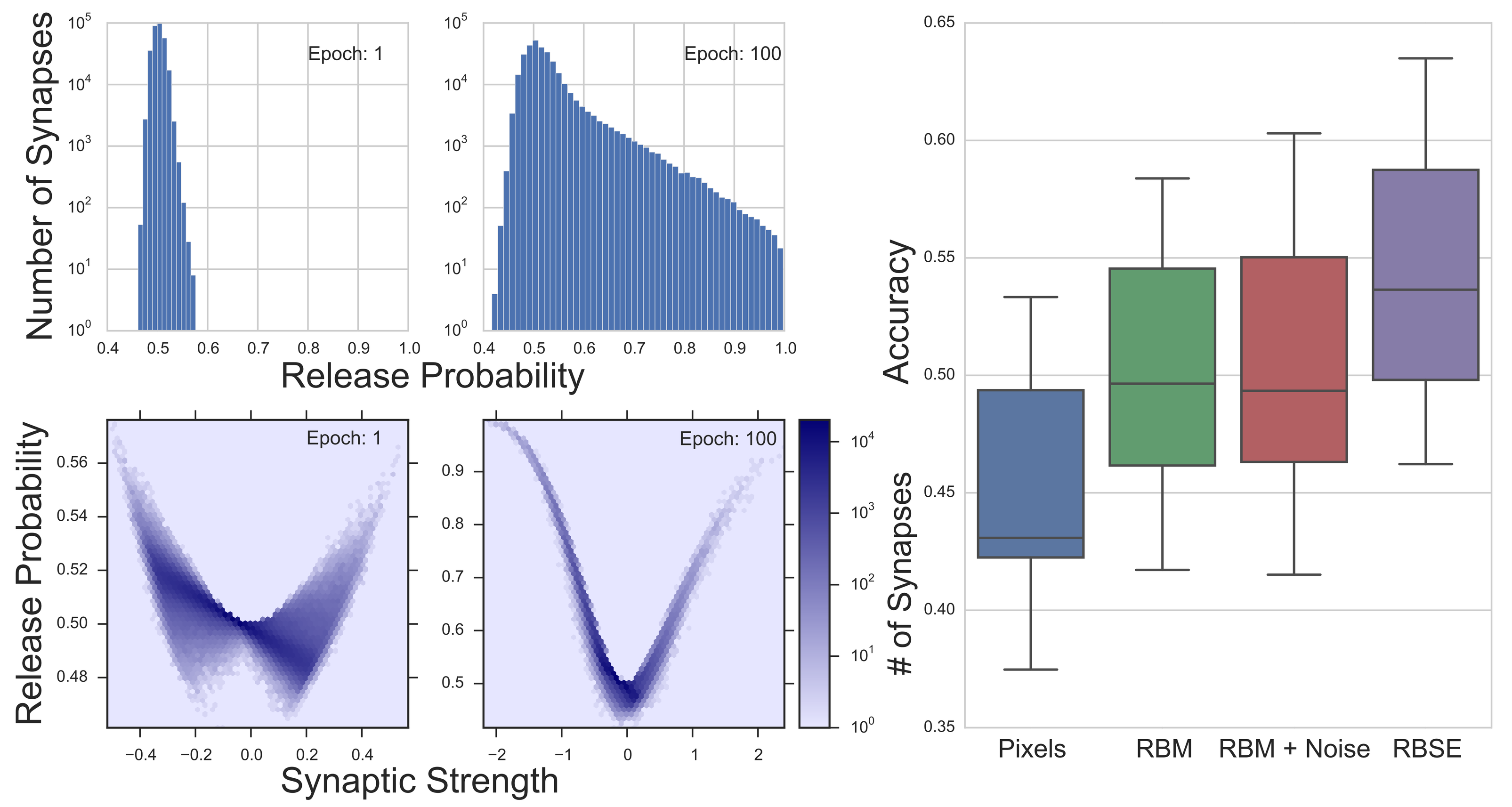


(a) Synaptic strengths (b) Synaptic reliability

Figure 3 : Synaptic connections learned by the RBSE from hand-written digits. Synapses are grouped for every hidden neuron into 28×28 images.

Results

- Probability of the synaptic vesicle release has evolved over the course of network training. We also observed correlation between reliability and strength of the synapses: **the strongest synapses became the most reliable**.



- A simple logistic classifier trained on the stochastic representations (features) produced by RBSE demonstrated better generalization on the hand-written digits dataset in the one-shot learning scenario.

Conclusions

- Noisy synapses lead to the notion of stochastic ensemble of generative models that can be trained with stochastic optimization techniques.
- The ensemble results from the synaptic stochasticity and is fitted to the data. Hence synaptic probabilities store the information on the data variability.
- The strongest synapses become very reliable while the rest remain unreliable.
- The knowledge stored in the stochastic synapses and learned in an unsupervised manner can leverage the subsequent classification performance in one-shot scenario which suggests that RBSEs capture richer representations.

References

[1] T. Branco and K. Staras, "The probability of neurotransmitter release: variability and feedback control at single synapses," *Nat. Rev. Neurosci.*, vol. 10, pp. 373–383, 2009.

[2] C. F. Stevens and Y. Wang, "Changes in reliability of synaptic function as a mechanism for plasticity," *Nature*, vol. 371, no. 6499, pp. 704–707, 1994.

[3] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural computation*, vol. 14, no. 8, pp. 1771–1800, 2002.

[4] M. Al-Shedivat, E. Neftci, and G. Cauwenberghs, "Learning non-deterministic representations with energy-based ensembles," *arXiv preprint arXiv:1412.7272*, 2014.

Computational and Systems Neuroscience (Cosyne) meeting, Salt Lake City, USA, March, 2015

alshedivat.maruan@gmail.com, {nemre, gert}@ucsd.edu