Spiking neural networks as superior generative and discriminative models

Luziwei Leng*, Mihai Petrovici*, Roman Martel, Ilja Bytschok
Oliver Breitwieser, Johannes Bill, Johannes Schemmel, Karlheinz Meier

13.11.2015

Abstract

An increasing number of experiments suggest that the brain performs stochastic inference when dealing with incomplete and noisy sensory information. This, in turn, has led to the development of various theoretical models that attempt to explain how this could be achieved with spiking neural networks. One candidate theory interprets spiking activity as sampling from distributions over binary random variables (Buesing et al., 2011) and has been shown to be compatible with the ensemble dynamics of noise-driven LIF neurons in the high-conductance state (Petrovici et al., 2013, 2015; Probst et al., 2015). Based on this theory, we constructed hierarchical LIF networks that sample from restricted Boltzmann distributions and compared their performance with conventional restricted Boltzmann machines (RBMs) on a commonly used dataset (MNIST). An important result is that LIF networks can achieve similar classification rates (95.1 % with 1994 neurons) as their machine-learning counterparts of equal size (95.2 %). In classical RBMs however, statistics are typically gathered by Gibbs sampling. This algorithm has a distinct disadvantage when dealing with high-dimensional multimodal distributions, where it often gets trapped in a local minimum due to deep troughs in the energy landscape that appear during training. It is for this reason that conventional RBMs that may perform very well as discriminative models are, at the same time, rather poor generative models of the learned data. While various methods exist that alleviate this problem (such as AST, see Salakhutdinov, 2010) they usually come at a highly increased computational cost. In the second part of our study, we show how short-term plasticity enables LIF networks to travel efficiently through the energy landscape and thereby attain a generative performance that far surpasses the one achievable by conventional Gibbs sampling. This distinct advantage of biological neural networks allows them to simultaneously become good generative and discriminative models of learned data.

References

L. Buesing, J. Bill, B. Nessler, and W. Maass. Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLoS Computational Biology*, 7(11):e1002211, 2011.

G. Fuhrmann, I. Segev, H. Markram, and M. Tsodyks. Coding of temporal information by activity-dependent synapses. *Journal of neurophysiology*, 87(1):140–148, 2002.

M. A. Petrovici, J. Bill, I. Bytschok, J. Schemmel, and K. Meier. Stochastic inference with deterministic spiking neurons. *arXiv preprint arXiv:1311.3211*, 2013.

M. A. Petrovici, I. Bytschok, J. Bill, J. Schemmel, and K. Meier. The high-conductance state enables neural sampling in networks in lif neurons. *annual conference of the OCNS 2015, to appear in BMC Neuroscience*, 2015.

D. Probst, M. A. Petrovici, I. Bytschok, J. Bill, D. Pecevski, J. Schemmel, and K. Meier. Probabilistic inference in discrete spaces can be implemented into networks of lif neurons. *Frontiers in computational neuroscience*, 9, 2015.

R. Salakhutdinov. Learning deep boltzmann machines using adaptive mcmc. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 943–950, 2010.

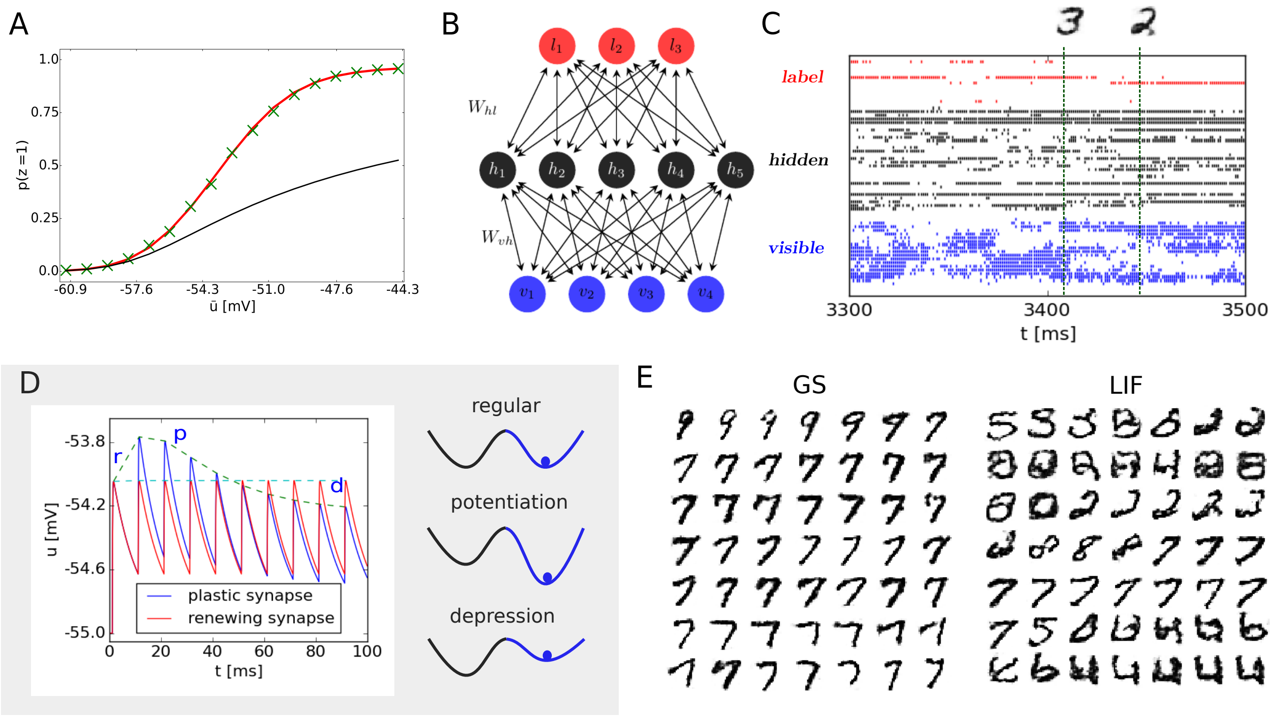

Figure 1: **(A) Activation function of the LIF neuron**: Activation function of an LIF neuron in the high-conductance state (HCS, green crosses) and theoretical prediction from (Petrovici et al., 2013, 2015; Probst et al., 2015) (red line) compared to a typical LIF response function (black line). In the HCS, the firing rate of an LIF neuron can be well approximated by a logistic function, thereby endowing the network with Glauber dynamics that can be mapped to Gibbs sampling from a Boltzmann distribution. **(B) Hierarchical LIF network structure (note the equivalence to an RBM)**: A network consisting of 784 visible, 1200 hidden and 10 label units was trained on the full MNIST dataset with the so-called coupled adaptive simulated tempering learning algorithm (Salakhutdinov, 2010). **(C) Selected spike trains from the LIF network**: Note the increasing sparseness of the activity in consecutive layers. The network produces different images when the activity in the label layer switches on from one neuron to another. **(D) Short-term plasticity as modeled by the Tsodyks-Markram (TSO) mechanism and its effect on the local energy landscape**: Renewing synapses (red line) would keep the average interaction between pairs of neurons constant, while plastic synapses with appropriate TSO (Fuhrmann et al., 2002) parameters first strengthen, then weaken the effective interaction. This causes a local change in the energy landscape, first deepening the energy trough and sharpening the produced image, followed by a local flattening of the energy landscape which pushes the network state into a different mode. **(E) Comparison between a sequence of images generated by the conventional RBM with Gibbs sampling (GS) and one generated by the LIF network with short-term synaptic plasticity**: Due to large variance in the energy landscape, Gibbs sampling becomes trapped in a local mode, therefore constantly generating the (approximately) same image. The LIF network is significantly better at mixing, producing a varied sequence of images.