Bayesian Computing with Spikes

A. Baumbach, M. A. Petrovici, I. Bytschok, D. Stöckel, L. Leng, O. Breitwieser, J. Schemmel, K. Meier

Heidelberg University, Kirchhoff-Institute for Physics

Electronic Vision(s)

The Manfred Stärk Foundation

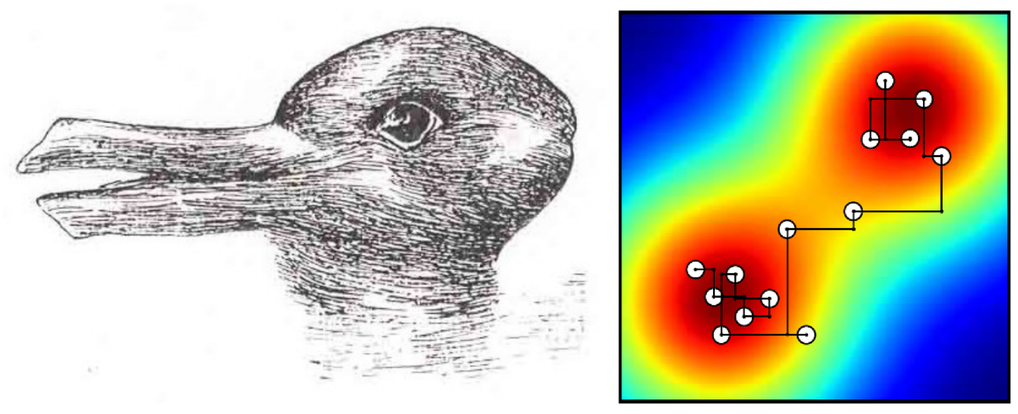
Sampling with Spiking Neurons

You are able to raed this snetence, even thgouh mnay of the wdros are mieepsllsd and yuor biarn has to dcdiee which wrod was maent to be wtreitn.

The fact that you can read the opening sentence is a sign of the brain's ability to infer knowledge based on partial, ambiguous or even false information.

Any such decision can be thought of as a probabilistic process. Here the used letters are correct, but their order is scrambled.

Evidence has been accumulating that the brain's internal representation of both its surrounding and decisions is sample based (e.g. [1]). An interesting research question therefore is: How can sample-based probabilistic inference be implemented in a network of spiking neurons?

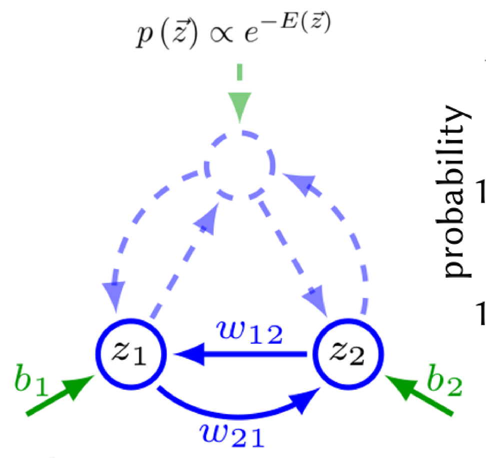


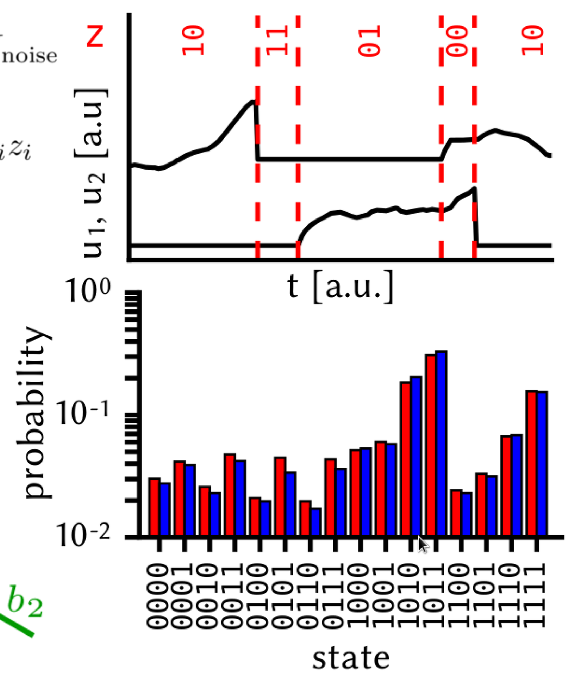
In the LIF sampling framework [2, 3] a neuron represents a binary random variable. The network samples from a distribution that is shaped by its synaptic connections (weights) and external inputs (biases).

$$\frac{du}{dt} C_m = g_l (E_l - u) + I_{\text{ext}} + I_{\text{noise}}$$

$$E(\vec{z}) = \sum_{i,j} W_{ij} z_i z_j + \sum_i b_i z_i$$

$$p(\vec{z}) \propto e^{-E(\vec{z})}$$

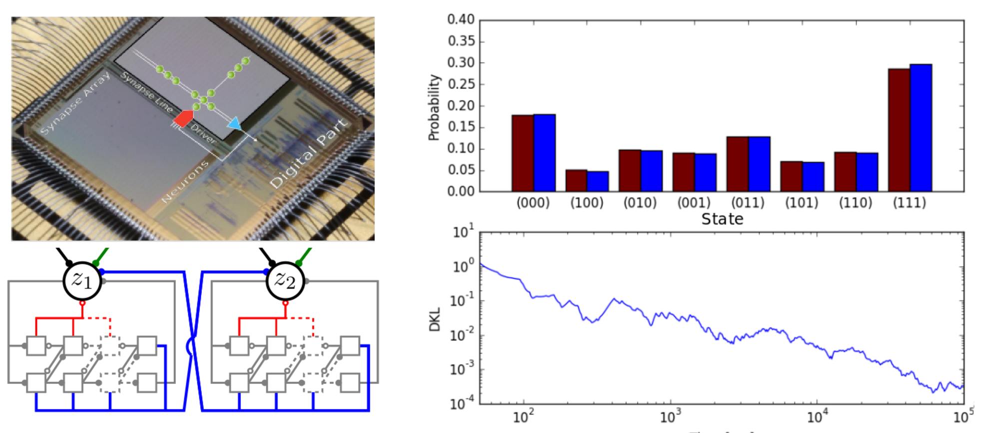




Accelerated Neuromorphic Sampling

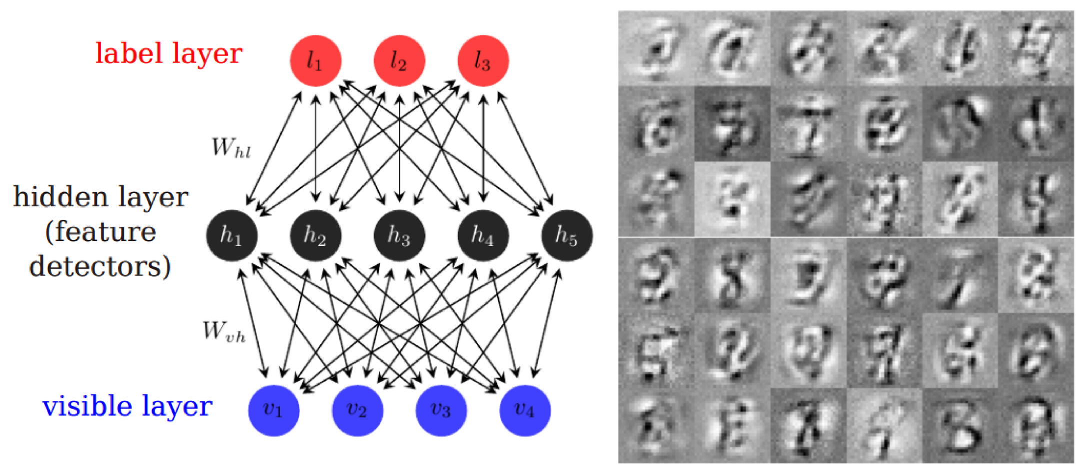
The LIF model is the de facto standard for neuromorphic hardware platforms (here: Spikey chip with 384 neurons, 100k synapses) [4].

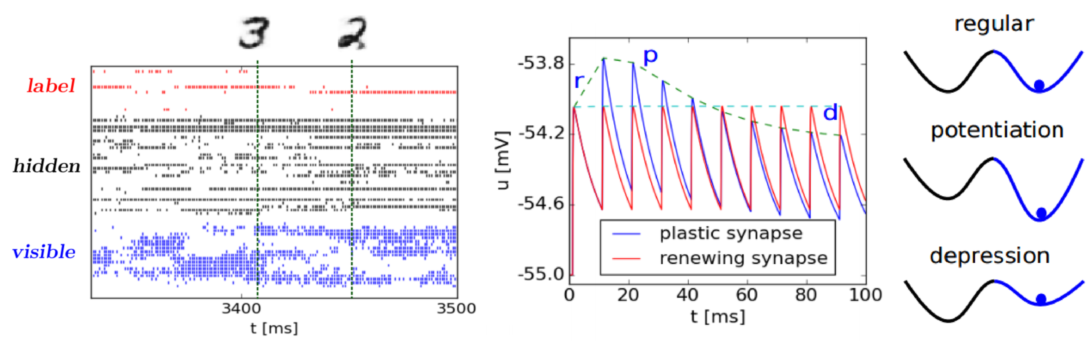
Replacing single units with small subnetworks enables robustness towards analog-hardware-inherent parameter restrictions and noise [5]. The hardware speed-up factor of 10^4 with respect to biological real-time provides a massive advantage for computation with sampled distributions.



Generative and Discriminative Models

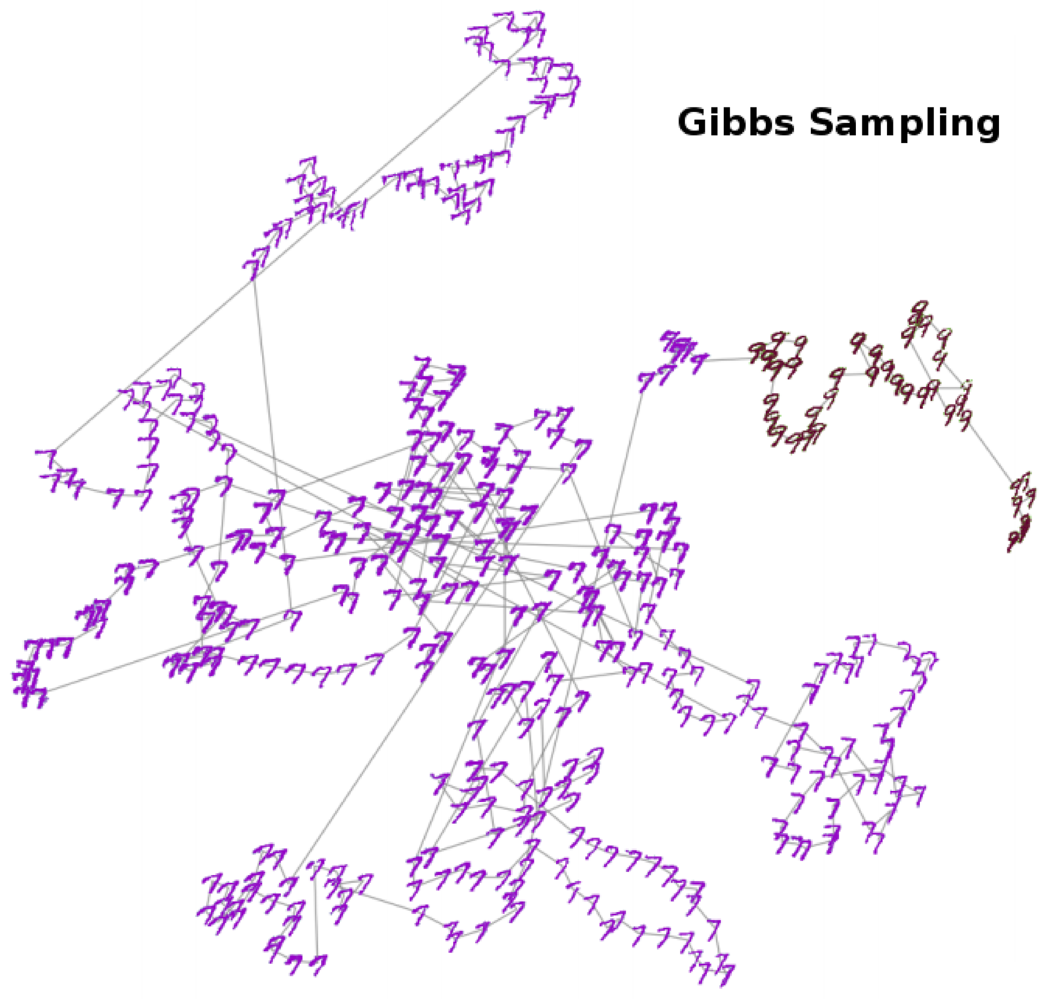
Adding hidden layers allows the representation of arbitrarily complex distributions. Immediate applications lie, for example, in the field of pattern recognition.

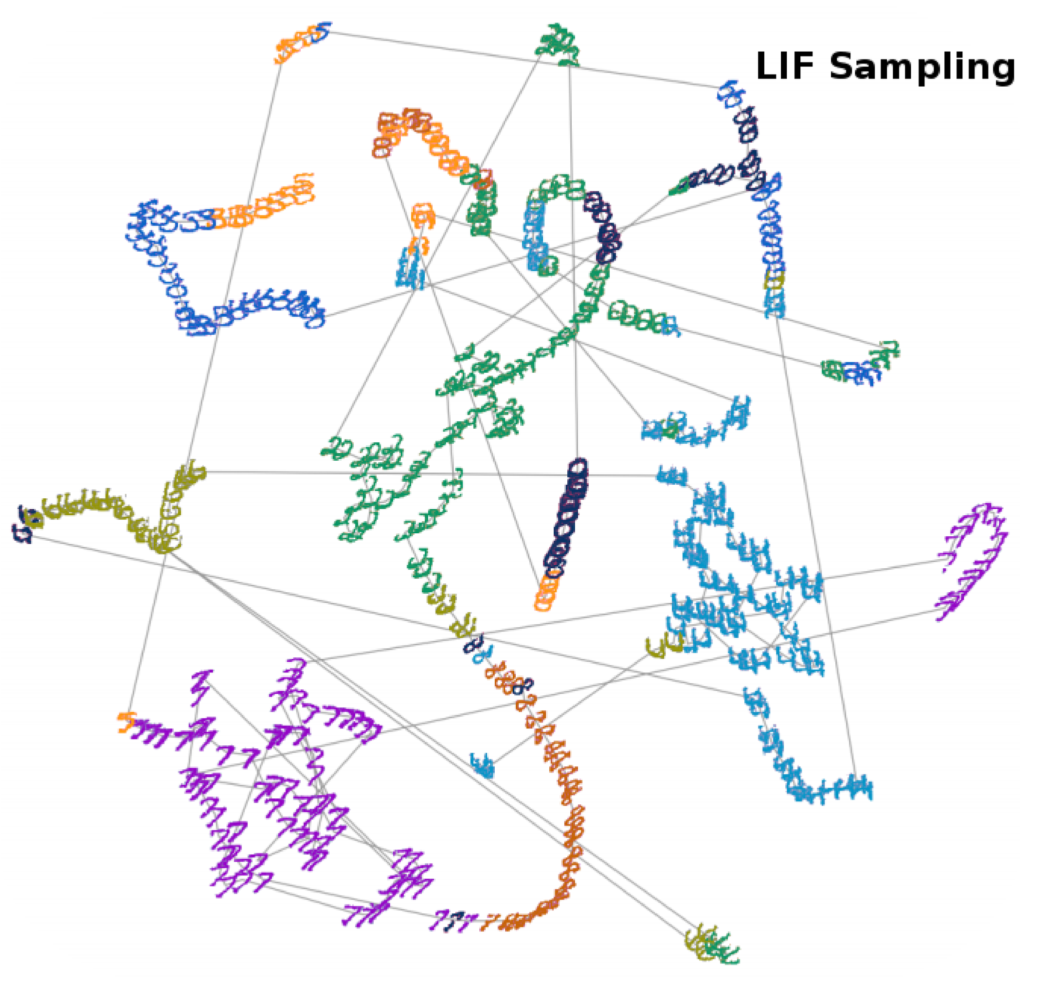




The LIF sampling framework enables the application of various learning algorithms. With a local learning rule, small networks of LIF neurons can achieve high classification performance (97% on MNIST [6]).

Additionally, these networks are, inherently, generative models of the learned data. This is essential for, e.g., pattern completion. While classical machine-learning networks often have difficulties with such tasks, spiking networks can profit from biological mechanisms such as short term plasticity to boost their performance.



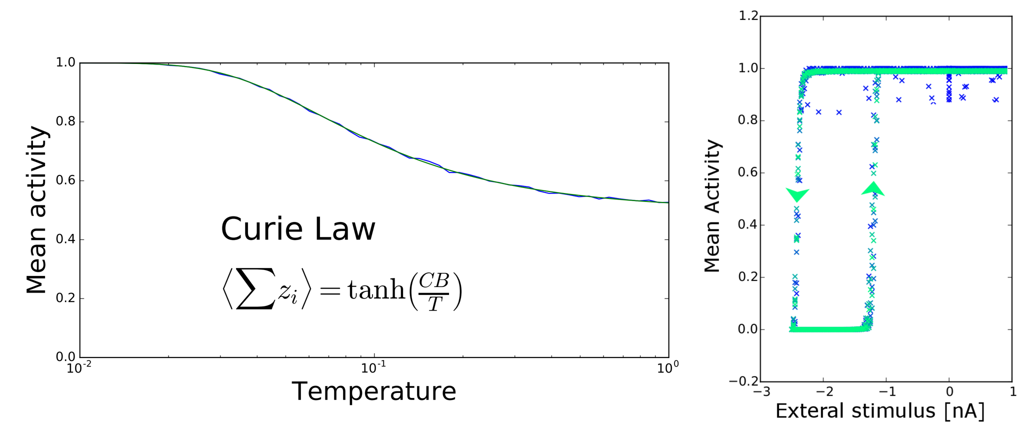


Ensemble Phenomena

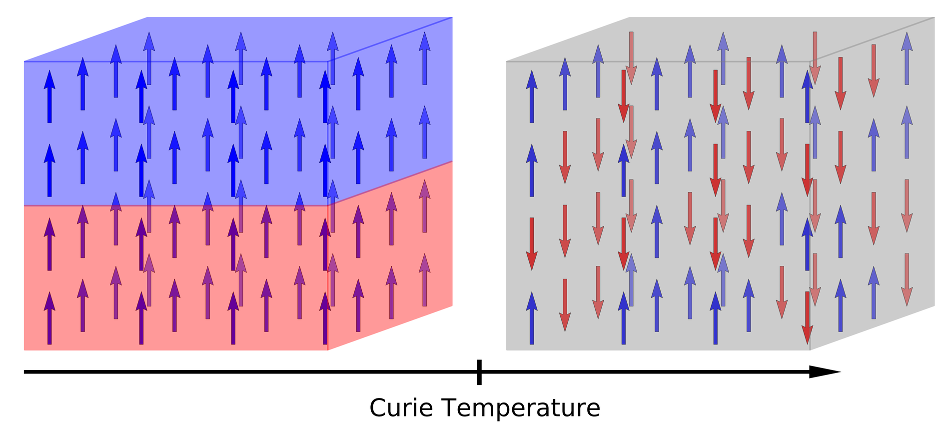
LIF sampling networks share the same hamiltonian (energy function) with well-studies systems from solid-state physics.

The Ising model, used to describe magnetic phenomena, is one such example. In the Ising model, the two-state units represent electron spins. Interactions between these spins correspond to synaptic weights in a neural network.

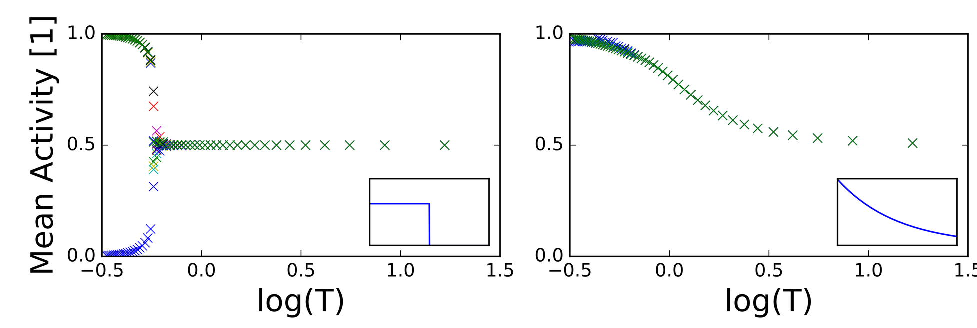
An exact parameter translation from the LIF domain to the Ising domain can be formulated, allowing the emulation of various well-studied phenomena from statistical physics with networks of spiking neurons, such as the Curie law or hysteresis.



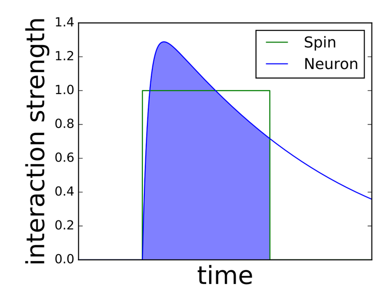
The arguably most interesting ensemble phenomena happen around so-called critical points, where ensembles undergo phase transitions, altering their macroscopic behavior. In ferromagnetic materials, such a transition happens at the Curie temperature, above which they become paramagnetic.



Intriguingly, equivalent LIF networks show no such phase transition. The reason lies in the long-tailed shape of synaptic interaction kernels in biological neural networks.



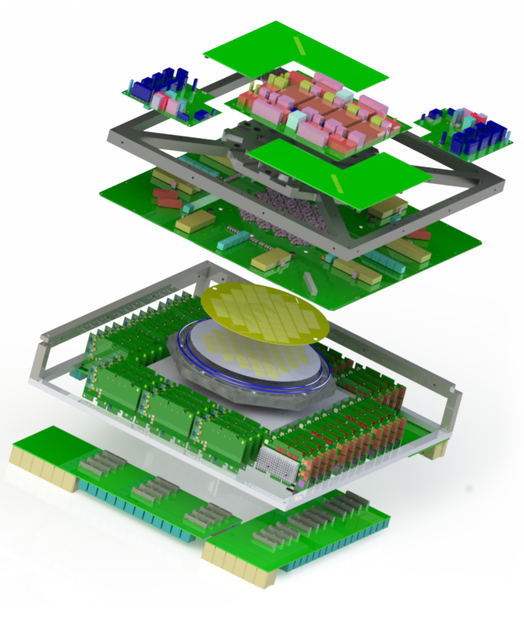
This calls into question the use of oversimplified abstract models for predicting spiking network behavior.



Outlook

The physical emulation of LIF sampling networks in large-scale accelerated analoge systems [7] enables a multitude of applications:

- Fast pattern recognition and completion
- Long-term learning
- Emulation of statistical physics phenomena
- (Quantum) annealing solutions of NP-hard problems



References

[1] J. Fiser, P. Berkes, G. Orbán, and M. Lengyel, "Statistically optimal perception and learning: from behavior to neural representations," *Trends in cognitive sciences*, vol. 14, no. 3, pp. 119–130, 2010.

[2] M. A. Petrovici, J. Bill, I. Bytschok, J. Schemmel, and K. Meier, "Stochastic inference with spiking neurons in the high-conductance state," *Phys. Rev. E*, vol. 94, p. 042312, Oct 2016.

[3] M. A. Petrovici, *Form Versus Function: Theory and Models for Neuronal Substrates*. Springer, 2016.

[4] T. Pfeil, A. Grübl, S. Jeltsch, E. Müller, P. Müller, M. A. Petrovici, M. Schmuker, D. Brüderle, *et al.*, "Six networks on a universal neuromorphic computing substrate.," *Front. in neurosci.*, vol. 7, pp. 11–11, 2012.

[5] M. A. Petrovici, D. Stöckel, I. Bytschok, J. Bill, T. Pfeil, J. Schemmel, and K. Meier, "Fast sampling with neuromorphic hardware," in *Advances in Neural Information Processing Systems (NIPS)*, vol. 28, 2015.

[6] L. Leng, M. A. Petrovici, R. Martel, I. Bytschok, O. Breitwieser, J. Bill, *et al.*, "Spiking neural networks as superior generative and discriminative models," in *Cosyne Abstracts 2016*, (Salt Lake City USA), 2015.

[7] J. Schemmel, J. Fieres, and K. Meier, "Wafer-scale integration of analog neural networks," in *2008 IEEE International Joint Conference on Neural Networks*, pp. 431–438, IEEE, 2008.

http://www.kip.uni-heidelberg.de/vision/publications/

andreas.baumbach@kip.uni-heidelberg.com