Event-driven contrastive divergence: neural sampling foundations

Emre Neftci [1]*, Srinjoy Das [1, 2], Bruno Pedroni [3], Kenneth Kreutz-Delgado [1, 2] and Gert Cauwenberghs [1, 3]

[1] Integrated Systems Neuroengineering Laboratory, Institute for Neural Computation, University of California, San Diego, La Jolla, CA, USA, [2] Electrical and Computer Engineering Department, University of California, San Diego, La Jolla, CA, USA, [3] Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA

A commentary on

Event-driven contrastive divergence for spiking neuromorphic systems
by Neftci, E., Das, S., Pedroni, B., Kreutz-Delgado, K., and Cauwenberghs, G. (2014). Front. Neurosci. 7:272. doi: 10.3389/fnins.2013.00272

In a recent Frontiers in Neuroscience paper (Neftci et al., 2014) we contributed an on-line learning rule, driven by spike-events in an Integrate and Fire (IF) neural network, that emulates the learning performance of Contrastive Divergence (CD) in an equivalent Restricted Boltzmann Machine (RBM) amenable to real-time implementation in spike-based neuromorphic systems. The event-driven CD framework assumes the foundations of *neural sampling* (Buesing et al., 2011; Maass, 2014) in mapping spike rates of a deterministic IF network onto probabilities of a corresponding stochastic neural network. In Neftci et al. (2014), we used a particular form of neural sampling previously analyzed in Petrovici et al. (2013)[1], although this connection was not made sufficiently clear in the published article. The purpose of this letter is to clarify this connection, and to raise the reader's awareness to the existence of various forms of neural sampling. We highlight the differences as well as strong connections across these various forms, and suggest applications of event-driven CD in a more general setting enabled by the broader interpretations of neural sampling.

In the Bayesian view on neural information processing, the cognitive function of the brain arises from its ability to encode and combine probabilities describing its interactions with an uncertain world (Doya et al., 2007). A recent neural sampling hypothesis has shed light on how probabilities may be encoded in neural circuits (Fiser et al., 2010; Berkes et al., 2011). In the neural sampling hypothesis, spikes are viewed as samples of a target probability distribution. From a modeling perspective, a key advantage of this view is that learning in spiking neural networks becomes more tractable than the alternative one, in which neurons encode probabilities, because one can borrow from well-established algorithms in machine learning (Fiser et al., 2010) (see Nessler et al., 2013 for a concrete example).

Merolla et al. (2010) demonstrated a Boltzmann machine using IF neurons. In this model, spiking neurons integrate Poisson-distributed spikes during a fixed time window set by a global rhythmic oscillation. A first-passage time analysis shows that the probability that a neuron spikes in the given time window follows a logistic sigmoid function consistent with a Boltzmann distribution. The particular form of rhythmic oscillation ensures that, even when neurons are recurrently

---

[1] A functionally equivalent formulation can be found in an earlier version posted on arXiv (Neftci et al., 2013).

coupled, the network produces a sample of a Boltzmann distribution for each oscillation cycle. Merolla et al. (2010) also suggest an alternative, more biologically plausible forms of learning induced by rhythmic oscillations that resemble the role of theta oscillations across large neuronal ensembles. Our event-driven CD rule is compatible with Merolla et al.'s sampler because it would simply result in updating weights at every cycle of the rhythmic oscillation.

Shortly after, Buesing et al. (2011) proved that abstract neuron models consistent with the behavior of biological spiking neurons (Jolivet et al., 2006) can perform Markov Chain Monte Carlo (MCMC) sampling of a Boltzmann distribution. Their sampler does not require global oscillations, although these could enable the sampling from multiple distributions within the same network (Habenschuss et al., 2013). To demonstrate the performance of the sampler, a Boltzmann machine was trained off-line using CD. Learning in this model was further extended to on-line updates in a precursor of event-driven CD (Pedroni et al., 2013).

An open question was whether neuron models that describe the biological processes of nerve cells endowed with deterministic action potential generation mechanisms can support stochastic sampling as described with the more abstract spiking forms in Buesing et al. (2011). An answer to this question is relevant for understanding how neural sampling can be instantiated in biological neurons, but also for implementing neural samplers on low-power neuromorphic implementations of spiking neurons (Indiveri et al., 2011). The stochastic nature of neural sampling suggests studying the behavior of neurons under noisy inputs. The diffusion model commonly referred to as the Ornstein-Uhlenbeck process (Van Kampen, 1992) has been the basis of a standard continuous-time stochastic neuron model since the first rigorous analysis of its behavior in Capocelli and Ricciardi (1971). Petrovici et al. (2013) discuss these issues and provide a rigorous link between deterministic neuron models (leaky integrate-and-fire with conductance-based synapses) and stochastic network-level dynamics, as can be observed *in vivo*. In particular, they identify how the high-conductance state caused by Poissonian background bombardment can provide the fast

membrane reaction time required for precise sampling. They provide analytical derivations of the activation function at the single-cell level as well as for the synaptic interaction and investigate the convergence behavior of the sampled distribution at the network level.

O'Connor et al. (2013) employ the Siegert approximation of IF neurons to compute CD updates. The Siegert or diffusion approximation expresses the firing rate of an IF neuron, as a function of input firing rates, under the assumption that all inputs are independent and Poisson distributed. After learning, the parameters of the learned Boltzmann machine are transferred to the equivalent network of IF neurons. Although the off-line CD learning in O'Connor et al. (2013) operated using firing rates rather than spikes, in its basic form, it is functionally equivalent and compatible with event-driven CD under the condition that spike times are uncorrelated.

Our work implements a biologically-inspired algorithm for the purposes of training Boltzmann machines (Neftci et al., 2014). We assumed a neuronal model consistent with biology and realizable in a neuromorphic implementation. Petrovici et al. (2013) provided a deeper physical and mathematical interpretation of neural sampling. Similarly to their approach, we considered the standard leaky IF neuron stimulated by non-capacitively summed pre-synaptic inputs obeying Poisson statistics.

The performance of event-driven CD on the MNIST handwritten digit recognition task was robust to spike probabilities that deviate slightly from the Boltzmann distribution, even though such distributions violate the assumptions of CD formulated for training RBMs. This suggests that event-driven CD provides a general learning framework for biologically-inspired spiking RBMs and is consistent with wide range of neural samplers.

Acknowledgment

References

Berkes, P., Orbán, G., Lengyel, M., and Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331, 83–87. doi: 10.1126/science.1195870

Buesing, L., Bill, J., Nessler, B., and Maass, W. (2011). Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput. Biol.* 7:e1002211. doi: 10.1371/journal.pcbi.1002211

Capocelli R. M., and Ricciardi, L. M. (1971). Diffusion approximation and first passage time problem for a model neuron. *Kybernetik* 8, 214–223. doi: 10.1007/BF00288750

Doya, K., Ishii, S., Pouget, A., and Rao, R. P. N. (2007). *Bayesian Brain Probabilistic Approaches to Neural Coding*. Cambridge, MA; London: MIT Press.

Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* 14, 119–130. doi: 10.1016/j.tics.2010.01.003

Habenschuss, S., Jonke, Z., and Maass, W. (2013). Stochastic computations in cortical microcircuit models. *PLoS Comput. Biol.* 9:e1003311. doi: 10.1371/journal.pcbi.1003311

Indiveri, G., Linares-Barranco, B., Hamilton, T.J., van Schaik, A., Etienne-Cummings, R., Delbruck, T., et al. (2011). Neuromorphic silicon neuron circuits. *Front. Neurosci.* 5:73. doi: 10.3389/fnins.2011.00073

Jolivet, R., Rauch, A., Lüscher, H.-R., and Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *J. Comput. Neurosci.* 21, 35–49. doi: 10.1007/s10827-006-7074-5

Maass, W. (2014). Noise as a resource for computation and learning in networks of spiking neurons. *Proc. IEEE* 102, 860–880. doi: 10.1109/JPROC.2014.2310593

Merolla, P., Ursell, T., and Arthur, J. (2010). The thermodynamic temperature of a rhythmic spiking network. arXiv:1009.5473

Neftci, E., Das, S., Pedroni, B., Kreutz-Delgado, K., and Cauwenberghs, G. (2013). Event-driven contrastive divergence for spiking neuromorphic systems. arXiv:1311.0966

Neftci, E., Das, S., Pedroni, B., Kreutz-Delgado, K., and Cauwenberghs, G. (2014). Event-driven contrastive divergence for spiking neuromorphic systems. *Front. Neurosci.* 7:272. doi: 10.3389/fnins.2013.00272

Nessler, B., Pfeiffera, M., Buesing, L., and Maass, W. (2013). Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Comput. Biol.* 9:e1003037. doi: 10.1371/journal.pcbi.1003037

O'Connor, P., Neil, D., Liu, S.-C., Delbruck, T., and Pfeiffer, M. (2013). Real-time classification and sensor fusion with a spiking deep belief network. *Front. Neurosci.* 7:178. doi: 10.3389/fnins.2013.00178

Pedroni, B., Das, S., Neftci, E., Kreutz-Delgado, K., and Cauwenberghs, G. (2013). "Neuromorphic adaptations of restricted boltzmann machines and deep belief networks," in *International Joint Conference on Neural Networks, IJCNN*. Available online at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6707067

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

Van Kampen, N. G. (1992). *Stochastic Processes in Physics and Chemistry, Vol. 1.* Amsterdam: Elsevier.