Ten thousand times faster: Classifying multidimensional data on a spiking neuromorphic hardware system.

Michael Schmuker[1,2], Daniel Brüderle[3], Sven Schrader[3], Martin P Nawrot[1,2]
email: m.schmuker@fu-berlin.de

1: Neuroinformatics & Theoretical Neuroscience, Institute of Biology, Freie Universität Berlin, Berlin, Germany
2: Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany
3: Electronic Visions, Kirchoff Institute for Physics, Universität Heidelberg, Germany

Motivation

- Neuromorphic computing is an emerging technology that aims at bioinspired high-performance computing with spiking neuronal networks.
- The FACETS/Brainscales neuromorphic hardware system runs networks of spiking neurons with a speedup of 10^4 [1].
- Our aim was to implement a network of spiking neurons that can be trained in a supervised fashion, and to run this network on neuromorphic hardware to classify multidimensional data.
- The structure of the first layers of neuronal processing in the olfactory system provides a well suited template for a neuronal architecture processing multidimensional data.

Challenges

Classifier circuit and learning rule

Challenge: Implement a supervised classifier that operates with spiking neurons.

Solution: A spiking network implemented in PyNN [2], running in the NEST simulator and on the FACETS/Brainscales hardware.

- A feature-encoding layer converges onto an association layer that has winner-take-most properties (Fig. 1).
- The network is trained in a supervised fashion, using a perceptron-like learning rule operating on firing rates (Fig. 1 caption).

Sampling data with virtual receptors

Challenge: Firing rates of spiking neurons can only represent a bounded and non-negative range of values. We need a suitable transformation mapping real values into that value range.

Solution: Virtual Receptors (VR). The response strength of a VR depends on its distance to the presented data point [3].

- We use a Neural Gas (NG) algorithm [4] to distribute virtual receptors in data space, like olfactory receptors sample chemical space (Fig. 2). Receptor response is computed as a function of the distance between data point and receptor.
- This transformation yields a bounded and non-negative representation of any real-valued data set. Dimensionality can be adjusted to exceed the number of original data dimensions (dimensional oversampling), enabling a sparser representation.

Decorrelation

Challenge: Virtual receptors provide correlated data, but the classifier learning rule works best with uncorrelated data.

Solution: Decorrelation through lateral inhibition in a preprocessing layer (see decorrelation layer in Fig. 1).

- Three kinds of inhibitory connectivity matrices were tested: NG-based (inhibitory connections between receptors given by the NG graph edges), correlation (inhibitory weight depends on correlation between receptors), and random lateral inhibition.
- Correlation-based lateral inhibition yields best decorrelation, followed by NG and random connectivity (Fig. 3).
- Benchmarking the impact of decorrelation on classifier performance shows an increase in accuracy with increasing lateral inhibition, but no clear preference for a specific method, probably a ceiling effect of the spiking classifier (Fig. 4).

Implementation on neuromorphic hardware

Challenge: Hardware neurons vary in their firing rate response (Fig. 5B). The classifier learns on output rate, so rate variation has negative impact on the classifier performance.

Solution: Calibrate the sensitivity of neuron groups (glomeruli) to achieve more homogeneous representation of input rates.

- We developed a calibration method that balances inhomogenities across model glomeruli (Fig. 5C).
- After calibration, the hardware implementation of the classifier reaches the same performance as in the simulator (Fig. 5D).

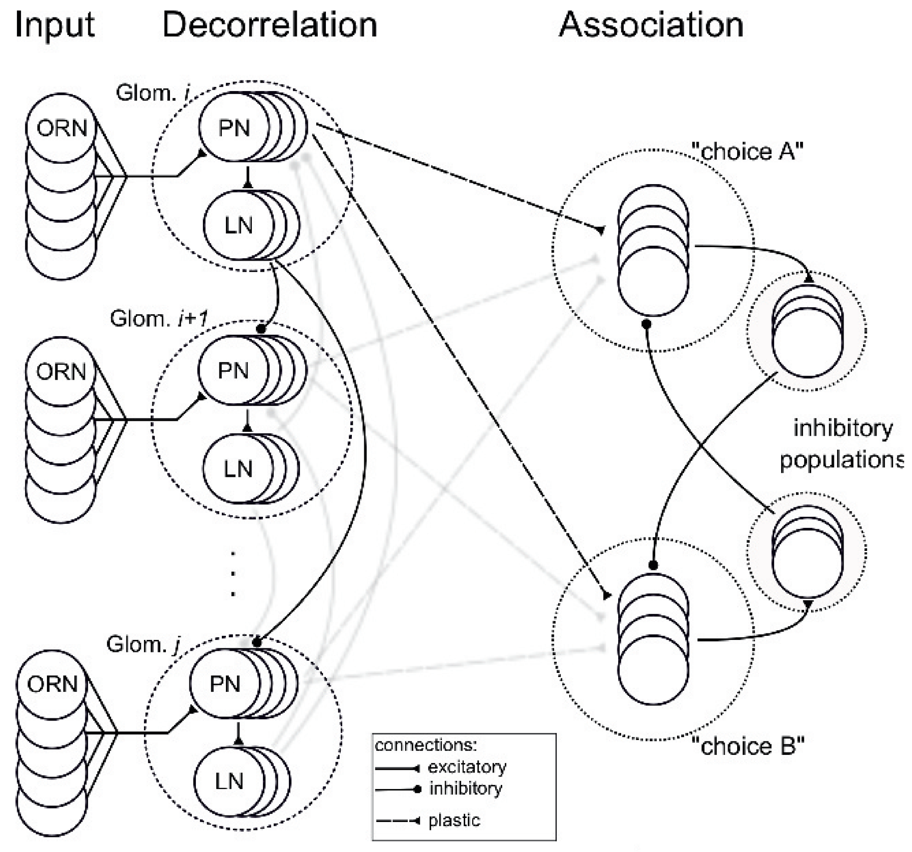


Fig. 1: Schematic of the neuromorphic classifier. Input neurons (ORNs) emit poisson spike trains with averag rates according to the numeric values of the input pattern. Synaptic weights between the decorrelation and association layers are subject to plasticity during classifier training (dashed lines). Classifier training algorithm:

1. Present labeled data point, i.e., set firing rates according to pattern.
2. Determine winner population in the association layer.
3. Update weights: if association was correct, increase weights of active synapses; decrease weights if association was incorrect.
4. Repeat until all training data points have been presented.

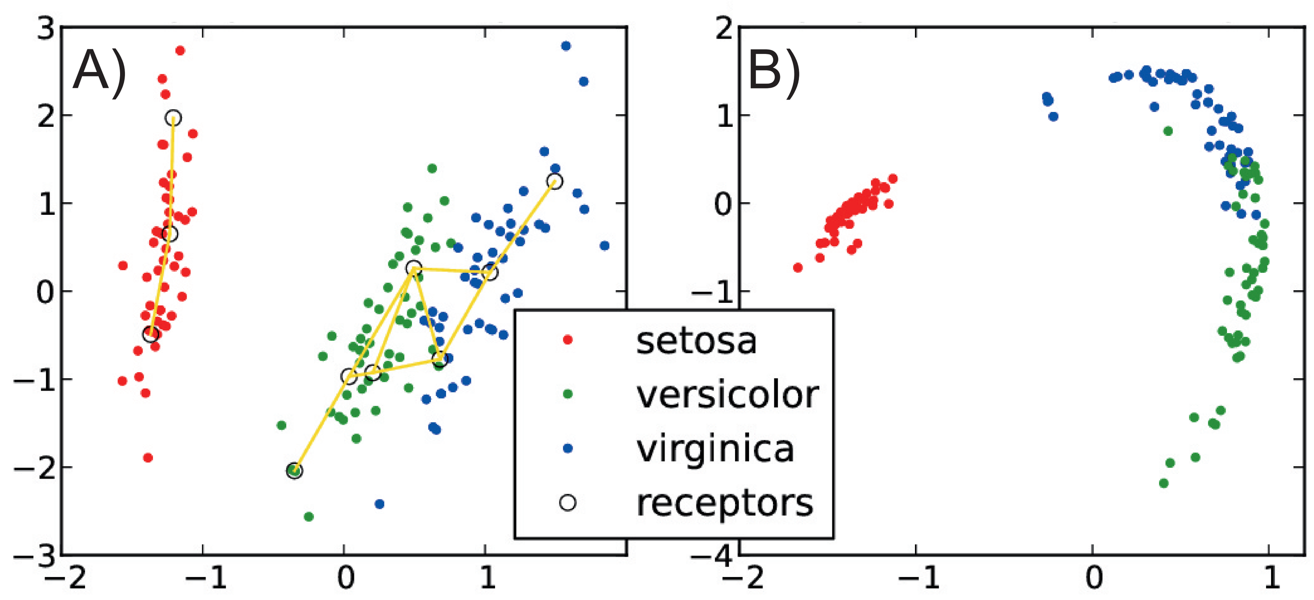


Fig. 2: Sampling Fisher's iris data set [5] with virtual receptors. **A)** Virtual receptors after training the neural gas (NG) (2D PCA projection of 4D space). Yellow lines represent edges in the NG graph. **B)** Data representation in virtual receptor space (2D projection of 10D space). The pronounced structure indicates a large amount of residual correlation.

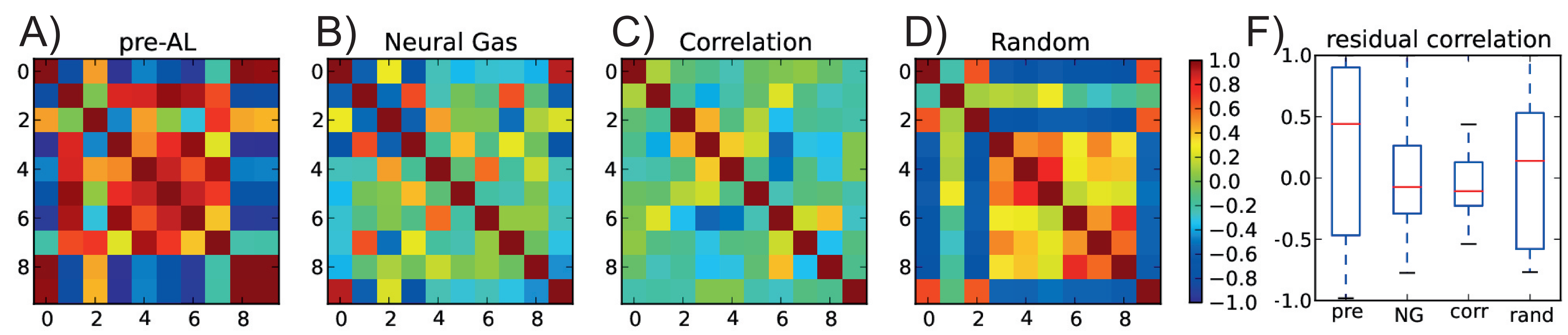
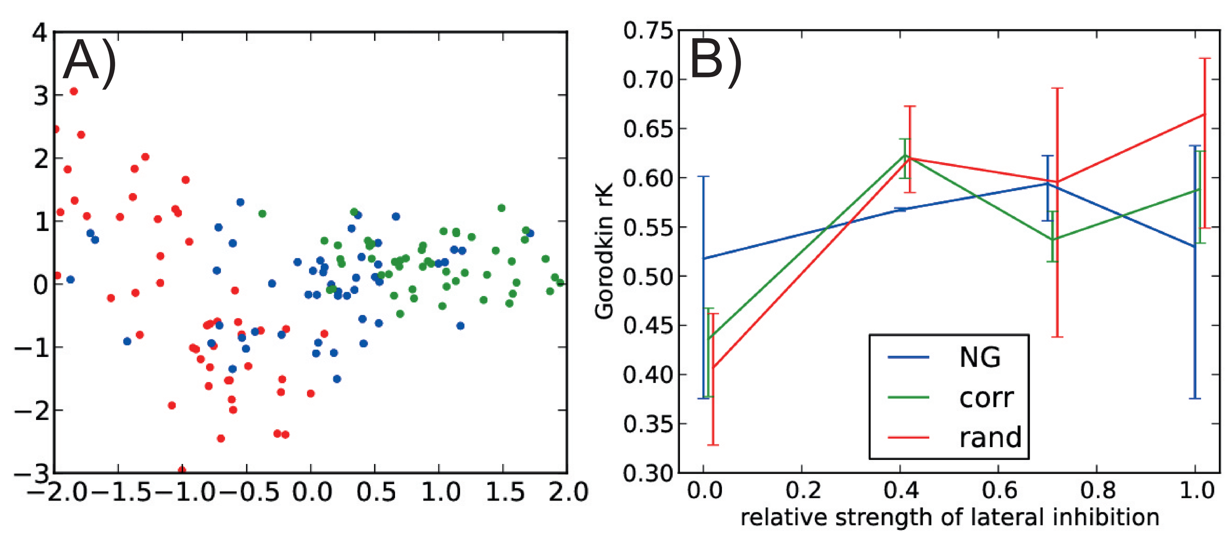


Fig. 3 Correlation matrix of 10 virtual receptors **A)** before processing, **B)** after processing with Neural-Gas-based lateral inhibition, **C)** correlation-based lateral inhibition, **D)** random lateral inhibition. **F)** Boxplots of the residual correlation for A-D.

Fig. 4 A) Data representation after correlation-dependent lateral inhibition (2D projection of 10D space). Class overlap may be an artefact from low-dimensional embedding - separability must be judged by a classifier. **B)** Effect of decorrelation on classifier performance (five-fold crossvalidated, Gorodkin's K-category correlation coefficient [6]). Error bars: min/max of three repetitions.



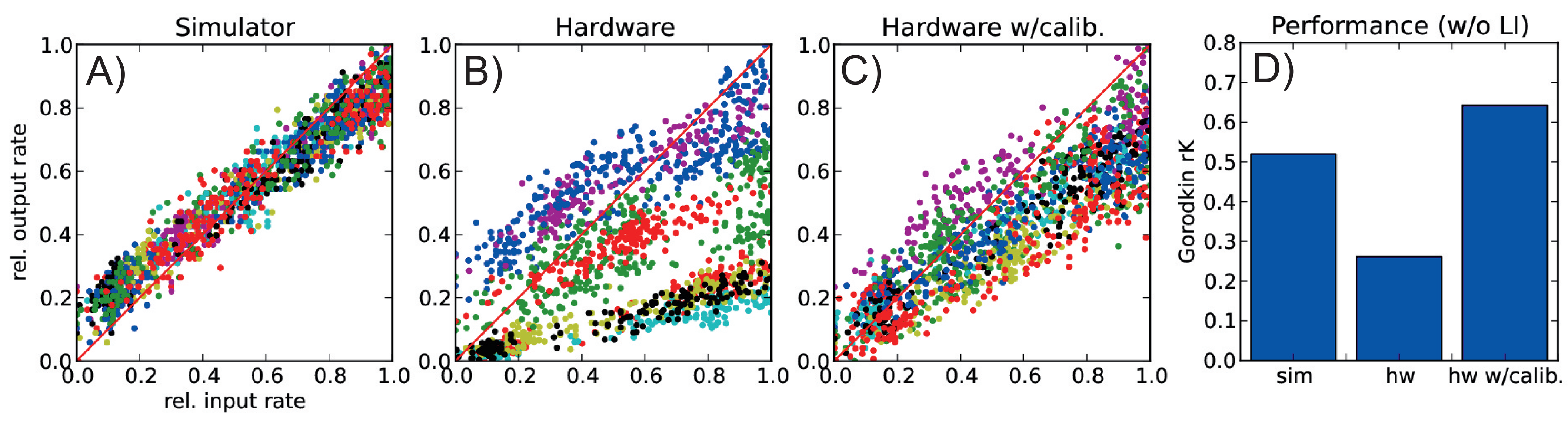


Fig. 5 A) Input-output relation (firing rate) in the simulator. Firing rate averaged over all neurons in a group (*glomerulus*). Colors denote different glomeruli. Each point corresponds to one stimulus presentation. **B)** Same setup on in hardware neurons, **C)** after calibration. **D)** Classifier performance on simulator and on hardware (without lateral inhibition).

Conclusions

- Virtual receptors provide a non-negative representation of any real-valued data set, suitable for processing with spiking neurons.
- Correlation-based lateral inhibition efficiently reduced residual correlation from the virtual receptor representation.
- Decorrelation improved performance of the spiking classifier.
- We successfully implemented the classifier on a neuromorphic hardware system with high speedup factor, an important step towards bioinspired high-performance computing.

References

[1] Brüderle D, Bill J, Kaplan B, Kremkow J, Meier K, Müller E & Schemmel J (2010). Simulator-like exploration of cortical network architectures with a mixed-signal VLSi system. In *Proc. IEEE ISCAS* 2010, p. 2784–8787. **[2]** Brüderle D, Müller E, Davison A, Muller E, Schemmel J & Meier K (2009). Establishing a novel modeling tool: a python-based interface for a neuromorphic hardware system. *Front Neuroinf.* 3, 17. **[3]** Schmuker M & Schneider G (2007). Processing and classification of chemical data inspired by insect olfaction. *PNAS* 104, 20285-20289. **[4]** Martinetz T & Schulten K (1991). A“ neural-gas” network learns topologies. In Artificial Neural Networks T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas, eds. (North-Holland: Elsevier B.V.), pp. 397-402. **[5]** Fisher RA (1936). The Use of Multiple Measurements in Taxonomic Problems. *Ann Hum Genet* 7: 179–188. **[6]** Gorodkin J (2004). Comparing two K-category assignments by a K-category correlation coefficient. *Comp Biol Chem.* 28, 367-374.

Funding: DFG SPP 1392: SCHM2472/1-1 - BMBF BCCN Berlin 01GQ1001D