

A distributed VLSI attractor network learning visual stimuli in real time and performing perceptual decisions

Massimiliano Giulioni Paolo Del Giudice¹

Italian National Institute of Health, Viale Regina Elena 299, 00161 Rome, Italy
{massimiliano.giulioni, paolo.delgiudice}@iss.infn.it

Introduction

Attractor dynamics makes the time evolution of a dynamical system collapse into a subspace of its available state space; in the simplest instance, the one which will concern us, the subspace is a set of discrete points (or trajectories constrained to their neighborhood). To the extent that the attractor endpoint depends on the initial conditions of the dynamics, the state space is partitioned into *basins* for each attractor state. In the cognitive metaphor, a class of stimuli (initial states falling into the same basin) is defined by the fact that all its members cause the *retrieval* of the same prototypical representation (the attractor state). The stimulus and the retrieved representation belong to the same state space; for the recurrent neural networks that we will consider, each point in the state space is a distribution of activity in the network: the one initially imposed by a stimulus evolves into a selective self-sustained activity pattern (i.e. one that depends on the initial state, but not on its details).

Attractor neural models have been initially developed and improved to account for a wide array of experimental evidence related to working memory. It is increasingly becoming clear that the dynamic scheme has a wider scope. Models based on bistable (two attractors) or multi-stable (multiple attractors) networks have been proposed as theoretical underpinnings for understanding perceptual decision mechanisms and processes of information integration, as well as multi-stable perception and binocular rivalry.

Motivated by such a multiplicity of putative roles that attractor dynamics might play in neural computation, and the related interest in versatile components for computation in neuromorphic systems, we undertook to develop VLSI neural networks, communication infrastructures, and strategies for experiments and analysis, suited to match this level of complexity.

The core of our systems is the neuromorphic chip described in detail in [8] and [9], comprising 128 Integrate-and-fire (IF) neurons and about 16384 plastic reconfigurable synapses.

¹and INFN, Sezione Roma1, P.le A. Moro 5, Rome, Italy

Implemented in the chip are the IF neurons with constant leakage described in [6], designed after [10]. Synapses implement the theoretical model proposed in [3], and the design is essentially the one we described in [9], to which we refer the reader for details. For the purpose of the present report we just mention that synapses are bi-stable (two states of potentiated or depressed efficacy), Hebbian (potentiation occurs for co-active pre- and post-synaptic neurons, depression for active pre- and inactive post-synaptic), spike-driven (and stochastic, to the extent that spike trains are highly irregular in the asynchronous states we consider), and possess a regulatory mechanism ('stop-learning') which inhibits further changes of a synapse whose post-synaptic neuron is already driven towards the activity state that the synaptic change would reinforce.

The communication infrastructure is based on the parallel asynchronous AER standard (Address-Event Representation, [2]), and the PCI-AER board [14] which, for a system of AER-compliant neuromorphic chips, allows to monitor their activity, to program the inter-chip connectivity and to inject synthetic spikes into the chips (e.g. to emulate external stimuli or further neural populations, not physically implemented).

One element of novelty in the present work is an improved infrastructure (to be described in detail elsewhere) to host multiple identical neural chips, which is flexible enough to allow the user to easily map different network topologies onto the physical chips [5].

Figure 1 illustrate the main elements of the setup used for the experiments reported in the present chapter: the board hosting two identical neural chips, the front-end electronics for control, the microcontroller and FPGA for managing USB communication with a PC and proper addressing of configuration commands to the chips; the retina chip [11] providing real-time stimuli to the on-chip neural network; the PCI-AER board referred to above.

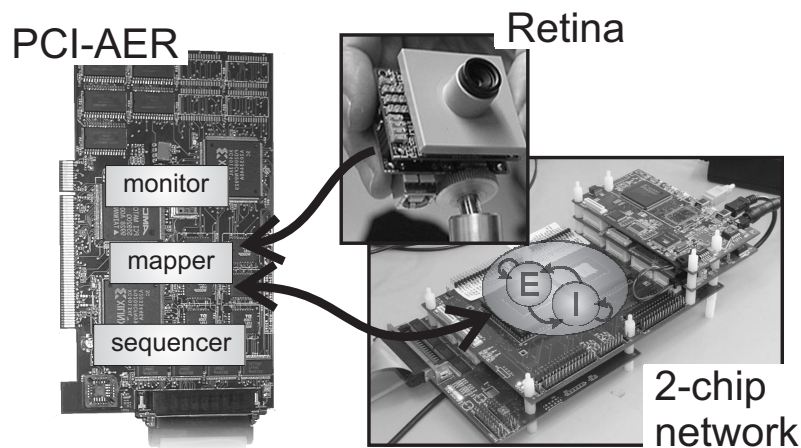


Figure 1. Hardware setup.

In the following three Sections we will summarize results obtained in the attempt to 1) generate and control attractor dynamics in the VLSI chip; 2) success-

fully *learn* in real time the synaptic structure needed to support attractor states as representations of real visual stimuli; 3) demonstrate attractor-based stochastic decision making by a richer architecture, implemented in the same VLSI chips.

Attractor dynamics in the VLSI chip

In the present Section we want to demonstrate the basic computational element, a recurrent neural network which, because of its self-excitation, is able to sustain states of elevated activity, elicited by a stimulus and persistent after its removal. We illustrate how the associated basic phenomenology can be obtained and controlled in a VLSI neuromorphic chip.

In the experiments reported here we take full advantage of the innovative reconfigurability properties of the chip, in particular: the synaptic connectivity graph can be set, allowing for very sparse as well as full-connected networks; each synapse can be set as excitatory or inhibitory; each synapse can be set as recurrent (connecting two neurons physically implemented on chip) or 'AER' (conveying spikes from other AER-compliant devices or from synthetic spike sources); each synapse can also be set as plastic (i.e. the above mechanisms governing the efficacy changes are active) or not.

Each pre-synaptic contact can receive either spikes from neurons on-chip or from external devices via the AER communication infrastructure.

In this Section we will refer to the network shown in Figure 2.

Input from the PC is intended to provide both stimuli and an adequate background activity for the on-chip populations. The only difference between the two on-chip excitatory populations is their recurrent connections: E_{att} has a higher level of self-excitation than E_{bkg} .

To implement such a network in hardware we first configure the on-chip synaptic matrix and then run a calibration procedure to set the chip bias levels for the VLSI neurons and synapses such that they correspond to the theoretical values of J , β and τ_{arp} . Setting the synaptic matrix presents no difficulties since the operation is based on a simple digital protocol handled by a dedicated microcontroller. Setting biases is a more demanding task which is performed comparing on-chip measures with simulations and theory predictions at various levels, from the characterization of single circuits to the measure of mean population behavior. For details on this procedures the reader is referred to [4].

In the experiments described in the present Section we set all the synapses as non plastic, as we only want to illustrate the attractor dynamics in a network whose synaptic connectivity is preset in order to support such states. For this purpose, all the recurrent excitatory synapses between neurons belonging to E_{att} are set to be potentiated, with respect to the ones between neurons in the E_{bkg} population and the ones connecting E_{att} to E_{bkg} .

Figure 3 illustrates some relevant aspects of the attractor dynamics the network robustly generates. The bottom traces in the two panels show the time course of the stimulation protocol. A stimulus is implemented through an increase in the excitatory spike rate from the external (AER) neural populations. In the left panel the network, starting from a low activity stable state, receives a small

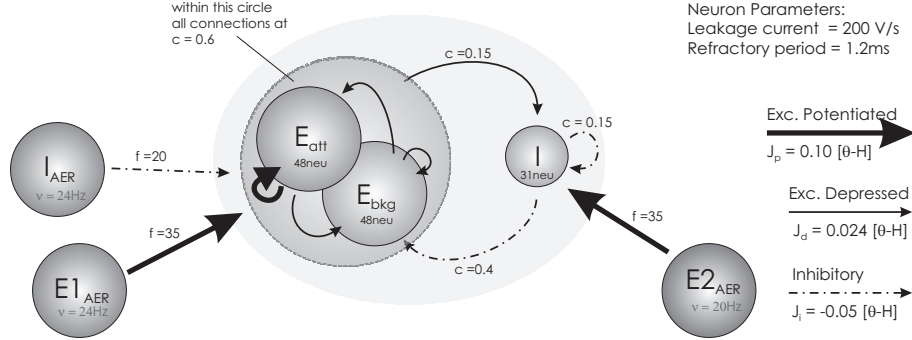


Figure 2. Network architecture. Two excitatory populations (E_{att} and E_{bkg}), each composed of 48 neurons, and one inhibitory population (I), composed of 31 neurons, are self- and reciprocally connected and receive external stimuli via the AER bus from three populations ($E1_{AER}$, $E2_{AER}$ and I_{AER}) simulated on a PC. Intra and inter-populations connectivity levels c are reported in the figure. c is the fraction of neurons in the target population with which a given neuron makes synaptic contacts. Synaptic efficacy values J are in units of $\theta - H$, where θ is the neuronal firing threshold and H is the reset potential. The leakage current $\beta = 200(\theta - H) \text{ s}^{-1}$ and the absolute refractory period is $\tau_{arp} = 1.2 \text{ ms}$

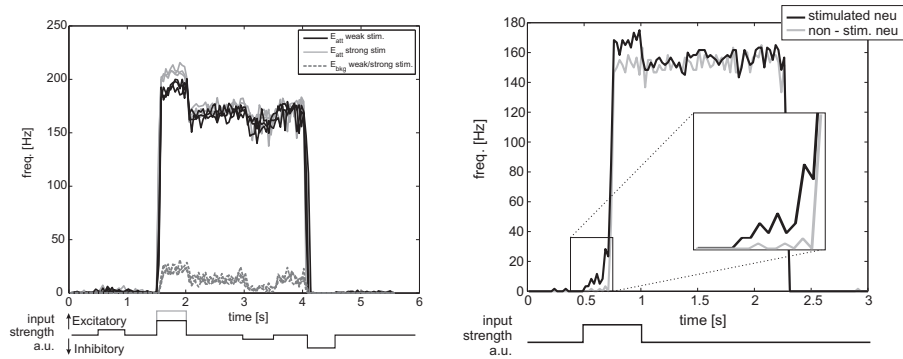


Figure 3. Left: mean population firing rate. Solid lines for the on-chip E_{att} population; dashed lines for on-chip E_{bkg} . Right: recruiting ability of attractor dynamics. Only half of the neurons belonging to E_{att} receive the external stimulus. Their mean firing rate is reported in black. In gray the mean firing rate of the remaining neurons of E_{att}

external input, to which it reacts slightly increasing its activity and falling back to the low activity state as the stimulus is removed. A larger stimulus to E_{att} excites the nonlinear response of the network (solid curves: firing rate of E_{att} neurons; dashed curves: firing rates of E_{bkg} neurons): the firing rate of stimulated neurons jumps to high values and, after the stimulus is released it relaxes to elevated values (which turn out to be in good agreement with the theoretical predictions, which we do not discuss here). This is the exquisite manifestation of the strong excitatory feedback in the network; the stimulus caused the network to make a transition from the low activity attractor state to the high activity one. It is noted that if a small (inhibitory) stimulation intervenes (at $t=3\text{s}$ in the figure) while the network reverberates in the high attractor state, the consequent

small decrease in the firing rate does not destabilize the high state, to which the network relaxes back after the stimulus ends. A larger inhibitory input (at $t=4s$ in the figure) brings the network back to the low attractor state. Attractor states are stereotyped activity patterns that should not depend on the details of the stimulus that ignited them. Indeed for a larger stimulus (the gray line starting at $t=1.5$ in the bottom trace), causing of course higher firing rates in the stimulated population as long as it is applied, the network relaxes to the same activity levels as in the first case, after the stimulus has been removed. Another manifestation of the attractor nature of the observed high state is illustrated in the right panel of Figure 3. In this case only half of the neurons belonging to E_{att} are stimulated (black curves). It is seen that, again because of the massive excitatory feedback, non-stimulated neurons (gray curves) are rapidly recruited and after the stimulus is removed the E_{att} population as a whole relaxes to the same attractor state as before.

In summary, Figure 3 illustrates both the property of *working memory* and *pattern completion* of our silicon attractor network.

Learning visual stimuli in real time

While it is interesting to show that the VLSI network demonstrated in the previous Section is able to exhibit robust attractor dynamics for a suitable preset synaptic structure, it is by no means obvious that such synaptic structure can be autonomously generated by the dynamic coupling between the ongoing, stimulus-induced neural activity and the consequent changes in the plastic synapses, which in turn affects the neural response to stimuli.

This is difficult to achieve for theoretical models (successful examples of autonomous build up of attractor states in model neural networks are still scarce, see [7] [1] [13], let alone for neuromorphic chips.

In this section we demonstrate the learning ability of our on-chip network in a simple, but non trivial, example.

The network architecture is similar to the one depicted in Figure 2 but here we use a total of 128 excitatory neurons and 50 inhibitory ones. The network is distributed over two identical chips hosted in the setup of Figure 1, the inter-chip connectivity being defined by a writable look-up-table in the PCI-AER board.

The connectivity levels are 0.25 for the excitatory neurons and 0.18 for the inhibitory ones. Connectivity between excitatory and inhibitory neurons is $c = 0.11$ in both directions.

External stimuli come from a neuromorphic retina [11] which encodes in patterns of AER spikes variations of light intensity in its visual fields. Spikes emitted by the retina are routed towards the chips through the PCI-AER board.

The entire system runs in real-time and both neural activities and synaptic states are monitored (again through the PCI-AER board).

The visual field of the retina is divided in 121 macropixels, each corresponding to a square of 10×10 retina pixels, and each macropixel is the input source to a single neuron in the on-chip network. Since each incoming spike affects the post-synaptic target neuron as a pulse with finite duration, to avoid overlapping of

successive pulses we set a fan out of each macropixel to 64 synapses of its single target excitatory neuron of the on-chip network. From this mapping it is possible to visualize the network firing rates to match the geometric arrangement of the macropixels, see the matrices in figure 4.

All synapses are set to be depressed at the beginning of the experiment; given a chosen stimulus structure, learning is expected to selective potentiate synapses connecting neurons which are co-active for that stimulus.

A T-shaped visual stimulus was chosen for the experiments illustrated in the present Section, which activates about half of the neurons in the network. Since the retina is sensitive only to temporal variations of luminosity contrast, the "T" is presented to the retina as a pattern of black dots randomly moving on a white background within the "T" contour. The portion of visual field not occupied by the "T" is filled with a sparse noisy background, and we set parameters such that a minimum difference of about 20 Hz was recorded among retina neurons "seeing" the "T" and those seeing just the background.

The implemented Hebbian synaptic dynamics is the one briefly described in the Introduction, and we refer the reader to references therein for details.

The learning protocol consists in repetitive presentations of the visual stimuli to the retina. Each presentation lasts about 1.5 seconds as shown in the top trace in figure 4 where we also report the network response. The initial rise in the activity is mainly due to rapid potentiation of some 'fast' synapses (in the chip, inhomogeneities and mismatch cause a wide dispersion of effective synaptic parameters, in the face of which we seek robust learning). When the stimulus is removed, the network goes back to a very low activity state.

As the stimulus is repeatedly presented, a selective synaptic structure builds up, such that the mature network is able, after the removal of the stimulus, to sustain a noisy, but recognizable, representation of the learnt "T" shape. Figure 4 illustrates the behaviour of the mature network ($t > 120$ s). The stimulus presented at $t=120.5$ s elicits a fast response of the selective excitatory sub-population, while the activity of the background sub-population show little increase; after stimulus removal at $t=122$ s the selective activity is kept elevated (because of the strong self-excitation generated by learning). After the high state persists for more than 2 seconds, it decays spontaneously back to the low state; indeed, finite-size fluctuations are large in such small systems, and affect the stability of attractor states in an important way.

We also checked (not shown) that: the mature network exhibits the expected pattern completion ability : upon presenting a degraded versions of the learnt T-shape, the network is able to retrieve the activity pattern corresponding to the complete "T"; learning is robust with respect to the choice of the initial condition for the synaptic matrix.

Stochastic decision making

We mentioned in the Introduction that the attractor picture has recently gained an increasingly wider scope; one relevant example is provided by attractor models proposed to account for experimental evidence (both in electrophysiology and

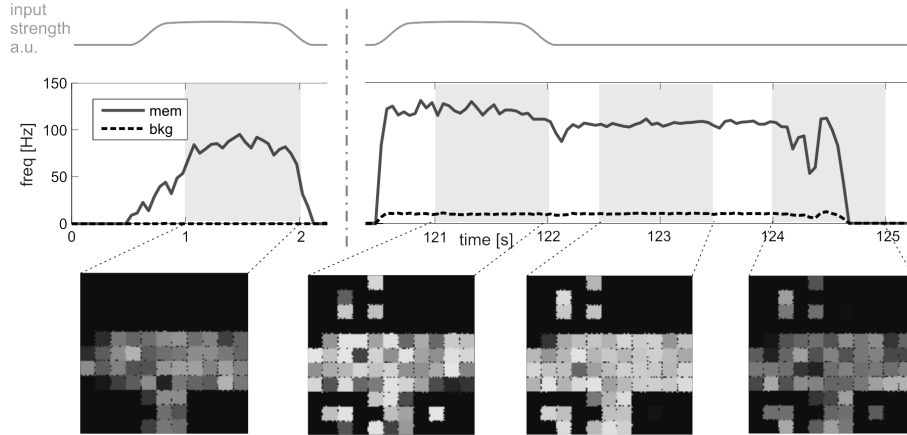


Figure 4. Learning. Solid line is the mean firing rate of those excitatory neurons receiving spikes from the retina macropixels activated by the "T". Dotted line is the mean firing rate of neurons "seeing" the background. Matrices are the reconstruction of the visual stimulus obtained from the firing rates of the on-chip excitatory neurons. Each matrix is computed by averaging the firing rate of single neurons over 1 second (white = 150Hz, black = 0Hz)

psychophysics) related to perceptual decision. In this final Section we show how, by reconfiguring our network to match the architecture proposed in [15].

In the typical experiment the subject is required to report an assessment as to the net direction of motion in a display in which dots move randomly on a screen up to varying degrees of coherence in the direction of motion towards two opposite visual targets (high coherence implying of course easier trials; at zero coherence the subject guesses at chance level). Both electrophysiology and psychophysics evidence suggest that, at least in some parietal areas, an integration of sensory information precedes the decision, the integration time reflecting the difficulty of the task, and that at parity of task difficulty the response time and quality expose high variability.

Attractor models proposed in this context and modifications thereof (see e.g. [16] [12]), share the basic architecture illustrated in Figure 5, which essentially implements an attractor-based noisy Winner-Take-All network. Two excitatory populations $E1$ and $E2$ receive input stimuli $S1$, $S2$ and compete through an inhibitory population. Once the competition is resolved (i.e. the decision is made) the representation of the choice is kept active in time as an attractor state.

The VLSI network implementing the architecture of Figure 5 is again distributed over two chips. Synaptic plasticity is disabled and we focus on the stochastic competitive dynamics in action.

Figure 6 describes the decision dynamics in the extreme case of zero coherence, when the decision is determined just by activity fluctuation in the integration phase. For illustrative purposes, the two panels provide different representations of the decision dynamics. In the left panel we show the time course of the average firing rates of the competing excitatory populations $E1$ and $E2$. It is seen that at the beginning of the stimulation period (gray strip) the symmetric input to $E1$ and $E2$ cause an equal increase of their activities (freq_{E1} and freq_{E2}), up to fluctuations that ultimately engage the inhibitory population that resolves

the competition. At this point freq_{E1} and freq_{E2} rapidly diverge, and after the end of the stimulation the winning (losing) population relax to the high (low) activity attractor states. The right panel describes the dynamics as a trajectory in the $(\text{freq}_{E1}, \text{freq}_{E2})$ phase plane (in log scale to make it easier to appreciate the noisy initial ramping of the system along the diagonal, before the competition is resolved).

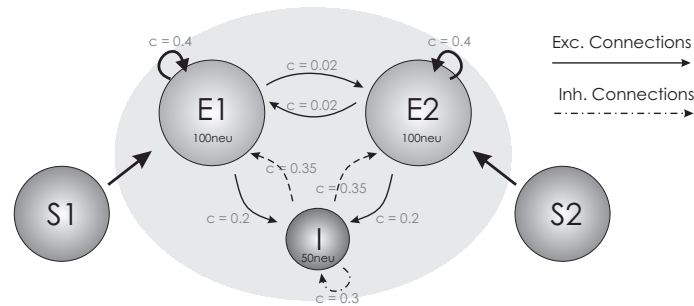


Figure 5. Network architecture for stochastic decision making

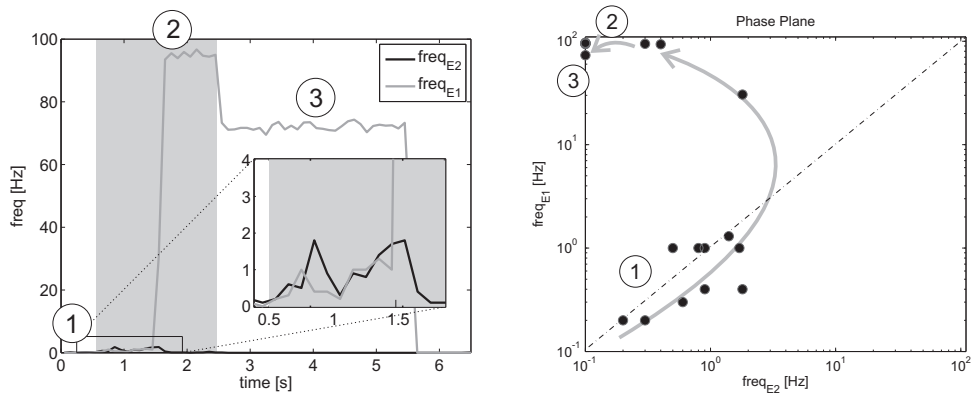


Figure 6. Left: time course of the average firing rates of the two competing excitatory populations during the decision process. Right: Decision dynamics as a trajectory in the $(\text{freq}_{E1}, \text{freq}_{E2})$ phase plane.

Conclusions

In the present Chapter we provided a preliminary account of a series of experiments aiming at demonstrating various features of attractor dynamics in VLSI neuromorphic chips of spiking neurons and plastic Hebbian synapses. The effort was motivated by the idea that attractor dynamics constitutes a versatile computational primitive for information processing in neural systems, and as such should be regarded as an important dynamic element to be implemented in the neuromorphic systems purporting to emulate the principles of information processing in the nervous system.

Acknowledgments

The work described in the the present chapter has been done at various stages in collaboration with the following colleagues, whom we like here to thank: Patrick Camilleri, Federico Corradi, Luca Federici, Vittorio Dante, Maurizio Mattia, Simeon Bamford, Jochen Braun, Davide Badoni, Mario Pannunzi. We warmly thank Tobi Delbruck for assistance in using the retina chip he kindly made available to us. This work has been supported in part by the EU project CORONET.

References

- [1] D. J. Amit and G. Mongillo. *Neural Computation*, 15:565, 2003.
- [2] K. A. Boahen. *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, 47(5):416–434, 2000.
- [3] J. M. Brader, W. Senn, and S. Fusi. *Neural Computation*, 19:2881–2912, 2007.
- [4] P. Camilleri, M. Giulioni, M. Mattia, J. Braun, and P. Del Giudice. In *Proc. IEEE International Joint Conference on Neural Networks*, 2010.
- [5] Luca Federici. *Master Thesis, unpublished*. University of Rome “Sapienza”, 2011.
- [6] S. Fusi and M. Mattia. *Neural Computation*, 11:633, 1999.
- [7] P. Del Giudice, S. Fusi, and M. Mattia. *Journal of Physiology Paris*, 97:659–681, 2003.
- [8] M. Giulioni. PhD thesis, University of Rome “Tor Vergata”, 2008.
- [9] M. Giulioni, P. Camilleri, V. Dante, D. Badoni, G. Indiveri, J. Braun, and P. Del Giudice. In *Proc. IEEE International Conference on Electronics Circuits and Systems*, pages 678–681, 2008.
- [10] G. Indiveri, E. Chicca, and R. Douglas. *IEEE Trans. on Neural Networks (in press)*, 17(1):211–221, 2006.
- [11] P. Lichtsteiner, C. Posch, and T. Delbruck. *IEEE Journal of Solid State Circuits*, 43(2):566–576, 2008.
- [12] D. Marti, G. Deco, M. Mattia, G. Gigante, and P. Del Giudice. *PLoS One*, 3(7):e2534, 2008.
- [13] M. Pannunzi, G. Gigante, M. Mattia, G. Deco, S. Fusi, and P. Del Giudice. in preparation. 2011.
- [14] P. Del Giudice V. Dante and A. M. Whatley. The neuromorphic engineer newsletter available at <http://ine-web.org/fileadmin/templates/docs/NME3.pdf>. 2005.
- [15] X. J. Wang. *Neuron*, 36:955968, 2002.
- [16] X. J. Wang. *Neuron*, 60(2):215–234, 2008.