

Energy Saving Accounts for the Suppression of Sensory Detail

Terry Bossomaier*, Lionel Barnett†, Vaenthan Thiruvardchelvan* and Herbert Jelinek*

*Centre for Research in Complex Systems

Charles Sturt University

Bathurst, Australia

Email: {tbossomaier,vthiru,hjelinek}@csu.edu.au

†Sackler Centre for Consciousness Science

School of Informatics

University of Sussex

Brighton, BN1 9QH, UK

Email: lionelb@sussex.ac.uk

Abstract—High functioning autistic people can exhibit exceptional skills with numbers, eidetic imagery and recall of concrete detail, as brought to popular attention in the film *Rain Man*. However, it now transpires that these skills are to some extent latent within all of us. We do not have access under normal circumstances to this concrete detail, yet brain stimulation experiments show that this detail exists in all of us. This paper proposes that one of the reasons for this lies in the brain’s need to conserve energy. Computer simulations using a spiking neural network support this hypothesis.

I. INTRODUCTION

The evidence from high functioning autistic individuals shows the overwhelming advantage of concept formation in the human brain. Such individuals tend to have weak concept formation but can have very powerful perception and memory for detail. The evolutionary significance is abundantly clear. What is less clear is *why the raw detail to which these people have access is not available to everybody else*. The surprising thing, revealed by direct brain stimulation, is that this detail is *not* destroyed on the way to conscious awareness, but is somehow *blocked from access*. This paper provides a novel solution to this conundrum.

Early indicators that some of this low-level detail might be accessible came from studies on victims of stroke and brain injury, where, for example, a person might discover the ability to draw realistically. Snyder and Mitchell [26] predicted that such access might be obtained using brain stimulation techniques in which the conceptual part of the brain was blocked, because concepts inhibit lower level detail [25].

It transpired that this was indeed the case. The direct brain stimulation techniques, *Transcranial Magnetic Stimulation* (TMS) and the more recent technique, *Transcranial Direct-Current Stimulation* (TDCS) can be used to turn-off part of the brain. By targeting the anterior temporal lobe in the left hemisphere—a brain area highly involved in concept formation and storage—it is possible to block access to concepts and thus release access to lower-level detail. In the first such study,

now nearly a decade ago, drawing and proof-reading [27] were found to be enhanced by TMS. So, for example, it is hard for many people to see the word “the” when it is repeated on a following line. Their ability to spot the error goes up when the meaning of the sentence is blocked by brain stimulation. Likewise numerosity [24] (rapidly estimating the number of objects in the field of view, and inspired by an incident in the film *Rain Man*), also goes up with TMS to the left anterior temporal lobe. Over the subsequent decade, a diverse range of higher-level cognitive phenomena have been shown to be enhanced through dis-inhibition with brain stimulation. False memory, where like objects may get mixed up in memory tests (e.g. chair instead of stool), can be reduced [2]. Even the ability to solve visual puzzles can be enhanced [4].

There are numerous arguments for why this might be the case, such as the possibility of computational overload, discussed further in Section IV. In this era of information overload, such an explanation is at first sight appealing, but is hard to quantify with our existing knowledge of the brain.

Closely linked to computational overload is the energy cost of neural computation. The human brain uses about 20% of the body’s energy [19] and various evolutionary changes, such as the appearance of meat in the diet, may have allowed the brain’s energy consumption to grow. Navarette et al. [15] show that in over 100 species of mammal, adipose deposits correlate negatively with encephalisation. Thus fat storage and increased brain size are orthogonal strategies for avoiding starvation.

Laughlin and Sejnowski [14] show that the brain’s overarching network structure is consistent with preserving energy. The energy required for the transmission of nerve impulses, or spikes, and synaptic transmission are very tightly optimized, approaching the thermodynamic limits within cellular constraints [13]. Neuronal spikes account for a significant fraction of neuronal energy usage [1].

The idea that the number of spikes might be kept to a minimum to save energy began with the idea of *sparse coding* in sensory systems [17], [23]. More recently, cells

have been observed which fire strongly when the subject is exposed to stimuli corresponding to a particular person, say Bill Clinton, and to very little else [18], [6]. They respond to the *concept*, and can be activated by pictures, voice or unique events. Obviously for most people such a cell would fire very infrequently. The alternative distributed representation might have many cells coding for all US presidents. All of these cells would be active for any president, thus making their average activity much higher.

However sparse coding is not the only way to reduce energy consumption by neurons using action potentials (APs). Changing the kinetics of the ion channels involved in generating the spike can reduce the energy requirements of the APs. Sengupta et al. [22] show that considerable differences in the relative cost of spike transmission versus the energy of synaptic transmission may be found, depending upon the exact ion channel kinetics, for example between giant squid neurons and those in mouse cortex.

This strong need to conserve energy suggests a possible explanation for why the raw sensory input is not accessible other than through external means such as TMS. *It is turned off to save energy.* Snyder et al [25] and Bossomaier and Snyder [3] propose a *concept model* for how inhibition mechanisms might generate the observed effects of TMS. The effect is to turn off the inhibitory mechanisms, dis-inhibiting their targets.

Inhibition is of course widespread in the brain, and the prefrontal cortex—the area with most development over other primates—abounds in inhibitory effects. But evidence is now emerging that even sensory perception in early areas such as primary visual area V1 depends upon top-down modulation, of which a large part is inhibitory [21], [7].

Feedback mechanisms are a common way of modulating input from lower processing areas of the cortex to higher processing areas. Visual processing streams provide a good example where higher order visual areas display an inhibitory top-down activity to lower visual processing areas such as V1 [21], [7]. However these models only consider connectivity patterns in the cortex related to visual processing. Jelinek and Elston [10] have shown that on a cellular level, processing complexity increases from V1 to prefrontal cortex, with layer-III pyramidal cell dendritic branching patterns becoming more complex and larger, thus requiring more energy. Higher visual processing areas deal more with conceptual phenomena by integrating simple information bits from lower processing areas.

Such top-down effects reduce activity at lower levels. Zhang et al [29] show that in inferotemporal cortex, activity corresponding to a particular object is vastly different depending upon whether attention is focussed on that object.

In this paper we show that spiking neural networks, even using the most basic approximation to the established Hodgkin-Huxley spike-dynamic equations [8], can exhibit significant energy savings in these inhibition models. The energy cost of neural computation is split between the generation of spikes and synaptic activity. The relative proportion varies across species [22], but the focus in this article is on the spike activity.

We consider two cases. The first implements our concept model of the previous paragraph. The second uses a Bayesian or attention approach to reduce energy costs even further. The essential feature of both models is the inhibition of inputs as soon as a concept has been activated.

II. THE SIMULATION MODEL

The simplest approximation to the Hodgkin-Huxley equations is the Leaky Integrate and Fire model. Izhikevich [9] points out that this neuron is capable of only a few of the 20 or so behaviors of which the full Hodgkin-Huxley model is capable. However, it is used here because if a very simple model can generate the behavior we observe, then so can any of the more complex models. This assures that the model is reasonably robust to parameter variations.

Eqn 1 shows the model for one neuron, where R is resistance, I the input current, u the membrane voltage and τ the time constant.

$$\frac{du}{dt} = -\frac{u}{\tau} + \frac{IR}{\tau} \quad (1)$$

The synaptic activation is represented by an alpha function with another time constant given by τ_s :

$$\varepsilon(t) = \frac{1}{\tau_s} e^{1-t/\tau_s} \quad (2)$$

The two simulation models use the same type of neuron, although the time constants are not the same.

A. Concept and Inhibition

In Model 1, we use a local inhibitory circuit, shown in figure 1. Since an eye fixation takes around 200msec [12], we assume this time represents the minimum time for which a concept would remain active. The inhibitory circuit requires around 20msec. It does not matter if the input spikes come in as a single volley or as some Poisson process; if the maximum spike rate is around 100 spikes per second, the concept cell can see about 2 spikes in 20msec, and should it see a spike from every cell, then it takes 40msec to turn the input cells off. This would represent an spike saving of around a factor of 5.

B. Model 2: Prior Knowledge and Intention

There is abundant evidence of the use of Bayesian information processing throughout sensory and cognitive processing. For the purposes of this paper, the implication is that only a small subset of feature detectors need to fire to recognize something, given the assumption that something is going to appear. Sometimes a single cue, such as somebody's hair color, might be enough to distinguish two people apart. In this case it is not necessary to wait for all feature cells to fire. Just a few cells may suffice, in which case the inhibition can start sooner. This is the essence of Model 2, illustrated in figure 2.

Now assume that we have attentional control or a mindset that one is going to see K5 or K7 represented by the cells labelled *prior* in figure 2. The facilitating cell is activated

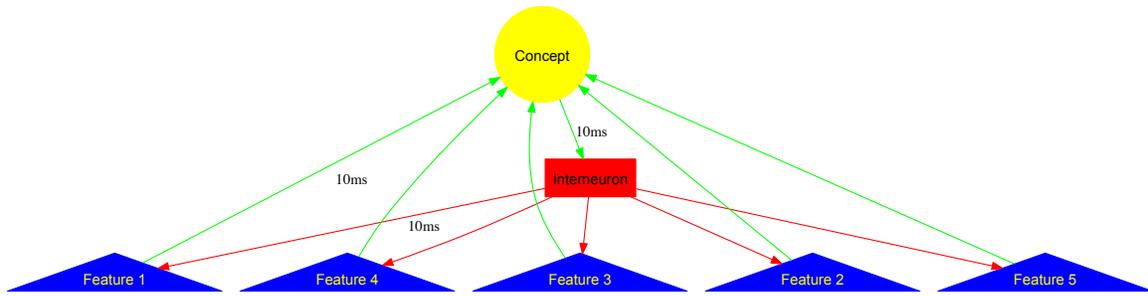


Fig. 1. The basic model. Sensory signals are the features in blue, of which there may be many more than 5. Connections from features to concept and concept to inhibitory interneuron are excitatory. Connections from the interneuron to the features are inhibitory

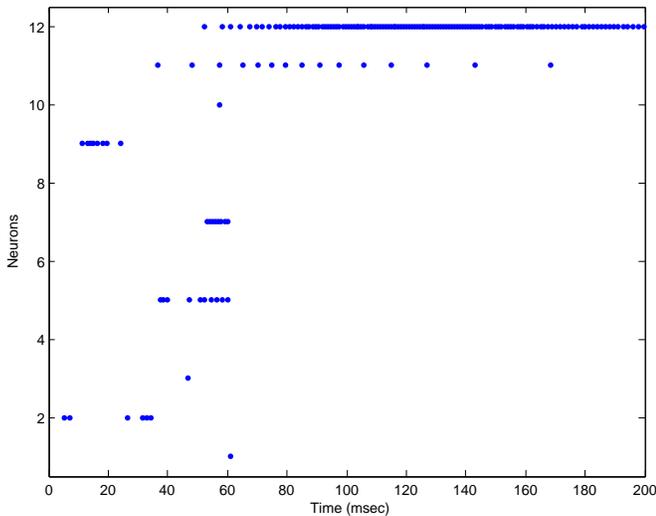


Fig. 3. Activity of the network of model 1. Cells are laid out along the y-axis. The top cell is the inhibitory interneuron, the next cell down is the concept and the remainder are the features. Each dot represents a spike event. The inhibition in neuron 12 sets i-n after the concept neuron has started to fire.

from higher up, but is agnostic as to whether K5 or K7 appears. It fires slowly with a long recovery time and brings a small subset of features closer to threshold. This costs a small number of spikes and synaptic events, since on average only one facilitating cell will fire. Now only this small number of features needs to be activated for the concept to trigger. But since these features lead over the remainder, only they will be allowed to fire.

III. RESULTS

Simulations were carried out in Matlab using the Biological Neural Network toolbox [20]. Figure 3 shows the spiking patterns for Model 1. The features are suppressed for the duration of activation of the concept, representing at least a substantial decrease in energy usage. Whereas the activity of the concept and inhibitory neurons are maintained throughout the simulation of 200msec, the activity of the feature neurons rapidly dies away. Without the inhibition, there firing would also be maintained. Figure 4 shows the average number of spikes in each neuron over 100 runs.

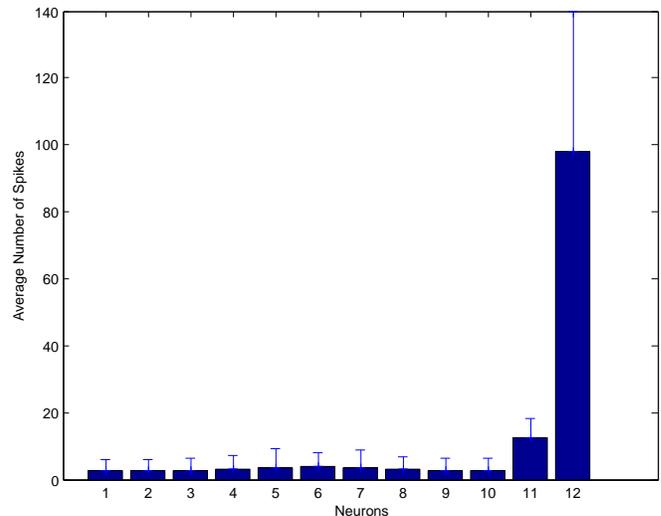


Fig. 4. The average number of spikes in for each neuron. Neurons 1–10 are the input features, neuron 11 the concept and neuron 12 the inhibitory neuron

Turning now to the case of the prior or attention neuron. This pre-activates some of the features as shown in figure 5. Figure 6 shows the average number of spikes over 100 runs of the simulation

In this article only one concept neuron appears, but a single prior could pre-activate any number of feature neurons, subserving more than one concept, but biasing the outcome to some subset of possible concepts which could occur in a given context.

IV. DISCUSSION

The conjecture that it is possible to reduce the spikes generated by a feature might seem surprising. But there is already substantial work demonstrating that a single spike per neuron may be enough for pattern recognition. Thorpe et al. [28] discovered that people can make *very* rapid decisions on whether pictures contain animals, so rapid that they are likely to be able to use only a single spike along the path from retina to associative cortex. Subsequent computational models demonstrate the feasibility of the single spike model.

The information overload argument suffers from a lack of understanding of what the brain can actually do on a large

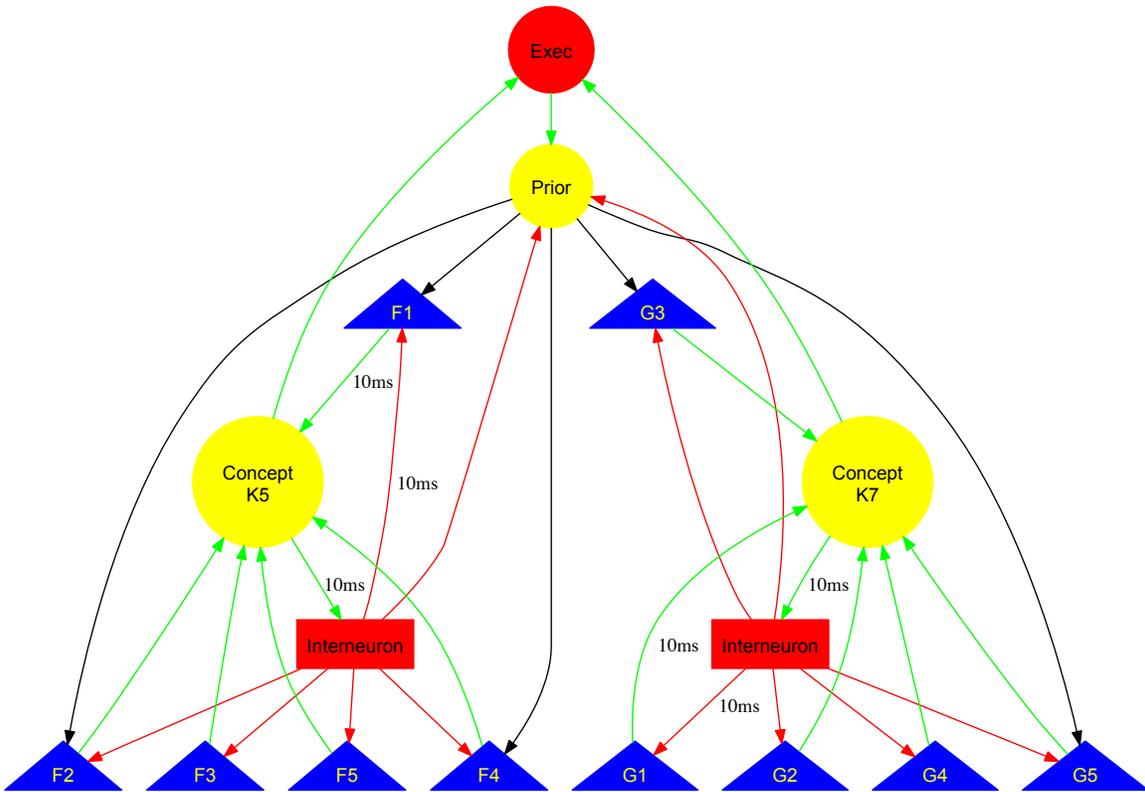


Fig. 2. The model with attention. Features, concepts and inhibitory interneuron are the same as in model 1. Here we have two concepts and a single attention/prior neuron. This has excitatory connections to a small number of feature detectors.

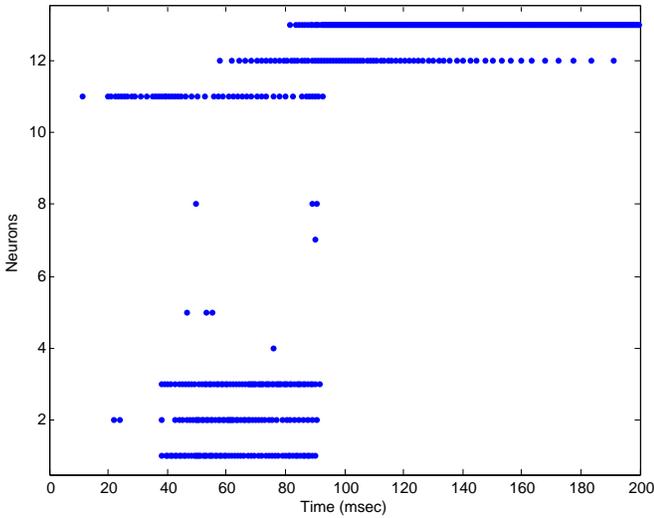


Fig. 5. Effective of an attention bias or prior assumption. The prior neuron (number 11) is already active and the three sensitized neurons fire first (1–3). Firing in the other feature detectors is suppressed.

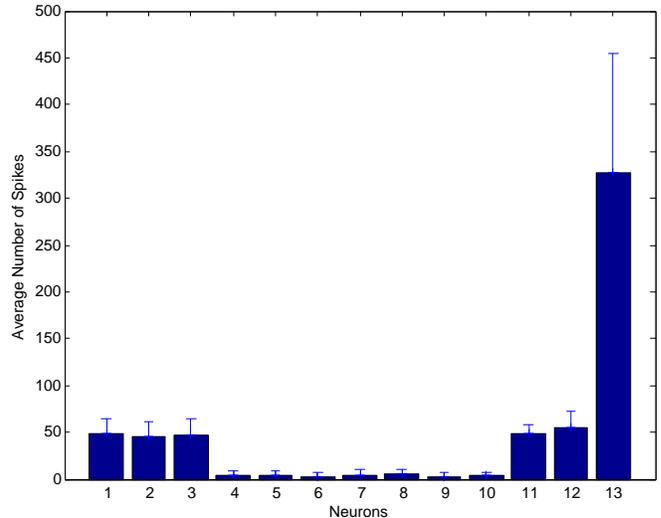


Fig. 6. Mean number of spikes. The prior neuron is number 11, the concept, 12 and the inhibition, 13.

scale. We know something about the capacity of simple neural networks, such as the number of patterns storable in a Hopfield net or the Vapnik-Chervonenkis Dimension of feedforward networks. But on the scale of the cortex we have only the most rudimentary of measures.

Darwin [5] famously remarked: *to suppose that the eye [...] could have been formed by natural selection, seems, I freely confess, absurd in the highest degree.* A century later, Nilsson and Pelger [16] showed that evolving an eye was actually relatively easy. By the same token, without a very good model of the computational limits of the brain, the information-

overload argument is hard to substantiate.

On the other hand, people are good at blocking out stimuli. The noise of a busy road, the drone of the engines in an aircraft cabin, the buzz of other speakers in a cocktail party – all demonstrate our remarkable capacity to shut out interference when we so desire. But this blocking is reversible and we *can* turn our attention to the distractions themselves. Kochlin [11] shows that the pre-frontal cortex can select one context and block others in choosing an action.

The blocking of sensory detail seems to be hardwired and is *not* switchable. To turn off this inhibition would require additional circuits to turn off conceptual information. In general, such circuits do not seem to have evolved, and external techniques such as TMS are required for their inhibition. This would make sense: strategies to save energy would be likely to have evolved much earlier than the expansion of the cortex and its sophisticated filters and control mechanisms.

REFERENCES

- [1] D. Attwel and S. Laughlin, “An energy budget for signaling in the grey matter of the brain. j cereb blood flow metab,” *J. Cereb. Blood Flow Metab.*, vol. 21, pp. 1133–1145, 2001.
- [2] P. Boggio, F. Fregni, C. Valasek, S. Ellwood, R. Chi, J. Gallate, A. Pascual-Leone, and A. Snyder, “temporal lobe cortical electrical stimulation during the encoding and retrieval phase reduces false memories,” *PLoS One*, vol. 4, no. 3, p. e4959, 2009.
- [3] T. Bossomaier and A. Snyder, “Absolute pitch accessible to everyone by turning off part of the brain?” *Organised Sound*, vol. 9, pp. 181–189, 2004. This paper provided a conceptual framework for understanding how the ability to label notes with their pitch on hearing them, known as absolute pitch, was inhibited by higher level musical concepts.
- [4] R. P. Chi and A. Snyder, “Facilitate insight by non-invasive brain stimulation,” *PLoS ONE*, vol. 6, no. 2, p. e16655, 02 2011. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0016655>
- [5] C. Darwin, *On the Origin of the Species*. John Murray, 1859.
- [6] K. Gaschler, “One person, one neuron?” *Scientific American*, vol. 17, pp. 77–82, 2006.
- [7] C. Gilbert, M. Ito, M. Kapadia, and G. Westheimer, “Interactions between attention, context and learning in primary visual cortex,” *Vision Research*, vol. 40, pp. 1217–1226, 2000.
- [8] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *Journal of Physiology*, vol. 117, pp. 500–544, 1952.
- [9] E. Izhikevich, “Which model to use for cortical spiking neurons,” *IEEE Trans. Neural Networks*, vol. 15, pp. 1063–1070, 2004.
- [10] G. Jelinek, H. and Elston, “Dendritic branching of pyramidal cells in the visual cortex of the nocturnal owl monkey: A fractal analysis,” *Fractals*, vol. 11, no. 4, pp. 391–396, 2003.
- [11] E. Kochlin, C. Ody, and F. Kouneiher, “The architecture of cognitive control in the human prefrontal cortex,” *Science*, vol. 302, no. 5648, pp. 1181–1185, 2003.
- [12] M. Land and B. Tatler, *Looking and Acting: Vision and Eye Movements during Natural Behaviour*. Oxford University Press, 2009.
- [13] S. Laughlin, R. d. Ruyter van Steveninck, and J. Anderson, “The metabolic cost of neural computation,” *Nature Neuroscience*, vol. 1, no. 1, pp. 36–41, 1998.
- [14] S. Laughlin and T. Sejnowski, “Communication in neural networks,” *Science*, vol. 301, no. 5641, pp. 1870–1874, 2003.
- [15] A. Navarette, C. v. Schaik, and K. Isler, “Energetics and the evolution of human brain size,” *Nature*, vol. 480, pp. 91–94, 2011.
- [16] D.-E. Nilsson and C. Pelger, “A pessimistic estimate of the time required for an eye to evolve,” *Proc. Royal Soc. Lond. B*, vol. 256, pp. 53–58, 1994.
- [17] B. Olshausen and D. Field, “Sparse coding with an overcomplete basis set: A strategy by v1?” *Vision Research*, vol. 37, no. 3, pp. 3311–3325, 1997.
- [18] R. Quiroga, L. Reddy, G. Kreiman, C. Koch, and L. Fried, “Invariant visual representation by single neurons in the human brain,” *Nature*, vol. 435, pp. 1102–1107, 2005.
- [19] M. Raichle and D. Gusnard, “Appraising the brain’s energy budget,” *PNAS*, vol. 99, no. 16, pp. 10237–10239, 2002.
- [20] A. Saffari, “Biological neural network toolbox for matlab.” [Online]. Available: <http://www.ymer.org/amir/software/biological-neural-networks-toolbox>
- [21] R. Schäfer, E. Vasilaki, and W. Senn, “Perceptual learning via modification of cortical top-down signals,” *PLoS Computational Biology*, vol. 3, no. 8, p. e165, 2007.
- [22] B. Sengupta, M. Stemmler, S. Laughlin, and J. Niven, “Action potential energy efficiency varies among neuron types in vertebrates and invertebrates,” *PLoS Comput Biol*, vol. 6, no. 7, p. e1000840, 07 2010. [Online]. Available: <http://dx.doi.org/10.1371/journal.pcbi.1000840>
- [23] E. Simoncelli and B. Olshausen, “Natural images statistics and neural representation,” *Annual Rev. Neurosci.*, vol. 24, pp. 1193–1216, 2001.
- [24] A. Snyder, H. Bahramali, T. Hawker, and D. Mitchell, “Savant-like numerosity skills revealed in normal people by magnetic pulses,” *Perception*, vol. 35, no. 6, pp. 837–845, 2006.
- [25] A. Snyder, T. Bossomaier, and D. Mitchell, “Concept formation: object attributes dynamically inhibited from conscious awareness,” *Journal of Integrative Neuroscience*, vol. 3, pp. 31–46, 2004.
- [26] A. Snyder and D. Mitchell, “Is integer arithmetic fundamental to mental processing?: The mind’s secret arithmetic,” *Proc. Royal Soc. London B*, vol. 266, pp. 587–592, 1999.
- [27] A. Snyder, E. Mulcahy, J. Taylor, D. Mitchell, P. Sachdev, and S. Gan-devi, “Savant-like skills exposed in normal people by suppressing the left fronto-temporal lobe,” *Journal of Integrative Neuroscience*, vol. 2, no. 2, 2003.
- [28] S. Thorpe, A. Delorme, and R. v. Rullen, “Spike-based strategies for rapid processing,” *Neural Networks*, vol. 14, pp. 715–725, 2001.
- [29] Y. Zhang, E. Meyers, N. Bichot, T. Serre, T. Poggio, and R. Desimone, “Object decoding with attention in inferior temporal cortex,” *PNAS*, vol. 108, pp. 8850–8855, 2011.