

Improvements of the Neural Network Simulation of the Vestibulo-Oculomotor Integrator

JEAN-PHILIPPE DRAYE,^{a,b,c} GUY CHERON,^{d,e}
GAËTAN LIBERT,^c AND EMILE GODAUX^d

^c*"Parallel Information Processing" Laboratory
Faculté Polytechnique de Mons
B-7000 Mons, Belgium*

^d*Department of Neurosciences
Faculty of Medicine
University of Mons-Hainaut
Mons, Belgium*

^e*Laboratory of Biomechanics
University of Brussels
Brussels, Belgium*

Connectionist models, also referred to as artificial neural networks, have focused much attention for the last few years on the principle of parallel distributed processing in the brain. Nevertheless, the main challenge of artificial neural networks simulation of real systems is to reach biological interpretability. For this purpose, the oculomotor system, particularly for its role in the horizontal vestibulo-ocular reflex (VOR), offers a number of simplifying features that are reasonably comprehensible.¹ In this system, the modifications to the vestibular signals by the vestibular nuclei (VN) and other neurons before they are transmitted to the motoneurons include a mathematical integration by the final neural integrator (NI); this one performs an essential function: the maintenance of position of both eyes in space. It has been localized in the nucleus prepositus hypoglossi (NPH).²

Several models of the NI were proposed, but to gain new insight into the nature of VN neurons, dynamic processing in the horizontal VOR was modeled using recurrent neural networks.³⁻⁵ The first neural network approach of the NI was proposed by Cannon and Robinson.^{3,4} Their hard-wired model can integrate a push-pull input signal without integrating the background rates and has the appealing property that localized lesions produced a decrease in the time constant of the entire network (a clinically observed phenomenon); the synaptic weights were explicitly specified. Later, Anastasio and Robinson proposed the first learning model for the NI⁵; nevertheless, this model lacks constraints on the synaptic weights (after training, it does not present the push-pull configuration that the real integrator has).

The purpose of the present investigation is not to develop a new model of the NI but rather to improve the biologically plausible features of the existing models. We based our work on the learning model of Anastasio-Robinson.⁵ This model simulates

^aAddress for correspondence: Jean-Philippe Draye, Faculté Polytechnique de Mons, Laboratoire "Processus, Informatique, Parallélisme," Rue de Houdain, 9, B-7000 Mons, Belgium. E-mail: JPD@PIP.FPMS.AC.BE

^bResearch assistant of the Belgian National Fund for Scientific Research.

the horizontal NI in the particular case of the dynamic processing of the horizontal VOR. It presents two afferent inputs from the horizontal canals, a fully connected hidden layer of 16 inhibitory units and two output units that represent the motoneurons of the median and lateral rectus muscles of the left eye (see FIG. 1). The improvements that we bring to the model include strong constraints on the synaptic weights (in order to respect the biological Dale's principle) and the introduction of an artificial distance in our network (to define a notion of proximity). Because of these modifications (the constraints on the weights introduces singular points during training), a new learning process was needed, and we chose a learning algorithm based on a general supervisor. This algorithm trains the network to leakily integrate (in the sense of the Newtonian calculus) the push-pull eye-velocity signals provided by the semicircular canals without integrating the background rate of 100 spikes/sec.

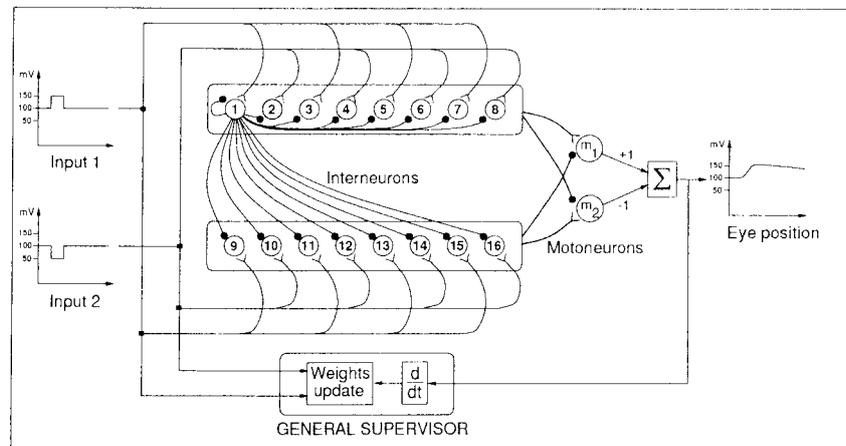


FIGURE 1. The modified artificial neural network for the modeling of the neural integrator. The 16 interneurons of the hidden layer are divided into two groups of eight and are fully connected with inhibitory connections (only the connections out of interneuron 1 are depicted). Each interneuron is connected to both motoneurons with a connection whose sign is given on the figure. The eye position is simply given by the difference between the motoneuron outputs. The general supervisor computes all the weight updates.

Once our network has been trained, we will be interested in the behavior of the hidden units of our model: Could the distribution of the artificial synaptic weights of these units be related to a particular organization depending on specific requirements? This task is not as simple as it might be; it is even widely assumed that the hidden units are of little use in illuminating the basic problem of how the brain processes signals. After training, we observed that the structure of the lateral connection weights of the hidden layer exhibits several clusters. The weights are obviously structured in conglomerates or groups of interneurons where the weights seem quite important. Other zones are quite flat and characterized by very low-value weights (see FIG. 2). The important thing is that the high-value weights are grouped in what we call clusters. A cluster can be defined as a particular group of adjoining interneurons that have strong and privileged connections with another neighborhood

of interneurons (the notion of neighborhood has a clear meaning with respect to the introduction of a distance in the network).

From the biological point of view, microelectrode recordings have shown that neighboring neurons often disclose similar patterns of electrical activity corroborating the existence of a functional clustering of the NI. For example, the goldfish NI presents very clear nonoverlapping compartments with proper functions (eye position or eye velocity integrators).⁶ Moreover, neurochemically defined clusters were

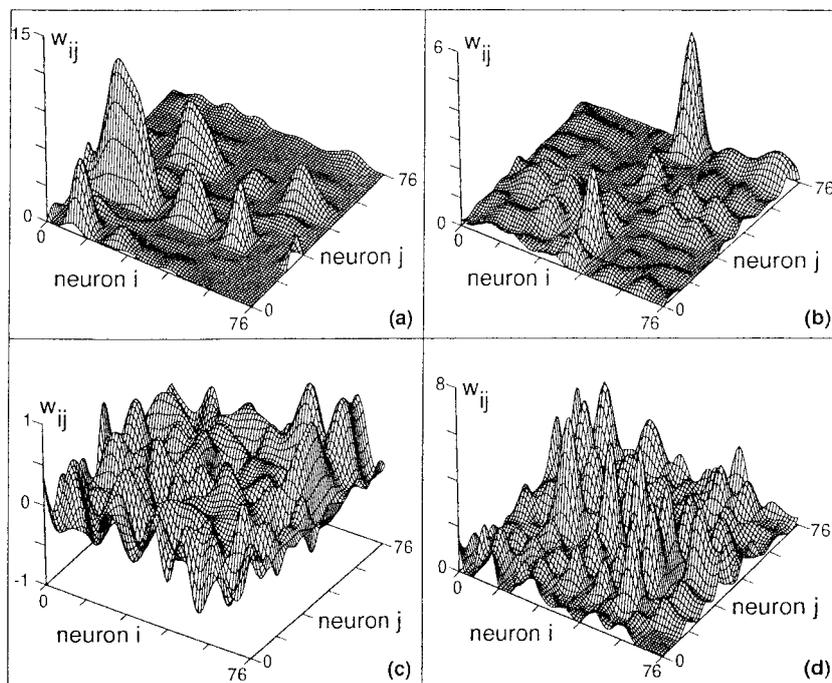


FIGURE 2. Surface plots of the weights distribution. The 16×16 weights surfaces were treated with cubic splines to get a better visualization. Values of the weights are plotted versus indexes i and j . Even if the lateral layer has inhibitory connections (except for c), the weights are plotted as positive values. (a and b) Two clustered structures of the weights distribution of our network. (c) The weight distribution of the Arnold-Robinson network trained with the general supervisor without any constraints on the sign of the weights (the sign can be positive or negative). (d) The weight distribution of the network where each interneuron has its own muscle.

found in the NPH of the cat. Like other iterated patterns of the brain circuitry (the columns and blobs in visual cortex, the barrels in the somatosensory cortex and the patches in the striatum, the clusters reflect one of the organizing principles of neural assemblies.⁷ This type of organization could therefore be considered to be expressions of different kinds of biological constraints. These can be genetic, biochemical, developmental (synaptic growth), and/or due to the information processing in the network. Artificial neural networks give us an opportunity to analyze this latter hypothesis by means of a computational model using supervised learning to mimic

the behavior of the NI. The spontaneous emergence of clusters in artificial neural networks, performing a well-defined physico-mathematical task (a temporal integration) is due to computational constraints, with a restricted space for solutions. Thus, information-processing constraints are a plausible factor in inducing the emergence of iterated patterns in biological neural networks.

REFERENCES

1. ROBINSON, D. A. 1992. *Behav. Brain Sci.* **15**: 644–655.
2. CHERON, G., E. GODAUX, J. M. LAUNE & B. VANDERKELEN. 1986. *J. Physiol. (Lond.)* **372**: 75–94.
3. CANNON, S. C. & D. A. ROBINSON. 1983. *Biol. Cybern.* **49**: 127–136.
4. CANNON, S. C. & D. A. ROBINSON. 1985. *Biol. Cybern.* **53**: 93–108.
5. ARNOLD, D. B. & D. A. ROBINSON. 1991. *Biol. Cybern.* **64**: 447–454.
6. PASTOR, A. M., R. R. DE LA CRUZ & R. BAKER. 1994. *Proc. Natl. Acad. Sci. USA* **91**: 807–811.
7. PURVES, D., D. R. RIDDLE & A.-S. LAMANTIA. 1992. *Trends Neurosci.* **15**: 362–368.