

A Computational Neurogenetic Model of a Spiking Neuron

Nikola Kasabov, Lubica Benuskova and Simej Gomes Wysoski
 Knowledge Engineering & Discovery Research Institute
 Auckland University of Technology
 Auckland, New Zealand
 E-mail: {nkasabov, lbenusko, swysoski} @aut.ac.nz

Abstract—The paper presents a novel, biologically plausible spiking neuronal model that includes a dynamic gene network. Interactions of genes in neurons affect the dynamics of the neurons and the whole network through neuronal parameters that change as a function of gene expression. The proposed model is used to build a spiking neural network (SNN) illustrated on a real EEG data case study problem. The paper also presents a novel computational approach to brain neural network modeling that integrates dynamic gene networks with a neural network model. Interaction of genes in neurons affects the dynamics of the whole neural network through neuronal parameters, which are no longer constant, but change as a function of gene expression. Through optimization of the gene interaction network, initial gene/protein expression values and ANN parameters, particular target states of the neural network operation can be achieved, and statistics about gene interaction matrix can be extracted. It is illustrated by means of a simple neurogenetic model of a spiking neural network (SNN). The behavior of SNN is evaluated by means of the local field potential, thus making it possible to attempt modeling the role of genes in different brain states, where EEG data is available to test the model. We use standard signal processing techniques like FFT to evaluate the SNN output to compare it with real human EEG data.

I. INTRODUCTION

With the advancement of molecular research technologies more and more data and information is being made available about the genetic basis of neuronal functions and diseases [1, 2]. This information can be utilized to create models of brain functions and diseases that include models of gene interaction. This area integrates knowledge from computer and information science, brain science, molecular genetics and we call it *computational neurogenetic modeling* (CNGM) [3, 4].

CNGM is a new area of research that has many open questions, some of them listed below:

- (1) Which real neuronal parameters are to be included in an ANN model and how to link them to activities of genes/proteins?
- (2) Which genes/proteins are to be included in the model and how to represent the gene interaction over time within each neuron?
- (3) How to integrate in time the activity of genes and neurons in an ANN model as it is known that neurons spike in millisecond intervals and the process of gene transcription and translation into proteins takes minutes or even hours?
- (4) How to integrate internal and external variables in a CNGM (e.g., genes and neuronal parameters with external signals acting on the brain) and how to treat various perturbations?
- (5) How to create and validate a CNG model in the presence of scarce data?
- (6) How to measure brain activity and the CNGM activity in order to validate the model?
- (7) What kind of useful information can be derived from CNGM?

Our approach is illustrated by means of a simple neurogenetic model of SNN. The behavior of SNN is evaluated by means of the local field potential (LFP), thus making it possible to attempt modeling the role of genes in different brain states, where EEG data is available to test the model. Support for this approach comes from recent studies that have shown that brain electrical oscillations are genetically determined and differ between individuals and families [5].

II. A BIOLOGICALLY PLAUSIBLE COMPUTATIONAL NEUROGENETIC MODEL OF A SPIKING NEURON

Here we propose a biologically plausible model of a spiking neuron that includes genes and proteins

that interact between each other and affect the spiking characteristics of the neuron.

In general, we consider two sets of genes – a set G_{gen} that relates to general cell functions and a set G_{spec} that defines specific neuronal information-processing functions (e.g. receptors, ion channels, etc.). The two sets form together a set $\mathbf{G}=\{G_1, G_2, \dots, G_n\}$. We assume that the expression level of each gene $g_j(t+\Delta t_j)$ is a nonlinear function of expression levels of all the genes in $\mathbf{G}(t)$, inspired by discrete models from [6, 7]:

$$g_j(t+\Delta t_j) = \sigma\left(\sum_{k=1}^n w_{jk} g_k(t)\right) \quad (1)$$

We work with normalized gene expression values in the interval $g_j(t) \in (0, 1)$. Each gene has its own delay Δt_j that represents the delay along the axis: $\{\text{genes}\}_{k=1}^n \rightarrow \text{proteins} \rightarrow \text{transcription factors (TFs)} \rightarrow \text{gene } j$ [8]. The coefficients $w_{ij} \in (-5, 5)$ are elements of the square matrix \mathbf{W} of gene interaction weights. These borders have been chosen experimentally to lead to various types of nonlinear dynamics, i.e. constant, periodic, quasi-periodic and chaotic. Initial values of gene expressions are small random values, i.e. $g_j(0) \in (0, 0.1)$.

In the current model we assume a simple scenario where: (1) one protein is coded by one gene; (2) relationship between the protein level and the gene expression level is linear; (3) protein levels lie between the minimal and maximal values (not necessarily being equal to 0 and 1, respectively). Thus, the protein level $p_j(t+\Delta t)$ is expressed by

$$p_j(t+\Delta t) = (p_j^{\max} - p_j^{\min})\sigma\left(\sum_{k=1}^n w_{jk} g_k(t)\right) + p_j^{\min} \quad (2)$$

The delay Δt corresponds to the delay caused by the gene j transcription and its initiation, mRNA translation into protein and posttranslational protein modifications [8]. In our model, some protein levels will be directly related to neuronal parameters P_j such that

$$P_j(t) = P_j(0)p_j(t) \quad (3)$$

where $P_j(0)$ is the initial value of the neuronal parameter at time $t = 0$. In such a way the gene/protein dynamics is linked to the dynamics of ANN. Some neuronal parameters and their correspondence to particular proteins are summarized in Table I. The choice of neural parameters depends on a task, which we want to simulate. In our case, let it be for instance a normal resting EEG signal.

Moreover, besides the genes coding for the proteins listed in Table I, we include in our gene network nine more genes that are not directly linked

to neuronal information-processing parameters. These genes are: c-jun, mGluR3, Jerky, BDNF, FGF-2, IGF-1, GALR1, NOS, S100beta. These other proteins and genes are known to have a regulatory effect upon genes and proteins that are directly linked to neural information-processing parameters. An example of a particular gene regulatory network (GRN) is given in figure 1.

TABLE I
NEURON'S PARAMETERS AND THEIR RELATED PROTEINS

Neuron's parameter	PROTEIN*
Amplitude and time constants of:	
Fast excitation	AMPA
Slow excitation	NMDAR
Fast inhibition	GABRA
Slow inhibition	GABRB
Firing threshold	SCN, KCN, CLC

*Abbreviations: AMPAR = (amino-methylisoxazole- propionic acid) AMPA receptor, NMDAR = (N-methyl-D-aspartate acid) NMDA receptor, GABRA = (gamma-aminobutyric acid) GABA receptor A, GABRB = GABA receptor B, SCN = Sodium voltage-gated channel, KCN = kalium (potassium) voltage-gated channel, CLC = chloride channel.

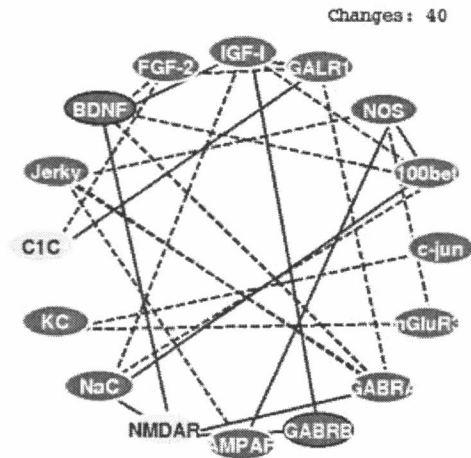


Fig. 1. Illustration of GRN. Solid (dashed) lines denote positive (negative) interactions between genes, respectively. For clarity of illustration, only the stronger of two-way connections are shown. Intensity of grey reflects the level of gene expression after 40 updates.

The CNGM model from formulas (1)–(3) is a general one and can be integrated with an SNN or any other neural network model. Unfortunately the model requires many parameters to be either known in advance or optimized during a model simulation. In the presented experiments we have made several

simplifying assumptions:

1. Each neuron has the same gene regulatory network (GRN), i.e. the same genes and the same interaction gene matrix \mathbf{W} .
2. Each GRN starts from the same initial values of gene expressions.
3. Feedbacks from neuronal activity or any other external factors to gene expression levels or protein levels are not explicitly considered.
4. Delays Δt are the same for all proteins. This is based on the fact that protein expression data are being gathered for all proteins of interest at the same time instant.

III. AN EXAMPLE OF USING THE CNG MODEL OF A NEURON FOR BUILDING SPIKING NEURAL NETWORK MODELS

The SNN used in our CNGM has been described elsewhere [3, 4]. Each postsynaptic potential has a fast and slow component, otherwise the model is based upon a classical Spike Response Model (SRM) [9].

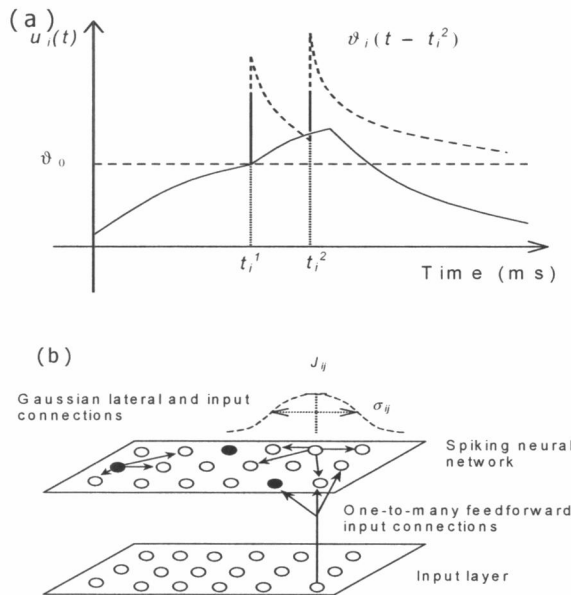


Fig. 2. (a) Spiking neuron model. When the membrane potential $u_i(t)$ of the i^{th} spiking neuron reaches the firing threshold $\vartheta_i(t)$ at time t_i^1 , the neuron fires an output spike. $\vartheta_i(t)$ rises after each output spike and decays back to the resting value ϑ_0 . (b) The SNN architecture. About 10–20% of $n = 120$ neurons are inhibitory neurons that are randomly positioned on the grid (filled circles). External input is a randomly Poisson with average frequency between 10–20 Hz.

Figure 2 illustrates the set up of the SNN. We keep the record of spiking activities of all neurons as well as record of the local field potential (LFP), the sum of many of which is in the brain proportional to EEG [10]. We define LFP as an average of all instantaneous membrane potentials, i.e. $\Phi(t) = (1/N) \sum u_i(t)$. For its analysis we use the fast Fourier method [11].

Frequency spectrum of LFP is divided into five frequency sub-bands, i.e. delta (0.1–3.5 Hz), theta (3.5–7.5 Hz), alpha (7.5–12.5 Hz), gamma 1 (12.5–18.0 Hz), gamma 2 (18.0–30.0 Hz), beta (30.0–50.0 Hz). For these frequency sub-bands we calculate the relative intensity ratios (RIR) over the relevant period of measurement. An example of a human resting interictal EEG is in figure 3 (obtained with permission from [12]). Temporal changes of RIRs for frequency sub-bands over the measurement time period are shown in figure 4.

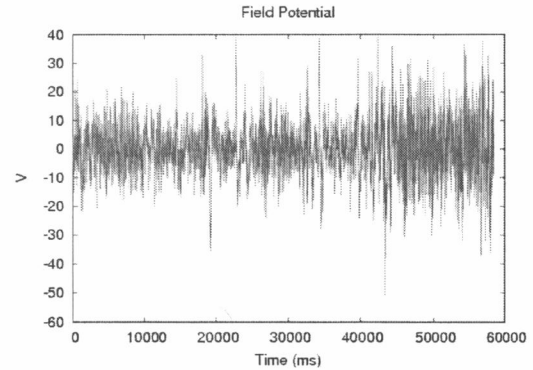


Fig. 3. Human interictal resting EEG signal.

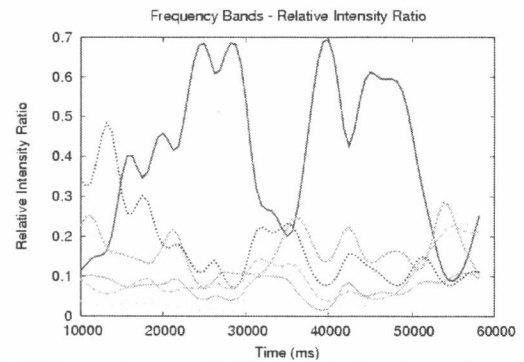


Fig. 4. Temporal course of RIRs for clinically relevant frequency sub-bands of the EEG signal from figure 3. Most of the time, the dominant sub-band is delta (0.5–3.5 Hz).

Thus, any EEG or LFP signal can be characterized by vector of five numbers expressing the average RIRs of particular frequency sub-bands over some time interval. For the EEG signal from figure 2, the average vector RIR (delta, theta, alpha, gamma 1,

$\gamma_2, \beta_2) = (0.56, 21, 0.16, 0.04, 0.007, 0.002)$. During our simulations we will try to find such gene networks that will lead to the SNN LFP with average RIR vector as closest as possible to the average RIR vector of the resting human EEG in terms of Euclidean distance.

Just for computational reasons, we will employ the delays Δt in equation (2) being equal to just 1s of the SNN time instead of minutes or tens of minutes of the real time [8]. Justification for this time compression in our simulations is illustrated in figures 5 and 6, where we show that the spectral characteristics of the SNN LFP do not change in time when neural parameters are constant (i.e. during the time interval Δt). Thus, with respect to spectral characteristics of the LFP signal, 1 s of SNN simulation with constant parameters can represent an arbitrarily long time interval of real time. This match allows us to update protein levels and the corresponding parameters values every 1s of SNN simulation instead of let us say every fifteen or even more minutes.

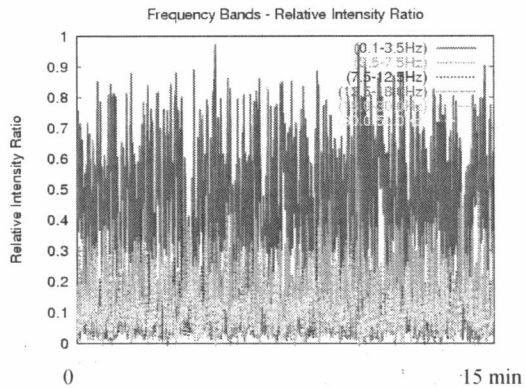


Fig. 5. 15 min of SNN spontaneous activity with constant parameters. The dominant frequency sub-band is delta.

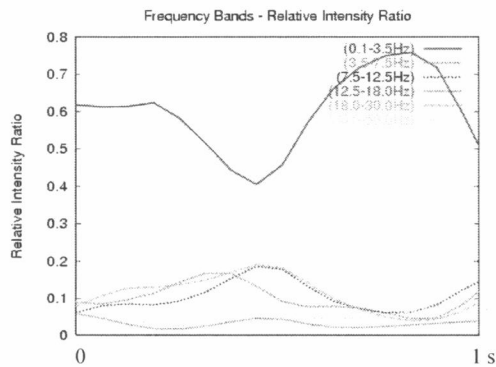


Fig. 6. The first second of the simulation from figure 5. The dominant frequency sub-band is still delta.

IV. OPTIMIZATION OF CNGM AND KNOWLEDGE DISCOVERY METHODS

We want to achieve a desired SNN output through optimization of the model 294 parameters. We are optimizing the interaction matrix \mathbf{W} between 16 genes, initial values of neural parameters, architectural parameters of SNN (except the total number of neurons, spike delays and probability of establishing a synaptic connection) and input frequency to the SNN. All model parameters and their value ranges to choose from during optimization are listed in Table II. These parameter ranges reflect considerable variations of values in real neurons and are taken from neurobiological data [13, 14]. We evaluate the LFP of the SNN by means of FFT in order to compare the SNN output with the EEG signal analyzed in the same way. It has been shown that LFPs in principle have the same spectral characteristics as EEG [15].

TABLE II
MODEL PARAMETERS AND THEIR VALUE RANGES

MODEL PARAMETERS	VALUE RANGE
Fast excitation: Amplitude rise / decay time constants	0.5 – 3.0
Slow excitation: Amplitude rise / decay time constants	1 – 5 / 5 – 10
Fast inhibition: Amplitude rise / decay time constants	0.5 – 4.0
Slow inhibition: Amplitude rise / decay time constants	10 – 20 / 30 – 50
Resting firing threshold, decay time constant / rise	4 – 8
Proportion of inhibitory neurons	5 – 10 / 20 – 30
Probability of external input firing	5 – 10
Peak/sigma of external input weight	20 – 80 / 50 – 150
Peak/sigma of lateral exc weights	19 – 25
Peak/sigma of lateral inh weights	5 – 50 / 2 – 5
Unit delay in e/i spike propagation	0.15 – 0.2
Probability of connection	0.011 – 0.019
Number of neurons	5 – 10 / 0.1 – 2
Gene interaction weight	5 – 14 / 2 – 8
Gene initial expression	10 – 60 / 4 – 10
Gene normalized expression	1 ms / 2 ms
Gene interaction weight	0.5
Gene initial expression	120
Gene normalized expression	(– 5, +5)
Gene interaction weight	(0.0, 0.1)
Gene normalized expression	(0.0, 1.0)

In order to find an optimal GRN within the SNN model so that the frequency characteristics of the LFP of the SNN model are similar to the brain EEG characteristics, we use the following simple optimization procedure:

1. Generate a population of N CNGMs, each with randomly generated values of coefficients for the GRN matrix \mathbf{W} , initial gene expression values $\mathbf{g}(0)$, initial values of SNN parameters $\mathbf{P}(0)$, and different connectivity (all these

parameters can have some predefined borders, see Table II);

2. Run each SNN over a period of time $T = 1$ min and record the LFP, $\Phi(t)$;
3. Calculate the spectral characteristics of the LFP using FFT;
4. Compare the spectral characteristics of SNN LFP to the characteristics of the target EEG signal. Evaluate the closeness of the LFP signal for each SNN to the target EEG signal characteristics. Find SNN models that match the EEG spectral characteristics better than other solutions. Let us say, the Euclidean distance between the RIR vectors be smaller than 0.1;
5. Analyze the GRN and the SNN parameters for significant gene patterns that cause the SNN model behavior.

Alternatively, we can use a genetic algorithm to find an optimal solution or solutions, but for illustration of our knowledge discovery method, the above simple optimization procedure will be sufficient.

$N = 400$ random gene interaction matrices \mathbf{W} s show almost a uniform distribution of interaction strengths between genes as can be seen in figure 7.

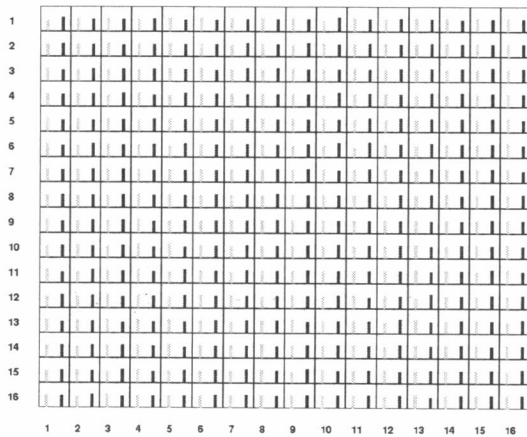


Fig. 7. Black and grey histograms show the percentage of positive and negative gene interactions, respectively, in 400 randomly generated interaction matrices \mathbf{W} .

Among these 400 random solutions, 15 \mathbf{W} s matrices led to an SNN LFP with spectral characteristics very close to the target EEG signal in terms of Euclidean distance between the RIR vectors belonging to the SNN LFP and human EEG being smaller than 0.1. Visual inspection of these 15 solutions also confirmed that the LFP signal and its spectral characteristics are similar to the target EEG signal. Distribution of gene interactions in these 15

\mathbf{W} matrices that led to the desired SNN output is no longer uniform (see figure 8).

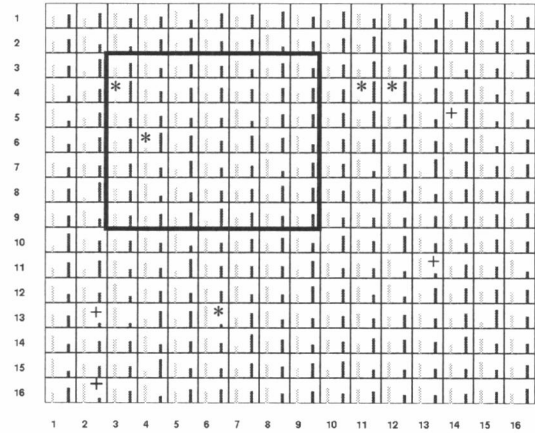


Fig. 8. Black and grey histograms show the percentage of positive and negative gene interactions, respectively, in 15 winner interaction matrices \mathbf{W} after experimental running of CNGM. * means $\alpha = 0.005$, and + means $\alpha = 0.025$ in the X^2 test.

To discover the knowledge, e.g. to find out what these solutions have in common, we have calculated how many times the interactions between genes become positive and how many times they become negative. We can use a basic frequency statistical analysis, for instance the X^2 -statistic, to make predictions about interactions between real genes in neurons. Let us consider the genes, which are directly related to information-processing parameters of neurons. In figure 8, they are outlined by a black rectangle. We have found that a statistically significant difference at greater than the 0.005 level of probability holds for interactions $w_{43} > 0$ and $w_{64} > 0$ (denoted by asterisks in black rectangle in figure 8 matrix). What that means is that the gene No. 3 (coding for the GABA_A receptor) when elevated in expression is accompanied by elevated expression of gene No. 4 (coding for GABA_B receptor), and that the gene No. 4 (coding for the GABA_B receptor) when elevated in expression is accompanied by elevated expression of gene No. 6 (coding for NMDA receptor). These interactions are intuitively reasonable since one can assume that when the inhibition is somehow enhanced, the system tries to keep balance by enhancing excitation as well, and vice versa. Table III summarizes all statistically significant deviations from uniformity among all the gene interactions in our sample of 15 \mathbf{W} s leading to LFP with frequency spectrum similar to a normal resting EEG. The meaning of these other interactions with genes that do not have any

assigned functions is not relevant for now, except as an illustration of expansion of our approach to genes that can have an indirect yet important effect.

TABLE III
STATISTICALLY SIGNIFICANT GENE INTERACTIONS

Significance α	Gene interaction $w(j, k)$
0.005 (*)	$w(4,3) > 0$, $w(6, 4) > 0$, $w(4, 11) > 0$, $w(4, 12) > 0$, $w(13, 6) < 0$
0.025 (+)	$w(5, 14) > 0$, $w(11, 13) < 0$, $w(13, 2) < 0$, $w(16, 2) < 0$

V. DISCUSSION

In real neural networks neuronal parameters that define the functioning of a neural network depend on genes and proteins in a complex way. Gene and protein expression values change due to internal dynamics of the gene/protein regulatory (interaction) network, initial conditions of the genes and external conditions. All this may affect gradually or quickly the functioning of the neural network as a whole. In our computer experiments, we have observed for example that different initial gene conditions can lead to the same outcome in terms of neuronal activity. Moreover, different types of gene interaction dynamics, i.e. be it constant, periodic, quasi-periodic or even chaotic, can lead to a similar LFP of an associated SNN model, provided some statistical distribution of gene interactions is maintained. Different statistics can be linked to different values in Table II. Thus, in the diseased brain, either altered initial conditions, mutated genes and/or altered interactions within GRN can lead to abnormalities in network activity. Realistic models of gene networks within neural networks should account for these processes.

In order to investigate these phenomena, we have set up a novel model of a CNGM that is simple and biologically plausible. Associated SNN uses principles from the simple spiking neuron models [9]. More detailed models of SNN can include for instance detailed ion receptor and channel kinetics and also multiple neuron compartments. It is possible to include also more than one brain areas in our CNG model. Particular neural network model should be chosen based on a problem, which we want to account for. On the example of a resting human EEG signal we have demonstrated our CNGM approach with the introduction of generic CNGM equations, optimization and knowledge discovery techniques.

ACKNOWLEDGMENT

The work has been supported by the NERF grant X0201 funded by FRST and by KEDRI funds (www.kedri.info).

REFERENCES

- [1] G. Marcus, *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexity of the Human Mind*, Basic Books, New York, 2004.
- [2] "The Nervous System," in *Genes and Disease*, National Centre for Biotechnology Information (NCBI), <http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowSection&rid=gnd.chapter.75>, 2003.
- [3] N. Kasabov and L. Benuskova, "Computational neurogenetics," *Journal of Computational and Theoretical Nanoscience*, 1, 1, pp. 47-61, 2004.
- [4] N. Kasabov, L. Benuskova and S. G. Wysocki, "Computational neurogenetic modelling: gene networks within neural networks," *Proc. IJCNN2004: IEEE Press*, Budapest, 2004, vol. 2, pp. 1203-1208.
- [5] C. E. M. vanBeijsterveldt and G. C. M. vanBaal, "Twin and family studies of the human electroencephalogram: a review and meta-analysis," *Biological Psychology*, 61, pp. 111-138, 2002.
- [6] D. C. Weaver, C. T. Workman and G. D. Stormo, "Modeling regulatory networks with weight matrices," *Proc. Pacific Symp. Biocomputing*, 1999, vol. 4, pp. 112-123.
- [7] L. F. A. Wessels, E. P. vanSomeren and M. J. T. Reinders, "A comparison of genetic network models," *Proc. Pacific Symp. Biocomputing*, 2001, vol. 6, pp. 508-519.
- [8] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell, "Nucleic acids, the genetic code, and the synthesis of macromolecules," in *Molecular Cell Biology*, W. H. Freeman & Co., New York, 2000, pp. 100-137.
- [9] W. Gerstner and W. M. Kistler, *Spiking Neuron Models*, Cambridge Univ. Press, Cambridge, MA, 2002.
- [10] J. W. Freeman, *Mass action in the nervous system*, Academic Press, New York, 1975.
- [11] R. Q. Quiroga, *Quantitative Analysis of EEG Signals: Time-Frequency Methods and Chaos Theory.*, Institute of Physiology and Institute of Signal Processing, Medical University Lübeck: Lübeck, 1998.
- [12] R. Q. Quiroga, *Dataset # 3: Tonic-clonic (Grand Mal) seizures*. <http://www.vis.caltech.edu/~rodri/data.htm>, 1998.
- [13] A. Destexhe, "Spike-and-wave oscillations based on the properties of GABA_B receptors," *J. Neurosci.*, 18, pp. 9099-9111, 1998.
- [14] F. Wendling, F. Bartolomei, J. J. Bellanger and P. Chauvel, "Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition," *Eur. J. Neurosci.*, 15, pp. 1499-1508, 2002.
- [15] A. Destexhe, D. Contreras and M. Steriade, "Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states," *J. Neuroscience*, 19, 11, pp. 4595-4608, 1999.