Design of experiments and Sobol’ sensitivity analysis of a hippocampus computational model
Amélie Aussel, Laure Buhry, Radu Ranta

To cite this version:
Amélie Aussel, Laure Buhry, Radu Ranta. Design of experiments and Sobol’ sensitivity analysis of a hippocampus computational model. 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2021, Oct 2021, Guadalajara/Virtual, Mexico. hal-03464025

HAL Id: hal-03464025
https://hal.science/hal-03464025
Submitted on 2 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Design of Experiments and Sobol’ sensitivity analysis of a hippocampus computational model

Amélie Aussel1,2, Laure Buhry1 and Radu Ranta2

Abstract — The hippocampus is a brain area involved in many memory processes. This structure can also be affected in neurological diseases such as mesial temporal lobe epilepsy. A better understanding of its electrophysiological activity could benefit both the neuroscientific and clinical communities. We proposed, in a previous paper, a detailed bio-realistic conductance-based mathematical model of more than thirty thousand neurons to reproduce the main oscillatory features of the healthy hippocampus during slow-wave sleep and wakefulness, from slow to very fast frequencies. One big challenge of this model is its parametrization. The aim of the present work is to combine neuroscientific expertise and systematic yet efficient exploration of the highly dimensional parameter space using well defined identification methods, namely the design of experiments and the Sobol’s sensitivity analysis.

Index Terms — computational neuroscience, hippocampus, parameter optimization, design of experiment (DOE), Sobol’ sensitivity analysis

I. INTRODUCTION

The hippocampus can exhibit oscillatory rhythms in a wide range of frequencies, depending on the cognitive task to perform or the vigilance state. For example, theta (5-10Hz) and gamma (30-100Hz) oscillations are typically coupled in spatial navigation and working memory tasks during active wakefulness [1], whereas delta oscillations (1-4Hz) and Sharp-Wave-Ripple (SWR) complexes (130-200Hz) play an important role in memory consolidation [2] and appear during slow-wave sleep and quiet wakefulness. In [3], we proposed a biologically detailed model of the hippocampus capable of reproducing these rhythms, and discussed possible mechanisms of transition between them. However, we were not able to provide a full parameter analysis at that time.

To analyze the influence of a set of parameters on a model, a naive approach would consist in simply running a large number of simulations, sampling the parameter space regularly. But when the parameter space is too large or too high-dimensional, which is the case in many computational neuroscience models (and in our hippocampal model as well), the time and computational resources needed to perform all these becomes unreasonably large.

In the present work, after briefly introducing our healthy hippocampus model, we propose a method based on Design of Experiments (DOE) techniques and Sobol’ sensitivity analysis [4]–[6]. In between optimization and brute-force, these methods allow to explore systematically, the role of the model’s main parameters (in red in figure 1A). We then use this parameter analysis to provide a functionality range that could reproduce either sleep or wakefulness oscillatory rhythms.

II. MATERIALS AND METHODS

A. Computational modeling of hippocampal oscillations

The model studied in this paper is directly inspired from our previous work [3]. Its complete structure is given Figure 1. We recall that the model covers multiple scales, from channel level mechanisms (Hodgkin-Huxley formalism, for classical K and Na channels, but also for Calcium-Activated Nonselective cation channel - CAN) to realistic anatomy and modulated connectivity (intra and inter structures connection probabilities and gains on synaptic conductances), including more than thirty thousand neurons to represent the four main regions of the hippocampal formation: the Dentate Gyrus (DG), CA3, CA1, and the entorhinal cortex (EC). The number of neurons of each type in each region were chosen in the same proportion as reported in the literature. However, the values of other parameters driving the model, although they have biological significance, cannot be assessed directly from biological experiments. They are nevertheless critical to the task at hand, namely explaining mechanisms involved in oscillatory pattern sleep-wake variations (SWR, theta-gamma) and transitions. It is therefore very important to have a good knowledge of the parameter space. The next subsections describe these ‘free’ parameters and the methods used for determining them.

B. Free parameters of the model

First, because our model includes less neurons than a real human hippocampus, each of the modeled neurons receive less synaptic contacts from its surroundings. Therefore, in order for these synaptic inputs to properly drive neural activity, the amplitude of each synaptic current should be increased. This is done by increasing the maximum conductances of both excitatory and inhibitory synapses: $g_{\text{max, e}}$ and $g_{\text{max, i}}$.

The maximum connection probabilities between different subregions are also critical. In order to avoid a too important increase in the dimension of the parameter space, we use a unique probability $p_{p\text{tri}}$ for all synaptic connections along the tri-synaptic loop (EC → DG → CA3 → CA1 → EC), and a unique connection probability $p_{\text{mono}}$ for synaptic connections made from the EC directly to CA3 and CA1.

Next, sleep-wake variations are known to modulate neurotransmitter concentrations, in particular the acetylcholine...
(ACh, see [7]). It has been shown that some synaptic transmissions are enhanced or decreased by a high ACh concentration. To represent these changes, we define three modulatory gains of the synaptic conductances: $G_e$ (increasing excitatory synaptic conductances in the DG), $g_e$ (decreasing excitatory synaptic conductances in the EC and CA3), and $G_i$ (increasing inhibitory synaptic conductances in the DG and CA1). ACh also plays a key role at the single cellular level in the activation of CAN currents in hippocampal pyramidal cells, therefore the conductance of the CAN channel $g_{CAN}$ will also be a studied parameter. See [3] for more details on the role of ACh in the model.

Finally, as neurons in the model do not fire in an organized manner on their own, an external stimulation mimicking inputs from other brain areas in sleep or wakefulness conditions has to be applied. Unlike what was presented in our previous work [3], the behavior of the model is here investigated under controlled stereotypical inputs. These inputs consist of a square wave current $I_{stim}$, starting at $t_0 = 250\,\text{ms}$, injected into all excitatory and inhibitory neurons of the EC following eq. 1. Its maximum amplitude $A_1$ and frequency $f_1$ are parameters to be studied:

$$I_{stim}(t) = \begin{cases} A_1 & \text{if } \{ \begin{array}{c} t > t_0 \\ \sin(2\pi f_1(t - t_0)) \geq 0 \end{array} \} \\ 0 & \text{otherwise.} \end{cases}$$

(1)

C. Output of the hippocampal model: LFP simulation

The output of the model is the Local Field Potential (LFP) generated by the network, as it would be measured by a macroscopic electrode placed inside the hippocampus [3].

Following the approach from [8], we considered the excitatory neurons as dipoles, with one extremity being the soma and the other being the location of the active synapse: either at its basal or apical dendrites depending on the synapse type (see [9]). The projection of a neuron’s soma on the stratum moleculare was computed as an estimation of the position of its apical dendrites, and its projection on the stratum oriens as an estimation of the position of its basal dendrites (Fig. 1B).

To mimic the electrode measure, we then averaged the LFP generated by each excitatory neuron on two sets of 144 points evenly distributed on a cylinder, and computed the difference between the two resulting signals, as in a bipolar sEEG montage (see figure 1B). Similarly to the intracerebral EEG recordings typically available to us, the simulated LFP was also bandpass filtered between 0.15Hz and 480Hz and downsampled to 1024Hz.

Since the hippocampus is able to produce both fast oscillations (gamma to fast ripple frequency range) and slow oscillations (delta to theta frequency range), we extract from the LFP output of the model two frequency characteristics of interest for biologists and clinicians:

- the peak frequency of the LFP spectrum in the 30-500Hz range $f_{fast}$ (from gamma to ripple and higher)
- the peak frequency of LFP spectrum in the 1-30Hz range $f_{slow}$ (including delta and theta frequency bands)

All the peaks reported in the Results section were highly significant in their respective frequency bands, as was assessed by computing their Z-scores (i.e. the number of standard deviations above the mean power of the considered frequency band): for the fast oscillations all Z-scores were above 7.5, and for the slow oscillations all Z-scores were above 4.5.

D. Design of experiments (DOE):

In the first part of our parameter analysis, we used a DOE technique (see [4]) to choose an optimal subset of simulations to run, and then fit a linear or polynomial function
on the observed behaviors. Here, we used a Box-Behnken design [5], which is made for quickly obtaining second-order modeling of the output by placing the experimental points (normalized between -1 and 1) on the middle of some edges of the parameter space’s hypercube, as well as points in the center of the cube (all factors set to 0).

We then used a least squares method to fit the data resulting from our exploration with a second order polynomial function of the parameters. Finally, we evaluated the significance of each of the factors used in this fitting with a Student’s t-test. A similar polynomial fitting was also applied to the simulations resulting from our Sobol’ global sensitivity analysis.

E. Sobol’ global sensitivity analysis:

Sobol’ methods (see [6]) are variance analysis methods, meaning that they aim at estimating the contributions of each parameter to the variance of the output. The analysis yields for each parameter, a so-called sensitivity index. Compared to DOE techniques, Sobol’ methods give a more complete vision of the influence of each parameter on the output, though they require to run more simulations.

Let us define $D$ the total variance of the output $y$, $D_i$ the variance of $y$ obtained by varying the parameter $x_i$ only, and more generally $D_z$ the variance of $y$ obtained by varying only the parameters in a subset $z$. The first order sensitivity index relative to the $i$-th parameter is then defined as $S_i = \frac{D_i}{D}$, and the sensitivity index relative to the subset $z$ of parameters is defined as $S_z = \frac{D_z}{D}$.

Provided that the input parameters are independent and the output is square-integrable, the variance $D_i$ of the output relative to the $i$-th parameter varied individually can be written as (see [10]):

$$D_i = \int_{[0,1]^N} \int_{[0,1]^{N-1}} y(x) y(x_i, x'_{N}) dx dx' - y_0^2,$$  \tag{2}

where all the parameters are normalized in the $[0,1]$ interval, $x = (x_1, ..., x_N)$, $x = (x_i, x'_{N})$ is a point of the parameter space, $x_i, x'_{N}$ denotes a vector with same component $x_i$ as in $x$ (the other components being $x'_{N}$), and $y_0$ is the mathematical expectation of $y$.

The variance $D_z$ relative to any set of parameters $z$ can be computed with a similar process:

$$D_z = \int_{[0,1]^N} \int_{[0,1]^{N-|z|}} y(x) y(z, x'_{N}) dx dz' - y_0^2.$$ \tag{3}

The integrals in this expression were approximated using a quasi Monte-Carlo method called Sobol’ sequence to define the sequence of simulations to run and obtain a fast convergence.

### III. Results

All the simulations were performed using the Brian2 libraries for Python ( [11]), on the Grid’5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (see https://www.grid5000.fr).

A. Parameter study with a Design of Experiments method

So as to get a first intuition on the behavior of the model, we conducted 170 simulations following a Box-Behnken design of experiments (DOE) matrix with our ten parameters normalized from the following ranges to the $[-1, 1]$ interval: $A_1 : [0.5 nA, 1.5 nA]$, $f_1 : [0Hz, 10Hz]$, $p_{tri} : [0.2, 0.7]$, $p_{mono} : [0.1, 0.5]$, $g_{max,e} : [50 pS, 70 pS]$, $g_{max,i} : [500 pS, 700 pS]$, $G_e : [1, 5]$, $g_e : [1, 5]$, $G_i : [1, 5]$. Recall that Box-Behnken design fits a second order polynomial model.

This model reproduces the values of $f_{fast}$ and $f_{slow}$ with good accuracy (a coefficient of determination $R^2 = 0.88$ and $R^2 = 0.95$ respectively).

In the approximation of $f_{slow}$, a few parameters seem to stand out, mostly the frequency of the input $f_1$ as well as the product of parameters $f_1 \cdot g_e$. The approximation of $f_{fast}$ on the other hand seem to involve more parameters, and the most important parameters are different from those for $f_{slow}$.

To simplify the models of $f_{fast}$ and $f_{slow}$, we tried to remove the least significant factors one by one and re-adjust the model iteratively (see Figure 2A).

For approximating $f_{slow}$, keeping only four parameters is enough to keep a $R^2$ value above 0.8. These four factors are: $f_1$, $f_1 \cdot g_e$, $f_1 \cdot p_{mono}$, and $f_1^2$. In other words, the slow frequency of the oscillations mostly follows the frequency of the input, though other parameters also influence it to a lesser extent. However, $f_{fast}$ appears as a more complex function of the network parameters: not less than 15 factors are needed to model it with a $R^2$ value above 0.8. These 15 factors include all of the initial ten parameters studied, except for the maximum conductance of inhibitory synapses $g_{max,i}$ and the cholinergic gain on excitatory synapses conductance $G_e$.

B. Parameter study with a Sobol’ method

To provide a better understanding of the parameters influencing the fast oscillations of the network in particular, we completed our previous analysis of $f_{fast}$ with a global sensitivity analysis following a Sobol’ method.

After performing 2200 simulations sampling the whole parameter space, the first-order as well as global sensitivity indices obtained for the output $f_{fast}$ are shown on Figure 2B. Even though $A_1$ looks like it is the only important factor in a first order approximation, global sensitivity indices reveal that other factors such as $f_1$, $g_e$ and $p_{mono}$ also carry a lot of information on the variance of this output. The gain put on increased excitatory synaptic conductances $G_e$ (in the dentate gyrus) and inhibitory synaptic conductances $G_i$ (in the dentate gyrus and CA1) have much less influence on $f_{fast}$ than the reduction of excitatory synaptic conductances $g_e$ (in the entorhinal cortex and CA3), in accordance with the common hypothesis stating that sharp-wave ripples complexes are initiated by recurrent excitatory connections in CA3 ([2]).
A- $R^2$ values of the models of $f_{\text{fast}}$ and $f_{\text{slow}}$ obtained with different number of factors, from a Box-Behnken DOE. Starting from the original model with all factors included, the least significant factor was removed iteratively. B- Sobol’ sensitivity indices (first order and global) for all our ten parameters for the study of $f_{\text{fast}}$, with confidence intervals at 95%. C- Mean and standard deviation of the set of parameters yielding fast oscillations in the gamma (30-90Hz) and ripple (130-200Hz) frequency range respectively. A star (*) next to a parameter’s name indicate that the gamma and ripple sets are statistically different ($p < 0.001$). D- Evolution of the $f_{\text{fast}}$ oscillatory frequency depending on the parameters $g_{\text{max},i}$, $g_e$, $p_{\text{mono}}$ and $g_{\text{CAN}}$. Colors indicate whether $f_{\text{fast}}$ is in the gamma band (blue), the ripple band (red), or in between (white).

A quadratic model can also be fit on the results of these simulations as was done with the Box-Behnken DOE, which yields a coefficient of correlation $R^2$ of 0.70 for the estimation of $f_{\text{fast}}$, with 30 out of the 65 factors being statistically significant ($p < 0.001$).

This analysis helped us find the appropriate set of parameters for reproducing either delta, theta, gamma or ripple oscillations, depending on the vigilance state. In the next section, we will focus on non input-related parameters.

C. Parameter choice for healthy hippocampus modeling

Representing typical oscillations of the sleep-wake cycle can be done by choosing the appropriate sets of parameters so as to reproduce biological observations, namely, $f_{\text{slow}}$ in the theta range and $f_{\text{fast}}$ in the gamma range for wakefulness and $f_{\text{slow}}$ in the delta range and $f_{\text{fast}}$ in the ripple range for slow-wave sleep respectively. We showed before the frequency $f_{\text{slow}}$ is mostly determined by the frequency $f_1$ of the input, which is why this section will be focused on the choice of parameters needed to obtain oscillations either in the gamma or ripple range.

From the simulations performed for the Sobol’ sensitivity analysis, we classified the parameters depending on the frequency range of the output $f_{\text{fast}}$ (gamma and ripple), and represented their distribution on Figure 2C. Four parameters, $p_{\text{tri}}$, $g_{\text{max},e}$, $g_e$ and $g_{\text{CAN}}$, have a significantly different distribution between the gamma and ripple bands ($p < 0.001$ in a WilcoxonMannWhitney test). Figure 2D shows the value of $f_{\text{fast}}$ depending on these four parameters on the simulations performed in the Sobol’ analysis.

Overall, a large set of parameters could be used to produce either sleep or wakefulness oscillations. However, not all of them would be biologically plausible. Among the parameters that can be used to tune the network’s behavior, some of them like $g_e$, $G_e$, $g_i$ and $g_{\text{CAN}}$ that reflect neurotransmitter concentrations could change their value to better reproduce either sleep or wakefulness oscillations. On the other hand, $p_{\text{tri}}$ and $p_{\text{mono}}$ represent the structural connectivity of the network, so they should not be changed between simulations once properly set. The same holds true for $g_{\text{max},e}$ and $g_{\text{max},i}$ which represent the basic conductances of synapses (i.e. without any synaptic plasticity or external neurotransmitter influence).

Using this analysis we were able to choose suitable sets of parameters to reproduce healthy sleep and wakefulness oscillations with fewer simulations compared to a regular sampling of the 10-dimensional parameter space.

Because $p_{\text{tri}}$ and $g_{\text{max},e}$ can been shown (Figure 2B) to have only limited influence on the output oscillations, we
chose to keep them at the mean value of the previous studied range, that is: $p_{\text{tri}} = 0.45$ and $g_{\text{max},e} = 60\text{pS}$. The value of $g_{\text{max},i}$ is slightly more influential, but as shown on Figure 2C the mean value $g_{\text{max},i} = 600\text{pS}$ is also appropriate to be able to reproduce both gamma and ripple oscillations.

The parameters $G_e$ and $G_i$, which have very limited influence on $f_{\text{fast}}$, represent the possible increases of the synaptic conductances due to the presence of neurotransmitters such as Acetylcholine in the network. Therefore, we chose to have $G_e = G_i = 1$ to represent the slow-wave sleep state (that is, the absence of Acetylcholine), and $G_e = G_i = 3$ (the mean value of the range we studied) for the representation of the wakefulness state, for biological plausibility reasons.

Then, the three main parameters that remain to be chosen are $p_{\text{mono}}$, $g_e$ and $g_{\text{CAN}}$. From Figure 2D, it can be seen the fastest oscillations (i.e. in the ripple range) can be obtained roughly when $g_e$ and $g_{\text{CAN}}$ are low, and $p_{\text{mono}}$ is rather high. Conversely, oscillations in the gamma range are obtained with higher $g_e$ and $g_{\text{CAN}}$ and low $p_{\text{mono}}$.

Because $p_{\text{mono}}$ should be kept the same during wakefulness and slow-wave sleep for biological plausibility, we propose to use the mean value of 0.3. For the other values, we chose to have $g_{\text{CAN}}$ set to its minimal value to represent sleep (0.5$\mu$S/cm$^2$) and to its highest value during wakefulness (25$\mu$S/cm$^2$), and $g_e$ to its lowest value during sleep ($g_e = 1$) and to its middle value ($g_e = 3$) during wakefulness to stay coherent with the other parameters representing the variation of synaptic conductances $G_e$ and $G_i$. These values, also represented on Figure 2D, are in agreement with our previous study in realistic inputs ([3]).

IV. CONCLUSION

In this work, we described methods based on Design of Experiments (DOE) techniques and Sobol’ sensitivity analysis to analyze the influence of parameters in a computational neuroscience model.

We applied this method to a model of the healthy hippocampus we had developed previously so as to explore further, more systematically, the role of its main parameters on the fast and slow frequency components of its oscillatory activity. With a DOE technique and second-order polynomial fitting, we found out by performing few simulations that the slow oscillations generated by the network are strongly dependant on the frequency of the input used to stimulate it, while the fast oscillations are more complex. We then used a Sobol’ sensitivity analysis to provide a deeper understanding of the fast oscillations.

The analysis of these simulations shows that our hippocampal model can have specific responses to stimuli from afferent structures or sensory modalities in accordance with its biological function, but these outputs also depend on structural and functional properties of the network. This analysis helped us find the appropriate set of parameters for reproducing either delta, theta, gamma or ripple oscillations, depending on the vigilance state. In total, 170 simulations were performed for the DOE and 2200 for the Sobol’ analysis, to be compared with the $3^{10} = 59,059$ simulations needed to get even a very sparse regular sampling of the parameter space.

Meaningful oscillatory rhythms can be reproduced, namely fast oscillations in the gamma or ripple frequency range, respectively. We found that only four parameters had a statistically significant distribution between the sets of previously ran simulations leading to gamma and ripple oscillations. In this reduced parameter space, we were able to identify biologically relevant sets of parameters for sleep and wakefulness behavior.

Such work could later be used as a reference when extending the model to other physiological rhythms, as well as pathological conditions such as epilepsy. A similar methodology could also be used in other detailed computational neuroscience models to help with the tuning of parameters.

REFERENCES


