Noname manuscript No.
(will be inserted by the editor)

An attempt to fit all parameters of a dynamical recurrent neural network from sensory neural spiking data

R. Ozgur DORUK · Kechen Zhang

Received: date / Accepted: date

Abstract A simulation based study on model fitting for sensory neurons from stimulus/response data is presented. The employed model is a continuous time recurrent neural network (CTRNN) which is a member of models with known universal approximation features. This feature of the recurrent dynamical neuron network models allow us to describe excitatory-inhibitory characteristics of an actual sensory neural network with any desired number of neurons. This work will be a continuation of [10] where the parameters associated with the sigmoidal gain functions are not taken into account. In this work, we will construct a similar framework but all parameters associated with the model are estimated. The stimulus data is generated by a Phased Cosine Fourier series having fixed amplitude and frequency but a randomly shot phase. Various values of amplitude, stimulus component size and sample size are applied in order to examine the effect of stimulus to the identification process. Results are presented in tabular and graphical forms at the end of this text. In addition a comparison of the results with previous researches including [10] will be presented.

Keywords Neural spiking · Inhomogeneous Poisson Point Processes · Maximum Likelihood Methods · Continuous Time Recurrent Neural Networks · Estimation of Sigmoidal Gain Parameters

1 Introduction

1.1 General Discussion

Theoretical and computational neuroscience is expected to be a major component of general neuroscience or neurobiology study in a near future. The signs of that developments first noted when the famous mathematical models such as Hodgkin-Huxley [15], Fitzhugh-Nagumo [12] and Morris-Lecar [21] were developed and refereed several times by numerous researchers worldwide. These models can adequately describe the bursting behaviors of firing neurons. They are highly nonlinear models and their main state variable is the membrane potential. In addition they are more or less involves certain physical components such as channel conductances etc (except Fitzhugh-Nagumo which only has membrane potential as a physical component). In addition one also has firing rate based models [17,20] which are thought to be of more truly a computational type. These are useful, if only the rate of firing and/or interspiking intervals (ISI) of the spikes are thought to be important in neural information coding [13]. The efficiency and usability of these models depend on the aim of the research and the limitations set by the simulation/experiment environment. In experiments related to computational neuroscience field, one such limitation may arise from the measurement capability. In vivo experiments, does not allow the

The computational tasks associated with this work are completed in the High Performance Computing Facilities (HPC) provided by TR-GRID/TRUBA Project operated by National Academic Computing Center of the Turkish Scientific and Technological Research Institution (TUBITAK-ULAKBIM) of Republic of TURKEY.

R.O.DORUK
Atilim University, Kizilcasar Mahallesi, Incek, Golbasi, 06836, Ankara, TURKEY
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: resat.doruk@atilim.edu.tr

K. Zhang
Johns Hopkins School of Medicine, Rutland Avenue, Traylor Building, No: 407, 21205, Baltimore, Maryland, USA

real time measurement of the membrane potential. An attempt to achieve this will likely to interrupt the propagation of action potentials due to a change in the axial membrane physical properties at the location of electrode. There is even the risk of a damage to the neuron. So, a feasible and safe way to collect data from a living neuron (in vivo) can be the placement of an electrode at a nearby location and by that one is able to record the firing instants of the interested neuron. Only disadvantage in doing that is the lack of continuous time dependent data. Another tackling feature of the neural spiking phenomena is that, it is not a deterministic event. The stochasticity of the ion channels [14] and synaptic noise led to the fact that the data transmitted along the neurons is a stochastic process. In studies such as [24], it can be noted that, this stochasticity of neural spiking obeys the famous Inhomogeneous Poisson Process at least for the sensory neurons. Those findings suggest the possible usage of a likelihood functions derived from point process theory [11,3]. This likelihood is a function of parameters and temporal location of individual spikes. Such approaches are expected to provide a better identification results than direct usage of the Poisson Mass Function instead [8]. The latter only requires the number of spikes rather than their locations.

There are certain challenges in this research. First of all, we will most probably not be able to have a reasonable estimate just from a single spiking response data set as we do not have a continuous response data. This is also demonstrated in the related kernel density estimation research such as [22,26,25, 16]. From these sources, one will easily note that repeated trials and superimposed spike sequences are required to obtain a meaningfully accurate firing rate information from the neural response data. In a realistic experimental environment, repeating the trials with same stimulus profile will not be appropriate as the corresponding responses of the repeated stimuli are found to be attenuated.

Because of this issue, a new stimulus should be provided at each excitation. This can be provided by choosing a fixed amplitude and frequency but randomly shot phase angle for our Fourier Series stimulus. Secondly, in the likelihood estimation, the complete data from the beginning will be used in the likelihood optimization. This will be a computational challenge as a very large data will be accumulated in each computation step. When considering an experiment we collect the data only by providing a random stimulus entry to the animal (experiment subject) and record the spike counts and locations. As animal is not involved in the computational part of the random stimuli based experiments an high performance computing (HPC) facility can be involved without a need of any wet experimental element. In this research, we are employing the high performance computing facilities (TRUBA/TR-GRID) of the National Academic Information Center (ULAKBIM) of Turkish Scientific and Technological Research Institution (TUBITAK).

1.2 Previous Studies

This work is a fairly novel attempt. There are very few studies in the literature that have a similar goal. Some examples can be given as [2,3,5,23,27]. The work in [2] presents a system identification study based on maximum likelihood estimation of the internal parameters of an integrate and fire neuron model. Likelihood function is derived from firing probabilities through local Bernoulli approximation. [5] aims at the detection of the functional relationships between neurons. Rather than modeling an individual neuron, it involves a characterization of the neural interactions through maximum likelihood estimation. [23] is somehow similar to [2]. A thorough explanation of maximum likelihood explanation is presented with an application to a linear-nonlinear Poisson cascade and an integrate and fire model generalized linear model. It also presents a comparison with traditional spike triggered average estimator. [27] presents a similar work to that of [2] and [23] with a different model. The model involve an estimation of a conditional intensity function modulated by an unobservable latent state-space process. Study also involves the identification of the latent process. Both estimation approaches are based on maximum likelihood method. [23] and [27] applies expectation maximization method in the solution of the maximum likelihood problems. For a more general discussion on the application of statistical techniques and their challenges in theoretical and computational neuroscience interested readers can apply to the reference [29].

This research has some common grounds with [2] and [23] due to the application of maximum likelihood method to a neural network identification problem. However the model used in this research is quite different from the ones in those sources. Instead of an integrate and fire model we prefer a more general continuous time recurrent neural network due to their universal approximation capability which is expected to be an advantage to model a multi-cellular region of the nervous system. In addition their dynamical properties are closer to network models such as Hodgkin-Huxley or Moris-Lecar equations.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27015v1 | CC BY 4.0 Open Access | rec: 29 Jun 2018, publ: 29 Jun 2018

Research such as [6,7] implements a generic neural network model which is of the a static feed-forward type. Based on all these, one can say that this study can be considered as a novel contribution to the neuroscience literature. In addition the work done in [2,3,5,23,27] are too elaborate in statistical theory with a very limited discussion on how to apply the theory to neuron modeling. This restricts the reproducibility of those research. This text concentrates also on how to apply the theory on the identification problem using computational tools such as MATLAB to increase its reproducibility.

1.3 Summary of This Work

In this research, we will perform a study similar to that of [10]. In that work, the parameters associated with the sigmoidal gain functions are assumed to be known. This brings a critical issue as in most of the cases the parameters of gain functions are not known or vaguely known. Because of that we will need to include them in the set of estimated parameters. This will increase the computational complexity but this difficulty can largely be solved thanks to the High-Performance-Computing facilities provided by TRUBA system. There will be total of 14 parameters to estimate. The mean values and percent errors of the estimates are presented in forms of tables whereas their variances are presented graphically to reveal its variation.

2 Materials & Methods

2.1 Continuous Time Recurrent Neural Networks (CTRNN)

The continuous time recurrent neural networks have a similar structure to that of the discrete time counterparts that are often met in artificial intelligence studies. These models can have any number of neurons and a mathematical model can be expressed as shown below [1]:

$$\tau_l \frac{dV_l}{dt} = -V_l + \sum_{j=1}^{n} W_{lj} g_j\left(V_j\right) + \sum_{k=1}^{m} C_{lk} I_k \tag{1}$$

Like that of [10], a second order CTRNN is examined in this work:

$$\begin{aligned} \tau_e \dot{V_e} &= -V_e + g_e(V_e) w_{ee} - g_i(V_i) w_{ei} + c_e I \\ \tau_i \dot{V_i} &= -V_i + g_e(V_e) w_{ie} - g_i(V_i) w_{ii} + c_i I \end{aligned} \tag{2}$$

In the above the subscripts e refers to excitatory neurons (or neurons having excitatory synapses), i refers to inhibitory neurons (or neurons having inhibitory synapses). The definitions of the variables and parameters are given in the following table (**Table 1**). In the estimation studies, it will be easier if the time constants are moved to the right:

$$\begin{aligned} \dot{V_e} &= \beta_e(-V_e + g_e(V_e) w_{ee} - g_i(V_i) w_{ei} + c_e I) \\ \dot{V_i} &= \beta_i(-V_i + g_e(V_e) w_{ie} - g_i(V_i) w_{ii} + c_i I) \end{aligned} \tag{3}$$

The functions $g_e(V_e)$ and $g_i(V_i)$ are sigmoidal gain functions and expressed as:

$$\begin{aligned} g_e\left(V_e\right) &= \frac{\Gamma_e}{1 + \exp\left(-a_e\left(V_e - h_e\right)\right)} \\ g_i\left(V_i\right) &= \frac{\Gamma_i}{1 + \exp\left(-a_i\left(V_i - h_i\right)\right)} \end{aligned} \tag{4}$$

where e and i refers to the excitatory and inhibitory neurons respectively. The definitions of the parameters in (4) are presented in **Table 1**. It should also be noted that, the weights in (3) are all assumed as positive coefficients and they have signs in the equation. So negative signs indicate that originating neuron is inhibitory (tend to hyper-polarize the other neurons in the network).

Table 1 The definitions of neuron variables and parameters

Variable or Parameter	Definition	Unit
I	External stimulus input	μA
V_e	Excitatory neuron membrane potential	mV
V_i	Inhibitory neuron membrane potential	mV
τ_e	Excitatory neuron time constant	s
τ_i	Inhibitory neuron time constant	s
β_e	Excitatory neuron reciprocal time constant $\left(\beta_e = \frac{1}{\tau_e}\right)$	s^{-1}
β_i	Inhibitory neuron reciprocal time constant $\left(\beta_i = \frac{1}{\tau_i}\right)$	s^{-1}
w_{ee}	Self excitation synaptic weight	mV·s
w_{ei}	Synaptic weight of excitatory neuron inhibition	mV·s
w_{ie}	Synaptic weight of inhibitory neuron excitation	mV·s
w_{ii}	Self inhibition synaptic weight	mV·s
c_e	Synaptic weight of external input to the excitatory neuron	kΩ
c_i	Synaptic weight of external input to the inhibitory neuron	kΩ
Γ_e	Maximum firing rate of the excitatory neuron	s^{-1}
Γ_i	Maximum firing rate of the inhibitory neuron	s^{-1}
a_e	Slope parameter (excitatory sigmoidal gain)	mV^{-1}
a_i	Slope parameter (inhibitory sigmoidal gain)	mV^{-1}
h_e	Soft threshold parameter (excitatory sigmoidal gain)	mV
h_i	Soft threshold parameter (inhibitory sigmoidal gain)	mV

2.2 Inhomogeneous Poisson spike model

The theoretical response of a neuron model will be the firing rate of excitatory or inhibitory neuron. In this research, we collect the data from excitatory neuron. So our response will theoretically be:

$$r_e = g_e\left(V_e\right) \tag{5}$$

In the actual environment, the neural spiking due to the firing rate $r_e\left(t\right)$ is available instead. While introducing this research, it is stated that this spiking events conform to an inhomogeneous Poisson process which is defined below:

$$\mathrm{Prob}\left[N\left(t+\Delta t\right)-N\left(t\right)=k\right]=\frac{e^{-\lambda}\lambda^{k}}{k!} \tag{6}$$

where

$$\lambda=\int_{t}^{t+\Delta t} r_e\left(\tau\right)d\tau \tag{7}$$

is the mean number of spikes based on the firing rate $r_e(t)$ which varies with time, and $N(\mathcal{T})$ indicates the cumulative total number of spikes up to time $\mathcal{T}$, so that $N\left(t+\Delta t\right)-N\left(t\right)$ is the number of spikes within the time interval $[t, t+\Delta t)$. In other words, the probability of having k number of spikes in the interval $(t, t+\Delta t)$ is given by the Poisson distribution above.

Consider a spike train $(t_1, t_2, \ldots, t_K)$ in the time interval $(0, T)$ (here $0 \le t_1 \le t_2 \le \ldots \le t_K \le T$ so t and Δt become $t = 0$ and $\Delta t = T$). Here the spike train is described by a list of the time stamps for the K spikes. The probability density function for a given spiking train $(t_1, t_2, \ldots, t_K)$ can be derived from the inhomogeneous Poisson process [11,3]. The result reads:

$$p\left(t_1, t_2, \ldots, t_K\right)=\exp\left(-\int_0^T r_e\left(t,\mathbf{x},\theta\right)dt\right)\prod_{k=1}^{K} r_e\left(t_k,\mathbf{x},\theta\right) \tag{8}$$

This probability density describes how likely a particular spike train $(t_1, t_2, \ldots, t_K)$ is generated by the inhomogeneous Poisson process with the rate function $r_e\left(t,\mathbf{x},\theta\right)$., Of course, this rate function depends implicitly on the network parameters and the stimulus used.

2.3 Maximum Likelihood Methods and Parameter Estimation

The network parameters to be estimated are listed below as a vector:

$$\theta = [\theta_1, \ldots, \theta_{14}] = [\beta_e, \beta_i, c_e, c_i, w_{ee}, w_{ei}, w_{ie}, w_{ii}, \Gamma_e, \Gamma_i, a_e, a_i, h_e, h_i] \tag{9}$$

which includes the time constants and all the connection weights in the excitatory-inhibitory network of (3). The maximum-likelihood estimation of network parameters is based on the likelihood function given by (8), which takes the individual spike timings into account. It is well known from estimation theory is that maximum likelihood estimation is asymptotically efficient, i.e., reaching the Cramér-Rao bound in the limit of large data size. To extend the likelihood function in (8) to the situation where there are multiple spike trains elicited by multiple stimuli, consider a sequence of M stimuli. This means that we drive the network in (3) M times by generating M different stimuli at each trial. If I_j and I_k are the stimuli for the j^{th} and k^{th} trials respectively for $j, k = 1, \ldots, M$, $I_j \neq I_k$ for all cases where $j \neq k$. Suppose the m-th stimulus ($m = 1, \ldots, M$) elicits a spike train with a total of K_m spikes in the time window $[0, T]$, and the spike timings are given by $S_m = \left(t_1^{(m)}, t_2^{(m)}, \ldots, t_{K_m}^{(m)}\right)$. By (8), the likelihood function for the spike train S_m is

$$p\left(S_m \mid \theta\right) = \exp\left(-\int_0^T r_e^{(m)}(t)\, dt\right) \prod_{k=1}^{K_m} r_e^{(m)}\left(t_k^{(m)}\right) \tag{10}$$

where $r_e^{(m)}$ is the firing rate in response to the m-th stimulus. Note that the rate function $r_e^{(m)}$ depends implicitly on the network parameters θ and on the stimulus parameters. The left-hand side of (10) emphasizes the dependence on network parameters θ, which is convenient for parameter estimation. The dependence on the stimulus parameters will be discussed in the next section.

We assume that the responses to different stimuli are independent, which is a reasonable assumption when the inter-stimulus intervals are sufficiently large. Under this assumption, the overall likelihood function for the collection of all M spike trains can be written as

$$L\left(S_1, S_2, \ldots, S_M \mid \theta\right) = \prod_{m=1}^{M} p\left(S_m \mid \theta\right) \tag{11}$$

By taking natural logarithm, we obtain the log likelihood function:

$$l\left(S_1, S_2, \ldots, S_M \mid \theta\right) = -\sum_{m=1}^{M} \int_0^T r_e^{(m)}(t)\, dt + \sum_{m=1}^{M} \sum_{k=1}^{K_m} \ln r_e^{(m)}\left(t_k^{(m)}\right) \tag{12}$$

Maximum-likelihood estimation of the parameter set is given formally by

$$\hat{\theta}_{ML} = \arg\max_{\theta} \left[l\left(S_1, S_2, \ldots, S_M \mid \theta\right)\right] \tag{13}$$

2.4 Stimulus

As discussed in **Section 1.1**, we will model the stimulus signal by a phased cosine Fourier series as shown below:

$$I = \sum_{n=1}^{N_U} A_n \cos\left(\omega_n t + \phi_n\right) \tag{14}$$

where A_n is the amplitude, $\omega_n = 2\pi f_0 n$ is the frequency of the n-th Fourier component in radians/sec, and ϕ_n is the phase of the component. Here the amplitude A_n and the base frequency f_0 (in Hz) are fixed but the phase ϕ_n will be a randomly chosen from a uniform distribution between $[-\pi,\ \pi]$ radians. The amplitude parameter A_n is fixed for all mode n as $A_n = A_{\max}$.

Table 2 The true values of the parameters of the network model in (3) and (4). These are the parameters to be estimated.

Parameter	Unit	True value (θ)
β_e	$1/s$	50
β_i	$1/s$	25
w_e	kΩ	1.0
w_i	kΩ	0.7
w_{ee}	mV·s	1.2
w_{ei}	mV·s	2.0
w_{ie}	mV·s	0.7
w_{ii}	mV·s	0.4
Γ_e	100	1/s
a_e	0.04	1/mV
h_e	70	mV
Γ_i	50	1/s
a_i	0.04	1/mV
h_i	35	mV

3 Results

3.1 Model Details

In this section, we will summarize the functional and numerical details of the neural network parameter estimation algorithm. In the example application, the algorithms presented in **Section 2.3** are applied to probe an excitatory-inhibitory network. In order to verify the performance of the parameter estimation we have to compare the estimates with their true values. So we will need a set of reference values of the model parameters in (3) and (4). These are shown in **Table 2**. This set of parameters (**Table 2**) allows the network to have a unique equilibrium state for each stationary input. This set of parameters (**Table 2** allows the network to have a unique equilibrium state for each stationary input. The numerical values of the parameters are exactly the same as that of [10] so that we can make a comparison between the results. In addition certain details of the model such as wide and narrow pulse responses, response to the Fourier series stimulus with different component sizes (N_U), mean firing rate response and dynamic range can be found in the same reference [10]. These will not be presented here in order to save space and not to dilute the interest of the readers.

3.2 Stimulus and Response

The stimulus to be used in this example is given in (14). In this application, A_n will be constant for all mode n ($n = 1 \dots N_U$). The phase angles ϕ_n will be assigned randomly (uniformly distributed between $[-\pi, \pi]$) at each iteration. It should be noted that we do not intend to provide a random stimulus here. The phase angles will be randomly drawn at the beginning of each iteration and stays constant till the new iteration starts. Due to the random assignment they will be different for each iteration. This will yield a different Fourier series stimulus at each iteration. In the case of an experiment, this is critical as the response of a neuron may cease after repeating the stimulation with the same stimulus profile a few times.

3.3 Spike Generation

As we have discussed in **Section 1.1**, we will not have any measurement of membrane potential $V_e(t)$ or $V_i(t)$. Instead, we will record the spike timings of the neuron and try to solve a maximum likelihood estimation of network parameters $\hat{\theta}$ using the likelihood function in (11). Because of that, the simulation needs a method to generate the spike timings of the neurons. As we know from [24] that the spikes obey an inhomogeneous Poisson distribution, one can obtain the spike timings from a simulation of an

inhomogeneous Poisson process of which event or firing rate is given by:

$$r_e(t) = g_e\left(V_e\right) = \frac{\Gamma_e}{1 + \exp\left(-a_e\left(V_e - h_e\right)\right)} \tag{15}$$

There are numerous methodologies to generate the Poisson events given the event rate $r_e(t)$. These ranging from discrete simulation [11] to thinning [18]. Discrete simulation may be beneficial when one solves the dynamical models by fixed step solvers such as Euler Integration or Runge-Kutta methods. The only disadvantage of this approach is that, it confines the spikes into discrete time bins.However, if one has a sufficiently small discrete time bin such as $\Delta t = 1$ ms, the statistical distribution of the spikes should approach to that of an Inhomogeneous Poisson Process [11]. Discrete simulation of neural spiking can be summarized as shown below:

1. Given the firing rate of any neuron as $r(t)$
2. Find the probability of firing at time t_i by evaluating $p_i = r(t_i)\Delta t$ where Δt is the integration interval. It should be as small as 1 ms.
3. Compute a random variable by drawing a sample from a distribution which is uniform between 0 and 1. Define this as $x_{rand} = U[0, 1]$ where U stands for uniform distribution.
4. If $p_i > x_{rand}$ fire a spike at $t = t_i$, else do nothing.
5. Collect spikes as $S = [t_1, \ldots, t_{N_s}]$ where N_s will be the number of spikes obtained at a single run of simulation.

3.4 Step-by-step description of the Problem and Simulation

The working principles in the example problem can be described in a step-by-step fashion as shown below:

1. A single run of simulation will last for $T_f = 3$ seconds.
2. The neuron model in (3) will be simulated at the true value of parameters which are given in **Table 2** and firing rate data is stored as $r_m(t)$ where m is the current number of simulation.
3. Firing rate data $r_m(t)$ is used to generate neural spikes S_m in the m^{th} run using the methodology defined in **Section 3.3**. This data will be used to compute the likelihood. The number of spikes will be K_m at the m^{th} run.
4. Repeat the simulation N_{it} times to obtain enough number of spikes.
5. The spiking data needed by (12) will be obtained at the 4^{th} step. However, the firing rate component of (12) should be computed at the current iteration of the optimization.
6. Run an optimization algorithm of which objective computes the firing rate at the current iterated value of the parameters but the spikes from **Step 4**.

3.5 Optimization Algorithm

Theoretically, any optimization algorithm ranging from gradient descent to derivative free simulated annealing can be utilized in computational parts of this research. Most of these algorithms are provided as ready made routines in the optimization and global optimization toolboxes of MATLAB. Regardless of the type of algorithm, all of the methods converge to a local optimum and requires an initial guess. As a result, one needs to start from multiple initial guesses to have a adequate amount of local optimum that will allow us to detect the global one. If we have a convex problem, different initial guesses are expected to converge to same local optimum and this is the desirable situation. However, this may not always be the case in problems similar to that of this research. In any case, the main criteria on the choice of the algorithms is the speed of convergence. Though we have a HPC computing facility we should choose the fastest algorithm as we need to collect a huge amount of data to conclude about the efficiency of the project. Some initial evaluations suggested that local optimizer routines provided by MATLAB's `fmincon` should be preferred concerning speed and computational resource considerations. The MATLAB provided global optimization algorithms such as genetic algorithms, simulated annealing or pattern search may also be utilized to solve the same problem. However they seem to be computationally more intensive and they are also designated officially as local optimizers. Thus one may need to repeat the trials a few times to find an optimum. This is not desired if we adapt this research to an experiment. The `fmincon` algorithm needed gradient information but it can be provided by itself through finite difference approximations. There will be 14 (this number equals to the number of cores in a local machine) initial

Table 3 Typical data related to the simulation scenario. **Table 4**

Parameter	Symbol	Value
Simulation Time	T_f	3 sec.
Number of Trials	N_{it}	100
# of Components in Stimulus	N_U	5
Method of Optimization	N/A	Interior-Point Gradient Descent (MATLAB)
# of True Parameters	Size(θ)	8
Stimulus Amplitude (μA)	$A_{\max}$	100
Base Frequency	f_0	3.333 Hz

Table 4 The data related to the analysis of the problem for different number of trials N_{it}, number of stimulus components N_U, stimulus amplitude $A_{\max}$

Parameter	Symbol	Value(s)
Number of Trials	N_{it}	25, 50, 100, 200, 400
# of Components in Stimulus	N_U	5, 10, 20, 30, 40, 50
Stimulus Amplitude (μA)	$A_{\max}$	25, 50, 100, 200, 400

guesses and each initial run will be performed on one core. The whole optimization will be run parallel by the `parfor` parallel for loop structure of MATLAB. The initial guesses are generated randomly from a uniform distribution.

3.6 Simulation data

The nominal data in the current problem are given in **Table 3**. In order to reveal the effect of different number of stimulus components N_U, amplitude level $A_{\max}$ and number of trials N_{it} we will repeat the problem for a set of different values of those parameters. The different values of those parameters are provided in **Table 4**. The initial levels of membrane potentials of excitatory and inhibitory neurons are $V_e(0) = 0$ and $V_i(0) = 0$. As we will most probably not know the true values of those conditions assumption of zero values should be sufficient. We will repeat the simulation 10 times for each case, so that we will have sufficient number of results to perform a statistical analysis.

3.7 Presentation of the results

In this section, the numerical results of the maximum likelihood estimation of the parameters of our neuron model in (3) using maximum likelihood estimation through the maximization of (12) against parameters in (9). The optimization is performed using the gradient based interior-point method provided by MATLAB's `fmincon` algorithm. All the cases in **Table 4** are examined under the conditions in **Table 3**. The overall results are presented in the following forms:

1. Mean estimated values are displayed as tables.
2. To be used in comparison, variation of percent estimation errors will be presented in graphical form.
3. Graphical results that presents the variances of estimates and whole variation of percent errors against changing number of stimulus components N_U, amplitude parameter $A_{\max}$, number of trials N_{it} and base frequency f_0 will also be given.

3.7.1 Variation of the estimates against stimulus component size (N_U)

One can see the variation of mean estimated values of the parameters in (9) and their associated percent estimation errors in **Tables 5** and **Figure 1**. The variation of their mean square estimation errors against varying N_U is available in **Figure 2**.

3.7.2 Variation of the estimates against stimulus amplitude parameter ($A_{\max}$)

One can see the variation of mean estimated values of the parameters in (9) and their associated percent estimation errors in **Tables 6** and **Figure 3**. The variation of their mean square estimation errors against varying $A_{\max}$ is available in **Figure 4**.

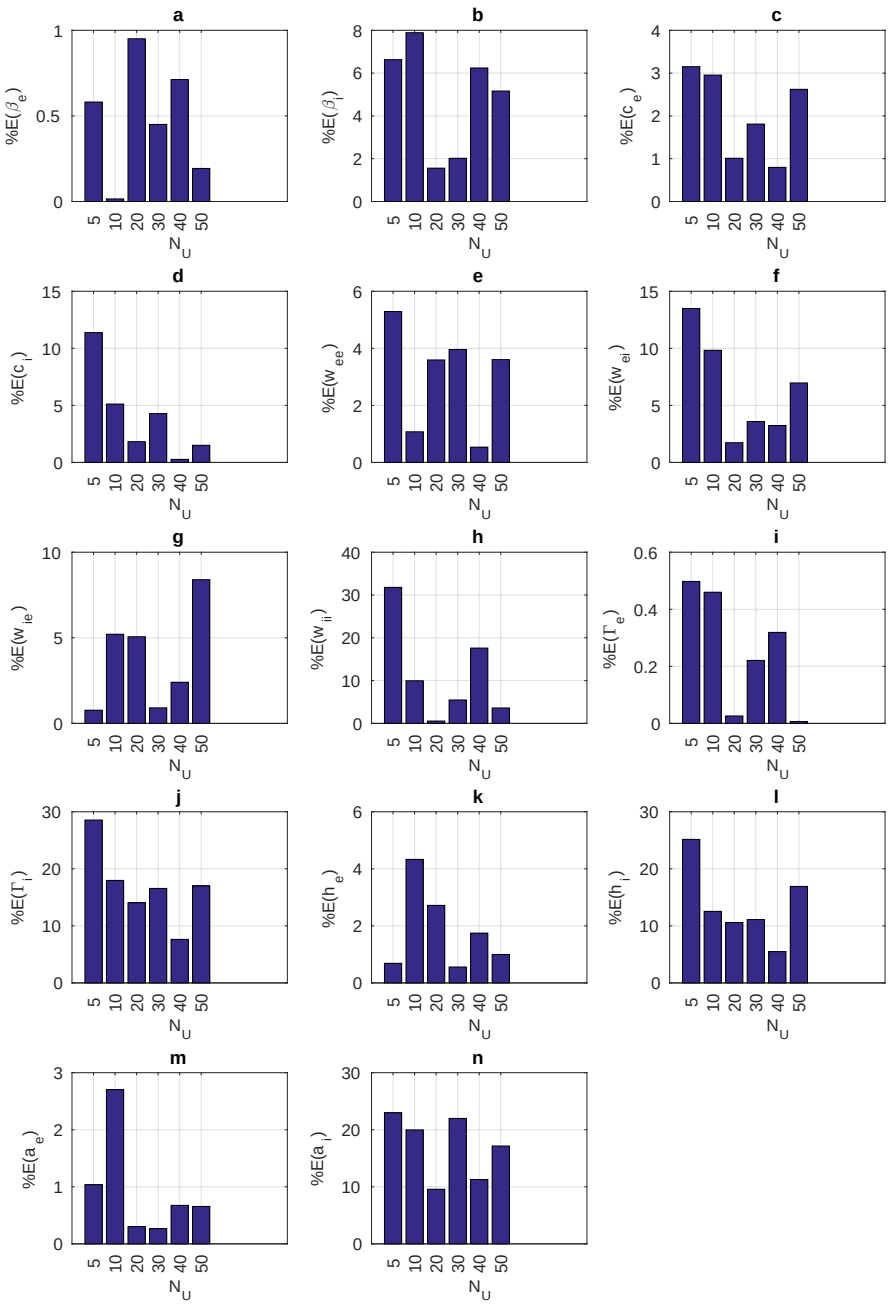

Fig. 1 Variation of the percent errors of estimation (%E) associated with each parameter in (9) against stimulus component size N_U. The other conditions are $A_{\max} = 100$ μA, $N_{it} = 100$ and $f_0 = 3.3333$ Hz. In this figure (a) %E of β_e, (b) %E of β_i, (c) %E of c_e, (d) %E of c_i, (e) %E of w_{ee}, (f) %E of w_{ei}, (g) %E of w_{ie}, (h) %E of w_{ii}, (i) %E of Γ_e, (j) %E of Γ_i, (k) %E of h_e, (l) %E of h_i, (m) %E of a_e and (n) %E of a_i.

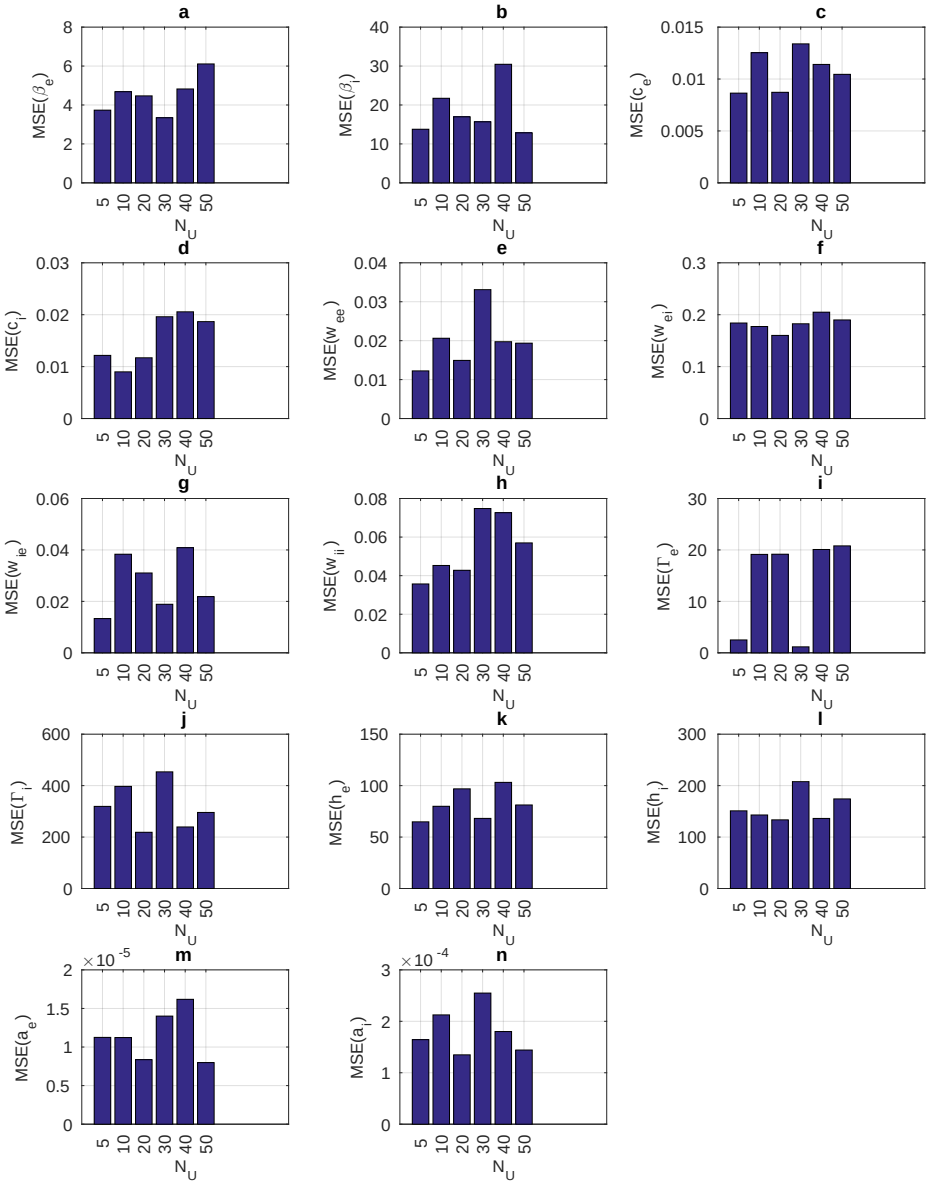

Fig. 2 Variation of the mean square errors of estimation (MSE) associated with each parameter in (9) against stimulus component size N_U. The other conditions are $A_{\max} = 100$ μA, $N_{it} = 100$ and $f_0 = 3.3333$ Hz. In this figure (a) MSE of β_e, (b) MSE of β_i, (c) MSE of c_e, (d) MSE of c_i, (e) MSE of w_{ee}, (f) MSE of w_{ei}, (g) MSE of w_{ie}, (h) MSE of w_{ii}, (i) MSE of Γ_e, (j) MSE of Γ_i, (k) MSE of h_e, (l) MSE of h_i, (m) MSE of a_e and (n) MSE of a_i.

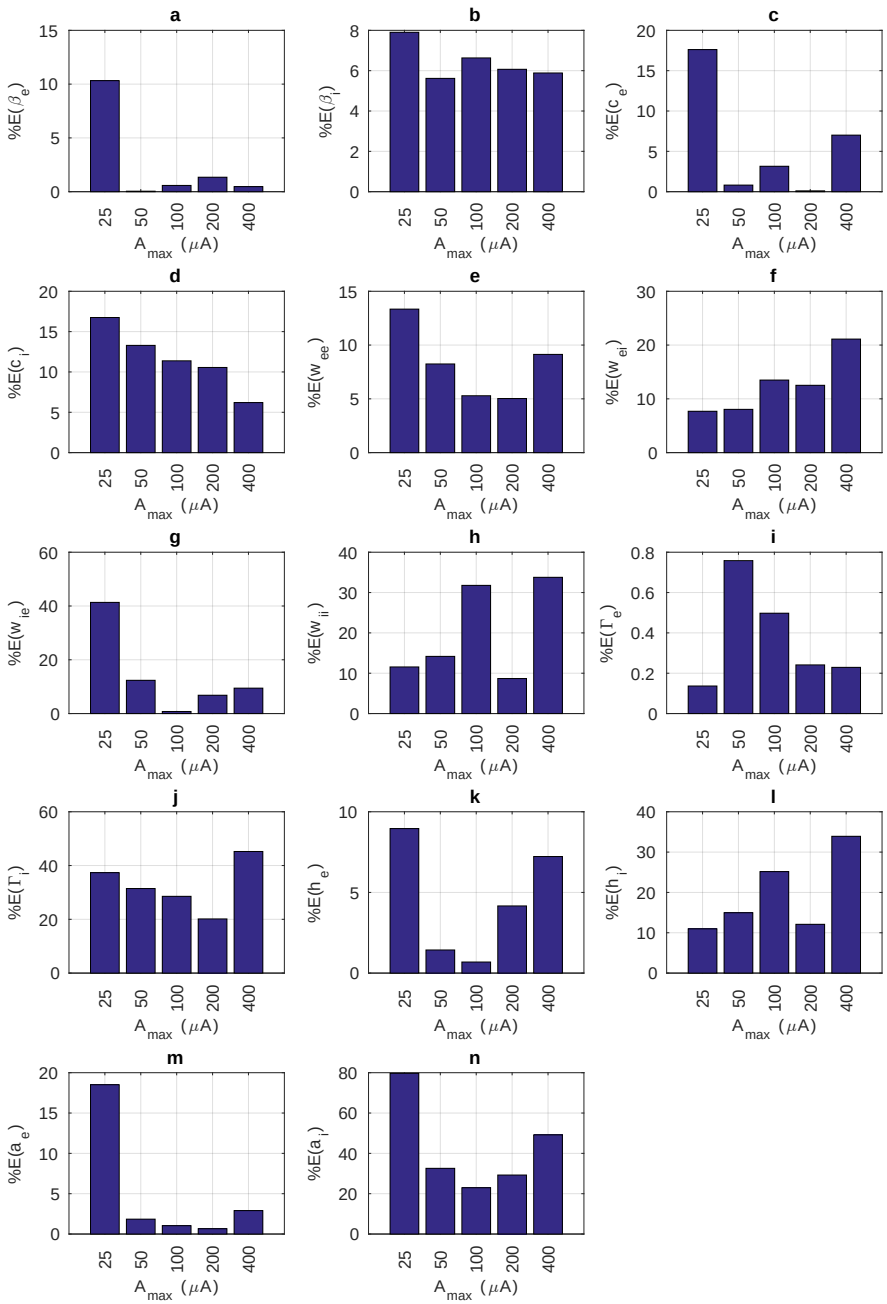

Fig. 3 Variation of the percent errors of estimation (%E) associated with each parameter in (9) against stimulus amplitude parameter $A_{\max}$. The other conditions are $N_U = 5$, $N_{it} = 100$ and $f_0 = 3.3333$ Hz. In this figure (a) %E of β_e, (b) %E of β_i, (c) %E of c_e, (d) %E of c_i, (e) %E of w_{ee}, (f) %E of w_{ei}, (g) %E of w_{ie}, (h) %E of w_{ii}, (i) %E of Γ_e, (j) %E of Γ_i, (k) %E of h_e, (l) %E of h_i, (m) %E of a_e and (n) %E of a_i.

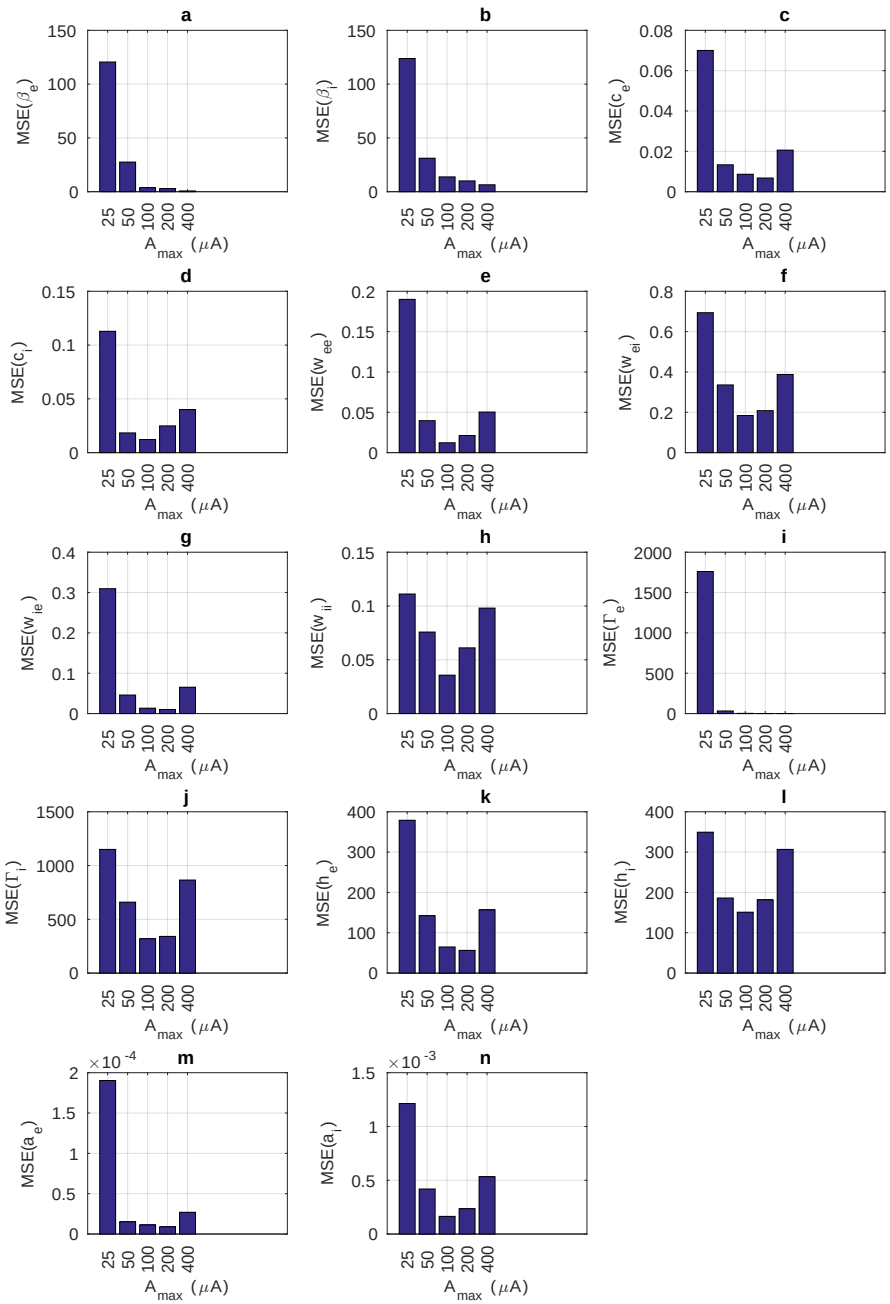

Fig. 4 Variation of the mean square errors of estimation (MSE) associated with each parameter in (9) against stimulus amplitude parameter $A_{\max}$. The other conditions are $N_U = 5$, $N_{it} = 100$ and $f_0 = 3.3333$ Hz. In this figure (a) MSE of β_e, (b) MSE of β_i, (c) MSE of c_e, (d) MSE of c_i, (e) MSE of w_{ee}, (f) MSE of w_{ei}, (g) MSE of w_{ie}, (h) MSE of w_{ii}, (i) MSE of Γ_e, (j) MSE of Γ_i, (k) MSE of h_e, (l) MSE of h_i, (m) MSE of a_e and (n) MSE of a_i.

Table 5 The values of the mean estimated parameters $\hat{\theta}$ against increasing stimulus component size N_U. The other conditions are $A_{\max} = 100$ μA, $N_{it} = 100$ and $f_0 = 3.3333$ Hz.

θ	$N_U = 5$	$N_U = 10$	$N_U = 20$	$N_U = 30$	$N_U = 40$	$N_U = 50$
β_e	49.709	50.008	49.524	49.775	50.356	49.903
β_i	23.342	23.027	24.610	24.495	23.440	23.709
c_e	1.032	1.030	1.010	1.018	1.008	1.026
c_i	0.620	0.664	0.687	0.670	0.698	0.711
w_{ee}	1.263	1.213	1.243	1.248	1.194	1.243
w_{ei}	1.730	1.803	1.965	1.928	1.935	1.861
w_{ie}	0.695	0.736	0.665	0.694	0.683	0.641
w_{ii}	0.527	0.440	0.402	0.378	0.330	0.414
Γ_e	99.502	100.460	99.974	100.221	100.319	99.994
Γ_i	64.276	58.984	57.024	58.281	53.817	58.503
h_e	70.481	73.031	68.097	70.391	71.222	70.698
h_i	26.194	30.610	31.295	31.113	33.079	29.081
a_e	0.040	0.039	0.040	0.040	0.040	0.040
a_i	0.049	0.048	0.044	0.049	0.045	0.047

Table 6 The values of the mean estimated parameters $\hat{\theta}$ against increasing stimulus amplitude $A_{\max}$ (in μA). The other conditions are $N_U = 5$, $N_{it} = 100$ and $f_0 = 3.3333$ Hz.

θ	$A_{\max} = 25$	$A_{\max} = 50$	$A_{\max} = 100$	$A_{\max} = 200$	$A_{\max} = 400$
β_e	55.161	50.023	49.709	50.673	50.235
β_i	23.024	23.595	23.342	23.483	23.528
c_e	0.824	0.992	1.032	0.999	1.070
c_i	0.583	0.607	0.620	0.626	0.657
w_{ee}	1.360	1.299	1.263	1.260	1.310
w_{ei}	2.154	1.839	1.730	1.750	1.578
w_{ie}	0.989	0.787	0.695	0.748	0.766
w_{ii}	0.446	0.457	0.527	0.435	0.535
Γ_e	99.863	99.242	99.502	99.759	99.771
Γ_i	68.675	65.724	64.276	60.073	72.614
h_e	63.730	68.997	70.481	72.914	75.057
h_i	31.159	29.762	26.194	30.766	23.129
a_e	0.047	0.041	0.040	0.040	0.039
a_i	0.072	0.053	0.049	0.052	0.060

Table 7 The values of the mean estimated parameters $\hat{\theta}$ against increasing sample size (number of iterations) N_{it}. The other conditions are $A_{\max} = 100$ μA, $N_U = 5$ and $f_0 = 3.3333$ Hz.

θ	$N_{it} = 25$	$N_{it} = 50$	$N_{it} = 100$	$N_{it} = 200$	$N_{it} = 400$
β_e	48.552	48.628	49.709	49.550	49.990
β_i	23.376	24.032	23.342	23.844	24.064
c_e	1.028	1.038	1.032	1.055	1.069
c_i	0.600	0.605	0.620	0.629	0.654
w_{ee}	1.313	1.283	1.263	1.294	1.293
w_{ei}	1.900	1.843	1.730	1.790	1.939
w_{ie}	0.724	0.704	0.695	0.687	0.709
w_{ii}	0.523	0.479	0.527	0.524	0.458
Γ_e	99.830	100.001	99.502	99.549	99.816
Γ_i	69.977	69.315	64.276	62.857	57.140
h_e	62.648	66.440	70.481	72.038	74.851
h_i	23.198	26.267	26.194	27.345	30.833
a_e	0.041	0.040	0.040	0.039	0.038
a_i	0.054	0.051	0.049	0.047	0.045

3.7.3 Variation of the estimates against sample size (N_{it})

One can see the variation of mean estimated values of the parameters in (9) and their associated percent estimation errors in **Tables 7** and **Figure 5**. The variation of their mean square estimation errors against varying N_{it} is available in **Figure 6**.

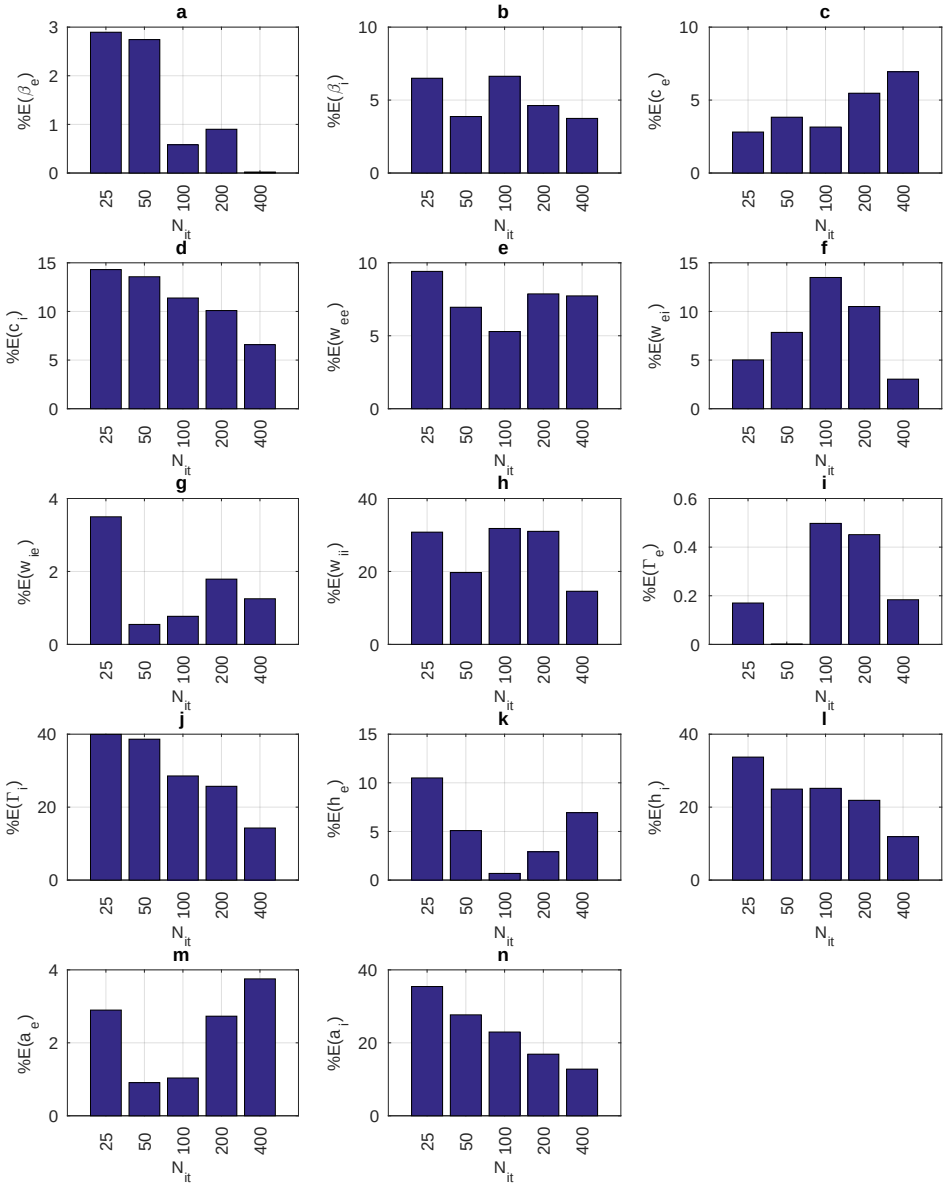

Fig. 5 Variation of the percent errors of estimation (%E) associated with each parameter in (9) against sample size N_{it}. The other conditions are $A_{\max} = 100$ μA, $N_U = 5$ and $f_0 = 3.3333$ Hz. In this figure (a) %E of β_e, (b) %E of β_i, (c) %E of c_e, (d) %E of c_i, (e) %E of w_{ee}, (f) %E of w_{ei}, (g) %E of w_{ie}, (h) %E of w_{ii}, (i) %E of Γ_e, (j) %E of Γ_i, (k) %E of h_e, (l) %E of h_i, (m) %E of a_e and (n) %E of a_i.

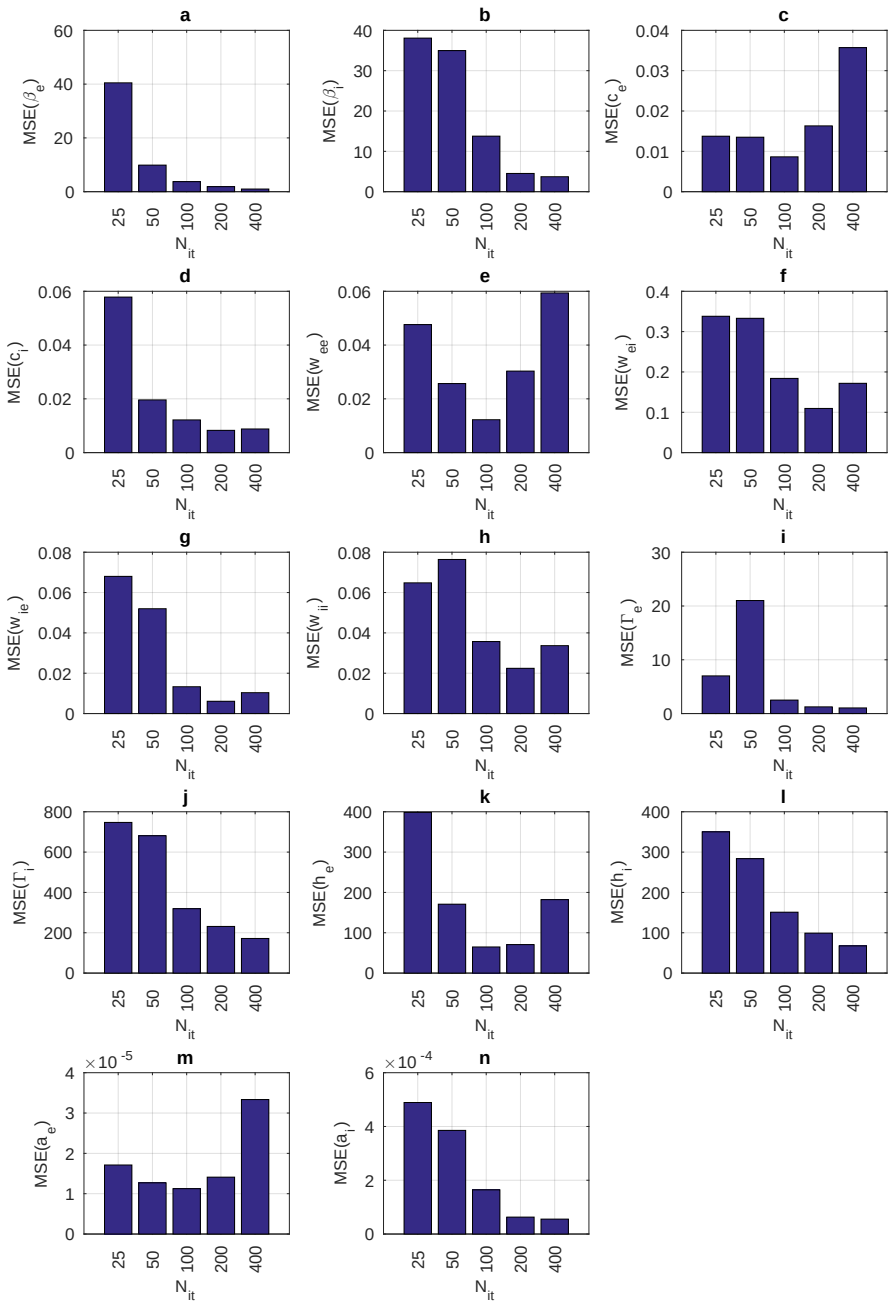

Fig. 6 Variation of the mean square errors of estimation (MSE) associated with each parameter in (9) against sample size N_{it}. The other conditions are $A_{\max} = 100$ μA, $N_U = 5$ and $f_0 = 3.3333$ Hz. In this figure (a) MSE of β_e, (b) MSE of β_i, (c) MSE of c_e, (d) MSE of c_i, (e) MSE of w_{ee}, (f) MSE of w_{ei}, (g) MSE of w_{ie}, (h) MSE of w_{ii}, (i) MSE of Γ_e, (j) MSE of Γ_i, (k) MSE of h_e, (l) MSE of h_i, (m) MSE of a_e and (n) MSE of a_i.

Table 8 The values of the mean estimated parameters $\hat{\theta}$ against increasing stimulus base frequency f_0 in Hz. The other conditions are $A_{\max} = 100$ μA, $N_U = 5$ and $N_{it} = 100$.

θ	$f_0 = 1/3$	$f_0 = 1$	$f_0 = 7/3$	$f_0 = 10/3$	$f_0 = 5$
β_e	52.173	50.712	51.040	49.709	50.320
β_i	24.457	25.128	23.659	23.342	23.971
c_e	0.985	1.001	1.002	1.032	1.002
c_i	0.649	0.688	0.705	0.620	0.695
w_{ee}	1.235	1.233	1.165	1.263	1.238
w_{ei}	1.851	1.901	1.972	1.730	1.913
w_{ie}	0.705	0.703	0.709	0.695	0.626
w_{ii}	0.408	0.352	0.284	0.527	0.436
Γ_e	100.028	100.090	99.783	99.502	99.769
Γ_i	57.370	55.332	49.659	64.276	58.040
h_e	72.101	71.536	72.096	70.481	69.871
h_i	32.657	34.530	34.311	26.194	32.624
a_e	0.040	0.040	0.040	0.040	0.040
a_i	0.048	0.045	0.042	0.049	0.048

3.7.4 Variation of the estimates against stimulus base frequency (f_0)

One can see the variation of mean estimated values of the parameters in (9) and their associated percent estimation errors in **Tables 8** and **Figure 7**. The variation of their mean square estimation errors against varying f_0 is available in **Figure 8**.

4 Discussion & Conclusion

4.1 General Discussion

In this paper, we have performed a theoretical (or simulation based) study the stimulus-response relationship of sensory neurons which generate a discontinuous noisy data. Knowing the fact that the response data obeys an inhomogeneous Poisson process, one will be able to fit a model through a point process maximum likelihood estimation (derived from local Bernoulli approximation). The stimulus is modeled as a phased cosine Fourier series of which fundamental amplitude and frequency are assigned as constant parameters but phases are shot randomly.

The simulations are repeated several times with different amplitude $A_{\max}$, frequency f_0, stimulus component size N_U and sample sizes N_{it} to examine the influence of their variation on estimation performance (percent and mean square estimation errors against true parameter values in **Table 2**).

This study has a similar framework to that of [10] except the fact that, we attempt to estimate the parameters of the sigmoidal gain functions ($\Gamma_e, \Gamma_i, a_e, a_i, h_e, h_i$) together with the reciprocal time constants (β_e, β_i) and network weights ($c_e, c_i, w_{ee}, w_{ei}, w_{ie}, w_{ii}$). Inclusion of more parameters to the estimation procedure is expected to improve the universal approximation capability of the CTRNN in consideration. However, that will bring an extra computational complexity and issues such as parameter confounding.

Here, we will first specifically evaluate the findings od this research. After its completion, we will attempt to compare the results with that of [10] and similar previous studies in the literature.

4.2 Evaluation of the Results

4.2.1 Estimation performance against changing stimulus component size N_U

Increasing, the stimulus component size N_U do not yield an overall improvement of the estimation performance. This result can be noted from the variation of the percent and mean square estimation errors (**Figures 1** and **2**). For some parameters such as c_i percent estimation errors decrease with increasing stimulus component size. Contrary to that, we have a reversed behavior associated with the parameter w_{ie}. For almost all parameters mean square errors of estimation do not display a considerable variation. Considering the computational complexity of the study, it is reasonable to keep N_U in the range 5-20.

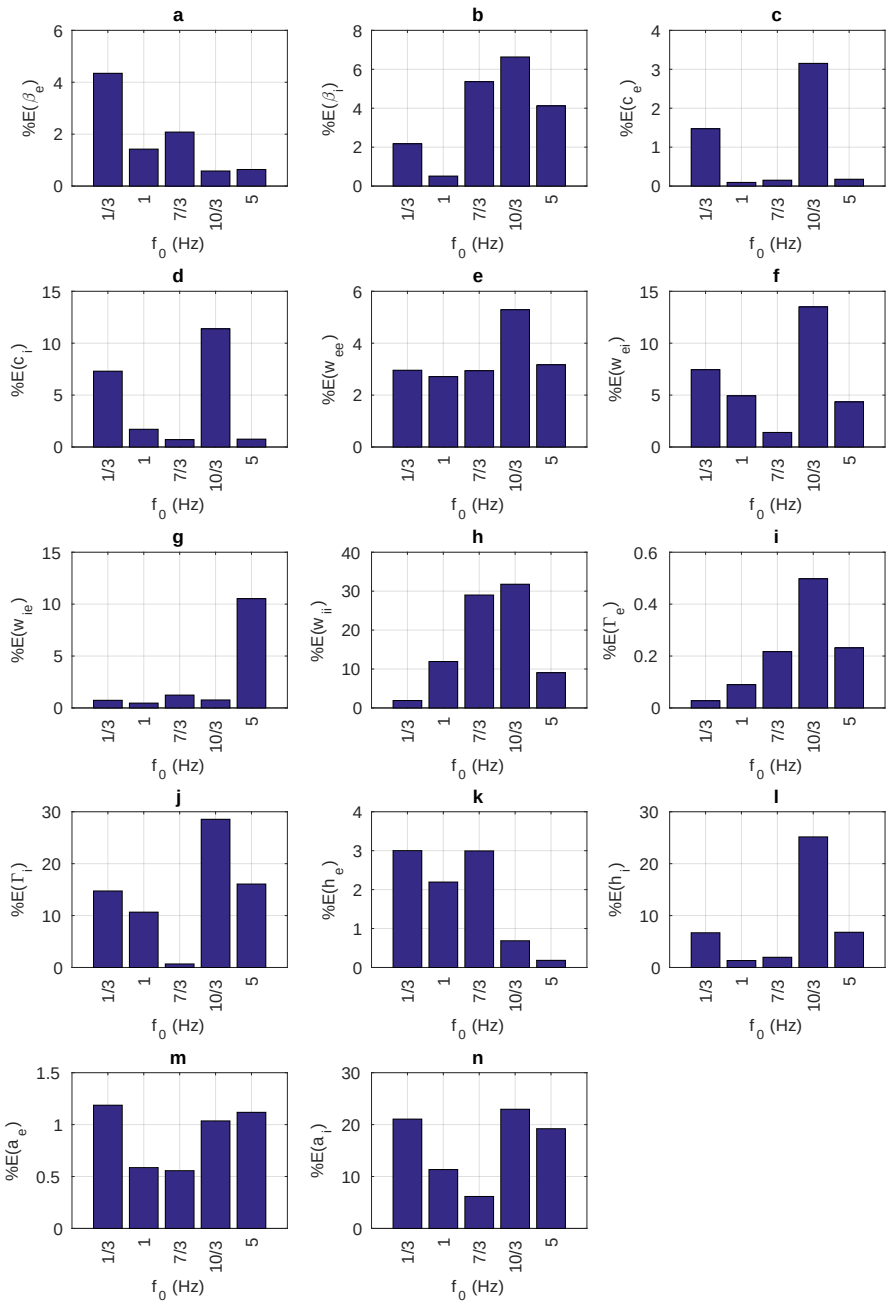

Fig. 7 Variation of the percent errors of estimation (%E) associated with each parameter in (9) against stimulus base frequency f_0 (in Hz). The other conditions are $A_{\max} = 100$ μA, $N_U = 5$ and $N_{it} = 100$. In this figure (a) %E of β_e, (b) %E of β_i, (c) %E of c_e, (d) %E of c_i, (e) %E of w_{ee}, (f) %E of w_{ei}, (g) %E of w_{ie}, (h) %E of w_{ii}, (i) %E of Γ_e, (j) %E of Γ_i, (k) %E of h_e, (l) %E of h_i, (m) %E of a_e and (n) %E of a_i.

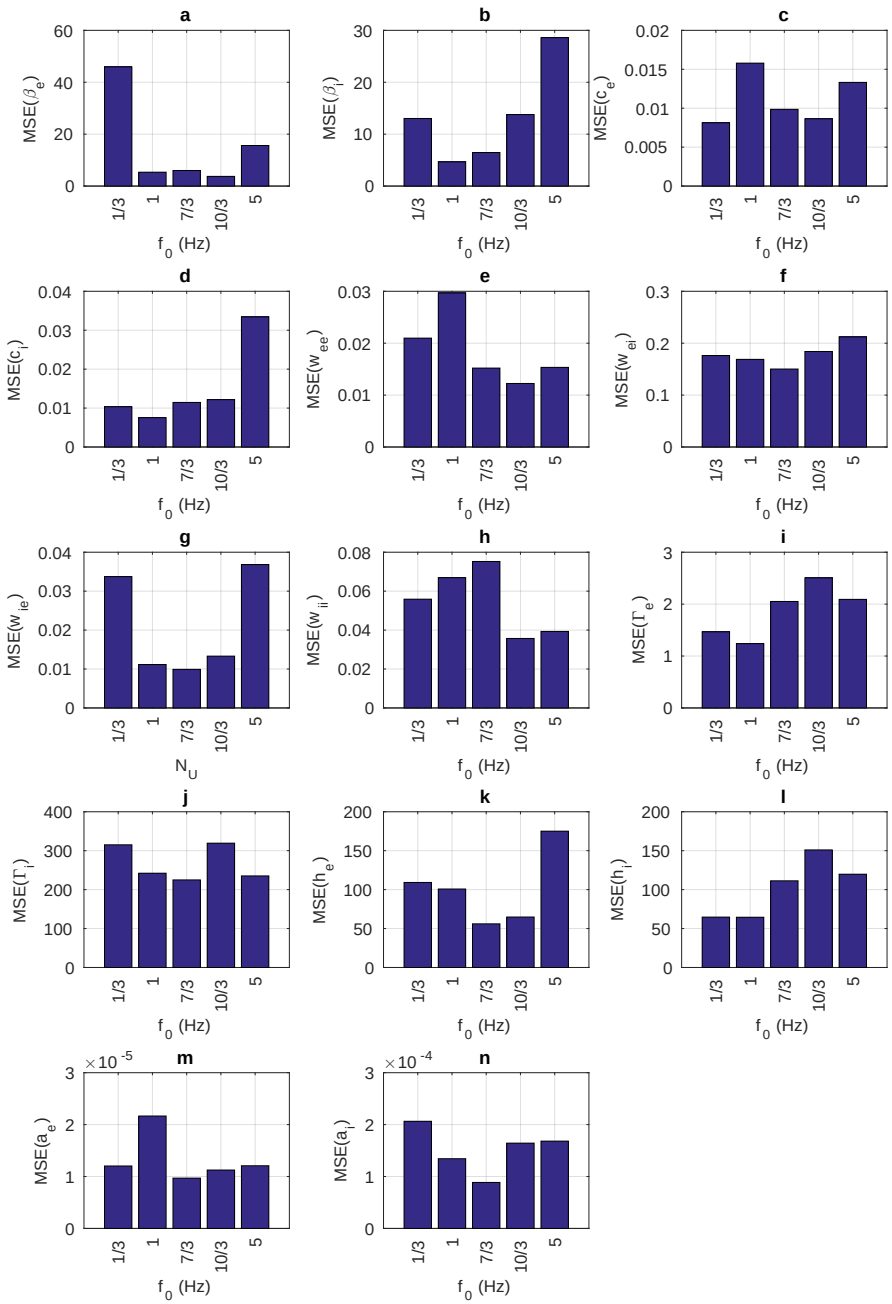

Fig. 8 Variation of the mean square errors of estimation (MSE) associated with each parameter in (9) against stimulus base frequency f_0 (in Hz). The other conditions are $A_{\max} = 100$ μA, $N_U = 5$ and $N_{it} = 100$. In this figure (a) MSE of β_e, (b) MSE of β_i, (c) MSE of c_e, (d) MSE of c_i, (e) MSE of w_{ee}, (f) MSE of w_{ei}, (g) MSE of w_{ie}, (h) MSE of w_{ii}, (i) MSE of Γ_e, (j) MSE of Γ_i, (k) MSE of h_e, (l) MSE of h_i, (m) MSE of a_e and (n) MSE of a_i.

4.2.2 Estimation performance against changing stimulus amplitude parameter $A_{\max}$

Concerning the stimulus amplitude one can make the following comments. **Figure 4** revealed that, there is a general improvement in the mean square errors of estimation when the stimulus amplitude parameter $A_{\max}$ is kept in the range 25-100 μA. Variation of the percent estimation errors (**Figure 3**) display a mixed behavior. For example percent error associated with the parameter w_{ee} shows an improvement whereas that of h_i degrades with increasing amplitude. Compilation of the two findings related to the errors suggests that, the stimulus amplitude parameter $A_{\max}$ should be kept in the range 25-50 μA.

4.2.3 Estimation performance against changing sample size N_{it}

An interesting result of this work is related to the sample size N_{it}. Theoretically speaking, when sample size increases the estimation errors are expected to decrease. A similar result is also noted in [10]. However, **Figures 5** and **6** shows that increasing sample size does not necessarily improve the estimation performance. Concerning the mean square errors, except for w_{ii} and Γ_e increasing the sample size in the range 25 - 100 improves the estimation (**Figure 6**) but not when this range is exceeded. Same comments can not be made for the percent estimation errors (**Figure 5**). The above situation may be linked to parameter confounding issue. That is an undesired phenomenon in statistics that can render the experimental data useless and ruin its results. [4,19] states that a parameter will be called a confounder (also called as confounding or lurking parameter) if it influences both an independent and a dependent variable. This will cause a spurious association between them. The major peculiarities related to a parameter confounding phenomenon will be increased variance and bias.

Inclusion of sigmoidal gain ((4)) parameters in the estimation procedure increases the level of parameter confounding. The problem in [10] also has a level of confounding but it is not as severe as in the case of this research. A related discussion on parameter confounding can be found in [9]. In that, the discussed model is same as that of [10] and the current research. Shortly, if we expand the second order CTRNN model in (3) as shown below:

$$\begin{aligned}\dot{V}_e &= -\beta_e V_e + \frac{\Gamma_e \beta_e w_{ee}}{1+\exp(-a_e(V_e-h_e))} - \frac{\Gamma_i \beta_e w_{ei}}{1+\exp(-a_i(V_i-h_i))} + \beta_e c_e I \\ \dot{V}_i &= -\beta_i V_i + \frac{\Gamma_e \beta_i w_{ie}}{1+\exp(-a_e(V_e-h_e))} - \frac{\Gamma_i \beta_i w_{ii}}{1+\exp(-a_i(V_i-h_i))} + \beta_i c_i I\end{aligned} \tag{16}$$

One can see that the triplet $[\beta_e, w_{ee}, \Gamma_e]$ and $[\beta_i, w_{ei}, \Gamma_i]$ has a similar effect on the dyanmics of V_e. Same situation appears for variable V_i due to the parameter triplets $[\beta_i, w_{ie}, \Gamma_e]$ and $[\beta_i, w_{ii}, \Gamma_i]$. In addition to those the threshold and slope parameters (h_e, a_e, h_i, a_i) have a lower but similar influence. These correlations should be behind the existing confounding issue.

4.2.4 Estimation performance against stimulus base frequency f_0

Concerning the variation of estimation performance against changing stimulus base frequency f_0, one should refer to **Figures 7** and **8**. The results do not seem to reveal a definite variation profile. Specifically speaking, to have comparably lower percent estimation errors for most parameters the base frequency f_0 should be relatively low ($1/3 - 7/3$) Hz. Concerning the mean square error, same discussion can be made for midrange frequencies ($7/3 - 10/3$) Hz. So it is reasonable to chose the frequency at $f_0 = 7/3$ Hz.

4.3 Comparison to other studies

It would be convenient to compare the findings of this research to that of [10]. In that, only 8 parameters $\theta_p = [\beta_e, \beta_i, c_e, c_i, w_{ee}, w_{ei}, w_{ie}, w_{ii}]$ are estimated. So we will need to compare the estimation performance related to these parameters.

4.3.1 Changing N_U

When the simulation conditions are set as $f_0 = 3.3333$ Hz, $N_{it} = 100$, $A_{\max} = 100$ μA and N_U is varied in the range $[5, 10, 20, 30, 40, 50]$, the mean square errors associated with the parameters θ_p stays smaller in [10]. Also, like in this work variation of N_U does not lead to a distinct profile in [10].

4.3.2 Changing $A_{\max}$

Concerning the stimulus amplitude $A_{\max}$, a comparison study will reveal that overall mean square errors associated with θ_p stays smaller in [10]. Conditions are $f_0 = 3.3333$ Hz, $N_{it} = 100$ and $N_U = 5$. In both cases $A_{\max}$ varied in the range $[25, 50, 100, 200, 400]$ and the amplitudes in the mid-range ($A_{\max}$ in the range 100-200 μA) seemed to provide smallest mean square error among all.

4.3.3 Changing N_{it}

In [10], increasing the number of samples leads to a decrease in the mean square errors. In this study, this behavior is not seen for all parameters (see **Section 4.2.3**). In that section, possible reasons are also discussed. Overall mean square errors are smaller in [10]. In both cases the conditions are same and they are $f_0 = 3.3333$ Hz, $A_{\max} = 100$ and $N_U = 5$.

4.3.4 Changing f_0

Concerning the base frequency of stimulus f_0 [10] shows a decreasing behavior in the mean square errors among a large range of frequencies ($1/3 \leq f_0 \leq 10/3$ Hz). In this research, a similar behavior can be seen for a few parameters such as β_e however for most parameters this behavior is not seen. Looking at **Figure 8**, it is reasonable to stay in the range $1 \leq f_0 \leq 10/3$ Hz. The conditions are same for both studies i.e. $N_{it} = 100$, $N_U = 5$ and $A_{\max} = 100$.

4.3.5 General comparison to previous research

It will be beneficial to compare the results of this study to a few different studies in the neuroscience literature. Examples are [27,30,28]. As done in this paper, those studies worked on estimation of neural model parameters by maximum likelihood methods. [27] estimates two models, one with 23 parameters and one with 4 parameters. In both attempts, some mean estimates appeared to deviate as much as 25% from their true parameters values. The error levels vary from parameter to parameter and lies in the range $[0.3\%, 24.7\%]$. However, most of the parameters have error levels larger than 10%. In [30], the error levels seems to be improved and they lie in the range $[0.3\%, 5\%]$. However, this model has fewer parameters (only three) and thus it might be a trivial result. In [28], some time dependent variables are being estimated using likelihood methods. Based on the results obtained in the mentioned research, one can say that the error levels vary with the region of the signals in consideration. Although the trials are repeated 120 times to perform model fitting, the percent estimation error stays around 30% or larger.

Based on the results of this research and the compiled ones above, we can deduce that:

1. The level of efficiency of the estimation algorithm is definitely acceptable. The estimation error levels are comparable to that of various researches performed before.
2. Large estimation percent errors appear only for few parameters in this research, whereas works such as [27] appear to have larger percent estimation errors for many parameters. Moreover, **Figure 7** suggests that a choice of frequency like $f_0 = 1$ Hz yields lower percent estimation errors than that of those studies.

Parameter confounding is likely the reason behind the larger mean square errors in this work.

Compliance with Ethical Standards

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Beer, R.D.: On the dynamics of small continuous-time recurrent neural networks. Adaptive Behavior **3**(4), 469–509 (1995)
2. Brillinger, D.R.: Maximum likelihood analysis of spike trains of interacting nerve cells. Biological cybernetics **59**(3), 189–200 (1988)

3. Brown, E.N., Barbieri, R., Ventura, V., Kass, R.E., Frank, L.M.: The time-rescaling theorem and its application to neural spike train data analysis. Neural computation **14**(2), 325–346 (2002)
4. Campbell, M.J., Machin, D., Walters, S.J.: Medical statistics: a textbook for the health sciences. John Wiley & Sons (2010)
5. Chornoboy, E., Schramm, L., Karr, A.: Maximum likelihood identification of neural point process systems. Biological cybernetics **59**(4), 265–275 (1988)
6. DiMattina, C., Zhang, K.: Active data collection for efficient estimation and comparison of nonlinear neural models. Neural computation **23**(9), 2242–2288 (2011)
7. DiMattina, C., Zhang, K.: Adaptive stimulus optimization for sensory systems neuroscience. Frontiers in neural circuits **7** (2013)
8. Doruk, O., Zhang, K.: Building a dynamical network model from neural spiking data: Application of poisson likelihood. arXiv preprint arXiv:1710.03071 (2017)
9. Doruk, R.O., Zhang, K.: Adaptive stimulus design for dynamic recurrent neural network models. arXiv preprint arXiv:1610.05561 (2016)
10. Doruk, R.O., Zhang, K.: Fitting of dynamic recurrent neural network models to sensory stimulus-response data. arXiv preprint arXiv:1709.09541, (Accepted by Journal of Biological Physics on 26 Mar 2018) (2017)
11. Eden, U.T.: Point process models for neural spike trains. Neural Signal Processing: Quantitative Analysis of Neural Activity pp. 45–51 (2008)
12. FitzHugh, R.: Impulses and physiological states in theoretical models of nerve membrane. Biophysical journal **1**(6), 445–466 (1961)
13. Gerstner, W., Kreiter, A.K., Markram, H., Herz, A.V.: Neural codes: firing rates and beyond. Proceedings of the National Academy of Sciences **94**(24), 12740–12741 (1997)
14. Herz, A.V., Gollisch, T., Machens, C.K., Jaeger, D.: Modeling single-neuron dynamics and computations: a balance of detail and abstraction. Science **314**(5796), 80–85 (2006)
15. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of physiology **117**(4), 500 (1952)
16. Koyama, S., Shinomoto, S.: Histogram bin width selection for time-dependent poisson processes. Journal of Physics A: Mathematical and General **37**(29), 7255 (2004)
17. Ledoux, E., Brunel, N.: Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. Frontiers in computational neuroscience **5** (2011)
18. Lewis, P.A., Shedler, G.S.: Simulation of nonhomogeneous poisson processes by thinning. Naval Research Logistics Quarterly **26**(3), 403–413 (1979)
19. McDonald, J.H.: Small numbers in chi-square and g-tests. Handbook of biological statistics pp. 86–89 (2014)
20. Miller, K.D., Fumarola, F.: Mathematical equivalence of two common forms of firing rate models of neural networks. Neural computation **24**(1), 25–31 (2012)
21. Morris, C., Lecar, H.: Voltage oscillations in the barnacle giant muscle fiber. Biophysical journal **35**(1), 193–213 (1981)
22. Nawrot, M., Aertsen, A., Rotter, S.: Single-trial estimation of neuronal firing rates: from single-neuron spike trains to population activity. Journal of neuroscience methods **94**(1), 81–92 (1999)
23. Paninski, L.: Maximum likelihood estimation of cascade point-process neural encoding models. Network: Computation in Neural Systems **15**(4), 243–262 (2004)
24. Shadlen, M.N., Newsome, W.T.: Noise, neural codes and cortical organization. Current opinion in neurobiology **4**(4), 569–579 (1994)
25. Shimazaki, H., Shinomoto, S.: A method for selecting the bin size of a time histogram. Neural Computation **19**(6), 1503–1527 (2007)
26. Shimazaki, H., Shinomoto, S.: Kernel bandwidth optimization in spike rate estimation. Journal of computational neuroscience **29**(1-2), 171–182 (2010)
27. Smith, A.C., Brown, E.N.: Estimating a state-space model from point process observations. Neural Computation **15**(5), 965–991 (2003)
28. Truccolo, W., Eden, U.T., Fellows, M.R., Donoghue, J.P., Brown, E.N.: A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. Journal of neurophysiology **93**(2), 1074–1089 (2005)
29. Wu, M.C.K., David, S.V., Gallant, J.L.: Complete functional characterization of sensory neurons by system identification. Annu. Rev. Neurosci. **29**, 477–505 (2006)
30. Zhou, G.T., Schafer, W.R., Schafer, R.W.: A three-state biological point process model and its parameter estimation. IEEE Transactions on Signal Processing **46**(10), 2698–2707 (1998)