Identification of a stochastic resonate-and-fire neuronal model via nonlinear least squares and maximum likelihood estimation

Jun Chen*
Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida 32817, USA
Email: jchenee2015@gmail.com
*Corresponding author

Peter Molnar
Department of Zoology, Savaria Campus, University of West Hungary, Karolyi Gaspar ter 4, Szombathely, H-9700, Hungary
Email: pmolnar@pminfonet.com

Aman Behal
Department of Electrical Engineering and Computer Science and NanoScience Technology Center, University of Central Florida, Orlando, Florida 32817, USA
Email: abehal@ucf.edu

Abstract: Recent work has shown that the resonate-and-fire neuronal model is both computationally efficient and suitable for large network simulations. In this paper, we examine the estimation problem of a resonate-and-fire neuronal model with stochastic firing threshold. The model parameters are divided into two sets. The first set is associated with the subthreshold behaviour and can be estimated by a least squares algorithm, while the second set includes parameters associated with the firing threshold and its identification is formulated as a maximum likelihood estimation problem. The latter is in turn solved by a simulated annealing approach that avoids local optima. The proposed identification approach is evaluated using both simulated and in-vitro data, which shows a good match between prediction by identified model and the actual data, concluding the efficiency and accuracy of the proposed approach.

Keywords: resonate-and-fire; neuronal model; stochastic threshold; parameter estimation; maximum likelihood; simulated annealing.


Biographical notes: Jun Chen received his BS in Automation from Zhejiang University, China in 2009 and PhD in Electrical Engineering from Iowa State University, USA in 2014. His research interests include systems and control, including modelling and identification, failure diagnosis and prognosis, discrete event and hybrid systems. He received Best Paper Award from IEEE Transactions on Automation Science and Engineering, Research Excellence Award from Iowa State University, and Outstanding Student Award from Zhejiang University.

Peter Molnar is an Electrophysiologist with extensive experience in cell culture and cell patterning methods and in the development of in vitro systems for drug screening applications. He worked for pharmaceutical companies for several years and led preclinical drug development projects. As an Assistant Professor at UCF, he has established an electrophysiological laboratory and worked on externally funded projects as PI and co-PI. At University of West Hungary he established an in vitro electrophysiological laboratory and did extensive teaching. He has published intensively concerning interaction of cardiac myocytes and neurons with surfaces, effect of surface modifications on physiology of cultured cells, cell patterning and in vitro toxin detection using cardiac myocytes and neurons as whole-cell biosensors.
1 Introduction

A fundamental issue in computational neuroscience is to characterise the relationship between the neural output voltage and the input current to the cell (Paninski et al., 2004). During the last decade, spiking neuron model has been extensively researched. While detailed Hodgkin-Huxley type neuron models can accurately reproduce the behaviour of most neurons, it requires the tuning of hundreds of parameters to fit experimental data and hence is computationally inefficient. Furthermore, such detailed dynamics restrict neuronal network simulations to only a handful of neurons at a time (Izhikevich et al., 2003).

For computational efficiency, a simple integrate-and-fire (IF) model (Burkitt, 2006) is widely used for large network simulations. However, since the IF model only has one variable, it cannot produce many types of firing patterns [e.g., bursting (Izhikevich, 2001)]. In Izhikevich et al. (2003), a quadratic model with an additional variable is employed to capture neuronal adaptation by accounting for the activation of $K^+$ ionic currents and inactivation of $Na^+$ ionic currents. The ability of this type of neuronal model with a recovery variable to qualitatively reproduce major types of firing pattern is shown in Izhikevich (2001) and Naud et al. (2008). Furthermore, the computational efficiency of such a spiking model is demonstrated in Izhikevich (2007).

In order to estimate parameters for a particular model from experimentally recorded voltage traces, many parameter estimation methods have been introduced. In Brette and Gerstner (2005), an adaptive exponential integrate-and-fire neuron model is manually tuned to fit a detailed Hodgkin-Huxley-based model. The model in Brette and Gerstner (2005) shows good results, but it is labor intensive; furthermore, its trial-and-error approach depends mainly on the researcher’s experience. Therefore, automated parameter search methods become necessary because it is unrealistic for one to process all the data comparisons by manual procedures (LeMasson and Maex, 2001; Vanier and Bower, 1999).

In Prinz et al. (2003), a database of single-compartment neuron model is constructed by exploring the entire parameter space, which is only practical when said space has low dimensions. In Weaver and Wearne (2006), a compartment model is enhanced by using simulated annealing for parameter estimation. In Lansky et al. (2006), Paninski et al. (2005) and Zhang and Feng (2002), various estimation formulations are presented for deterministic and stochastic IF models. In Zhang et al. (2009), an expression for the probability distribution of the interspike interval of a leaky integrate-and-fire (LIF) model is derived, and maximum likelihood estimates (MLE) of the input information are developed for an LIF neuron from a set of recorded spike trains. In Jabalameli and Behal (2015), the identification problem is studied for a multi-timescale adaptive threshold neuronal model introduced in Kobayashi et al. (2009). In our previous work (Zhi et al., 2012), a framework for a linear-in-the-parameters representation of the aforementioned quadratic model of Izhikevich et al. (2003) is developed, which allows the identification of system parameters via persistent excitation and a weighted least squares estimation approach.

An assumption made in the model of Izhikevich et al. (2003) prevents the approach in Zhi et al. (2012) from being applicable to arbitrary experimental data. While the quadratic model is known to be able to reproduce biologically meaningful firing patterns, however, it can only reproduce these patterns qualitatively, not quantitatively – specifically, the model cannot quantitatively reproduce the upstroke/downstroke of the spike unless the parameters are assumed to be voltage-dependent (Izhikevich, 2007). Since we are interested only in reproducing a spiking pattern and not the shape of the spike itself, the spike shape-related quadratic term is dropped so that the resulting ‘resonate-and-fire’ model only described the sub-threshold dynamics of the system. This allows to narrow down the interest zone to the now-linear sub-threshold region, thus making the upstroke/downstroke irrelevant to the identification of the system. Furthermore, the resulting linearity of the sub-threshold model allows for it to be treated analytically, which is useful for deriving closed form solutions and computing the update law. However, an additional challenge introduced is the
Identification of a stochastic resonate-and-fire neuronal model

2 System model

In this paper, we study a resonate-and-fire neuronal spiking model, which can be described by the state equations

\[
\frac{dv}{dt} = k_1 v + k_2 - k_3 u + k_3 i \tag{1}
\]

\[
\frac{du}{dt} = a(bu - v) \tag{2}
\]

where \(v\) denotes the membrane potential and is the only system output, \(u\) is an immeasurable membrane recovery state variable which provides a negative feedback to \(v\), while \(i\) denotes injected current and/or synaptic current. The parameters \(k_3\) and \(a\) denote the time scale of the two state variables, the parameter \(b\) is the level of sub-threshold adaptation, while the parameters \(k_1\) and \(k_2\) are linked to the sub-threshold behaviour of the neuron. When the membrane potential \(v\) hits a threshold, denoted as \(V_t\), the neuron is said to fire a spike, and the state variables are reset according to

\[
\begin{align*}
\text{if } v &= V_t, \text{ then } \\
&\begin{cases}
  v \rightarrow c \\
  u \rightarrow u + d
\end{cases}
\end{align*} \tag{3}
\]

where \(V_t\) is the threshold, \(c\) denotes the post-spike reset value of the membrane potential, while \(d\) denotes the amount of spike adaptation of the recovery variable.

Remark 1: Although it is formulated differently here, the model presented here is the same as the one discussed in Young (1937) or FitzHugh (1966). Note that (1) and (2) do not model the post threshold dynamics of the cell; these equations only describe the sub-threshold behaviour and the spikes are modelled as spiking train using (3). Unlike in Jolivet et al. (2004), where a triangular pulse is used to mimic the shape of the spike, we ignore the entire spiking behaviour, using only a straight line to indicate the occurrence of a spike.

Remark 2: Note that equation (3) introduces discontinuity into the system modelled as (1), (2) and (3). Such discontinuity makes traditional identification techniques (e.g., least squares) not applicable to the identify parameters in all system equations (1), (2) and (3). Therefore, in this paper, we proposed a two-stage approach that separately identifies parameters in (1) to (2) and (3). See next section for more details.

Note that the model above only describes the sub-threshold behaviour, and a voltage threshold needs to be defined to indicate the initiation of a spike. In Izhikevich (2001), a fixed threshold is used. However, empirical data suggests that the voltage threshold for a spike depends not only on the instantaneous value of the voltage, but also on the rate of voltage change. The variation in the spike threshold could also be a function of the instantaneous firing rate (Azouz and Gray, 2000). Using an ion-channel-based detailed neuronal model, the probability of firing as a function of voltage was used to identify a specific value that could be considered as a fixed threshold, but no such value could be clearly identified. Figure 1 shows the probability of firing as a function of voltage, showing no obvious ‘jump point’ in firing probability that can be defined as a ‘threshold’. The data used in Figure 1 was generated from a detailed ion-channel-based spiking model of McCormick et al. (1993) with parameters from Destexhe et al. (1998). Therefore, Figure 1 suggests that the threshold to have a stochastic distribution rather than a deterministic value. Stochastic spike thresholds have been previously proposed and proved in Bershadskii et al. (2003) and White et al. (1998). Others have used a stochastic term in the membrane dynamics to represent signal noise, e.g., see...
Lansky et al. (2006). In fact this is computationally identical to subtracting the noise from the potential and adding it to the threshold (Izhikevich et al., 2003). In this paper, the threshold is assumed to be a random variable with i.i.d. Gaussian distribution, i.e., \( V_t \sim N(\mu, \sigma) \).

Figure 1  Firing probability v.s. membrane voltage (see online version for colours)

Notes: For each voltage value, if there are \( N \) points at this value in the dataset and \( n \) of them directly followed by a spike, the firing probability is defined as \( n/N \). 

Source: Data was generated from a detailed ion-channel-based spiking model of McCormick et al. (1993) with parameters from Destexhe et al. (1998)

3 The estimation problem

Our goal is to estimate the following set of unknown parameters associated with the resonate-and-fire neuron model:

\[
\theta \triangleq (k_1, k_2, k_3, a, b, c, d, m, \sigma) = (\theta_l, \theta_i)
\]

where the partitions \( \theta_l \) and \( \theta_i \) will be explicitly defined below. It is assumed that the input excitation and membrane potential recordings are available for measurement. Given the model of (1) to (3), the system parameters to be estimated are divided into two sets; specifically, a set \( \theta_l \triangleq (k_1, k_2, k_3, a, b) \), which is associated with the linear dynamics of (1) and (2), and another set \( \theta_i \triangleq (c, d, m, \sigma) \), which is linked to the after-spike resetting and threshold distribution parameters. In this section, we propose a two stage approach to estimate both \( \theta_l \) and \( \theta_i \). In the first stage, the measurements of sub-threshold voltage are used to estimate \( \theta_i \), while in the second stage, the estimated \( \theta_i \) together with the measured interspike intervals (ISIs, computed from measured voltage recording) are utilized to identify \( \theta_l \).

3.1 Stage I: Sub-threshold estimation

The linear system represented by (1) and (2) can be analytically treated and solved, which is particularly useful when using gradient-based or least squares-based estimation techniques, thanks to the availability of the derivative of the objective function with respect to the parameters. Based on (1) and (2), the solution for the sub-threshold membrane potential can be expressed as

\[
v(t) = [k_2 - k_3 u(t_0)] f_1(\theta_l, t) + k_3 (f_2(\theta_i, t) \ast \delta) + v(t_0) f_2(\theta_i, t) + k_2 \sigma f_3(\theta_i, t)
\]

where \( u(t_0) \) and \( v(t_0) \) are initial values for \( u \) and \( v \), the symbol \( \ast \) denotes the convolution operator on two functions, while the functions \( f_1(\theta_l, t) \), \( f_2(\theta_i, t) \), and \( f_3(\theta_i, t) \) in the solution are explicitly given in Appendix A. Because of the dependence of \( v(t) \) on \( u(t_0) \) and \( v(t_0) \) as can be seen in (4), \( v(t) \) depends not only on the parameter set \( \theta_i \) but also on the reset values \( c \) and \( d \). However, at this stage, the effect of \( c \) and \( d \) can be set aside by utilising a spike-free continuous subsequence of the measured data lying beyond an initial transient period. Functionally, this is equivalent to a model with initial conditions \( u(t_0) = v(t_0) = 0 \) in (4), resulting in voltage dynamics that are independent of \( \theta_i \) in the sub-threshold region, depending only on \( \theta_l \) such that \( v(t) \) can be compactly expressed as follows

\[
v = f(t, i; \theta_l).
\]

Given input \( i, f \) is a nonlinear function on the parameter set \( \theta_l \). The estimate for \( \theta_l \) is made via a nonlinear least squares estimation-based technique (Kelley, 1999). Nonlinear least squares is a form of least squares used to fit observations to a model that is nonlinear in the parameters. Given a discrete observed data sequence of length \( N \), the objective function to be minimised is

\[
S = \sum_{i=1}^{N} (v_i - \hat{v}_i)^2
\]

where \( \hat{v}_i \triangleq f(i; t; \hat{\theta}_l) \) denotes the estimated voltage at time \( i \) and \( v_i \) is the measured voltage at time \( i \). The parameter update law is given by

\[
\begin{align*}
\hat{\theta}_l^{i+1} &= \hat{\theta}_l^i + \Delta \hat{\theta}_l \\
\Delta \hat{\theta}_l &= (J^\prime J)^{-1} J^\prime \Delta v
\end{align*}
\]

where \( \Delta v = v - \hat{v} \) is a \( n \times 1 \) vector defining the error between the measured and the estimated voltage while \( J \) is the \( N \times M \) Jacobian matrix, with elements given by

\[
J_{ik} = \frac{\partial \hat{v}_i}{\partial \theta_{ik}}.
\]

Here, \( M = 5 \), corresponds to the number of parameters in the set \( \theta_l \). Furthermore, \( \hat{\theta}_l \) denotes the estimate of the \( k^{th} \) element of the parameter vector \( \theta_l \). Note that the Jacobian matrix \( J \) is a function of the estimated parameters and! updates from one iteration to the next. The convergence of the parameter estimates to their actual values is ensured by utilising a persistently exciting input current injection as we have demonstrated in our earlier work in Zhi et al. (2012).
3.2 Stage II: Estimation of reset and threshold distribution parameters

During this stage, the sub-threshold parameters $\hat{\theta}_i$ estimated from the previous step are utilised in the reconstruction of the membrane potential – specifically, we can write
\[
\hat{v} = g(t, I, \hat{\theta}_i, c, d).
\]
(8)

Given $i$ and $\hat{\theta}_i$, it can be seen from (4) that $\hat{v}$ is a linear function of the unknown reset parameters $c$ and $d$. We remark that in Stage II, the reset parameters (i.e., the initial conditions) cannot be ignored because these parameters strongly determine the rate of occurrence of spikes. The insight here is that the reconstructed membrane potential is allowed to evolve according to (8) and the profile likelihood function can be expressed as follows

\[
L(\theta_i) = \log \left( \prod_i \int_{v_i}^{\infty} G(V_i; m, \sigma) dV_i \right) \times \prod_j \int_{-\infty}^{s_j} G(V_i; m, \sigma) dV_i \\
= \sum_i \log \int_{v_i}^{\infty} G(V_i; m, \sigma) dV_i \\
+ \sum_j \log \int_{-\infty}^{s_j} G(V_i; m, \sigma) dV_i
\]

where $G(V_i; m, \sigma)$ is the Gaussian probability density function of the threshold $V_i$ with mean $m$ and variance $\sigma^2$. $L(\theta_i)$ could be referred to as ‘log-likelihood’ as a more appropriate term, but for simplicity, it is referred to simply as ‘likelihood’. Based on the likelihood function (9), the maximum likelihood estimation (MLE) parameters are given by
\[
\hat{\theta}_i = \arg \max_{\theta_i} L(\theta_i).
\]
(10)

Before choosing a technique to solve (10), it is prudent to first study the uniqueness of the maximum of (9). Since the likelihood function is the log multiplication of a series of basic functions, a sufficient condition for the log concavity of the overall function of (9) is for each of the underlying functions to be log concave. If this condition is met, then the likelihood function has a unique maximum. In Appendix B, we show that this sufficient condition is not satisfied. Therefore, being unable to rule out the possibility that the likelihood function has more than one maximum, it is imprudent to utilise a gradient-based method, which is very likely to get stuck at a local maximum. Instead, in the next section, we explore the simulated annealing method.

3.3 Simulated annealing

Unlike gradient-based law or least squares which might get stuck at a local maximum, simulated annealing (SA) is designed for global optimisation (Kirkpatrick et al., 1983; Corana et al., 1987). Thus, we use the simulated annealing technique to find the maximum of the previously formulated likelihood function of (9). This particular technique has its origin in the metallurgic industry where an annealing technique that involves the controlled heating and cooling of a material is used to minimise the energy of its crystals. After selecting an initial point, the algorithm randomly selects a point $s'$ in the neighborhood of the old point $s$ at each iteration. The newly selected point is considered or rejected as being a ‘better’ point depending on the probability function $P$

\[
P = \begin{cases} 
1 & \text{if } L(s') > L(s) \\
e^{-1/T} & \text{otherwise}
\end{cases}
\]

where $T$ is denoted as the ‘temperature’. The possibility to accept the worse point provides the algorithm the capability of getting away from the local maximum. The SA algorithm is allowed to move randomly in the entire parameter space by initialising the ‘temperature’ with a large value $T_0$. Then the parameter gradually decreases. The ‘cooling’ algorithm for the temperature $T$ used here is given by

\[
T = T_0 (1 - n/N)^2
\]

where $n$ is the current iteration, and $N$ is the user-specified maximum number of iterations, after which the algorithm will stop. A detailed description of the algorithm is as follows. Note that the initial point $s_0$ and initial temperature $T_0$ are randomly selected, while the maximum number of iterations $N$ is user-specified.

Setp 0 Initialise the algorithm by selecting a starting point $s_0$, a starting temperature $T_0$, and a maximum number of iterations $N$. Set a boundary for each direction if needed. Then compute the likelihood for the initial point.

Setp 1 A new point $s'$ in the neighborhood of $s$ is generated. Projection applies if the new point exceeds the preset boundary.

Setp 2 The likelihood of the new point $s'$ is computed. The probability function (11) is called, and $s$ is replaced with $s'$ if $P$ is larger than a randomly chosen number between 0 and 1. Otherwise, $s$ remains untouched.
226  J. Chen et al.

Setp 3  The current iteration \( n \) is increased and the temperature \( T \) is updated.

Setp 4  If \( n < N \), go back to Step 2, otherwise stop the search.

4   Data acquisition and processing methods

We applied the aforementioned two-stage estimation strategy to two types of data, namely, simulated data and in-vitro recording from embryonic rat motoneurons. For generating simulated data, an ion-channel-based spiking model of McCormick et al. (1993) with parameters from Destexhe et al. (1998) was used in the NEURON (Hines and Carnevale, 1997) simulation environment\(^1\). Please refer to McCormick et al. (1993), Destexhe et al. (1998), Hines and Carnevale (1997) and Hodgkin and Huxley (1952) for details of this model. Using the current clamp mode of the simulation, we injected pre-generated white noise current to stimulate the cell and collected membrane potential data sampled at 50 kHz.

For the experimental part of the research, primary cultures of embryonic rat motoneurons were prepared according to NIH guidelines and in agreement with the Institutional Animal Care and Use Committee (IACUC) approved protocol. Rat spinal motoneurons were dissected from day 14 (E14) embryos as published earlier (Das et al., 2003). Cells were dissociated with trypsin (Invitrogen, 0.05%) and centrifuged for 15 min at 500 g. Motoneurons were purified with immunopanning using antibodies (antibody 192, 1:2 dilution, ICN Biomedicals, Akron, OH) recognising the low affinity NGF receptor expressed only by ventral motoneurons at this age. Purified motoneurons were plated on 22 mm\(^2\) ornithine/laminin-coated coverslips at a density of 200 cells/mm\(^2\) in Neurobasal (Gibco-BRL) medium supplemented with B27 (2% v/v; Invitrogen), L-glutamine (0.5 mM), 2-mercaptoethanol (25 \( \mu \)M), glial cell line-derived neurotrophic factor (1 ng/ml CNTF; cell sciences) L-glutamate (25 \( \mu \)M) was added to the culture medium during the first five days of growth.

Conventional whole-cell patch clamp recordings were performed on the culture cells between day 7 and 14 in culture. The extracellular solution was Neurobasal culture medium, the pH was adjusted to 7.3 with HEPES. Patch pipets were prepared from borosilicate glass (BF150-86-10; Sutter, Novato, CA) with a Sutter P97 pipet puller and filled with intracellular solution (in mM: K-glucuronate 140, EGTA 1, MgCl\(_2\) 2, Na\(_2\)ATP 2, phosphocreatine 5, phosphocreatine kinase 2.4 mg, Hepes 10; pH 7.2). The resistance of the electrodes was 6–8 M\(\Omega\). Current clamp experiments were performed with the Multiclamp 700A amplifier (Axon, Union City, CA). Signals were digitised at 10 kHz with an Axon Digidata 1322A interface. Data recording and initial analysis were performed with pClamp 10 software (Axon). White noise current was injected in current clamp mode from stimulus files at resting membrane potential. The amplitude of the current signal was adjusted to evoke sub-and-suprathreshold (action potential) responses from the cells. Due to measurement error, there can be a noise contained in the experimental data. To account this, a low-pass filter was applied to remove the high frequency component, before feeding the data into the proposed estimation algorithms. In particular, an 8th order Butterworth low-pass filter with a cut-off frequency of 600 Hz was utilised. The filtered data was then processed to determine the location of spikes in the spike train. A sufficiently long piece of subthreshold data (i.e., a data subsequence without any spikes) was also selected for the first stage (i.e., sub-threshold estimation) of the proposed two-stage estimation strategy.

5   Results and discussion

In this section, target data refers to measurements that are generated either by Hodgkin-Huxley model or experimental data, while predicted data refers to that generated by the model (1), (2) and (3) with estimated parameters.

5.1   Detailed Hodgkin-Huxley model (simulation)

In this section, the results of the proposed method are presented for the simulated data generated from Hodgkin-Huxley model. From the subsequence of data streams shown in Figure 2(a), one can see that the sub-threshold traces of the target and predicted data are very close, showing that dynamics (1) and (2) yield good approximations in the sub-threshold region. Figure 2(b) shows that the predicted spike trains closely follow the target with the same adaptation rate. The proposed method was evaluated ten times. It was seen that the estimated parameters led to 16.3 spikes in average, while the target train had 15 spikes under the same input stimulus, thus, there was an average overprediction of 1.3 spikes. Figure 3 shows the evaluation results under a step input current injection. It can be seen that the prediction closely follows the first three spikes in the target train but misses the fourth one, although the prediction still successfully replicates the spike adaptation pattern encoded in the target train.

5.2   Experimental results

In this section, the results of the proposed method are presented for the experimental data obtained from in-vitro recordings from embryonic rat motoneurons. As in the case of the detailed Hodgkin-Huxley model, the results presented here for evaluation use previously unseen data, i.e., the data stream utilised for model evaluation is different from that utilised for generating the estimated model. Figure 4(a) shows the sub-threshold approximation using the estimated parameters from which it is clear to see that the linear model dynamics of (1) and (2) are a good approximation in the sub-threshold regime. Figure 4(b) shows a subsequence of the target and predicted spike trains, where the prediction misses one spike but locates the remaining spikes correctly. The statistics after ten evaluations of the experimental data revealed that 16 spikes were predicted while the target data only had 14, i.e., the mean error was two spikes.
5.3 Discussion

In our formulation, we assume that the sub-threshold trace could be defined as a ‘linear zone’ where the dynamics (1) and (2) are a good approximation. This sub-threshold region is upperbounded by the variable threshold. Results shown in the last section indicate the validity of this assumption, since the error between the prediction and the target in Figures 2 and 4 is below 5%. Since the linear model is capable of replicating only the sub-threshold dynamics, the reference voltage trace used for the nonlinear least squares estimation should not contain any supra-threshold activities. This is ensured by choosing a sufficiently long data subsequence between spikes, with starting and ending points far away from the nearest spike. By sufficiently long subsequence, it is meant that the data should be rich enough to lead to sub-threshold parameter convergence. Furthermore, by ensuring that the starting point of the data is far away from a spike, one can ignore the transient artifacts related to the initial conditions (i.e., reset) immediately following a spike.

Another assumption made in this formulation is that all the stochastic components in the system, including system noise (either caused by measurement noise or by the environment of the cell) as well as the variability of the threshold, can be captured by the threshold distribution that we have employed. Our results suggest that such a simplification leads to overprediction of the firing rate. However, the results are still acceptable – the errors are small in percentage and the predicted spike trains have remarkably similar spiking patterns as the targets that they are intended to emulate. We believe that a model with more sophisticated assumptions or additional components could address the overprediction issue, yet this would result in a larger computation cost.

Computation of the likelihood function (9) is the major cost of computation in the proposed two-stage strategy. The original likelihood function is defined over all time instants. The simulated data is sampled at 50kHz and the experimental data is sampled at 10kHz. This results in a large size of $|t_i|$ as defined in Section 3.2. In practice, considering the fact that $G(V_t; m, \sigma)$ is a Gaussian probability density function which is practically zero beyond $3\sigma$ distance from the mean $m$, the voltage trace is preprocessed to make $|t_i|$ only contain points that are not too far away from the mean. This requires us to have some a priori knowledge about the parameters before we delve into the estimation. This is deemed acceptable since by looking at the reference train, one can easily draw reasonable bounds for the parameters $m$ and $\sigma$. 
6 Conclusions

In this paper, a systematic data-driven approach was proposed for parameter estimation in stochastic resonate-and-fire neuronal model. Different from literature, the spiking threshold was modelled as a random process, inspired by empirical data. Further, Gaussian distribution was assumed for the random threshold. A two-stage estimation problem was formulated, first stage of which is a nonlinear least squares estimator for identifying parameters related to sub-threshold behaviour utilising measurement data that do not involve any spiking, while the second stage is a maximum likelihood estimator for identifying parameters related to spiking pattern. Both simulated and experimental data (obtained from embryonic rat motoneurons) are used to validate the proposed estimation approach, both showing excellent match between measurement and the prediction of identified model. Future research directions include (1) improved problem formulation that allows analytic optimisation technique for finding optimal parameters, (2) improved metrics to evaluate prediction error that uses the expectation of number of spikes (since spiking is a stochastic process) and (3) implementation of the proposed approach as a generic computational tool.

Acknowledgements

The study was supported by Award # R15NS062402 from the National Institute of Neurological Disorders And Stroke (NINDS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS or the NIH.

References


**Notes**

1 For the interested reader, we note that code for this model is available at [http://senselab.med.yale.edu/senselab/modeldb/ShowModel.asp?model=3817](http://senselab.med.yale.edu/senselab/modeldb/ShowModel.asp?model=3817).
Appendix

Appendix A

The functions \(f_1(\theta_t, t), f_2(\theta_t, t),\) and \(f_3(\theta_t, t)\) in (4) are explicitly defined as follows

\[
f_1(t) = \frac{1}{\sqrt{\Delta}} \left\{ \exp \left( -\frac{(a-k_1)}{2} + \Delta t \right) - \exp \left( -\frac{(a-k_1)}{2} - \Delta t \right) \right\}
\]

\[
f_2(t) = \left( \frac{k_1 + a}{2\sqrt{\Delta}} + \frac{1}{2} \right) \exp \left( -\frac{(a-k_1)}{2} + \Delta t \right) + \left( \frac{k_1 + a}{2\sqrt{\Delta}} + \frac{1}{2} \right) \exp \left( -\frac{(a-k_1)}{2} - \Delta t \right)
\]

\[
f_3(t) = \frac{1}{a(k_3b-k_1)} \left\{ \frac{(a-k_1)}{2a} (k_3b-k_1) \sqrt{\Delta} \exp \left( -\frac{(a-k_1)}{2} + \Delta t \right) - \frac{(a-k_1)}{2a} (k_3b-k_1) \sqrt{\Delta} \exp \left( -\frac{(a-k_1)}{2} - \Delta t \right) \right\}
\]

where \(k_1, k_2, a,\) and \(b\) are parameters that have previously been defined while the constant \(\Delta\) is defined as follows

\[
\Delta = (a-k_1)^2 - 4a(k_3b-k_1).
\]

Appendix B

In this appendix, we will show that \(\int_{-\infty}^{\infty} G(V_t; m, \sigma) \, dv_t\) is not log-concave. We begin by defining

\[
I = \int_{-\infty}^{\infty} G(V_t; m, \sigma) \, dv_t
\]

\[
F = \log I
\]

The Hessian matrix of \(F\) is

\[
H = \begin{bmatrix}
\frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} & \frac{\partial^2 F}{\partial x_1 \partial x_3} & \frac{\partial^2 F}{\partial x_1 \partial x_4} \\
\frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} & \frac{\partial^2 F}{\partial x_2 \partial x_3} & \frac{\partial^2 F}{\partial x_2 \partial x_4} \\
\frac{\partial^2 F}{\partial x_3 \partial x_1} & \frac{\partial^2 F}{\partial x_3 \partial x_2} & \frac{\partial^2 F}{\partial x_3^2} & \frac{\partial^2 F}{\partial x_3 \partial x_4} \\
\frac{\partial^2 F}{\partial x_4 \partial x_1} & \frac{\partial^2 F}{\partial x_4 \partial x_2} & \frac{\partial^2 F}{\partial x_4 \partial x_3} & \frac{\partial^2 F}{\partial x_4^2}
\end{bmatrix}
\]

where \(x_i\) denotes the \(i^{th}\) component of the vector \(x\) which is defined as follows

\[
x = \theta_t = [c, d, m, \sigma]^T.
\]

Since \(H\) is a symmetric matrix, we define the unique elements as follows

\[
\frac{\partial^2 F}{\partial x_1^2} = -\left( \frac{\partial v}{\partial c} \right)^2 G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} + G(v; m, \sigma) \right]
\]

\[
\frac{\partial^2 F}{\partial x_1 \partial x_2} = -\frac{\partial^2 F}{\partial x_2 \partial x_1} = \frac{\partial v}{\partial c} \frac{\partial v}{\partial d} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} + G(v; m, \sigma) \right]
\]

\[
\frac{\partial^2 F}{\partial x_1 \partial x_3} = -\frac{\partial^2 F}{\partial x_3 \partial x_1} = \frac{\partial v}{\partial c} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} - \int_{-\infty}^{\infty} \frac{G(y; m, \sigma)}{\sigma^2} \frac{y-m}{\sigma^2} \, dy \right]
\]

\[
\frac{\partial^2 F}{\partial x_2 \partial x_2} = -\frac{\partial^2 F}{\partial x_3 \partial x_3} = \frac{\partial v}{\partial d} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} + G(v; m, \sigma) \right]
\]

\[
\frac{\partial^2 F}{\partial x_1 \partial x_4} = \frac{\partial^2 F}{\partial x_4 \partial x_1} = \frac{\partial v}{\partial c} \frac{\partial v}{\partial m} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} - \int_{-\infty}^{\infty} \frac{G(y; m, \sigma)}{\sigma^2} \frac{y-m}{\sigma^2} \, dy \right]
\]

\[
\frac{\partial^2 F}{\partial x_2 \partial x_4} = \frac{\partial^2 F}{\partial x_4 \partial x_2} = \frac{\partial v}{\partial d} \frac{\partial v}{\partial m} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} - \int_{-\infty}^{\infty} \frac{G(y; m, \sigma)}{\sigma^2} \frac{y-m}{\sigma^2} \, dy \right]
\]

\[
\frac{\partial^2 F}{\partial x_3 \partial x_4} = \frac{\partial^2 F}{\partial x_4 \partial x_3} = \frac{\partial v}{\partial m} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} - \int_{-\infty}^{\infty} \frac{G(y; m, \sigma)}{\sigma^2} \frac{y-m}{\sigma^2} \, dy \right]
\]

\[
\frac{\partial^2 F}{\partial x_1 \partial x_1} = \frac{\partial^2 F}{\partial x_3 \partial x_3} = \frac{\partial v}{\partial c} \frac{\partial v}{\partial m} G(v; m, \sigma) \frac{I}{\sigma^2} \left[ \int \frac{v-m}{\sigma^2} - \int_{-\infty}^{\infty} \frac{G(y; m, \sigma)}{\sigma^2} \frac{y-m}{\sigma^2} \, dy \right]
\]
\[
\frac{\partial^2 F}{\partial x_1 \partial x_4} - \frac{\partial^2 F}{\partial x_3 \partial x_4} = \frac{1}{T^2} \left\{ I \int_{-\infty}^{v} \left[ \left( -\frac{1}{\sigma} G(y; m, \sigma) \right) \left( \frac{y - m}{\sigma^2} \right) \right. \\
+ \left. \frac{(y - m)^2}{\sigma^3} G(y; m, \sigma) \left( \frac{y - m}{\sigma^2} \right) \right] dy \\
- \left( \int_{-\infty}^{v} G(y; m, \sigma) \left( \frac{y - m}{\sigma^2} \right) dy \right) \\
\int_{-\infty}^{v} \left( -\frac{1}{\sigma} G(y; m, \sigma) \right) dy \\
+ \frac{(y - m)^2}{\sigma^3} G(y; m, \sigma) \right\} \\
\frac{\partial^2 F}{\partial x_1^2} \frac{1}{T^2} \left\{ I \int_{-\infty}^{v} \frac{G(y; m, \sigma)}{\sigma^2} \right. \\
\left. \left( 2 - \frac{5(y - m)^2}{\sigma^2} + \frac{(y - m)^4}{\sigma^4} \right) dy \\
- \left[ \int_{-\infty}^{v} \left( -\frac{1}{\sigma} G(y; m, \sigma) \right. \\
\left. + \frac{(y - m)^2}{\sigma^3} G(y; m, \sigma) \right] dy \right)^2 \} \\
\]