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.

Reference to this paper should be made as follows: Chen, J., Molnar, P. and Behal, A. (2017) 'Identification of a stochastic resonate-and-fire neuronal model via nonlinear least squares and maximum likelihood estimation', *Int. J. Modelling, Identification and Control*, Vol. 28, No. 3, pp.221–231.

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.

Aman Behal received his five-year integrated MTech degree in Electrical Engineering from IIT Bombay, Bombay, India in 1997 and PhD in Controls and Robotics from Clemson University, Clemson, SC in 2001. From 2002 to 2003, he was a Post-Doctoral Fellow in Bioengineering at Clemson University. In 2003, he joined the Faculty of the Electrical and Computer Engineering Department at Clarkson University, Potsdam, NY. In 2006, he joined the University of Central Florida, Orlando, where he is currently an Assistant Professor jointly with the Department of Electrical Engineering and Computer Science and the NanoScience Technology Center. He has authored three research monographs, 34 journal publications, and nearly 50 conference articles. He serves as an Associate Editor on the IEEE Control Systems Society's Conference Editorial Board. His research has been funded by the National Science Foundation, National Institutes of Health, the Department of Energy, and the National Multiple Sclerosis Society.

This paper is a revised and expanded version of a paper entitled 'Maximum likelihood parameter estimation in a stochastic resonate-and-fire neuronal model' presented at 2011 International Conference on Computational Advances in Bio and Medical Sciences, Orlando, FL, USA, February 2011.

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

requirement to separate estimation of the threshold and the post-spike reset parameters.

In this paper, the unknown threshold of the resonate-and-fire model is considered to be a stochastic variable. By considering the threshold to be the only stochastic component, the membrane potential at sub-threshold levels is thus deterministic and solvable, from which an estimate of the subthreshold parameters can be made. Under the assumption that both the input to the cell and the resulting membrane potential trace are available to the identification algorithm, the parameters for sub-threshold dynamics are estimated using a nonlinear least squares approach while the parameters associated with the threshold distribution as well as the post-spike reset of the state variables are estimated using simulated annealing to maximise the likelihood of the observed spiking pattern. Simulation and experimental results using *in vitro* motoneuron data demonstrate the efficiency and accuracy of the proposed estimation approach. To the best of our knowledge, this is the first paper that deals with parameter estimation in resonate-and-fire model with both sub-threshold measurements and firing trains, distinguishing itself from literature where only spiking location are considered. The methodology proposed in this paper provides a systematic approach for data-driven parameter estimation in the aforementioned model, which achieving excellent numerical results as demonstrated by experimental data. Note that data-driven identification is a very important topic in literature, with wide application in power systems, fault prediction, etc. (Shamsudin and Chen, 2012; Guo et al., 2015; Du et al., 2015; Mahmoud and Qureshi, 2012; Chen and Kumar, 2014, 2015a, 2015b).

A preliminary version of this paper has been presented as Chen et al. (2011). This paper extends (Chen et al., 2011) by providing detailed derivation and proof, as well as applying the proposed methodology in *in-vitro* data. The remainder of this paper is organised as follows. In Section 2, a resonate-and-fire neuron model with stochastic threshold is presented. Section 3 provides technical details of the the estimation process are provided. In Section 4, the methodology used for the acquisition of the experimental data is described while the results obtained from both simulated and experimental data follow in Section 5 along with a discussion of the results. Conclusions are drawn in Section 6.

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

$$\text{if } v = V_t, \text{ then } \begin{cases} v \to c \\ u \to u + d \end{cases} \tag{3}$$

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.

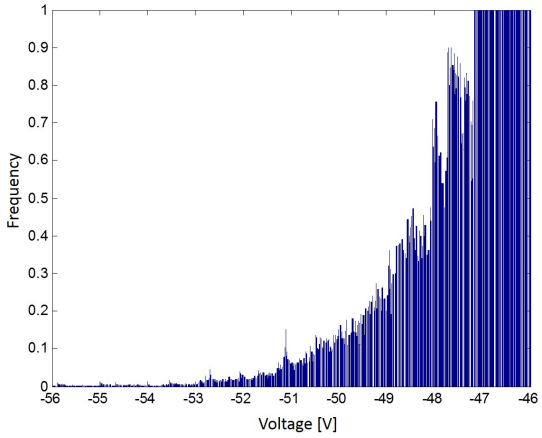
Remark 1: Although it is formulated differently here, the model presented here is the same as the one discussed in Young (1937) or FitzHugh (1966). Note that (1) and (2) do not model the post threshold dynamics of the cell; these equations only describe the sub-threshold behaviour and the spikes are 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., $V_t \sim N(m, \sigma)$.

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 database and n of them directly followed by a spike, the 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:

$$\theta \triangleq (k_1, k_2, k_3, a, b, c, d, m, \sigma) = (\theta_l, \theta_t)$$

where the partitions θ_l and θ_t 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 $\theta_t \triangleq (c, d, m, \sigma)$, which is linked to the after-spike resetting and threshold distribution parameters. In this section, we propose a two stage approach to estimate both θ_l and θ_t. In the first stage, the measurements of sub-threshold voltage are used to estimate θ_l, while in the second stage, the estimated θ_l together with the measured interspike intervals (ISIs, computed from measured voltage recording) are utilised to identify θ_t.

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

$$v(t) = [k_2 - k_3 u(t_0)] f_1(\theta_l, t) + k_3 (f_2(\theta_l, t) * i) + v(t_0) f_2(\theta_l, t) + k_2 a f_3(\theta_l, t) \quad (4)$$

where where $u(t_0)$ and $v(t_0)$ are initial values for u and v, the symbol $*$ denotes the convolution operator on two functions, while the functions $f_1(\theta_l, t)$, $f_2(\theta_l, t)$, and $f_3(\theta_l, t)$ 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 θ_l 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 θ_t in the sub-threshold region, depending only on θ_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 θ_l. The estimate for θ_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

$$\hat{\theta}_l^{t+1} = \hat{\theta}_l^t + \Delta\hat{\theta}_l \quad (5)$$

$$\Delta\hat{\theta}_l = (J'J)^{-1} J' \Delta v \quad (6)$$

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

$$J_{ik} = \frac{\partial \hat{v}_i}{\partial \hat{\theta}_{l_k}}. \quad (7)$$

Here, $M = 5$, corresponds to the number of parameters in the set θ_l. Furthermore, $\hat{\theta}_{l_k}$ denotes the estimate of the k^{th} element of the parameter vector θ_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 $\hat{\theta}_l$ estimated from the previous step are utilised in the reconstruction of the membrane potential – specifically, we can write

$$\hat{v} = g\left(t, I, \hat{\theta}_l, c, d\right). \tag{8}$$

Given i and $\hat{\theta}_l$, 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 unknown parameters $c, d, m,$ and σ are adjusted so as to maximise the likelihood of occurrence of the observed firing pattern. During this stage, it is assumed that the only output measured from the cells is the location of the spikes (i.e., spiking time). Since the spikes from real neurons have finite width, the spike time t_j is defined to be the time of the peak of each spike. During the interval $t \in (t_{j-1}, t_j)$, the cell is not firing. All data points during the intervals $\bigcup_j (t_{j-1}, t_j)$ can be denoted as t_i, since the data being examined has been discretised. The likelihood function for the parameter set θ_t can therefore be defined as the probability that the stochastic threshold V_t is below the reconstructed membrane potential $\hat{v}$ at spike times t_j and above $\hat{v}$ at non-spike times t_i. Mathematically, the profile likelihood function can be expressed as follows

$$\begin{aligned} L(\theta_t) &= \log\left(\prod_i \int_{\hat{v}_i}^{\infty} G(V_t; m, \sigma) dV_t \right. \\ &\quad \left. \times \prod_j \int_{-\infty}^{\hat{v}_j} G(V_t; m, \sigma) dV_t \right) \\ &= \sum_i \log \int_{\hat{v}_i}^{\infty} G(V_t; m, \sigma) dV_t \\ &\quad + \sum_j \log \int_{-\infty}^{\hat{v}_j} G(V_t; m, \sigma) dV_t \end{aligned} \tag{9}$$

where $G(V_t; m, \sigma)$ is the Gaussian probability density function of the threshold V_t with mean m and variance σ^2. $L(\theta_t)$ 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

$$\hat{\theta}_t = \arg\max_{\theta_t} L(\theta_t). \tag{10}$$

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

$$P = \begin{cases} 1 & \text{if } L(s') > L(s) \\ e^{-1/T} & \text{otherwise} \end{cases} \tag{11}$$

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_o and initial temperature T_o are randomly selected, while the maximum number of iterations N is user-specified.

Setp 0 Initialise the algorithm by selecting a starting point s_o, a starting temperature T_o, 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.

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 $\times$ 22 mm^2 ornithine/laminincoated coverslips at a density of 200 cells/mm^2 in Neurobasal (Gibco-BRL) medium supplemented with B27 (2% v/v; Invitrogen), L-gluramine (0.5 mM), 2-mercaptoethanol (25 μM), glial cell line-derived neurotrophic factor (1 ng/ml CNTF; cell sciences) L-glutamate (25 μM) was added to the culture medium during the first five days of growth.

Conventional whole-cell path 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-gluconate 140, EGTA 1, MgC12 2, Na2ATP 2, phophocreatine 5, phophocreatine kinase 2.4 mg, Hepes 10; pH:7.2). The resistance of the electrodes was 6–8 MΩ. 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.

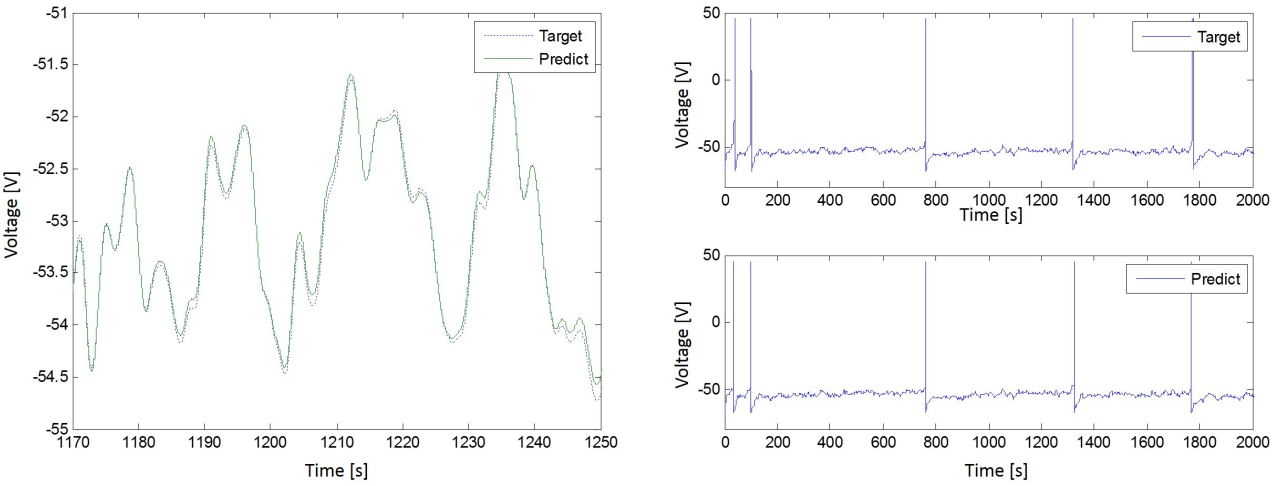
5.1 Detailed Hodgkin-Huxley model (simulation)

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

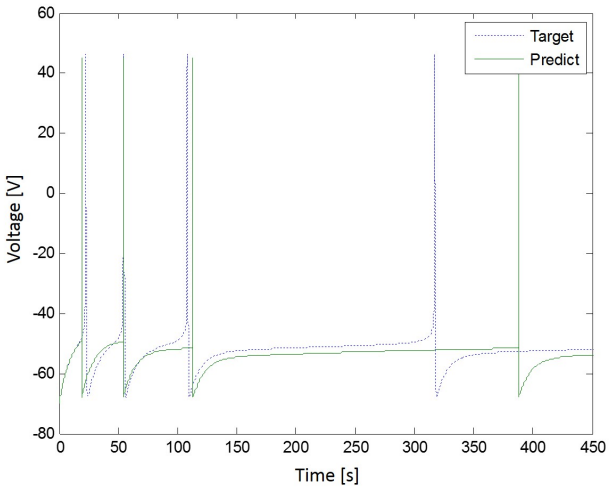
In this section, the results of the proposed method are presented for the experimental data obtained from *in-vitro* recordings from embryonic rat motoneurons. As in the case of the detailed Hodgkin-Huxley model, the results presented here for evaluation use previously unseen data, i.e., the data stream 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.

Figure 2 Estimation results using simulated reference data, (a) comparison of sub-threshold dynamics, target data is generated under the same type of input as the one used in the reference data generation (b) comparison of spike train, target data is generated under the same type of input as the one used in the reference data generation (see online version for colours)



(a) (b)

Figure 3 Estimation results using simulated reference data: prediction and target data under DC current (see online version for colours)



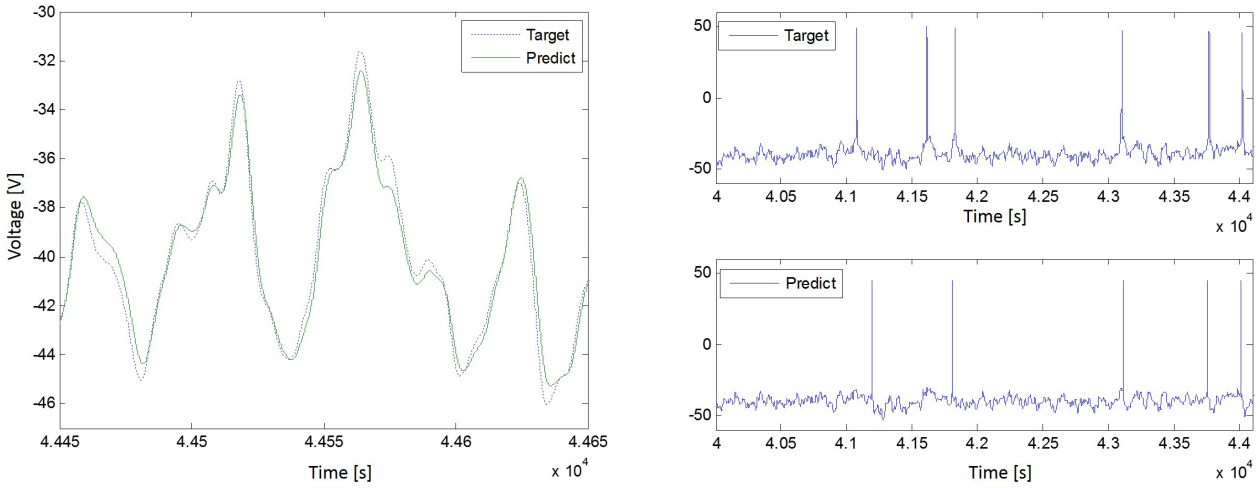
5.3 Discussion

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 3σ 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 σ.

Figure 4 Estimation results using *in-vitro* reference data, (a) comparison of sub-threshold dynamics, target data used here has not been used in estimation process before (b) comparison of spike trains, target data used here has not been used in estimation process before (see online version for colours)



(a) (b)

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

Azouz, R. and Gray, C.M. (2000) 'Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons *in vivo*', *Proceedings of the National Academy of Sciences*, Vol. 97, No. 14, pp.8110–8115.

Bershadskii, A., Dremencov, E. and Yadid, G. (2003) 'Short-term memory and critical clusterization in brain neurons spike series', *Physics Letters A*, Vol. 313, No. 1, pp.158–161.

Brette, R. and Gerstner, W. (2005) 'Adaptive exponential integrate-and-fire model as an effective description of neuronal activity', *Journal of Neurophysiology*, Vol. 94, No. 5, pp.3637–3642.

Burkitt, A.N. (2006) 'A review of the integrate-and-fire neuron model: I. homogeneous synaptic input', *Biological Cybernetics*, Vol. 95, No. 1, pp.1–19.

Chen, J. and Kumar, R. (2014) 'Pattern mining for predicting critical events from sequential event data log', in *International Workshop on Discrete Event Systems*, France, May, pp.1–6.

Chen, J. and Kumar, R. (2015a) 'Stochastic failure prognosability of discrete event systems', *IEEE Transactions on Automatic Control*, Vol. 60, No. 6, pp.1570–1581.

Chen, J. and Kumar, R. (2015b) 'Fault detection of discrete-time stochastic systems subject to temporal logic correctness requirements', *IEEE Transactions on Automation Science and Engineering*, Vol. 12, No. 4, pp.1369–1379.

Chen, J., Suarez, J., Molnar, P. and Behal, A. (2011) 'Maximum likelihood parameter estimation in a stochastic resonate-and-fire neuronal model', in *Computational Advances in Bio and Medical Sciences (ICCABS), IEEE 1st International Conference on*, Orlando, FL, February, pp.57–62.

Corana, A., Marchesi, M., Martini, C. and Ridella, S. (1987) 'Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm corrigenda for this article is available here', *ACM Transactions on Mathematical Software (TOMS)*, Vol. 13, No. 3, pp.262–280.

Das, M., Molnar, P., Devaraj, H., Poeta, M. and Hickman, J.J. (2003) 'Electrophysiological and morphological characterization of rat embryonic motoneurons in a defined system', *Biotechnology Progress*, Vol. 19, No. 6, pp.1756–1761.

Destexhe, A., Contreras, D. and Steriade, M. (1998) 'Mechanisms underlying the synchronizing action of corticothalamic feedback through inhibition of thalamic relay cells', *Journal of neurophysiology*, Vol. 79, No. 2, pp.999–1016.

Du, Z., Li, X. and Mao, Q. (2015) 'A new online hybrid learning algorithm of adaptive neural fuzzy inference system for fault prediction', *International Journal of Modelling, Identification and Control*, Vol. 23, No. 1, pp.68–76.

FitzHugh, R. (1966) *Mathematical Models of Excitation and Propagation in Nerve*, Publisher Unknown.

Guo, Y., Guo, L.Z., Billings, S.A. and Wei, H-L. (2015) 'Identification of nonlinear systems with non-persistent excitation using an iterative forward orthogonal least squares regression algorithm', *International Journal of Modelling, Identification and Control*, Vol. 23, No. 1, pp.1–7.

Hines, M.L. and Carnevale, N.T. (1997) 'The neuron simulation environment', *Neural Computation*, Vol. 9, No. 6, pp.1179–1209.

Hodgkin, A.L. and Huxley, A.F. (1952) 'A quantitative description of membrane current and its application to conduction and excitation in nerve', *The Journal of Physiology*, Vol. 117, No. 4, p.500.

Izhikevich, E.M. et al. (2003) 'Simple model of spiking neurons', *IEEE Transactions on Neural Networks*, Vol. 14, No. 6, pp.1569–1572.

Izhikevich, E.M. (2001) 'Resonate-and-fire neurons', *Neural networks*, Vol. 14, No. 6, pp.883–894.

Izhikevich, E.M. (2007) *Dynamical Systems in Neuroscience*, MIT Press.

Jabalameli, A. and Behal, A. (2015) 'A constrained linear approach to identify a multi-timescale adaptive threshold neuronal model', in *IEEE International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)*, Miami, FL, October, pp.1–6.

Jolivet, R., Lewis, T.J. and Gerstner, W. (2004) 'Generalized integrate-and-fire models of neuronal activity approximate spike trains of a detailed model to a high degree of accuracy', *Journal of Neurophysiology*, Vol. 92, No. 2, pp.959–976.

Kelley, C.T. (1999) *Iterative Methods for Optimization*, Vol. 18, Siam.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P. et al. (1983) 'Optimization by simulated annealing', *Science*, Vol. 220, No. 4598, pp.671–680.

Kobayashi, R., Tsubo, Y. and Shinomoto, S. (2009) 'Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold', *Frontiers in Computational Neuroscience*, September, Vol. 3, DOI: 10.3389/neuro.10.009.2009.

Lansky, P., Sanda, P. and He, J. (2006) 'The parameters of the stochastic leaky integrate-and-fire neuronal model', *Journal of Computational Neuroscience*, Vol. 21, No. 2, pp.211–223.

LeMasson, G. and Maex, R. (2001) 'Introduction to equation solving and parameter fitting', *Computational Neuroscience: Realistic Modeling for Experimentalists*, CRC Press, London, pp.1–23.

Mahmoud, M.S. and Qureshi, A. (2012) 'Model identification and analysis of small-power wind turbines', *International Journal of Modelling, Identification and Control*, Vol. 17, No. 1, pp.19–31.

McCormick, D.A., Wang, Z. and Huguenard, J. (1993) 'Neurotransmitter control of neocortical neuronal activity and excitability', *Cerebral Cortex*, Vol. 3, No. 5, pp.387–398.

Naud, R., Marcille, N., Clopath, C. and Gerstner, W. (2008) 'Firing patterns in the adaptive exponential integrate-and-fire model', *Biological Cybernetics*, Vol. 99, Nos. 4–5, pp.335–347.

Paninski, L., Pillow, J.W. and Simoncelli, E.P. (2004) 'Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model', *Neural Computation*, Vol. 16, No. 12, pp.2533–2561.

Paninski, L., Pillow, J. and Simoncelli, E. (2005) 'Comparing integrate-and-fire models estimated using intracellular and extracellular data', *Neurocomputing*, Vol. 65, pp.379–385.

Prinz, A.A., Billimoria, C.P. and Marder, E. (2003) 'Alternative to hand-tuning conductance-based models: construction and analysis of databases of model neurons', *Journal of Neurophysiology*, Vol. 90, No. 6, pp.3998–4015.

Shamsudin, S.S. and Chen, X.Q. (2012) 'Identification of an unmanned helicopter system using optimised neural network structure', *International Journal of Modelling, Identification and Control*, Vol. 17, No. 3, pp.223–241.

Vanier, M.C. and Bower, J.M. (1999) 'A comparative survey of automated parameter-search methods for compartmental neural models', *Journal of Computational Neuroscience*, Vol. 7, No. 2, pp.149–171.

Weaver, C.M. and Wearne, S.L. (2006) 'The role of action potential shape and parameter constraints in optimization of compartment models', *Neurocomputing*, Vol. 69, No. 10, pp.1053–1057.

White, J.A., Klink, R., Alonso, A. and Kay, A.R. (1998) 'Noise from voltage-gated ion channels may influence neuronal dynamics in the entorhinal cortex', *Journal of Neurophysiology*, Vol. 80, No. 1, pp.262–269.

Young, G. (1937) 'Note on excitation theories', *Psychometrika*, Vol. 2, No. 2, pp.103–106.

Zhang, P. and Feng, J. (2002) 'Ideal observer of single neuron activity', *Neurocomputing*, Vol. 44, pp.243–247.

Zhang, X., You, G., Chen, T. and Feng, J. (2009) 'Maximum likelihood decoding of neuronal inputs from an interspike interval distribution', *Neural Computation*, Vol. 21, No. 11, pp.3079–3105.

Zhi, L., Chen, J., Molnar, P. and Behal, A. (2012) 'Weighted least-squares approach for identification of a reduced-order adaptive neuronal model', *Neural Networks and Learning Systems, IEEE Transactions on*, Vol. 23, No. 5, pp.834–840.

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.

Appendix

Appendix A

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

$$\begin{aligned}
f_1(t) =& \frac{1}{\sqrt{\Delta}}\left\{\exp\left(\frac{-(a-k_1)+\sqrt{\Delta}}{2}t\right)\right. \\
&\left.-\exp\left(\frac{-(a-k_1)-\sqrt{\Delta}}{2}t\right)\right\} \\
f_2(t) =& \left(\frac{k_1+a}{2\sqrt{\Delta}}+\frac{1}{2}\right)\exp\left(\frac{-(a-k_1)+\sqrt{\Delta}}{2}t\right) \\
&+\left(-\frac{k_1+a}{2\sqrt{\Delta}}+\frac{1}{2}\right)\exp\left(\frac{-(a-k_1)-\sqrt{\Delta}}{2}t\right) \\
f_3(t) =& \frac{1}{a(k_3 b-k_1)} \\
&-\frac{(a-k_1)+\sqrt{\Delta}}{2a(k_3 b-k_1)\sqrt{\Delta}}\exp\left(\frac{-(a-k_1)+\sqrt{\Delta}}{2}t\right) \\
&-\frac{-(a-k_1)+\sqrt{\Delta}}{2a(k_3 b-k_1)\sqrt{\Delta}}\exp\left(\frac{-(a-k_1)-\sqrt{\Delta}}{2}t\right)
\end{aligned}$$

where k_1, k_3, a, and b are parameters that have previously been defined while the constant Δ is defined as follows

$$\Delta = (a-k_1)^2 - 4a(k_3 b - k_1).$$

Appendix B

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

$$\begin{aligned}
I &= \int_{-\infty}^{v} G(V_t; m, \sigma)\, dV_t \\
F &= \log I
\end{aligned}$$

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

$$\begin{aligned}
\frac{\partial^2 F}{\partial x_1^2} =& -\left(\frac{\partial v}{\partial c}\right)^2 \frac{G(v; m, \sigma)}{I^2}\left[I\frac{v-m}{\sigma^2} + G(v; m, \sigma)\right] \\
\frac{\partial^2 F}{\partial x_2 \partial x_1} =& \frac{\partial^2 F}{\partial x_1 \partial x_2} \\
=& -\frac{\partial v}{\partial c}\frac{\partial v}{\partial d}\frac{G(v; m, \sigma)}{I^2}\left[I\frac{v-m}{\sigma^2} + G(v; m, \sigma)\right] \\
\frac{\partial^2 F}{\partial x_3 \partial x_1} =& \frac{\partial^2 F}{\partial x_1 \partial x_3} \\
=& \frac{\partial v}{\partial c}\frac{G(v; m, \sigma)}{I^2}\left[I\frac{v-m}{\sigma^2} - \right. \\
& \left.\int_{-\infty}^{v} G(y; m, \sigma)\left(\frac{y-m}{\sigma^2}\right) dy\right] \\
\frac{\partial^2 F}{\partial x_4 \partial x_1} =& \frac{\partial^2 F}{\partial x_1 \partial x_4} \\
=& \frac{\partial v}{\partial c}\frac{G(v; m, \sigma)}{I^2}\left[\left(-\frac{1}{\sigma} + \frac{(v-m)^2}{\sigma^3}\right) I \right. \\
& - \int_{-\infty}^{v}\left(-\frac{1}{\sigma}G(y; m, \sigma)\right. \\
& \left.\left. + \frac{(y-m)^2}{\sigma^3}G(y; m, \sigma)\right) dy\right] \\
\frac{\partial^2 F}{\partial x_2^2} =& -\left(\frac{\partial v}{\partial d}\right)^2 \frac{G(v; m, \sigma)}{I^2}\left[I\frac{v-m}{\sigma^2} + G(v; m, \sigma)\right] \\
\frac{\partial^2 F}{\partial x_3 \partial x_2} =& \frac{\partial^2 F}{\partial x_2 \partial x_3} \\
=& \frac{\partial v}{\partial d}\frac{G(v; m, \sigma)}{I^2}\left[I\frac{v-m}{\sigma^2}\right. \\
& \left. - \int_{-\infty}^{v} G(y; m, \sigma)\left(\frac{y-m}{\sigma^2}\right) dy\right] \\
\frac{\partial^2 F}{\partial x_4 \partial x_2} =& \frac{\partial^2 F}{\partial x_2 \partial x_4} \\
=& \frac{\partial v}{\partial d}\frac{G(v; m, \sigma)}{I^2}\left[\left(-\frac{1}{\sigma} + \frac{(v-m)^2}{\sigma^3}\right) I \right. \\
& - \int_{-\infty}^{v}\left(-\frac{1}{\sigma}G(y; m, \sigma)\right. \\
& \left.\left. + \frac{(y-m)^2}{\sigma^3}G(y; m, \sigma)\right) dy\right] \\
\frac{\partial^2 F}{\partial x_3^2} =& \frac{1}{I^2}\left\{I\int_{-\infty}^{v}\left[G(y; m, \sigma)\left(\frac{y-m}{\sigma^2}\right)^2\right.\right. \\
& \left. - G(y; m, \sigma)\frac{1}{\sigma^2}\right] dy \\
& \left. - \left[\int_{-\infty}^{v} G(y; m, \sigma)\left(\frac{y-m}{\sigma^2}\right) dy\right]^2\right\}
\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 F}{\partial x_4 \partial x_3} =& \frac{\partial^2 F}{\partial x_3 \partial x_4} \\ =& \frac{1}{I^2}\left\{ I \int_{-\infty}^{v} \left[\left(-\frac{1}{\sigma} G(y;m,\sigma) \right.\right.\right. \\ & \left. + \frac{(y-m)^2}{\sigma^3} G(y;m,\sigma) \right) \left(\frac{y-m}{\sigma^2} \right) \\ & \left. -G(y;m,\sigma) \frac{2\,(y-m)}{\sigma^3} \right] dy \\ & - \left(\int_{-\infty}^{v} G(y;m,\sigma) \left(\frac{y-m}{\sigma^2} \right) dy \right) \\ & \int_{-\infty}^{v} \left(-\frac{1}{\sigma} G(y;m,\sigma) \right. \\ & \left. + \frac{(y-m)^2}{\sigma^3} G(y;m,\sigma) \right) dy \Bigg\}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 F}{\partial x_4^2} =& \frac{1}{I^2}\left\{ I \int_{-\infty}^{v} \frac{G(y;m,\sigma)}{\sigma^2} \right. \\ & \left(2 - \frac{5(y-m)^2}{\sigma^2} + \frac{(y-m)^4}{\sigma^4} \right) dy \\ & - \left[\int_{-\infty}^{v} \left(-\frac{1}{\sigma} G(y;m,\sigma) + \right. \right. \\ & \left. \left. \left. \frac{(y-m)^2}{\sigma^3} G(y;m,\sigma) \right) dy \right]^2 \Bigg\}\end{aligned}$$

By appropriately choosing an invertible matrix C, we can diagonalise H to obtain the matrix D as follows

$$D = C^T H C = \begin{bmatrix} \left(I \frac{v-m}{\sigma^2} + G \right) & 0\,0\,0 \\ 0 & -1\,0\,0 \\ 0 & 0\,1\,0 \\ 0 & 0\,0\,0 \end{bmatrix}.$$

It is clear to see that D has both positive and negative eigenvalues; thus, it is indefinite. implying that $\int_{-\infty}^{v_i} G\left(V_t; m, \sigma\right) dV_t$ is not log-concave. We note here that it is not essential to explicitly find the matrix C. In fact, the diagonal matrix D was obtained through an appropriate sequence of row and column operations.