Entropy-based Iterative Learning Estimation for Stochastic Non-linear Systems and Its Application to Neural Membrane Potential Interaction

Qichun Zhang and Xuewu Dai

Abstract— **This paper investigates the unknown parameter estimation problem for a class of stochastic non-linear systems which are subjected to non-Gaussian noises. Motivated by the iterative learning identification, an iterative learning estimation algorithm has been presented with an entropy-based performance criterion. Since the data sets from each iteration are collected and stored in memory, the entropy of the estimation error can be approximated by kernel density estimation and the optimal vector-valued estimated parameter can be obtained using the gradient descent searching while the learning rate is able to be selected properly to guarantee the convergence of the presented algorithm. In order to verify the effectiveness of this algorithm, one solution to biological neural membrane potential interaction characterization problem is given via a numerical simulation using the presented algorithm.**

I. INTRODUCTION

System identification is a significant research problem in engineering and a lot of impressive results have been obtained [1]. However, most of the existing results are based on the linear systems or static systems. Motivated by iterative learning control [2][3], many algorithms of iterative learning identification have been presented [4]. Comparing to other identification algorithms, the iterative learning approach uses more information for parameter estimation which comes from each iteration. In this paper, this extra information can be used to overcome the non-linear dynamics of the stochastic systems.

Basically, most of the dynamic systems can be described by differential equations. As an example, the following equation is given as a system model.

$$\dot{x} = f(x) + \alpha g(x) \tag{1}$$

where α is a real constant as the unknown parameter. f and g are general non-linear functions of system state x. Basically, we can transform this problem to a stabilization problem where the estimated state can be stated as

$$\dot{\hat{x}} = f(\hat{x}) + \hat{\alpha} g(\hat{x}) \tag{2}$$

In addition, the estimation error can be expressed as $\tilde{x} = x - \hat{x}$ such that the dynamic of the estimation error is formulated as follows:

$$\dot{\tilde{x}} = f(x) - f(\hat{x}) + \alpha g(x) - \hat{\alpha} g(\hat{x}) \tag{3}$$

It shows that the optimal estimation can be achieve if $\tilde{x}$ is stabilized. However, it is very difficult to search $\hat{\alpha}$ without any further assumption, for instance Lipschitz condition.

Notice that even the simplest case is difficult to solve following this approach, we have to develop a novel approach. Motivated by the iterative learning identification and stochastic distribution optimisation [5], we can re-investigate this problem from the view of batch because most of the modelling training is repeatable. Meanwhile, the random noise affects the performance of the estimation. Most of the existing results use Gaussian distribution as an assumption, we are trying to release this assumption which also forms the purpose of this paper.

As a summery, a generalised format for piratical systems with non-Gaussian noises are investigated. To find out the unknown parameter in the model, an iterative learning algorithm has been developed under the entropy-based performance criterion. As a practical application, the biological neural membrane potential interaction can be characterized using the presented algorithm which validated the effectiveness of this algorithm. Moreover, the convergence analysis and some potential extensions are also illustrated in this paper. In particular, the contributions of this paper can be further summarised as follows: 1) a novel iterative learning algorithm is given for non-linear stochastic systems subjected to non-Gaussian noises, 2) the convergence of the algorithm has been analysed while the sufficient conditions are obtained, and 3) the neural membrane potential interaction can be characterised following the presented algorithm.

The rest of this paper has been organized as follows: Section II presents the essential preliminaries including the system model formulation and the identification objective. The iterative learning identification algorithm and its convergence analysis are given in Section III and Section IV, respectively. In Section V, a numerical simulation is given to verify the presented algorithm which is based on the neural membrane potential interaction

characterisation application. Section VI presents some potential extensions of this presented algorithm. Finally, this paper is concluded by Section VII.

II. FORMULATION

Following the discussion in previous section, this paper investigates a class of generalized non-Gaussian stochastic non-linear systems which are expressed by the following discrete-time model.

$$x_{k+1} = f(x_k) + \Theta g(x_k) + w_k$$
$$y_k = c^T x_k + v_k \tag{4}$$

where k denotes the sampling index. $y_k \in \mathbb{R}$, $v_k \in \mathbb{R}$, $x_k \in \mathbb{R}^n$ and $w_k \in \mathbb{R}^n$ stand for the system output, measurement noise, vector-valued system state and process noise. f and g are general non-linear functions. Θ denotes the unknown parametric matrix. Notice that the distributions of the random noises w_k and v_k are non-Gaussian and all the variables are with the appropriate dimensions.

It is clear that the objective is to find out the unknown parameter Θ if we have the measured data $\bar{y}$. Since the estimated state is denoted as $\hat{y}$, then the estimation error is stated as

$$\tilde{y}_k = \bar{y}_k - \hat{y}_k, k = 1, 2 \cdots k_t \tag{5}$$

where k_t is pre-specified positive integer. Furthermore, we can use d to stand for the iteration index, then the data set for each batch can be descried as the following vector with k_t elements.

$$\tilde{y}_d = [\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{k_t}]^T \tag{6}$$

which can be further stated as

$$\tilde{y}_d = \bar{y} - C\hat{x}_d \tag{7}$$

where

$$C = \begin{bmatrix} c^T & & \\ & \ddots & \\ & & c^T \end{bmatrix}$$
$$\hat{x}_d = [\hat{x}_{d,1}^T, \ldots, \hat{x}_{d,k_t}^T]^T \tag{8}$$

In other words, the objective can be achieved if there exist any positive integer k_d such that $d \leq k_d$ then $\tilde{y}_d$ is bounded and close to zero vector. Notice that the identification problem is investigated in this paper, therefore, we can further assume that the investigated system (4) is stable which leads to the bounded system states and system output.

III. ITERATIVE LEARNING ESTIMATION

The structure of the algorithm is shown in Fig 1 which is the standard procedure of the iterative learning strategy. To implement this algorithm, the performance criterion should be selected firstly which is motivated by [6]
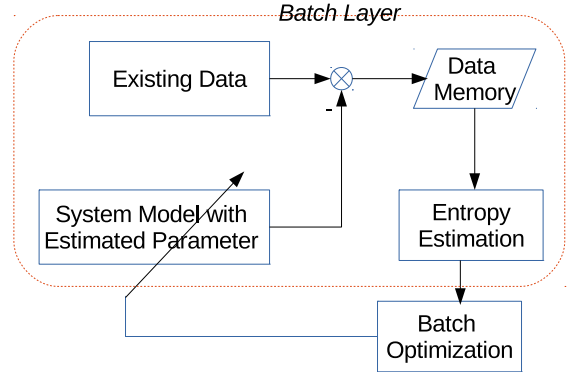

Fig. 1. Block diagram of the presented iterative learning strategy.

Notice that the batch-based estimation error vector is affected by non-Gaussian random noise vector, the classical statistical tools such as variance, covariance, etc, cannot be used to characterize the randomness. Thus the entropy of the estimation error has to be considered here.

$$H(\tilde{y}_d) = -\log_b \int_{\Omega_d} \gamma_d^2(\alpha_d) \, d\alpha_d \tag{9}$$

where H denotes the entropy operation, γ_d and α_d stand for the probability density function and random variate of the data set of estimation error while Ω_d is the sampling space of this random variate.

Therefore, the complete performance criterion can be selected as follows.

$$J_d = E(\tilde{y}_d^T \tilde{y}_d) + H(\tilde{y}_d) \tag{10}$$

where E denotes the mean-value operation.

Equivalently, this performance criterion can be restated using the concept of information potential [7] if b is pre-specified as $0 < b < 1$.

$$J_d = E(\tilde{y}_d^T \tilde{y}_d) + \int_{\Omega_d} \gamma_d^2(\alpha_d) \, d\alpha_d \tag{11}$$

Moreover, the information potential can be approximated by the measured data combining the kernel density estimation [8], [9]. Then the performance criterion is rewritten by

$$J_d = E(\tilde{y}_d^T \tilde{y}_d) + P(\tilde{y}_d) \tag{12}$$

where

$$P(\tilde{y}_d) = \frac{1}{(nk_t)^2} \sum_{i=1}^{nk_t} \left[\sum_{j=1}^{nk_t} G_\sigma(\tilde{y}_{d,i} - \tilde{y}_{d,j}) \right]^2 \tag{13}$$

while $\tilde{y}_{d,i}$ and $\tilde{y}_{d,j}$ denote any element in the estimation error vector with $i, j = 1, 2, \cdots, nk_t$. G_σ is the probability density function of Gaussian distribution.

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{14}$$

Once the value of the presented performance criterion can be calculated for any d-th iteration, the estimated parameter is updated by the following equation.

$$\hat{\Theta}_d = \hat{\Theta}_{d-1} - \xi_d \frac{\partial J_{d-1}}{\partial \Theta}\bigg|_{\Theta = \hat{\Theta}_{d-1}} \qquad (15)$$

where ξ_d denotes the learning rate for d-th interation.

For some specified situations, some elements of Θ are zero or equal to other elements such as diagonal matrix, symmetric matrix, etc. Then the vectorization operation can be used to reduce the computational complexity in (15).

Remark 1: The entropy is equivalent to variance if the random noise obeys Gaussian distribution, therefore this performance criterion can be considered as an extension of the minimum variance criterion.

IV. CONVERGENCE

One important problem still remains after the iterative learning law has been given which is the selection problem of the learning rate ξ_d. Using the similar approach in [10] and [11], the convergence of the presented algorithm is analysed with the proper ξ_d by the following theorem.

Theorem 1: The iterative learning estimation algorithm (15) is convergent with the entropy-based performance criterion (10) if there exists a learning rate matrix such that the following two conditions hold.

$$\left(C^T C \bar{M} - \bar{y}^T C \right) \left(\Delta_d - \xi_d \frac{\partial J_{d-1}}{\partial \hat{\Theta}_{d-1}} g\left(\bar{M}\right) \right) < 0 \qquad (16)$$

$$\sum_{i,j=1}^{nk_t} (\tilde{y}_{d,i} - \tilde{y}_{d,j}) C \left(\Delta_{\max} - \xi_d \frac{\partial J_{d-1}}{\partial \hat{\Theta}_{d-1}} g\left(\bar{M}\right) \right) < 0 \qquad (17)$$

where

$$\Delta_d = f\left(\bar{M}\right) - f\left(\hat{x}_{d-1}\right) + \hat{\Theta}_{d-1}\left(g\left(\bar{M}\right) - g\left(\hat{x}_{d-1}\right)\right)$$
$$\Delta_{max} = f\left(\bar{M}\right) - f\left(\hat{x}_{d-1,max}\right)$$
$$+ \hat{\Theta}_{d-1,max}\left(g\left(\bar{M}\right) - g\left(\hat{x}_{d-1,max}\right)\right) \qquad (18)$$

while *max* denotes the index number of the maximum element in the vector $\tilde{y}_d$. $\bar{M}$ is the upper boundary of estimated system states.

Proof: The following inequality holds for any d-th iteration if the presented algorithm is convergent.

$$J_d - J_{d-1} \leq 0 \qquad (19)$$

which leads to the following two inequalities if we substitute (12) and (13) into inequality (19):

$$\tilde{y}_d^T \tilde{y}_d - \tilde{y}_{d-1}^T \tilde{y}_{d-1} < 0 \qquad (20)$$

and

$$P\left(\tilde{y}_d\right) - P\left(\tilde{y}_{d-1}\right) < 0 \qquad (21)$$

Notice that inequality (20) can be restated by

$$\frac{\partial \tilde{y}_d^T \tilde{y}_d}{\partial d} < 0 \qquad (22)$$

Since the estimation error is expressed by Eq. (7), inequality (22) can be restated by

$$\frac{\partial \tilde{y}_d^T \tilde{y}_d}{\partial d} = \frac{\partial (\bar{y} - C\hat{x}_d)^T (\bar{y} - C\hat{x}_d)}{\partial d}$$
$$= \frac{\partial \left(\bar{y}^T \bar{y} - 2\bar{y}^T C\hat{x}_d + \hat{x}_d^T C^T C\hat{x}_d \right)}{\partial d}$$
$$= 2\left(C^T C\hat{x}_d - \bar{y}^T C \right) \frac{\partial \hat{x}_d}{\partial d} < 0 \qquad (23)$$

As a result, we can obtain

$$\frac{\partial \hat{x}_d}{\partial d} = \hat{x}_d - \hat{x}_{d-1}$$
$$= \left(f\left(\hat{x}_d\right) - f\left(\hat{x}_{d-1}\right) - \hat{\Theta}_{d-1}g\left(\hat{x}_{d-1}\right) + \hat{\Theta}_d g\left(\hat{x}_d\right)\right)$$
$$= \Delta_d - \xi_d \frac{\partial J_{d-1}}{\partial \hat{\Theta}_{d-1}} g\left(\hat{x}_d\right) \qquad (24)$$

Substituting (24) into (23), the condition (16) can be obtained.

Similarly, inequality (21) can also be expressed by

$$\frac{\partial P\left(\tilde{y}_d\right)}{\partial d} < 0 \qquad (25)$$

Using the kernel density estimation (13), inequality (25) can be further expressed as

$$\frac{\partial P\left(\tilde{y}_d\right)}{\partial d} = \frac{\partial}{\partial d} \frac{1}{(nk_t)^2} \sum_{i=1}^{nk_t} \left[\sum_{j=1}^{nk_t} G_\sigma\left(\tilde{y}_{d,i} - \tilde{y}_{d,j}\right) \right]^2$$
$$= \frac{2}{(nk_t)^2} \sum_{i=1}^{nk_t} \left[\sum_{j=1}^{nk_t} G_\sigma\left(\tilde{y}_{d,i} - \tilde{y}_{d,j}\right) \frac{\partial G_\sigma\left(\tilde{y}_{d,i} - \tilde{y}_{d,j}\right)}{\partial d} \right] \qquad (26)$$

Notice that the kernel function is selected as Gaussian distribution (14), we can have the following equation.

$$\frac{\partial G_\sigma\left(\tilde{y}_{d,i} - \tilde{y}_{d,j}\right)}{\partial d} = \frac{\tilde{y}_{d,j} - \tilde{y}_{d,i}}{\sqrt{2\pi}\sigma^3}$$
$$\times \exp\left(\frac{\tilde{y}_{d,j} - \tilde{y}_{d,i}}{2\sigma^2} \right) \frac{\partial \left(\tilde{y}_{d,i} - \tilde{y}_{d,j}\right)}{\partial d} \qquad (27)$$

Next, there always exists a real positive number M makes the following inequality holds.

$$\frac{\partial \left(\tilde{y}_{d,i} - \tilde{y}_{d,j}\right)}{\partial d} < \frac{\partial \tilde{y}_{d,\max}}{\partial d} - \frac{\partial \tilde{y}_{d,\min}}{\partial d} < M\frac{\partial \tilde{y}_{d,\max}}{\partial d} \qquad (28)$$

Moreover, similar to (24), we obtain

$$\frac{\partial \tilde{y}_{d,\max}}{\partial d} = \frac{\partial \left(\bar{y}_{\max} - C\hat{x}_{d,\max}\right)}{\partial d} = -C\frac{\partial \hat{x}_{d,\max}}{\partial d} \qquad (29)$$

where

$$\frac{\partial \hat{x}_{d,\max}}{\partial d} = \Delta_{\max} - \xi_d \frac{\partial J_{d-1}}{\partial \hat{\Theta}_{d-1}} g\left(\hat{x}_{d,\max}\right) \qquad (30)$$

Since Gaussian distribution and natural exponential function are positive function, therefore the inequalities above result in the following inequality.

$$\sum_{i,j=1}^{nk_t} (\tilde{y}_{d,i} - \tilde{y}_{d,j}) C \left(\Delta_{\max} - \xi_d \frac{\partial J_{d-1}}{\partial \hat{\Theta}_{d-1}} g(\hat{x}_{d,\max}) \right) < 0 \tag{31}$$

When $d = 1$, the initial values of the estimated parameters are bounded while the maximum element of the estimated parameters can be found. The learning rate can be confirmed while the performance criterion will decrease and the estimated parameters will go close to the true values. For each iteration $d = 2, 3, \cdots$, the estimation boundary will become smaller than the initial boundary which means that the $\hat{x}_{d,\max}$ is bounded. Denoting the maximum value of the upper boundary of $\hat{x}_{d,\max}$ as a real positive vector $\bar{M}$, then the conditions (16) and (17) are implementable while the proof is completed combining the inequalities as the conditions of the theorem. ∎

In the end, the procedure of the presented algorithm can be summarized by the following flow chart.
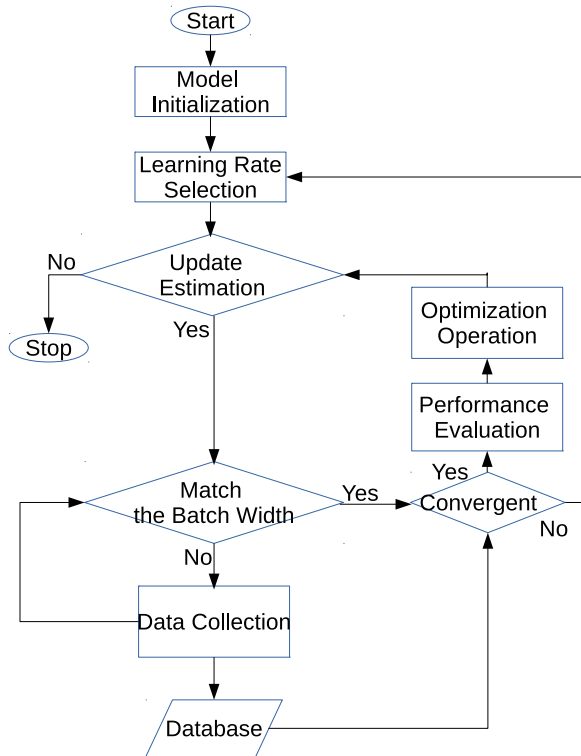

Fig. 2.   Flow chart of the presented iterative learning strategy.

Remark 2: Notice that (16) and (17) in Theorem 1 are posterior conditions, therefore the estimated parameter would not be updated if the collected data cannot satisfy this posterior condition for any iteration.

V. ONE APPLICATION IN BIOLOGICAL NEURAL SYSTEMS

As a typical biological system, the mechanism description of membrane potential generation for nerve fibre is modelled by Hodgkin-Huxley equation [12]. This model can be used to describe the individual response of the axon. Moreover, the extended Hodgkin-Huxley equation [13], [14], [15] is developed to describe the multi-axon behaviour where the mutual coupling factor matrix is introduced into the model to characterized the neural interaction. As a numerical example, the model is given for two coupled axons as follows.

$$c_{m_1} \frac{\partial V_{m_1}}{\partial t} = -\sum I_{ion_1} - g_{12} (V_{m_1} - V_{m_2})$$
$$c_{m_2} \frac{\partial V_{m_2}}{\partial t} = -\sum I_{ion_2} - g_{12} (V_{m_2} - V_{m_1}) \tag{32}$$

where the current of the ionic channels can be further stated by

$$I_{Na} = m^3 (V_m, t) h(V_m, t) \bar{g}_{Na} (V_m - V_{Na})$$
$$I_{Ka} = n^4 (V_m, t) \bar{g}_{Ka} (V_m - V_{Ka})$$
$$I_{leak} = \bar{g}_{leak} (V_m - V_{leak}) \tag{33}$$

while the parameters are given in [12]. In particular, the initial value of the membrane potential is $-60Vm$, $c_m = 0.01$, $V_{Na} = 55.17$, $V_{Ka} = -72.14$, $V_{leak} = -49.42$, $g_{Na,max} = 1.2$, $g_{Ka,max} = 0.36$, $g_{leak,max} = 0.03$.

The measured trans-membrane potential is formulated by the following equation.

$$\bar{V}_{m,k} = \sum_{i=1}^{2} V_{m_i,k} \tag{34}$$

where k stands for the sampling index and $V_{m_i,k}$ is the sampling of the membrane potentials in Eq.(32).

Based on the discretization operation, the membrane potential can be represented by the following model.

$$V_{m,k+1} = f(V_{m,k}) + \Xi g(V_{m,k}) + w_k$$
$$\bar{V}_{m,k} = \sum_{i=1}^{2} V_{m_i,k} + v_k \tag{35}$$

where

$$\Xi = \begin{bmatrix} 0 & g_{12} \\ g_{12} & 0 \end{bmatrix} \tag{36}$$

while the measured trans-membrane potential can be obtained with mutual coupling factor $g_{12} = 0.001$ which can be treated as a benchmark data set. Basically, we can obtain the data set and the mutual coupling factor is unknown in practice. w_k and v_k denote the modelling error including the non-Gaussian noises.

Following the presented algorithm, the results are shown by the following figures while the initial value of the mutual coupling factor is chosen as zero. Figure 3 shows that the response curves of these two axons. It has been shown that there is a time difference between two

spikes, therefore, the measured membrane potential is not in standard shape of membrane potential such as the curve in Fig.4. Rearranging this curve as a benchmark data set and selecting a proper learning rate $\xi_d = 4e{-}10$, the performance of the algorithm is illustrated by Fig. 5 and Fig. 6 where the value of the performance criterion decreases by the iterations. Particularly, the estimation error curves are convergent to zero in Fig. 6. Using the learning formula, the estimated coupling factor is obtained recursively and the curve in Fig. 7 demonstrates that the estimated coupling factor becomes closer and closer to the true value. In the end, this simulation shows that the presented iterative learning estimation algorithm is effective to characterize the neural interaction while the unknown mutual coupling factor can be estimated properly.
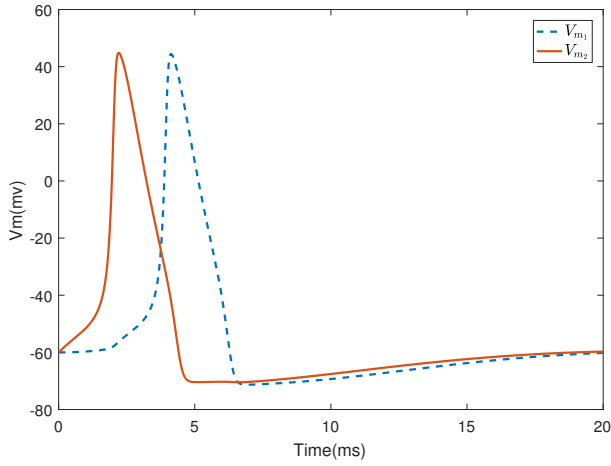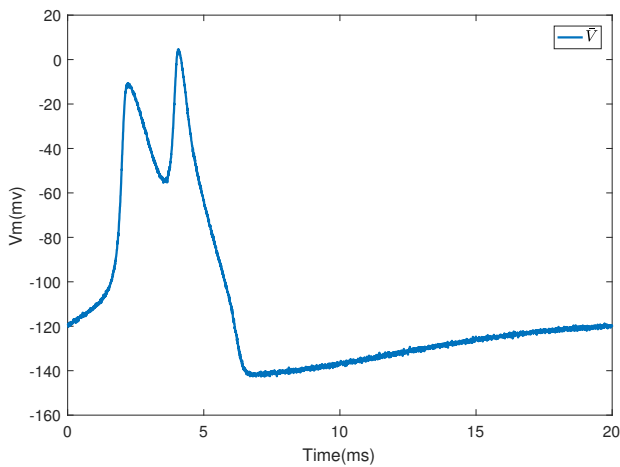

Fig. 5. The curve of the performance criterion.

Fig. 3. The response of two coupled axons.

Fig. 6. The estimation error sets with various batch index.

Fig. 4. The measured membrane potential.
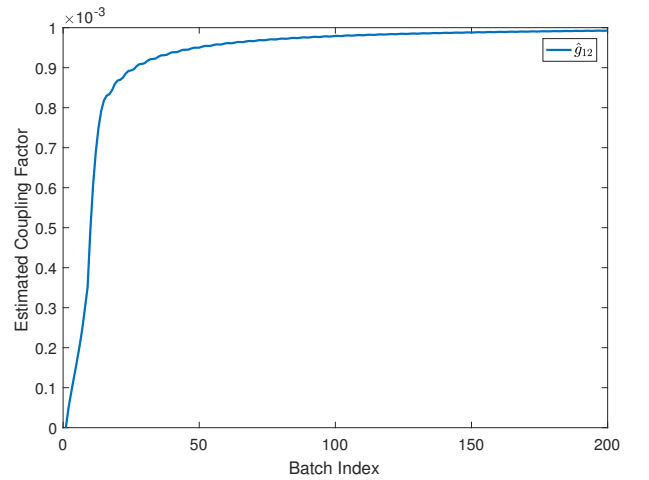
Fig. 7. The estimation of the mutual coupling factor.

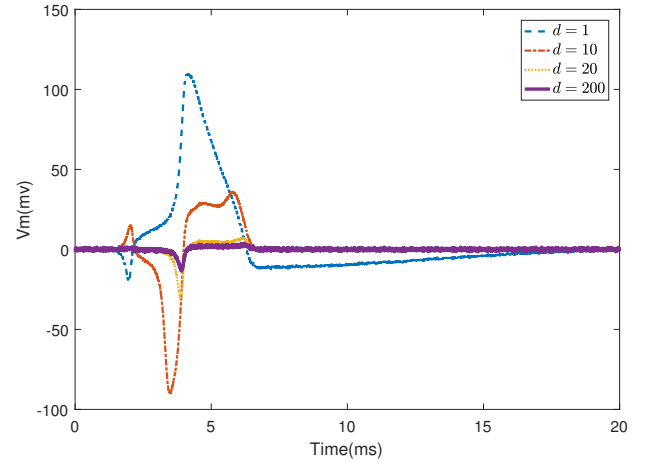Notice that there exists a error around $4ms$ which is very sensitive due to the depolarization. However, the estimation will not be affected if the stimuli of the membrane is not big enough to achieve action potential.

VI. DISCUSSIONS

In this paper, we assume that the unknown parameter is a real constant matrix which is time-invariant. However, a lot of piratical systems are with time-varying unknown parameter. Fortunately, this similar approach can be adopted if another learning operation can be added into the presented algorithm.

$$\hat{\Theta}_{d,k} = \hat{\Theta}_{d,k-1} - \varepsilon \frac{\partial J_{d,1:k-1}}{\partial \Theta}\bigg|_{\Theta=\hat{\Theta}_{d,k=1}} \qquad (37)$$

where $J_{d,1:k-1}$ denotes the subset of the estimation error and $\varepsilon > 0$ is the pre-specified learning rate. In other words, this iterative learning operation can be considered as the inner loop of the presented algorithm.

Another potential extension is for the partial differential equation. Following the similar idea, the estimation data set would be multi-dimensional rather than a vector which implies that the estimation error is subjected by multiple random variate. Therefore, the probability density function γ_d in the performance criterion should be replaced by the joint probability density function. The performance of the iterative learning estimation would be affected by the dependence of these random variates. Moreover, this problem can be simplified if the partial differential equation can be restated by the B-spline or RBF neural network decoupling model[16], [17], [18]. Then the estimation objective is to minimize the weight of the B-spline decoupling model using iterative learning approach. Note that this framework can also be adopted for industrial practical processes while the data learning can be combined with the observer [19] and filter [20] design to estimate the unmeasurable system variables.

VII. CONCLUSIONS

Parameter identification problem has been investigated in this paper. Different from the existing results, the stochastic non-linear systems are subjected to non-Gaussian noises. Based on the formulation of the problem description, the iterative learning algorithm is presented to estimate the unknown parameter with the entropy-based performance criterion. Furthermore, the convergence of the presented algorithm is analysed. Due to the fact that a lot of practical systems can be described by the non-Gaussian stochastic non-linear systems, we apply this algorithm to the biological neural system as a validation. The neural interaction widely exists among the axons, the time-invariant coupling factor matrix is used to characterize the mutual couplings based on the measured data. Following the presented algorithm, the estimated coupling factor is very close to the true value. Therefore, the simulation results illustrate the correctness and effectiveness of the presented algorithm. As a further discussion, the potential extensions of this algorithm are also given in this paper which can be considered as the future works particularly.

REFERENCES

[1] L. Ljung, *System identification*. Wiley Online Library, 1999.
[2] H.-S. Ahn, Y. Chen, and K. L. Moore, "Iterative learning control: Brief survey and categorization," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1099–1121, 2007.
[3] H. Wang and P. Afshar, "Ilc-based fixed-structure controller design for output pdf shaping in stochastic systems using lmi techniques," *IEEE transactions on automatic control*, vol. 54, no. 4, pp. 760–773, 2009.
[4] Y. Chen, C. Wen, H. Dou, and M. Sun, "Iterative learning identification," in *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, vol. 5. IEEE, 1997, pp. 4702–4707.
[5] M. Ren, Q. Zhang, and J. Zhang, "An introductory survey of probability density function control," *Systems Science & Control Engineering*, vol. 7, no. 1, pp. 158–170, 2019.
[6] H. Wang, "Minimum entropy control of non-gaussian dynamic stochastic systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 398–403, 2002.
[7] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
[8] J. C. Principe, *Information theoretic learning: Rényi's entropy and kernel perspectives*. Springer Science & Business Media, 2010.
[9] Q. Zhang and A. Wang, "Decoupling control in statistical sense: Minimized mutual information algorithm," *International Journal of Advanced Mechatronic Systems*, vol. 7, no. 2, pp. 61–70, 2016.
[10] P. Afshar, F. Yang, and H. Wang, "Ilc-based minimum entropy filter design and implementation for non-gaussian stochastic systems," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 4, pp. 960–970, 2012.
[11] P. Afshar, H. Wang, and T. Chai, "An ilc-based adaptive control for general stochastic systems with strictly decreasing entropy," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 471–482, 2009.
[12] A. L. Hodgkin and A. F. Huxley, "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, 1952.
[13] Q. Zhang and F. Sepulveda, "A statistical description of pairwise interaction between the nerve fibres," in *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*. IEEE, 2017, in press.
[14] ——, "A model study of the neural interaction via mutual coupling factor identification," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*. IEEE, 2017, pp. 3329–3332.
[15] ——, "On the conduction of nerve signals along coupled axons using a pairwise statistical description," in *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2017 9th International Conference on*, vol. 1. IEEE, 2017, pp. 183–186.
[16] H. Wang, *Bounded dynamic stochastic systems: modelling and control*. Springer Science & Business Media, 2012.
[17] Q. Zhang and F. Sepulveda, "Modelling and control design for membrane potential conduction along nerve fibre using b-spline neural network," in *Advanced Computational Methods in Life System Modeling and Simulation*. Springer, 2017, pp. 53–62.
[18] ——, "Rbfnn-based modelling and analysis for the signal reconstruction of peripheral nerve tissue," in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 2017, pp. 474–479.
[19] X. Dai, Z. Gao, T. Breikin, and H. Wang, "High-gain observer-based estimation of parameter variations with delay alignment," *IEEE Transactions on Automatic Control*, vol. 57, no. 3, pp. 726–732, 2011.
[20] Q. Zhang and L. Hu, "Probabilistic decoupling control for stochastic non-linear systems using ekf-based dynamic set-point adjustment," in *2018 UKACC 12th International Conference on Control (CONTROL)*. IEEE, 2018, pp. 330–335.