Hierarchical algorithm for the identification of parameter estimation of linear system

1st Mohammad Abu Jami'in
Politeknik Perkapalan Negeri Surabaya
Surabaya, Indonesia
jammysby@gmail.com

2nd Khairul Anam
Teknik Ekektro
Universitas Jember
Jember, Indonesia
khairul@unej.ac.id

3rd Riries Rulaningtyas
Teknik Biomedik
Universitas Airlangga
Surabaya, Indonesia
riries-r@fst.unair.ac.id

4th Mohammaderik Echsony
Teknik Komputer Kontrol
Politeknik Negeri Madiun
Madiun, Indonesia
erik-sony@pnm.ac.id

***Abstract*—A novel technique to identification of autoregressive moving average (ARMA)systems is proposed to increase the accuracy and speed of convergence for the system identification. The convergence speed of recursive least square algorithm (RLS) is solved under differential equations that needs all necessary information about the asymptotic behavior. Using RLS estimation, the convergence of parameters is able to the true values if the data of information vector growing to infinite. Therefore, the convergence of the parameters of the RLS algorithm takes time or needs a large number of sampling. In order to improve the accuracy and convergence speed of the estimated parameters, we propose a technique that modifies the QARXNN model by running two steps to identify the system hierarchically. The proposed method performs two steps: first, the system is identified by least square error (LSE) algorithm. Second, performs multi-input multi-output feedforward neural networks (MIMO-NN) to refine the estimated parameters by updating the parameters based on the residual error of LSE. The residual error by using LSE is set as target output to train NN. Finally, we illustrate and verify the proposed technique with an experimental studies. The proposed method can find the estimated parameters faster with δ= 0.935129 % in tenth sampling. The results is almost consistence which the accuracy of the identified parameters δ did not change significantly with the increasing number of sampling or the number of data points.**

***Index Terms*—System identification, hierarchical algorithm, quasi-linear ARX neural network, convergence speed, parameter estimation.**

I. Introduction

An identification system is a method for obtaining a mathematical model through measurements data. A techniques to identify a system plays an important role in technological development. It is widely applied in modern life such as biomedical engineering [1], control systems [2], signal processing [3], [4], image processing [5], [6], communication [7], and power system [8]. In a model based control, the performance of model affects the performance of a controlled system. The inaccuracy of parameters of a system will affect the accuracy and failure of a control system. This is because the law of the control system utilizes the parameters of an estimated system. Therefore, the accuracy of the identified parameters and speed of convergence are important instead of function approximation. The parameters of system can be used to estimate the dynamic behavior of system such as stability, the performance of dynamic response, controllability, observability, damping and natural frequency of system. Moreover, the calculation of the control signal is derived from the system parameters being controlled.

Many researchers have proposed various techniques to improve the performance index of convergence speed and parameter accuracy. The convergence of RLS estimation is a function over time. The analysis of convergence is solved by differential equation that needs all information about the asymptotic behavior of system [9]. In the deterministic case, the convergence of RLS tends to the true parameters when the data of information vector goes to infinity [10]. To improve the performance index under least square (LS) algorithm , some researchers propose a a method such as dichotomous coordinate descent recursive least square (DCD-RLS) algorithm [11], a technique of refined instrumental variable for continuous systems (RIVC) [12], A portable alternating least squares algorithm applied for factorization of parallel matrix [13].

A quasi linear-ARX neural network (QARXNN) consists of linear and nonlinear subsystem modeling [14], [15]. In nonlinear system identification, QARXNN is used for mapping the system with linear correlation between the regression vector and its coefficients. Nonlinear sub-model performed by multi-input multi-output feedforward neural networks (MIMO-NN) is used to parameterize the regression vectors. moreover, the coefficients has useful advantages: 1) easy to derive control law of nonlinear system under the inverse of the modeled system [16]–[19], 2) the dynamic analysis of nonlinear system can be approached under the law of linear system [17], [20], [21], 3) the coefficients can be used to analyze and check the stability nonlinear system [17], [18], [20].

In this paper, the modification of learning process and a technique to update parameters under quasi Linear-ARX Neural Network Model is proposed, namely hierarchical algorithm. The proposed method perform two steps identification hierarchically. First, the system is identified under least square error algorithm (LSE). Second, refining the estimated parameters based on the residual error of LSE algorithm. In the first step, we identify the system using linear algorithm to obtain the estimated parameters. The error of LSE algorithm is refined in the second step by running MIMO-NN and the estimated parameter is updated simultaneously. Thus, we

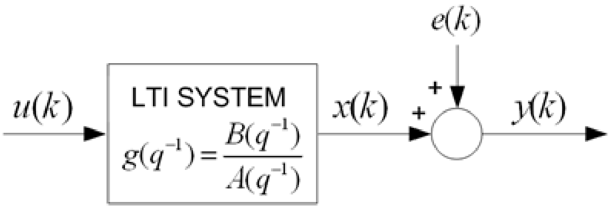

Fig. 1. LTI systems with coloured noises

sharpen the search of the estimated parameters of LSE method by running MIMO-NN training. The convergence which is a function of time on the LSE method is replaced through MIMO-NN training. We capture the dynamic behavior of the system by running MIMO-NN to update parameters of LSE algorithm. Finally, we demonstrate the proposed technique with the experiments and numerical simulations.

II. Problem Description

Assume that the unknown nth-order of discrete time linear system is presented by:

$$A(z^{-1})y(k) = B(z^{-1})u(k) + \omega(k) \tag{1}$$

where

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}$$
$$B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \cdots + b_m z^{-m}$$

The system (1) is shown in Fig. 1. The symbol z^{-l} is the notation of delay. The input and output of system are denoted by $u(k)$, $y(k)$.The variable of $\omega(k)$ denotes a stochastic white noise with zero mean and variance σ_ω^2. The order of system is denoted by n for denumerator and m is for numerator, which are assumed to be known. The $u(k)$ and $\omega(k)$ are independence and uncorrelated statistically. For initial conditions, the system is set as $y(k)$ = 0, $u(k)$ = 0, for $k < 0$.

In matrix equation system (1) is presented by

$$y(k) = \phi^T(k)\theta + \omega(k) \tag{2}$$

where

$$\theta = [a_1, a_2, ..., a_n, b_1, b_2, ..., b_m]^T \in R^{n+m}$$
$$\phi(k) = [-y(k-1)... - y(k-n)\; u(k-1)...u(k-m)]^T$$

θ denotes the parameter of syustem and $\phi(k)$ is a regression vector composed of delayed input-output data, respectively.

Assumption 1. The identified system is assumed stable satisfied by $A(z^{-1}) = 0$ with $\mid z \mid < 1$.

Assumption 2. The training data of input $u(k)$ and output $y(k)$ are bounded.

The system identification under LS algorithm is performed to minimize the index stated as

$$J_N(\theta) = \sum_{k=1}^{N} (y(k) - \phi^T(k)\theta)^2. \tag{3}$$

Under ergodic theorem, the asymptotic behaviors of LS estimation [22] are

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N}\phi^T(k)\phi(k) &= \lim_{N\to\infty} \hat{R}_{\phi\phi}(N) \\ &= R_{\phi\phi} = E[\phi^T(k)\phi(k)] \end{aligned} \tag{4}$$

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N}\phi^T(k)y(k) &= \lim_{N\to\infty} \hat{R}_{\phi y}(N) \\ &= R_{\phi y} = E[\phi^T(k)y(k)] \end{aligned} \tag{5}$$

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N}\phi^T(k)\omega(k) &= \lim_{N\to\infty} \hat{R}_{\phi\omega}(N) \\ &= R_{\phi\omega} = E[\phi^T(k)\omega(k)]. \end{aligned} \tag{6}$$

$\hat{\theta}_{LS}(N)$ is the estimated parameter, which can be described as

$$\hat{\theta}_{LS}(N) = \hat{R}_{\phi\phi}^{-1}(N)\hat{R}_{\phi y}(N) \tag{7}$$

If the number of sampling N reach to infinity then [22]:

$$\lim_{N\to\infty} \hat{\theta}_{LS}(N) = \theta + R_{\phi\phi}^{-1} R_{\phi\omega}. \tag{8}$$

By minimizing of performance index (3) the estimated parameters θ will be [22] :

$$\hat{\theta}_{LS}(N) = [\sum_{k=1}^{N} \phi^T(k)\phi(k)]^{-1}[\sum_{k=1}^{N} \phi^T(k)y(k)] \tag{9}$$

$\hat{R}_{\phi\phi}(N)$ denotes a covariance matrix which contains of the pass data input-output of system. related to (9) it is indispensable that $\theta_{LS}(N) \to \theta$ will be to the true value if the sampling number of $N \to \infty$ [23]. For $k = 1, 2, \cdots, N$, equation (2) can be stated in matrix equation as

$$y = \phi^T\theta + \omega \tag{10}$$

where

$$\begin{aligned} y &= [y(1), y(2), \cdots, y(N)]^T \\ \omega &= [\omega(1), \omega(2), \cdots, \omega(N)]^T \\ \Phi &= [\phi(1), \phi(2), \cdots, \phi(N)]^T \end{aligned}$$

The elements of Φ contains the regression vector until N sampling number written by
$\phi(1) = [-y(k-1), \cdots, -y(k-n), u(k-1), \cdots, u(k-m)]^T$
$\phi(N) = [-y(N-1), \cdots, -y(N-n), u(N-1), \cdots, u(N-m)]^T$. Running the identification process will be started after the number of sampling N equal or larger than $2m+2n$ when m and n the order of system's structure. [24].

In many applications it is important to estimate the parameter vector θ recursively (or on-line or sequentially) as more information becomes available. Several researchers develop the identification technique under recursive LS (RLS) algorithm to improve parameter accuracy and convergence speed [7], [11], [25]–[29]. The algorithms of RLS estimation is stated as [30].

$$\begin{aligned} \hat{\theta}_{LS}(k) &= \hat{\theta}_{LS}(k-1) - P(k)\phi(k) \\ &\quad (y(k) - \phi^T(k)\hat{\theta}_{LS}(k-1)) \end{aligned} \tag{11}$$

$$P(k) = P(k-1) - \frac{P(k-1)\phi(k)\phi^T(k)P(k-1)}{1 + \phi^T(k)P(k-1)\phi(k)} \tag{12}$$

where $P(0) = p_0 I$, p_0 is a positive constant number.

III. Quasi-ARX Neural Network Model (QARXNN)

Quasi linear-ARX Neural Networks model is divided into linear macro-part sub-model and nonlinear core part sub-model. The macro-part sub-model is to derive a nonlinear system into linear correlation where the coefficients of Taylor is the parameters for the regression vector. The core-part sub-model is to provide nonlinear coefficients for the input regression vector performed by multi input multi output neural networks. A nonlinear system stated as,

$$y(k) \quad = \quad g(\phi(k)) \tag{13}$$

where $g(\cdot)$ is a nonlinear function, $\phi(k) = [y(k-1) \cdots y(k-n_y)\ u(k-1) \cdots u(k-n_u)]^T$ denotes input vector, $y(k) \in R$ denotes the output and $k = 1, 2, \cdots$ denotes the sampling of time. Using Taylor expansion series, a nonlinear system can be derived as a linear correlation between the input vector and its coefficients. Hence, the system in (13) can expressed as a linear-like model [14], [15], [31]:

$$y(k) \quad = \quad \phi^T(k)\aleph(\xi(k)). \tag{14}$$

where $\aleph(\xi(k)) = [a_{(1,k)} \cdots a_{(n_k,t)}\ b_{(1,k)} \cdots b_{(n_u,k)}]^T$ is a nonlinear function of core part sub-model to parameterize the regression vector and $\xi(k) = [y(k-1) \cdots y(k-n_y)\ u(k-2) \cdots u(k-n_u)\ \nu(k)]^T$ is the input for core-part sub-model which virtual input $\nu(k)$ is added. The MIMO neural network or fuzzy model can be adopted as core-part sub-model. With NN set as a core-part sub-model, QARXNN model can be presented as

$$\begin{aligned} y(k) &= \phi^T(t)\aleph(\xi(k)) \\ \aleph(\xi(k), \Omega) &= W_2 \Gamma W_1(\xi(k)) + \theta \qquad (15) \\ &= \delta(\xi(k)) + \theta \qquad (16) \end{aligned}$$

where, $\Omega = \{W_1, W_2, \theta\}$ are set of network parameters, Γ is an operator of sigmoidal element for hidden nodes of core-part sub-model.

IV. Hierarchical Algorithm

Incorporating to hierarchical algorithm, a linear system is presented by

$$\begin{aligned} y(k) &= y_0 + \phi^T(k)\theta(\phi(k)) + \omega(k). \qquad (17) \\ \theta(\phi(k)) &= [a_1, a_2, ..., a_n, b_1, b_2, ..., b_m]^T \in R^{n+m} \\ \phi(k) &= [-y(k-1) ... - y(k-n)\ u(k-1) ... u(k-m)]^T \end{aligned}$$

$\omega(k)$ denotes a stochastic white noise with zero mean and variance σ^2_ω. The $\theta(\phi(k))$ denotes the estimated parameters where $\phi(k)$ is set as the input variable. The hierarchical processes for the updating of the estimated parameters is shown in Fig. 2. In the first step, we perform LSE algorithm to estimate the parameters denoted by $\theta_{LS}(N)$. The residual error of LSE e_{LS} is set as output for MIMO-NN of QARXNN model to estimate $\Delta\theta$ performed in the second step. QARXNN model is performed to increase the accuracy of the estimated parameters shown in Fig. 3. Finally, the estimated parameter is update by summing $\theta_{LS}(N)$ and $\Delta\theta$. At first step, LS algorithm is used to identify θ by minimizing a cost function of (18) in surface sub-model. The surface sub-model performed by using LS algorithm is presented in (19).

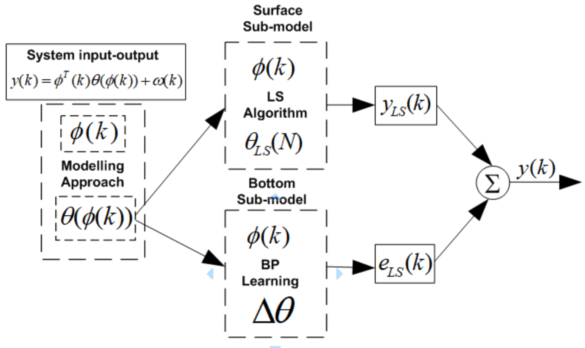

Fig. 2. The steps of hierarchical algorithm

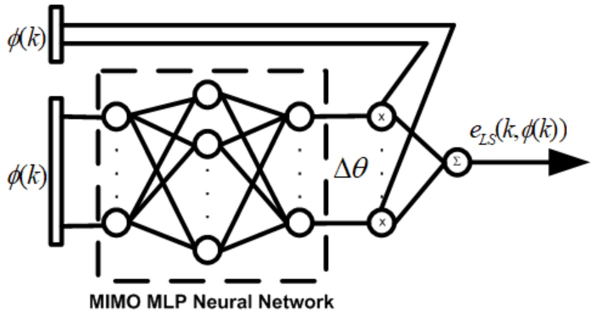

Fig. 3. Bottom sub-model incorporating to QARXNN.

$$J_N(\theta) = \sum_{k=1}^{N} (y(k) - \phi^T(k)\theta)^2. \tag{18}$$

$$y_{LS}(k) = \phi^T(k)\theta_{LS}(N). \tag{19}$$

Analytical minimisation of (18) leads to the least square (LS) estimate of θ as [23]:

$$\theta_{LS}(N) = [\sum_{k=1}^{N} \phi^T(k)\phi(k)]^{-1} [\sum_{k=1}^{N} \phi^T(k)y(k)] \tag{20}$$

where, N is a number of sampling in time moving window.

The residual error of LSE estimation is stated as,

$$e_{LS}(k) \quad = \quad y(k) - y_{LS}(k) \tag{21}$$

By substituting (19) to (21), we have (22) called as a bottom sub-model. It will be implemented under MIMO-NN of QARXNN model shown in Fig. 3.

$$\begin{aligned} e_{LS}(k) &= \phi^T(k)\theta - \phi^T(k)\theta_{LS}(N) \\ &= \phi^T(k)(\theta - \theta_{LS}(N)) \\ &= \phi^T(k)\Delta\theta(\phi(k)) \qquad (22) \end{aligned}$$

The $\Delta\theta^T(\phi(k))$ is a residual parameters which is the output of bottom sub-model. Hence. we update the estimated parameters $\hat{\theta}(k)$ by,

$$\hat{\theta}(k) = \theta_{LS}(N) + \Delta\theta(\phi(k)) \tag{23}$$

the estimated output of system will be

$$\hat{y}(k) = \phi^T(k)\hat{\theta}(k) \tag{24}$$

Remark 1: Using hierarchical algorithm, the convergence of LS algorithms in surface sub-model is improved by performing MIMO-NN in bottom sub-model. The residual errors of surface sub-model e_{LS} is used to update the estimated parameters $\hat{\theta}(k)$ under bottom sub-model by summing $\theta_{LS}(N)$ and $\Delta\theta$ in (23).

The LS algorithm is able to reach the true parameters when sampling data measurement tends to infinity. Thus, the longer memory of the past data of the information vector will be. Therefore, the convergence performance of LS algorithm is slow. Our motivation is to achieve faster convergence and accuracy with limited memory span by time moving window without sacrificing the simplicity of LS algorithm.

The proposed hierarchical algorithm, the estimated parameters is the summing between the surface sub-model under LS algorithm and the bottom sub-model performed by neural network. By performing the surface sub-model we can get the LS parameter estimate and the residual parameter of LS is performed using NN. The output of hierarchical learning for parameter estimation s is stated as,

$$\hat{y}(k) = \phi^T(k)\theta_{LS}(N) + \phi^T(k)\Delta\theta(\phi(k)). \tag{25}$$

Incorporating to MIMO-NN, the bottom sub-model is expressed as

$$\Delta\theta(\phi(k)) = W_2\Gamma W_1((\phi(k)) \tag{26}$$

the set of network parameters is denoted by $W_1, W_2 \in R^{(n+m)\text{x}(n+m)}$ that is the weight matrix at the the first second layer.

V. Learning Steps

We divide the system modeling into two sub-models. The surface sub-model will be performed using LSE algorithm and the bottom sub-model is done by MIMO-NN. The target output of surface sub-model is calculated by $s(k) = y(k) - e_{LS}$ and the target output for bottom sub-model is $b(k) = y(k) - y_{LS}(k)$. The target output for training of two sub-models are defined as,

$$SM1 \quad s(k) = \phi(k)\theta_{LS}(N). \tag{27}$$

$$SM2 \quad b(k) = \phi(k)\Delta\theta^T(\phi(k)). \tag{28}$$

The output of $SM1$ $s(k)$ is performed under LSE algorithm and the output of $SM2$ $b(k)$ is performed by MIMO-NN of quasi linear-ARX model. The learning processes of hierarchical algorithm are presented as

1) For initial condition, set e_{LS}=0 and set $i = 1$,i is the training sequence.
2) Estimate $\theta_{LS}(N)$ by using LSE algorithm for SM1.
3) Calculate the output of surface sub-model $s(k)$ in SM1. Set $s(k)$ as $y_{LS}(k)$ and calculate $b(k) = y(k) - y_{LS}(k)$. Use $b(k)$ as the target output for SM2
4) Estimate $\Delta\theta^T(\phi(k))$ using MIMO-NN of quasi linear-ARX model.
5) Update the estimated parameters $\hat{\theta}(k)$ using (23).
6) stop if a predetermined condition has been met such as the number of training or accuracy. Stop if the predetermined conditions are meet, otherwise go to 3). set $i = i + 1$.

VI. Example

Consider an autoregressive moving average system taken in [30] presented as

$$y(k) = \phi^T(k)\theta + \vartheta(k) \tag{29}$$

the parameters of the identified system are stated as

$$A(z^{-1})y(k) = B(z^{-1})u(k) + \vartheta(k) \tag{30}$$

$$A(z^{-1}) = 1 + a_1z^{-1} + a_2z^{-2} = 1 - 1.50z^{-1} + 0.60z^{-2}$$
$$B(z^{-1}) = b_1z^{-1} + b_2z^{-2} = 0.4z^{-1} + 0.3z^{-2}.$$

The parameter of θ and the regression vector $\phi(k)$ are defined by

$$\theta = [a_1, a_2, b_1, b_2]^T$$
$$\phi(k) = [-y(k-1), -y(k-2), u(k-1), u(k-2)]^T.$$

The $u(k)$ is an input of system with the zero mean and unit variance. The $\vartheta(k)$ is a white noise with zero mean and variance σ^2. The performance of the identification results are measured with the RMS error in (31) versus sampling times k. The performance of parameter accuracy is presented by δ in (32) versus sampling times.

$$RMS = \sqrt{\frac{\sum_{k=1}^{N}(y_p(k) - y(k))^2}{N}}. \tag{31}$$

$$\delta == \frac{\|\hat{\theta}(k) - \theta\|}{\|\theta\|}. \tag{32}$$

The output of system is mixed with noise signal perturbation ϑ of 20% or source to noise ratio (SNR) 13.98 dB.

$$SNR = 20\log\sqrt{(\frac{\sum_{k=1}^{N}x(k)^2}{\sum_{k=1}^{N}e(k)^2})}. \tag{33}$$

The SM2 is performed by MIMO-NN with the structure parameters is set as follows: n_u=2 and n_y=2, the input node $n = 4$ is the sum of n_u=2 and n_y=2, the number of training = 50. The 500 samplings of input-output data sequence shown in Fig. 4. The results of system identification is presented by the accuracy of output shown Fig. 5 and the accuracy of the estimated parameters shown in in Fig. 6.

The performance of estimated parameter is compared with the other measures shown in Table I.

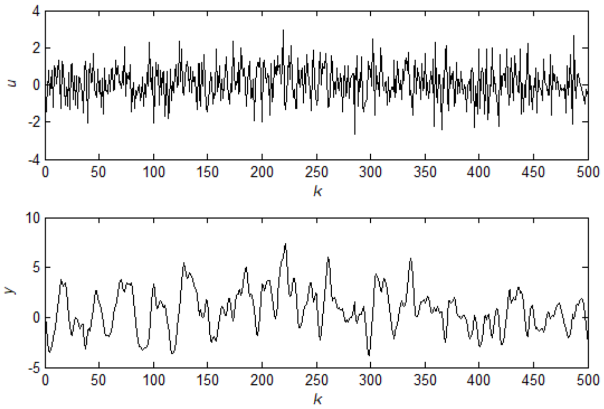

Fig. 4. The Input and Output of System

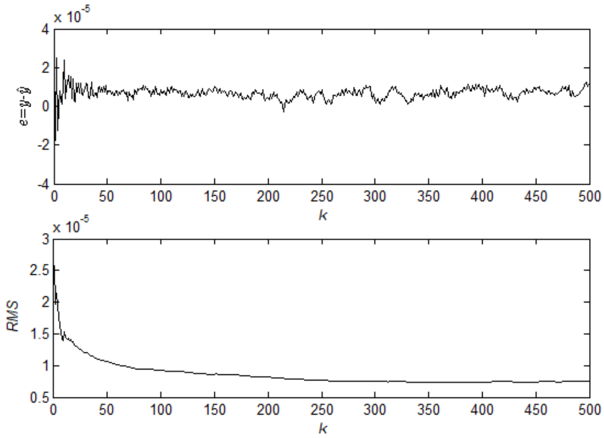

Fig. 5. Error and RMS error of the identified system

VII. Result and Discussion

In this novel, the fast convergence of the estimated parameter is discussed under two steps identification processes. In the first step, we identify the parameter of the system using LSE algorithm in surface sub-model. In the second step, a bottom sub-model is used to refine the estimated parameters using modified QARXNN model. The proposed algorithm has better results compared with the other measures. Based on the results of simulation the proposed method can find the estimated parameters faster compared with the others shown in Table I. We can get the estimated parameters with δ= 0.935129 % in tenth sampling. The results is almost consistence which the accuracy of the identified parameters δ did not change significantly with the increasing number of sampling or the number of pass input-output data.

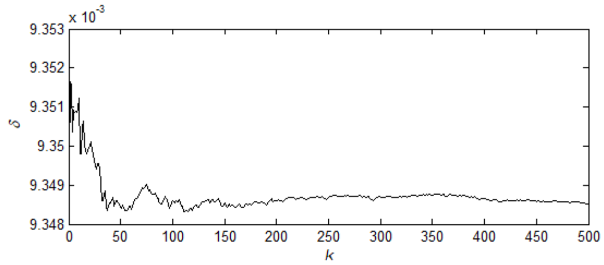

Fig. 6. Parameter estimation error

TABLE I
δ INDEX PERFORMANCE OF PARAMETER ESTIMATION.

k	Proposed	V-RLS	V-MILS [30]
10	0.935129	-	-
100	0.934860	2.52165	1.69763
200	0.934861	2.67837	0.81472
300	0.934869	1.65349	0.60837
400	0.934866	0.99951	0.60769
500	0.934853	0.91778	0.53914

In the first step, the parameters of system is identified under LSE algorithm. The LSE algorithm updated the parameters based on the error by $e_{LS}(k) = (y(k) - \phi^T(k)\hat{\theta}_{LS}(k-1))$. However, the performance of LS algorithm is low in accuracy and slow convergence. The LS algorithm reaches the convergence when the data of information vector goes to infinity. In order to improve the accuracy and convergence speed of system identification the error of e_{LS} is refined using neural network MIMO-NN injected to QARXNN. Under hierarchical algorithm with the proposed method, the convergence by time using LS can be improved by the number of training.

References

[1] H.-L. Wei, Y. Zheng, Y. Pan, D. Coca, L.-M. Li, J. E. W. Mayhew, and S. A. Billings, "Model estimation of cerebral hemodynamics between blood flow and volume changes: A data-based modeling approach," *IEEE Trans. on Biomedical Engineering*, vol. 56(6), pp. 1606–1616, 2009.

[2] Q. Sang and G. Tao, "Gain margins of adaptive control systems," *IEEE Trans. on Automatic Control*, vol. 55(1), pp. 104–115, 2010.

[3] W. Bobillet, R. Diversi, E. Grivel, R. Guidorzi, M. Najim, and U. Soverini, "Speech enhancement combining optimal smoothing and errors-in-variables identification of noisy AR processes," *IEEE Trans on Signal Processing*, vol. 55(12), pp. 5564–5578, 2007.

[4] F. Al-Obeidat, B. Spencer, and O. Alfandi, "Consistently accurate forecasts of temperature within buildings from sensor data using ridge and lasso regression," *Future Generation Computer Systems*, 2018.

[5] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Trans. on Image Processing*, vol. 17(6), pp. 887–896, 2008.

[6] H. Lu, Y. Li, T. Uemura, H. Kim, and S. Serikawa, "Low illumination underwater light field images reconstruction using deep convolutional neural networks," *Future Generation Computer Systems*, vol. 82, pp. 142–148, 2018.

[7] T. K. Akino, "Optimum-weighted RLS channel estimation for rapid fading MIMO channels," *IEEE Trans. on Wireless Communications*, vol. 7(11), pp. 4248–4260, 2008.

[8] N. Zhou, D. J. Trudnowski, J. W. Pierre, and W. A. Mittelstadt, "Electromechanical mode online estimation using regularized robust RLS methods," *IEEE Trans. on Power Systems*, vol. 23(4), pp. 1670–1680, 2008.

[9] L. Ljung, "On positive real transfer functions and the convergence of some recursive schemes," *IEEE Trans. on Auitomatic Control*, vol. AC-22(4), pp. 539–551, 1977.

[10] K. Nassiri-Toussi and W. Ren, "On the convergence of least squares estimates in white noise," *IEEE Trans. on Automatic Control*, vol. 39(2), pp. 364–368, 1994.

[11] M. Algreer, M. Armstrong, and D. Giaouris, "Active online system identification of switch mode DCDC power converter based on efficient recursive DCD-IIR adaptive filter," *IEEE Trans. on Power Electronics*, vol. 27(11), pp. 4425–4435, 2012.

[12] X. Liu, J. Wang, and W. Zheng, "Convergence analysis of refined instrumental variable method for continuous-time system identification," *IET Control Theory Appl.*, vol. 5(7), pp. 868–877, 2011.

[13] J. Chen, J. Fang, W. Liu, T. Tang, and C. Yang, "clmf: A fine-grained and portable alternating least squares algorithm for parallel matrix factorization," *Future Generation Computer Systems*, 2018.

[14] J. Hu, K. Kumamaru, and K. Hirasawa, "A Quasi-ARMAX approach to modelling of non-linear systems," *Int. J. Control*, vol. 74(18), pp. 1754–1766, 2001.

[15] M. A. Jami'in, I. Sutrisno, and J. Hu, "Lyapunov learning algorithm for quasi-ARX neural network to identification of nonlinear dynamical system," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics (Seoul)*, 2012, pp. 3141–3146.

[16] J. Hu and K. Hirasawa, "A method for applying neural networks to control of nonlinear systems," *Neural Information Processing: Research and Development, Springer, Berlin, GERMANY*, pp. 351–369, 2004.

[17] M. A. Jami'in, I. Sutrisno, J. Hu, N. B. Mariun, and M. H. Marhaban, "Quasi-arx neural network based adaptive predictive control for nonlinear systems," *IEEJ Trans. on Electrical and Electronic Engineering*, vol. 11(1), pp. 83–90, 2016.

[18] M. A. Jami'in, J. Hu, and E. Julianto, "A lyapunov based switching control to track maximum power point of wecs," in *Procedings of the International Joint Conference on Neural Networks (IJCNN 2016)*, 2016, pp. 3883–3888.

[19] M. A. Jami'in, Yuyun, and E. Julianto, "Hierarchical algorithms of quasi-linear arx neural networks for identification of nonlinear systems," *Engineering Letters*, vol. 25(3), pp. 321–328, 2017.

[20] M. A. Jami'in, I. Sutrisno, J. Hu, N. B. Mariun, and M. H. Marhaban, "An adaptive predictive control based on a quasi-arx neural network model," in *The 13th International Conference on Control, Automation, Robotics and Vision, ICARCV 2014*, 2014, pp. 253–248.

[21] M. Jami'in, I. Sutrisno, and J. Hu, "Maximum power tracking control for a wind energy conversion system based on a Quasi-ARX neural network model," *IEEJ Trans. on Electrical and Electronic Engineering*, vol. 10(4), pp. 368–375, 2015.

[22] M. Davis and R. Vinter, *Stochastic modelling and control*. Chapman and Hall, London, 1985.

[23] C. B. Feng and W.-X. Zheng, "Robust identification of stochastic linear systems with correlated noise," *IEE Procedings-D*, vol. 138(5), pp. 484–492, 1991.

[24] V. Strejc, "Least squares parameter estimation," *Automatica*, vol. 16, pp. 535–550, 1980.

[25] F. Ding, T. Chen, and L. Qiu, "Bias compensation based recursive least-squares identification algorithm for MISO systems," *IEEE Trans. on Circuits and Systems-II: Express Briefs*, vol. 53(5), pp. 349–353, 2006.

[26] D.-Z. Feng and W. X. Zheng, "Fast RLS-Type algorithm for unbiased equation-error adaptive IIR filtering based on approximate inverse-power iteration," *IEEE Trans. on Signal Processing*, vol. 53(11), pp. 4169–4185, 2005.

[27] S. Fattah, W.-P. Zhu, and M. Ahmad, "Identification of autoregressive moving average systems based on noise compensation in the correlation domain," *IET Signal Process.*, vol. 5(3), pp. 292–305, 2011.

[28] H.-F. Chen, "New approach to recursive identification for ARMAX systems," *IEEE Trans. on Automatic Control*, vol. 55(4), pp. 868–879, 2010.

[29] A. Ghoreyshi and H. L. Galiana, "Simultaneous identification of oculomotor subsystems using a hybrid system approach: Introducing hybrid extended least squares," *IEEE Trans. on Biomedical Engineering*, vol. 57(5), pp. 1089–1098, 2010.

[30] F. Ding, P. X. Liu, and G. Liu, "Multiinnovation least-squares identification for system modeling," *IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 40(3), pp. 767–778, 2010.

[31] L. Wang, Y. Cheng, and J. Hu, "A quasi-ARX neural network with switching mechanism to adaptive control of nonlinear systems," *SICE Journal of Control, Measurement, and System Integration*, vol. 3(4), pp. 246–252, 2010.