Control of a Non Observable Double Inverted Pendulum Using a Novel Active Learning Method based State Estimator

Alireza Ghatreh Samani, Saeed Bagheri Shouraki, Reza Eghbali, Mohammad Ghomi Rostami
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
alirezasamani1988@gmail.com, bagheri-s@sharif.edu, eghbali@ee.sharif.edu, ghomirostami@ee.sharif.edu

***Abstract*—In this paper a novel fuzzy approach exploiting Active Learning Method is employed in order to estimate the immeasurable states required to control a non-observable double inverted pendulum. Active Learning Method (ALM) is a fuzzy modeling method which exploits Ink Drop Spread (IDS) as its main engine. IDS is a universal fuzzy modeling technique which is very similar to the way human brain processes different phenomena. The ALM system is trained by the data obtained from Linear Quadratic Regulator (LQR) controller. LQR uses an optimal control approach which under certain conditions guarantees robustness. Instead of an expert's knowledge, the LQR controller output is used as a priori knowledge to train ALM. The application of ALM method is then investigated in conditions where some states of system like the upper angle of the pendulum and its angular velocity are not available and the proposed system is not observable. The fact of practical non-observablity of this system obliges us to use an open-loop state estimator to estimate the missing states. Instead, a novel state estimator using ALM is introduced which shows practical superiority in estimation.**

Keywords- Double inverted pendulum; Active Learning Method; Ink Drop Spread; Linear Quadratic Regulator; State estimator

I. INTRODUCTION

Inverted pendulum and double inverted pendulum always served as a benchmark for evaluating various control approaches. Controlling a double inverted pendulum is more investigated because of its higher complexity and its further applications in the everyday life. The complexity of this control problem grows further when all the angles of the pendulum rods cannot be monitored by sensors and the proposed system is not observable. These situations frequently occur in real applications; For instance, when we try to carry a load by a Segway robot we model it by a double inverted pendulum and consider the load as the upper rod connected to the first rod, but in fact we cannot use any sensors to monitor the deviation amount of the load and the system is non-observable. This fact forces us to use observers and state estimators in order to estimate the states which cannot be directly sensed. Because the system is non-observable, only classic open loop observers can be applied which may lead to instability due to the undesirable position of poles and zeros existing in the state space model. In the proposed method, we utilize a novel fuzzy observer which estimates the immeasurable states based on Active Learning Method (ALM) [1]. To utilize the ALM to estimate the states we first control the fully observable system by using a Linear Quadratic Regulator (LQR) in order to obtain the data necessary for training ALM [2]. After controlling the observable system, we use the controller output and system states to train the ALM inference engine which allows us to estimate the missing states. The estimated states can then be applied to the same LQR controller which was used previously. This paper is organized as follows: Firstly, dynamic and linearized equations of double inverted pendulum are investigated. Secondly the LQR controller is introduced which is then applied to the linearized system. Thirdly, the ALM inference system of the fuzzy observer is introduced which will be constructed based on the results of classic LQR controller which serve as a priori knowledge. Finally, the simulations and results are presented.

II. DYNAMIC AND LINEARIZED EQUATIONS OF A DOUBLE INVERTED PENDULUM

The double inverted pendulum system illustrated in Fig.1 consists of two rods with uniformly distributed weights which are connected to each other by a joint and the lower one is connected to a cart via another joint. Here, we assume that the two rods have the same weight and length. The notations are specified as follows:

m_1 = mass of first rod
m_2 = mass of second rod (equal to m_1)
m = mass of the cart
l = distance between center of gravity of each rod and joints
x = the horizontal distance from the origin
θ_1 = the angle between the first rod and the vertical axis
θ_2 =the angle between the second rod and the vertical axis

978-0-7695-4308-6/10 $26.00 © 2010 IEEE
DOI 10.1109/EMS.2010.17

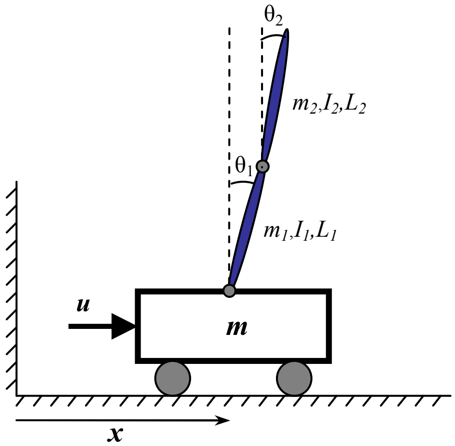

Figure1. Double inverted pendulum on a cart

u = the control force
I_1 = the inertial momentum of first rod
I_2 = the inertial momentum of second rod
L_1 = the length of the first rod
L_2 = the length of the second rod

According to Bogdanov et al. [3] the system equations can be summarized by using Lagrange equations as follows:

$$
\begin{aligned}
u \ &= \left(\sum m_i\right) \ddot{x} + (m_1 l_1 + m_2 L_1)\cos(\theta_1)\,\ddot{\theta}_1 \\
&\quad + m_2 l_2 \cos(\theta_2)\,\ddot{\theta}_2 - (m_1 l_1 + m_2 L_1)\sin(\theta_1)\,\dot{\theta}_1^2 \\
&\quad - m_2 l_2 \sin(\theta_2)\,\dot{\theta}_2^2 \\
0 \ &= (m_1 l_1 + m_2 L_2)\cos(\theta_1)\,\ddot{x} \\
&\quad + (m_1 l_1^2 + m_2 L_1^2 + I_1)\ddot{\theta}_1 \\
&\quad + m_2 L_1 l_2 \cos(\theta_1 - \theta_2)\,\ddot{\theta}_2 \\
&\quad + m_2 L_1 l_2 \sin(\theta_1 - \theta_2)\,\dot{\theta}_2^2 \\
&\quad - (m_1 l_1 + m_2 L_1)\, g \sin(\theta_1) \\
0 \ &= m_2 l_2 \cos(\theta_2)\,\ddot{x} + m_2 L_1 l_2 \cos(\theta_1 - \theta_2)\,\ddot{\theta}_1 \\
&\quad + (m_2 l_2^2 + I_2)\ddot{\theta}_2 - m_2 L_1 l_2 \sin(\theta_1 - \theta_2)\,\dot{\theta}_1^2 \\
&\quad - m_2 l_2\, g \sin(\theta_2)
\end{aligned}
$$

Because we used identical uniformly distributed masses, L_1 and L_2 are exactly twice the magnitude of l. If we choose the state vector as $\mathrm{x} = (x\ \dot{x}\ \theta_1\ \dot{\theta}_1\ \theta_2\ \dot{\theta}_2)$ by linearization around zero we can obtain the following state space equations:

$$\dot{\mathrm{x}} = A\mathrm{x} + B\mathrm{u} \tag{1}$$

where A and B stand for

$$
A = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -\frac{135 m_1 g}{56 m_1 + 97 m} & 0 & -\frac{135 m_1 g}{56 m_1 + 97 m} & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & -\frac{9(7m + 11 m_1) g}{l_1(56 m_1 + 97 m)} & 0 & -\frac{9(2m + m_1) g}{l_1(56 m_1 + 97 m)} & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & -\frac{27(2m + m_1) g}{l_2(56 m_1 + 97 m)} & 0 & -\frac{3(19m + 11 m_1) g}{l_2(56 m_1 + 97 m)} & 0
\end{pmatrix} \tag{2}
$$

$$
B = \begin{pmatrix}
0 \\
\frac{97}{56 m_1 + 97 m} \\
0 \\
-\frac{45}{l_1(56 m_1 + 97 m)} \\
0 \\
-\frac{3}{l_1(56 m_1 + 97 m)}
\end{pmatrix} \tag{3}
$$

III. LINEAR QUADRATIC REGULATOR CONTROLLER

Linear Quadratic Regulator (LQR) is a controller which is based on minimizing a control cost function [3,4]. This function denoted by J_{LQR} is usually the integral of a combination of control input and outputs over time:

$$J_{LQR} = \int_0^\infty \|\mathrm{z}(t)\|^2 + \rho\|u(t)\|^2 \ dt \tag{4}$$

In (4), the operand $\|.\|$ denotes the Euclidian norm; $\mathbf{z}(t)$ denotes a combination of output signals at time t and is defined by the equation below:

$$\mathrm{z} = G\mathrm{x} + Hu \tag{5}$$

In (5), G and H are matrixes of appropriate dimensions which define which states and control inputs will be present in z(t).

In (4), $u(t)$ denotes the control output and ρ is a factor by which we can decide which one of $u(t)$ or z(t) will be more minimized through the rest of the process. A large ρ emphasizes the role of $u(t)$ in (4) which results in minimizing $u(t)$ greatly at the expense of bigger z(t). Also a small ρ will result the controlled output, z(t), to be very small at the expense of large $u(t)$. Since we choose z to be exactly our states we wish to minimize z(t) as much as possible and consequently a small ρ is chosen. According to Kalman's inequality theorem because $H'G = 0$ the system stays stable [4]. Below, we used a more general form of (4):

$$J_{LQR} = \int_0^\infty \mathrm{z}(t)' Q \mathrm{z}(t) + \rho u(t)' R u(t)\, dt \tag{6}$$

where Q and R are symmetric positive-definite square cost matrixes with appropriate dimensions. The initial choice of Q and R is done based on Bryson's rule [5] which provides satisfactory results for our purposes and in this particular problem eliminate the need for any further attempt to obtain these matrixes. Based on this rule, Q and R are diagonal matrixes such that:

$$Q_{ii} = \frac{1}{\text{maximum acceptable value of } z_i^2} \tag{7}$$

$$R_{ii} = \frac{1}{\text{maximum acceptable value of } u_j^2} \tag{8}$$

According to Bogdanov et al. the control input, $u(t)$, will be a combination of states [3]:

$$u = -K\text{x} \tag{9}$$

Where the matrix K stands for:

$$K = (H'QH + \rho R)^{-1}(B'P + H'QG) \tag{10}$$

And the matrix P above can be obtained by solving the following Riccati equation:

$$A'P + PA + G'QG - (PB + G'QH)(H'QH + \rho R)^{-1}(B'P + H'QG) = 0 \tag{11}$$

After controlling the system by LQR controller, all the simulation data obtained are used to train the ALM observer. These data comprise all system states and control output.

IV. ACTIVE LEARNING METHOD

Active Learning Method (ALM) is a universal fuzzy modeling method [6].This method has two main features that make it very similar to the way that human mind processes different phenomena [6]. ALM mainly deals with phenomena modeled as Multi Input Single Output (MISO) systems and assumes that the proposed MISO systems comprise several Single Input Single output (SISO) systems, so ALM only deals with one input and one output at a time. This feature makes ALM very similar to the human brain. In addition, ALM uses a novel fuzzy interpolation method called Ink Drop Spread (IDS) [6,7] to construct the individual SISO systems. To model and estimate a MISO system, ALM decomposes it into many SISO systems. Then as shown in Fig. 2, each SISO system is expressed in a data plain created by gathering data from system when it is fed by different inputs. Each data plain provides two distinct features of the given data; the first feature is a curve called narrow path that is extracted from IDS method which will be discussed in the next section. The proposed narrow path indicates output behavior regarding to only one input; In other words, each narrow path determines the best function describing the relation output and one specific input. The other feature is called the spread of the points around narrow path. Spread demonstrates the deviation of the data points from the narrow path. In other words, spread determines the degree of the dependency of the output on each specific input[6]. As discussed later, this degree can be used to obtain the firing rate of each fuzzy rule [6].

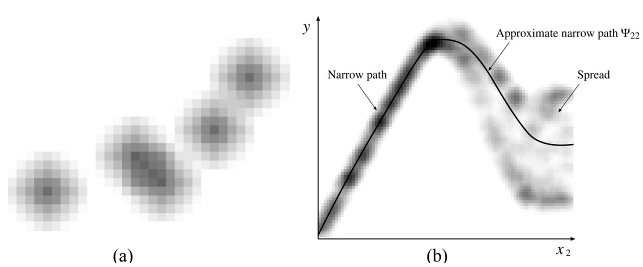

Figure 2. IDS plane: (a) Five data points spread on a plane (Courtesy of Murakami [8]) (b) Ink spreads and narrow path (Courtesy of Murakami[8])

A. Ink Drop Spread

IDS is the core part of ALM which extracts narrow paths and spreads [8]. In IDS we assume that every point in the input-output plain acts as a light source. The illuminating pattern of the light source has conical shape and is considered as a spatial fuzzy membership function. The illuminating patterns of different points are superposed forming new bright areas [9,10]. This concept is illustrated in Fig. 3. Finally, the narrow path is extracted by applying the center of gravity defuzzification method. After the extraction of narrow path, the spread value for every point in the narrow path is obtained by measuring the deviation of data points around the narrow path.

B. Combination of Fuzzy Rules

According to Shouraki et al. [9] ALM fuzzy rules have a general form shown below and the fuzzy scaling factors for the ruleswill be obtained by using the difference between the membership degree of the points on the narrow path and the membership degree of their neighbor points on the plane:

IF : x_1 is A_{1i} , … AND x_n is A_{nm}
THEN y is $\beta_{1i}U_{1i}$, … OR $\beta_{nm}U_{nm}$ (12)

where A_{nm} stands for the mth domain of nth input, U_{nm} stands for the narrow path related to nth input in its mth domain and the fuzzy number β_{nm} is representative of the deviations of the data points from the narrow path.

The IDS method may be used as a processing engine for the combination rule which can also be represented mathematically as follows:

IF : x_1 is A_{1i} , … AND x_n is A_{nm}

$$\text{THEN: } y \text{ is } \frac{\frac{1}{\alpha_{1i}}U_{1i}}{\frac{1}{\alpha_{1i}}+\cdots+\frac{1}{\alpha_{nm}}}, \ldots, \text{ OR } \frac{\frac{1}{\alpha_{1i}}U_{nm}}{\frac{1}{\alpha_{1i}}+\cdots+\frac{1}{\alpha_{nm}}} \tag{13}$$

where α_{nm} is the average of the spread function in mth domain of nth input calculated by center of area method.

V. DISCRIPTION OF ALM OBSERVER

In our double inverted pendulum system we can measure x, $\dot{x}$, θ_1, $\dot{\theta}_1$ and u by sensors, but we have no sensors to sense θ_2 and $\dot{\theta}_2$ which leads to system non-observability. According to the previous sections this situation is common when we want to control a Segway robot with a load that can slip on the top of it. Practically no sensor can be established to measure the angular deviation of loads with arbitrary shapes.

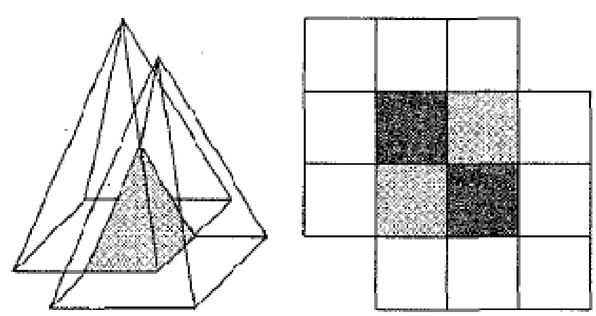

Figure 3. Each data point has a pyramidal fuzzy membership [10]

Classic closed loop observers cannot be used in this case because the stability of the system is only guaranteed under the condition of observability and then, a nonlinear fuzzy observer helps us in estimation of non-measurable states.

A. ALM Observer Training

As shown in Fig. 4, ALM observer is trained to estimate θ_2 and $\dot{\theta}_2$ by use of other states and system input. The training data are obtained from the states of the system after utilizing LQR control. Five narrow paths and spreads for θ_2 are extracted regarding to the x, $\dot{x}$, θ_1, $\dot{\theta}_1$ and u. Then we have extracted another five narrow paths and spread functions for $\dot{\theta}_2$ regarding to the same set of states and system input (u). Exploiting ALM fuzzy membership functions, θ_2 and $\dot{\theta}_2$ are estimated in each sample time and are then provided for the LQR state feedback controller. The results obtained from this step have been used to train the ALM algorithm again, therefore new narrow paths and spread were extracted and then by means of LQR method a new feedback gain is also calculated.

B. Impelementation of ALM Observer

According to the procedure explained in section IV, once ALM is trained, it can be used to estimate a system. In our system ALM observer estimates θ_2 and $\dot{\theta}_2$ in each sample time from the previous amount of x, $\dot{x}$, θ_1, $\dot{\theta}_1$ and u. Each of these states when projected to its narrow path gives an estimated amount for θ_2 and $\dot{\theta}_2$ ($\hat{\theta}_2$ and $\hat{\dot{\theta}}_2$) .These different amounts are then combined with a weighted average explained in section IV. At last the estimated states and other states are fed to the LQR controller.

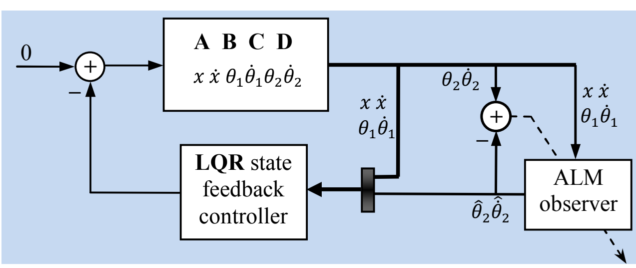

Figure 4. Block diagram of simulated control system

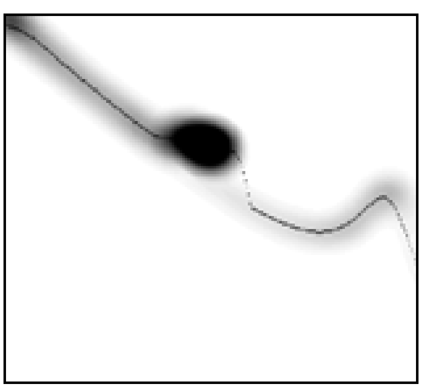

Figure 5. An example of a narrow path in a two- dimensional data plane

VI. Simulation and Result

To illustrate the performance of the ALM estimator we have made a comparison between the simulation results for the system that uses ALM observer and the system which uses Kalman filter estimator [11]. In both situations, the system is linearized around the zero bias point using (1), (2) and (3). Here, we assume m_1=2 kg, m= 20 kg and l= 0.3 m for masses and lengths.

We have chosen the maximum acceptable amount of each angle to be 0.1 (Rad) and the maximum acceptable angular velocity to be 0.02 (Rad/s). We also selected the relatively small amount of 0.15 for ρ in order to obtain smaller output values. Using these amounts the cost matrixes Q and R in (6) are calculated using (7) and (8). The gain matrix, K, is calculated using (10) and (11) and is used to obtain the control output mentioned in (9). All the simulation are conducted using Simulink and MATLAB. Fig. 6 compares the estimation of θ_2 with ALM state estimator, Kalman filter and original θ_2 obtained from a fully observable system. Fig. 7 illustrates the response of the system when ALM estimated states have been used as the input to the LQR state feedback with zero reference signal for θ_1 and θ_2. Finally, Fig. 8 presents the responses when we have used Kalman filter instead of ALM estimator.

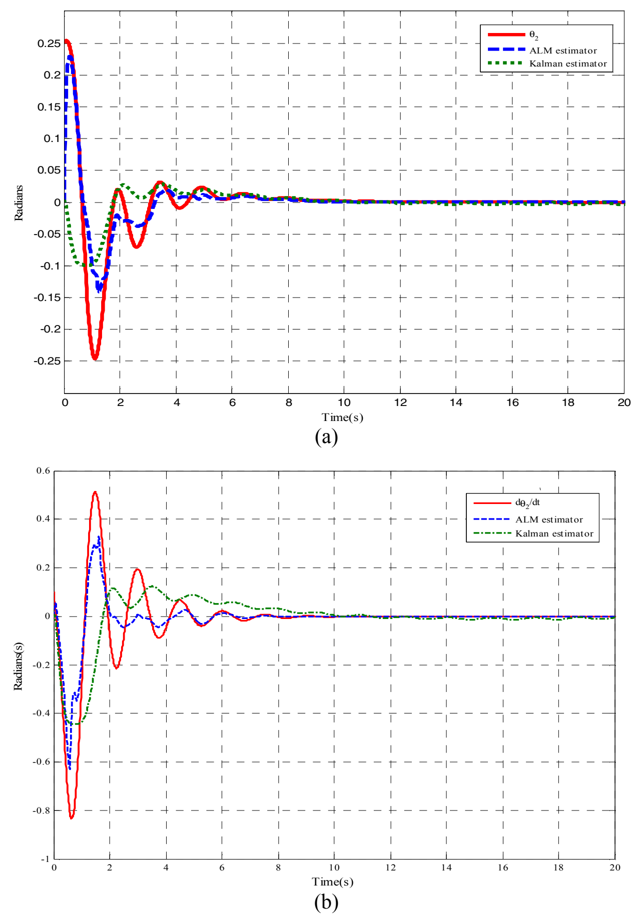

Figure 6. Comparison between ALM state estimator and Kalman state estimator: (a) $\boldsymbol{\theta_2}$ observation (b) $\boldsymbol{\dot{\theta}_2}$ observation

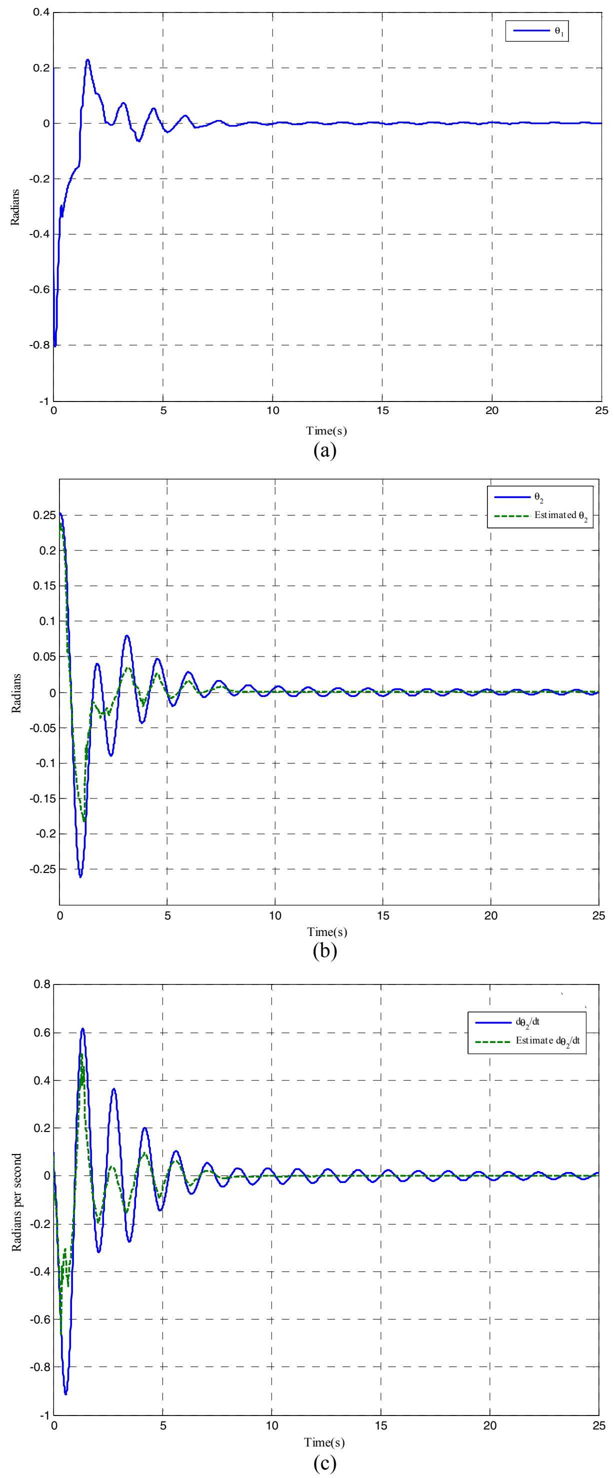

Figure 7. System control using ALM state estimator: (a) controlled $\boldsymbol{\theta_1}$ (b) controlled $\boldsymbol{\theta_2}$ (c) controlled $\boldsymbol{\dot{\theta}_2}$

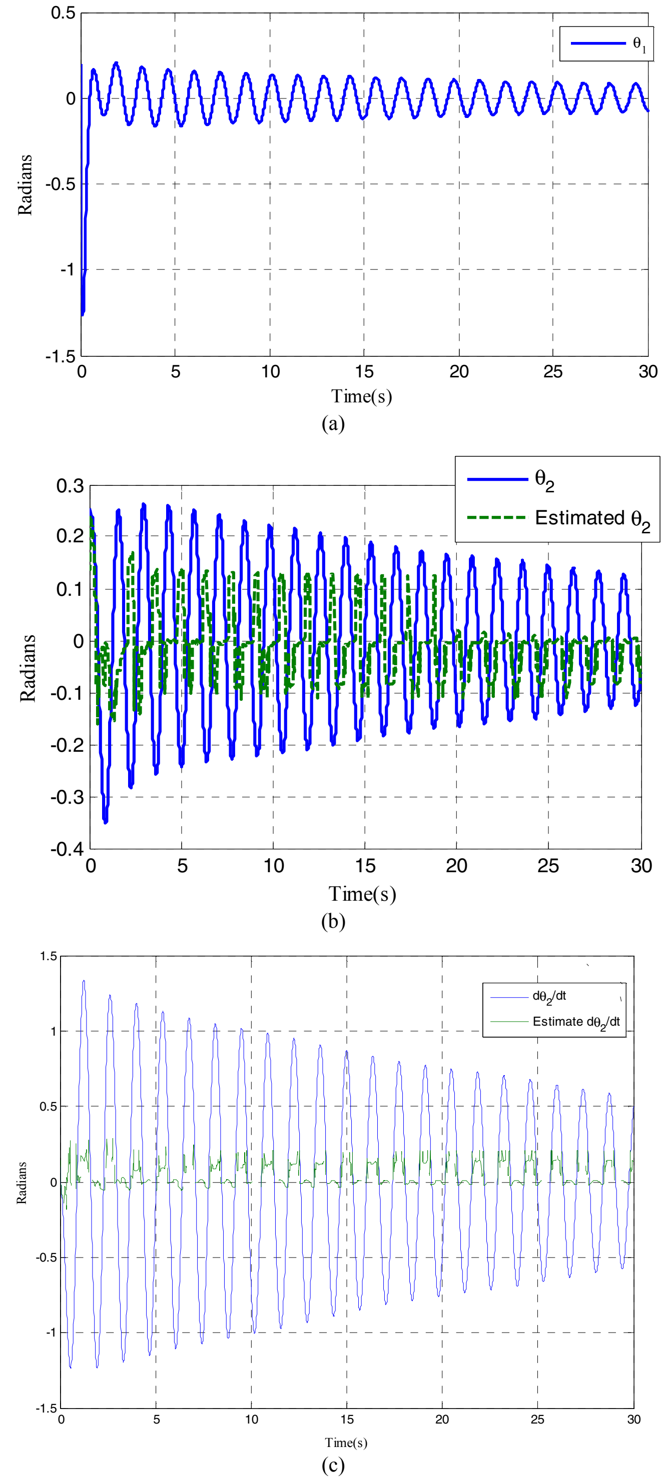

Figure 8. System control using Kalman estimator: (a) controlled $\boldsymbol{\theta_1}$ (b) controlled $\boldsymbol{\theta_2}$ (c) controlled $\boldsymbol{\dot{\theta}_2}$

VII. Conclusion

The obtained results show the supremacy of the ALM state estimator in both convergence time and estimation error. The proposed ALM state estimator is a powerful means which shows more accurate results under the condition of non-observability, a condition in which many classic observers cannot be applied due to instability. In addition, ALM state estimator provides a further degree of robustness because of obtaining its prior knowledge from

LQR approach which owns inherent robustness as discussed in previous sections. Overall, ALM observer shows superior functionality and circumvents the non-observability problem associated with practical applications.

ACKNOWLEDGMENT

The authors thank Mr. Hamidreza Mousavi and Mr. Mohsen Firouzi for their important comments and encouragement. We also thank all the people in the Sharif University of Technology Artificial Creatures Laboratory (ACL) for their great support.

REFERENCES

[1] S.A. Shahdi and S.B. Shouraki, "Supervised active learning method as an intelligent linguistic controller and its hardware implementation," Proc. AIA Symp, Artificial Intelligence and Applications (AIA 2002), ACTA Press, Sep. 2002, pp. 453-458.

[2] D. Bender and A. Laub, "The linear-quadratic optimal regulator for descriptor systems," IEEE Trans. on Automatic Control, Los Angeles, CA, Aug. 1987, Vol. AC-32, No. 8, pp. 672-688, doi: 10.1109/TAC.1987.1104694

[3] A. Bogdanov, "Optimal control of a double inverted pendulum on a cart," OHSU Technical Report CSE-04-006, Dec. 2004.

[4] R.E Kalman "Contributions to the theory of optimal control," Bol. Soc. Math. Mexicana, 1960, Vol. 5, pp. 102-119.

[5] A.E. Bryson, "Applied Linear Optimal Control, Examples and Algorithms," Cambridge University Press, ISBN: 0521012317, July 2002, pp. 201-223.

[6] S. B. Shouraki and N. Honda, "Outlines of a Soft Computer for Brain Simulation," Proc. IIZUKA '98, pp. 545-550.

[7] M. Murakami, N. Honda and J. Nishino, "A study on the modeling ability of the IDS method: A soft computing technique using pattern-based information processing," Int. Journal of Approximate Reasoning, Vol. 45, No. 3, Aug. 2007, pp. 470-487, doi:10.1016/j.ijar.2006.06.022.

[8] M. Murakami and N. Honda, "Performance of the IDS Method as a Soft Computing Tool," IEEE Trans. on Fuzzy Systems, Dec. 2008, Vol. 16, pp. 1582-1596, doi: 10.1109/TFUZZ.2008.2005693.

[9] S. B. Shouraki, "A novel fuzzy approach to modeling and control and its hardware implementation based on brain functionality and specifications," Ph.D. dissertation, 2000, University of Electro-Communications, Tokyo, Japan.

[10] Y. Sakurai, N. Honda and J. Nishino, "Acquisition of control knowledge of nonholonomic system by Active Learning Method," Proc. IEEE Symp, International Conference on Systems, Man and Cybernetics. IEEE Press, Oct. 2003, pp. 2400-2405, doi: 10.1109/ICSMC.2003.1244243.

[11] Greg Welchand and Gary Bishop, "An Introduction to the Kalman Filter," Technical Report 95-041,University of North Carolina at Chapel Hill Chapel Hill, NC 27599-3175, July 2006.