Parameter Estimation and Convergence Analysis for a Class of Canonical Dynamic Systems by Extended Kalman Filter

Ping Wei, Bin Xia, Xionglin Luo
Department of Automation
China University of Petroleum, Beijing
Beijing, China
e-mail: helloweiping@163.com, xiasubin@163.com, luoxl@cup.edu.cn

Abstract—**There were many researches about the parameter estimation of canonical dynamic systems recently. Extended Kalman filter (EKF) is a popular parameter estimation method in virtue of its easy applications. This paper focuses on parameter estimation for a class of canonical dynamic systems by EKF. By constructing associated differential equation, the convergence of EKF parameter estimation for the canonical dynamic systems is analyzed. And the simulation demonstrates the good performance.**

Keywords-parameter estimation; canonical dynamic system; EKF; covergence

I. INTRODUCTION

Parameter estimation is important for system modeling. There are several classical parameter estimation methods such as recursive least squares [1], Newton iteration [2] and hierarchical gradient algorithm [3]. Recently, the parameter estimation algorithm based on the combination of state and least squares parameter estimation was proposed for a class of canonical dynamic systems [4]. Later, the least squares iterative parameter estimation algorithm with decomposition was developed for the canonical dynamic systems [5]. The canonical dynamic systems can also be transformed into linear regression equation with linear Kalman filter for parameter estimation [6].

Nonlinear Kalman filter is another popular parameter estimation method based on approximation [7-9]. For instance, EKF can estimate both state variables and the parameters together, since the new augmented dynamic system becomes nonlinear by extending the parameters to state variables [8, 9]. Furthermore, the effectiveness of EKF algorithm was shown for applications in practical systems. For example, by combining the noise estimator with EKF algorithm, the performance of the soft sensor system was improved [10]. Besides, the advantage of EKF parameter estimation was represented in nonlinear dynamic models of biochemical networks [11].

In this paper, EKF parameter estimation will be used for a class of canonical dynamic systems as mentioned in [4,5]. In addition, the convergence of EKF parameter estimation algorithm for the canonical dynamic systems can be analyzed, since the stability of EKF parameter estimation algorithm was studied in details by constructing associated differential equation [8].

This paper is arranged as follows. EKF parameter estimation algorithm for a class of canonical dynamic systems is described in Section II. The convergence of EKF parameter estimation for the canonical dynamic systems is presented in Section III. In section IV, the simulation is given. Finally, a conclusion on the work of the paper is shown in Section V.

II. EKF PARAMETER ESTIMATION FOR A CLASS OF CANONICAL DYNAMIC SYSTEMS

Consider the following canonical dynamic system [4,5]

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) + w(t), \\ y(t) = Cx(t) + v(t). \end{cases} \tag{1}$$

where, $w(t)$ and $v(t)$ are uncorrelated white noises with zero mean and variance of Q_v and R_v respectively.

$$A = \begin{bmatrix} -a_1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n-1} & 0 & \dots & 1 \\ -a_n & 0 & \dots & 0 \end{bmatrix},\ B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},\ C = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}.$$

Extending state vector, let $z(t)=[x^T(t)\ a^T(t)\ b^T(t)]^T$. Where $x(t)=[x_1(t) \dots x_n(t)]^T$, $a(t)=[a_1(t) \dots a_n(t)]^T$, $b(t)=[b_1(t) \dots b_n(t)]^T$.

Using the augmented vector $z(t)$, system (1) can be rewritten as

$$\begin{cases} z(t+1) = f(z(t),u(t)) + \begin{pmatrix} w(t) \\ 0 \end{pmatrix}, \\ y(t) = Hz(t) + v(t). \end{cases} \tag{2}$$

where

$$f(z(t),u(t)) = \begin{bmatrix} A(a(t))x(t) + B(b(t))u(t) \\ a(t) \\ b(t) \end{bmatrix},$$

$H=[C\ 0 \dots 0]$.

Therefore, parameters and state variables of system (1) are both included in state variables of system (2). Then, EKF algorithm for system (2) is described as follows [8].

$$
\begin{cases}
\hat{z}(t+1) = f(\hat{z}(t), u(t)) + N(t)[y(t) - H\hat{z}(t)], \\
N(t) = F(\hat{z}(t), u(t))\overline{P}(t)H^{\mathrm{T}}[H\overline{P}(t)H^{\mathrm{T}} + R_v]^{-1}, \\
\overline{P}(t+1) = F(\hat{z}(t), u(t))\overline{P}(t)F^{\mathrm{T}}(\hat{z}(t), u(t)) + \overline{Q}_v \\
\qquad\qquad - N(t)[H\overline{P}(t)H^{\mathrm{T}} + R_v]N^{\mathrm{T}}(t).
\end{cases} \tag{3}
$$

where

$$
F(\hat{z}(t)) = \frac{\partial}{\partial z} f(z,u)\Big|_{z=\hat{z}(t)} = \begin{bmatrix} A(\hat{a}(t)) & -X(\hat{x}(t)) & U(u(t)) \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix},
$$

$$
A(\hat{a}(t)) = \begin{bmatrix} -\hat{a}_1(t) & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{a}_{n-1}(t) & 0 & \dots & 1 \\ -\hat{a}_n(t) & 0 & \dots & 0 \end{bmatrix}, \ \overline{Q}_v = \begin{bmatrix} Q_v & 0 \\ 0 & 0 \end{bmatrix},
$$

$$
X(\hat{x}(t)) = \begin{bmatrix} \hat{x}_1(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \hat{x}_1(t) \end{bmatrix}, \ U(u(t)) = \begin{bmatrix} u(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & u(t) \end{bmatrix}.
$$

Denote

$$
N(t) = \begin{bmatrix} K(t) \\ L_1(t) \\ L_2(t) \end{bmatrix}, \ \overline{P}(t) = \begin{bmatrix} P_{xx}(t) & P_{ax}(t) & P_{bx}(t) \\ P_{ax}^{\mathrm{T}}(t) & P_{aa}(t) & P_{ab}(t) \\ P_{bx}^{\mathrm{T}}(t) & P_{ab}^{\mathrm{T}}(t) & P_{bb}(t) \end{bmatrix},
$$

$$
A_t = A(\hat{a}(t)), \ B_t = B(\hat{b}(t)) = \begin{bmatrix} \hat{b}_1(t) & \cdots & \hat{b}_n(t) \end{bmatrix}^{\mathrm{T}},
$$

$$
X_t = X(\hat{x}(t)), \ U_t = U(u(t)), \ S_t = H\overline{P}(t)H^{\mathrm{T}} + R_v.
$$

Thus, the algorithm (3) is rewritten as

$$
\begin{cases}
\hat{a}(t+1) = \hat{a}(t) + L_1(t)[y(t) - C\hat{x}(t)], \\
\hat{b}(t+1) = \hat{b}(t) + L_2(t)[y(t) - C\hat{x}(t)].
\end{cases} \tag{4}
$$

$$
\begin{cases}
L_1(t) = P_{ax}^{\mathrm{T}}(t)C^{\mathrm{T}}S_t^{-1}, \\
L_2(t) = P_{bx}^{\mathrm{T}}(t)C^{\mathrm{T}}S_t^{-1}.
\end{cases} \tag{5}
$$

$$
\begin{cases}
P_{aa}(t+1) = P_{aa}(t) - L_1(t)S_tL_1^{\mathrm{T}}(t), \\
P_{bb}(t+1) = P_{bb}(t) - L_2(t)S_tL_2^{\mathrm{T}}(t).
\end{cases} \tag{6}
$$

$$
\hat{x}(t+1) = A_t\hat{x}(t) + B_tu(t) + K(t)[y(t) - C\hat{x}(t)]. \tag{7}
$$

$$
\begin{cases}
P_{ax}(t+1) = (A_t - K(t)C)P_{ax}(t) - X_tP_{aa}(t) + U_tP_{ab}^{\mathrm{T}}(t), \\
P_{bx}(t+1) = (A_t - K(t)C)P_{bx}(t) - X_tP_{ab}(t) + U_tP_{bb}(t).
\end{cases} \tag{8}
$$

where

$$
K(t) = [A_tP_{xx}(t)C^{\mathrm{T}} - X_tP_{ax}^{\mathrm{T}}(t)C^{\mathrm{T}} + U_tP_{bx}^{\mathrm{T}}(t)C^{\mathrm{T}}]S_t^{-1},
$$

$$
S_t = CP_{xx}(t)C^{\mathrm{T}} + R_v,
$$

$$
P_{ab}(t+1) = P_{ab}(t) - L_1(t)S_tL_2^{\mathrm{T}}(t),
$$

$$
\begin{aligned}
P_{xx}(t+1) = {} & A_tP_{xx}(t)A_t^{\mathrm{T}} - X_tP_{ax}^{\mathrm{T}}(t)A_t^{\mathrm{T}} + U_tP_{bx}^{\mathrm{T}}(t)A_t^{\mathrm{T}} \\
& - A_tP_{ax}(t)X_t + A_tP_{bx}(t)U_t + X_tP_{aa}(t)X_t \\
& - U_tP_{ab}^{\mathrm{T}}(t)X_t - X_tP_{ab}(t)U_t \\
& + U_tP_{bb}(t)U_t - K(t)S_tK^{\mathrm{T}}(t) + Q_v.
\end{aligned}
$$

Then, the process of parameter estimation is expressed clearly in terms of algorithm (4)-(8).

Remark 1. P_{xx} denotes the covariance of the estimation error of the state variable x. The covariance of the estimation error between parameter and state variable x are P_{ax} and P_{bx} respectively. P_{aa} and P_{bb} represent the covariance of the estimation error of a and b separately. P_{ab} is the covariance of the estimation error between parameter a and b.

III. Convergence Analysis

Denote $\widetilde{P}_{ax} = tP_{ax}$, $\widetilde{P}_{bx} = tP_{bx}$, $\widetilde{P}_{ab} = tP_{ab}$, $\widetilde{P}_{aa} = tP_{aa}$, $\widetilde{P}_{bb} = tP_{bb}$, $\widetilde{L}_1 = tL_1$, $\widetilde{L}_2 = tL_2$. When t approaches infinity, from the matrix inversion lemma[9], (4) can be rewritten as

$$
\begin{cases}
\hat{a}(t+1) = \hat{a}(t) + \dfrac{1}{t}\widetilde{L}_1(t)[y(t) - C\hat{x}(t)], \\
\hat{b}(t+1) = \hat{b}(t) + \dfrac{1}{t}\widetilde{L}_2(t)[y(t) - C\hat{x}(t)],
\end{cases} \tag{9}
$$

and (6) can be approximated as

$$
\begin{cases}
\widetilde{P}_{aa}^{-1}(t+1) = \widetilde{P}_{aa}^{-1}(t) + \dfrac{1}{t}\Big\{\widetilde{P}_{aa}^{-1}(t)\widetilde{L}_1(t)S_t \\
\qquad \Big[R_v + CP_{xx}(t)C^{\mathrm{T}}\Big]^{-1}S_t\widetilde{L}_1^{\mathrm{T}}(t)\widetilde{P}_{aa}^{-1}(t) - \widetilde{P}_{aa}^{-1}(t)\Big\}, \\
\widetilde{P}_{bb}^{-1}(t+1) = \widetilde{P}_{bb}^{-1}(t) + \dfrac{1}{t}\Big\{\widetilde{P}_{bb}^{-1}(t)\widetilde{L}_2(t)S_t \\
\qquad \Big[R_v + CP_{xx}(t)C^{\mathrm{T}}\Big]^{-1}S_t\widetilde{L}_2^{\mathrm{T}}(t)\widetilde{P}_{bb}^{-1}(t) - \widetilde{P}_{bb}^{-1}(t)\Big\}.
\end{cases} \tag{10}
$$

Denote

$$
\xi(t) = \begin{bmatrix} \hat{a}(t) \\ \hat{b}(t) \\ \mathrm{Col}\,\widetilde{P}_{aa}^{-1}(t) \\ \mathrm{Col}\,\widetilde{P}_{bb}^{-1}(t) \end{bmatrix}, \ \varphi(t) = \begin{bmatrix} \hat{x}(t) \\ \mathrm{Col}\,\widetilde{P}_{ax}(t) \\ \mathrm{Col}\,\widetilde{P}_{bx}(t) \end{bmatrix},
$$

$$
M = \begin{bmatrix} A_t - K(t)C & 0 & 0 \\ \alpha_1(\widetilde{P}_{aa}(t), t) & A_t - K(t)C & 0 \\ \alpha_2(\widetilde{P}_{ab}(t), t) & 0 & A_t - K(t)C \end{bmatrix},
$$

$$
N = \begin{bmatrix} B_t & K(t) \\ \beta_1(\widetilde{P}_{ab}(t), t) & 0 \\ \beta_2(\widetilde{P}_{bb}(t), t) & 0 \end{bmatrix}, \ e(t) = \begin{bmatrix} u(t) \\ y(t) \end{bmatrix},
$$

$$Q(t,\xi,\varphi)=\begin{bmatrix}\widetilde{L}_1(t)[y(t)-C\hat{x}(t)]\\ \widetilde{L}_2(t)[y(t)-C\hat{x}(t)]\\ \mathrm{Col}\left\{\widetilde{P}_{aa}^{-1}(t)\widetilde{L}_1(t)S_t\left[R_v+CP_{xx}(t)C^{\mathrm{T}}\right]^{-1}\right.\\ \left.S_t\widetilde{L}_1^{\mathrm{T}}(t)\widetilde{P}_{aa}^{-1}(t)-\widetilde{P}_{aa}^{-1}(t)\right\}\\ \mathrm{Col}\left\{\widetilde{P}_{bb}^{-1}(t)\widetilde{L}_2(t)S_t\left[R_v+CP_{xx}(t)C^{\mathrm{T}}\right]^{-1}\right.\\ \left.S_t\widetilde{L}_2^{\mathrm{T}}(t)\widetilde{P}_{bb}^{-1}(t)-\widetilde{P}_{bb}^{-1}(t)\right\}\end{bmatrix}.$$

Therefore, (9)(10) can be rewritten as

$$\begin{cases}\xi(t+1)=\xi(t)+\dfrac{1}{t}Q(t,\xi,\varphi),\\ \varphi(t+1)=M\varphi(t)+Ne(t).\end{cases}$$

where α_1, α_2, β_1, β_2 are obtained from (8). "Col" denotes some way to convert a matrix to column vector [8].

Provided that the parameters were kept constant a and b respectively as t approaches infinity. Then the following stable values of P_{xx}, S_t and $K(t)$ can be obtained.

$$\overline{P}_{xx}(a,b)=A(a)\overline{P}_{xx}(a,b)A^{\mathrm{T}}(a)+Q_v-\overline{K}(a,b)\overline{S}(a,b)\overline{K}^{\mathrm{T}}(a,b),$$
$$\overline{S}(a,b)=C\overline{P}_{xx}(a,b)C^{\mathrm{T}}+R_v\,,$$
$$\overline{K}(a,b)=A(a)\overline{P}_{xx}(a,b)C^{\mathrm{T}}\overline{S}^{-1}(a,b)\,.$$

Thus the process of $\bar{\hat{x}}$ is defined by constant parameters.

$$\bar{\hat{x}}(t+1)=A(a)\bar{\hat{x}}(t)+B(b)u(t)+\overline{K}(a,b)\bar{\varepsilon}(t;a,b)\qquad(11)$$

where

$$\bar{\varepsilon}(t;a,b)=y(t)-C\bar{\hat{x}}(t;a,b)\qquad(12)$$

Next, denote

$$\begin{cases}\overline{w}_1(t+1;a,b)=[A(a)-\overline{K}(a,b)C]\overline{w}_1(t;a,b)-X(\bar{\hat{x}}(t;a,b)),\\ \overline{w}_2(t+1;a,b)=[A(a)-\overline{K}(a,b)C]\overline{w}_2(t;a,b)+U(u(t)).\end{cases}\qquad(13)$$

$$\begin{cases}\overline{\psi}_1(t;a,b)=\overline{w}_1^{\mathrm{T}}(t;a,b)C^{\mathrm{T}},\\ \overline{\psi}_2(t;a,b)=\overline{w}_2^{\mathrm{T}}(t;a,b)C^{\mathrm{T}}.\end{cases}\qquad(14)$$

Taking no account of the correlation between parameter a and b as t approaches infinity, we can see processes $\overline{w}_1(t;a,b)\widetilde{P}_{aa}$ and $\overline{w}_2(t;a,b)\widetilde{P}_{bb}$ are processes $\widetilde{P}_{ax}$ and $\widetilde{P}_{bx}$ separately from (8) and (13). Comparing (5) and (14), processes $\widetilde{L}_1$, $\widetilde{L}_2$ are processes $\widetilde{P}_{aa}\overline{w}_1^{\mathrm{T}}(t;a,b)C^{\mathrm{T}}\overline{S}^{-1}(a,b)$ and $\widetilde{P}_{bb}\overline{w}_2^{\mathrm{T}}(t;a,b)C^{\mathrm{T}}\overline{S}^{-1}(a,b)$ respectively.

With the assumption that C1-C4[6] are satisfied, define $\widetilde{P}_{aa}^{-1}=R_a$, $\widetilde{P}_{bb}^{-1}=R_b$ and

$$\begin{cases}f(a)=\mathrm{E}\left\{\overline{\psi}_1(t;a,b)\overline{S}^{-1}(a,b)\bar{\varepsilon}(t;a,b)\right\},\\ f(b)=\mathrm{E}\left\{\overline{\psi}_2(t;a,b)\overline{S}^{-1}(a,b)\bar{\varepsilon}(t;a,b)\right\}.\end{cases}\qquad(15)$$

$$\begin{cases}G(a)=\mathrm{E}\left\{\overline{\psi}_1(t;a,b)\overline{S}^{-1}(a,b)\overline{\psi}_1^{\mathrm{T}}(t;a,b)\right\},\\ G(b)=\mathrm{E}\left\{\overline{\psi}_2(t;a,b)\overline{S}^{-1}(a,b)\overline{\psi}_2^{\mathrm{T}}(t;a,b)\right\}.\end{cases}$$

Thereby, the associated differential equation is

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}\tau}a(\tau)=R_a^{-1}(\tau)f(a(\tau)),\\ \dfrac{\mathrm{d}}{\mathrm{d}\tau}b(\tau)=R_b^{-1}(\tau)f(b(\tau)),\\ \dfrac{\mathrm{d}}{\mathrm{d}\tau}R_a(\tau)=G(a(\tau))-R_a(\tau),\\ \dfrac{\mathrm{d}}{\mathrm{d}\tau}R_b(\tau)=G(b(\tau))-R_b(\tau).\end{cases}\qquad(16)$$

Thus, the convergence of $\{\hat{a}(t)\}$ and $\{\hat{b}(t)\}$, generated from the algorithm (4)-(8), are in accordance with the stability of the associated differential equation (16). The convergence conclusion about the algorithm (4)-(8) can be explicated as follows.

Convergence conclusion: Under the conditions of C1-C4, supposing the norms of all eigenvalues of matrix $[A(a)-\overline{K}(a,b)C]$ were smaller than one, the recursive sequence of $\left\{(\hat{a}^{\mathrm{T}},\hat{b}^{\mathrm{T}})^{\mathrm{T}}\right\}$ converges to $(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}$ (the true value and local stable stationary point of (16)) with probability one as t approaches infinity. Furthermore, if the process noise Q_v=0, the estimated sequence $\left\{(\hat{a}^{\mathrm{T}},\hat{b}^{\mathrm{T}})^{\mathrm{T}}\right\}$ converges to $(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}$ (the true value and global stable stationary point of (16)) with probability one as t approaches infinity.

Proof. If $(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}$ is the true value of system (1), the true system can be written as

$$\begin{cases}x(t+1)=A_0x(t)+B_0u(t)+w(t),\\ y(t)=Cx(t)+v(t).\end{cases}$$

With the independence of $v(t)$ and $\overline{\psi}_i(t;a,b)\overline{S}^{-1}(a,b)$ (i=1, 2), from (12) and (15),

$$\begin{cases}f(a)=\mathrm{E}\left\{\overline{\psi}_1(t;a,b)\overline{S}^{-1}(a,b)C\left[x(t;a_0,b_0)-\bar{\hat{x}}(t;a,b)\right]\right\},\\ f(b)=\mathrm{E}\left\{\overline{\psi}_2(t;a,b)\overline{S}^{-1}(a,b)C\left[x(t;a_0,b_0)-\bar{\hat{x}}(t;a,b)\right]\right\}.\end{cases}$$

It is easy to know that the true value $(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}$ is the stationary point of equation (16) without the dependence between $\overline{\psi}_i(t;a,b)\overline{S}^{-1}(a,b)$ (i=1,2) and $\bar{\varepsilon}(t;a,b)$. The norms of all eigenvalues of matrix $[A(a)-\overline{K}(a,b)C]$ should be smaller than one by (13)-(15). It can be obtained that

$$\mathrm{E}\,\bar{\varepsilon}(t;a_0,b_0)=0\,.\qquad(17)$$

Next present the asymptotical stability of stationary point. Construct a Lyapunov function,

$$V(a,b)=\frac{1}{2}\mathrm{E}\left\{\left[\bar{\varepsilon}(t;a,b)\right]^2\right\}>0\,.\qquad(18)$$

Then, the derivative of $V(a,b)$ is

$$\frac{\mathrm{d}V(a,b)}{\mathrm{d}\tau}=\mathrm{E}\left\{\bar{\varepsilon}(t;a,b)\cdot\frac{\partial}{\partial a}\bar{\varepsilon}(t;a,b)\right\}R_a^{-1}f(a)+\mathrm{E}\left\{\bar{\varepsilon}(t;a,b)\cdot\frac{\partial}{\partial b}\bar{\varepsilon}(t;a,b)\right\}R_b^{-1}f(b).\qquad(19)$$

From (11)(12)

$$
\begin{cases}
\dfrac{\partial}{\partial a}\bar{\hat{x}}(t+1;a,b) = [A(a)-\bar{K}(a,b)C]\dfrac{\partial}{\partial a}\bar{\hat{x}}(t;a,b) \\
\qquad - X(\bar{\hat{x}}(t)) + \dfrac{\partial}{\partial a}\bar{K}(a,b)\cdot\bar{\varepsilon}(t;a,b), \\
\dfrac{\partial}{\partial b}\bar{\hat{x}}(t+1;a,b) = [A(a)-\bar{K}(a,b)C]\dfrac{\partial}{\partial b}\bar{\hat{x}}(t;a,b) \\
\qquad + U(u(t)) + \dfrac{\partial}{\partial b}\bar{K}(a,b)\cdot\bar{\varepsilon}(t;a,b).
\end{cases} \tag{20}
$$

$$
\begin{cases}
\dfrac{\partial}{\partial a}\bar{\varepsilon}(t;a,b) = -C\cdot\dfrac{\partial}{\partial a}\bar{\hat{x}}(t;a,b), \\
\dfrac{\partial}{\partial b}\bar{\varepsilon}(t;a,b) = -C\cdot\dfrac{\partial}{\partial b}\bar{\hat{x}}(t;a,b).
\end{cases} \tag{21}
$$

Supposed that $\bar{\varepsilon}(t;a,b)$ is independent of the partial derivates of $\bar{K}(a,b)$ and $\bar{\hat{x}}(t;a,b)$. Comparing (13) with (20), from(14)(17)(21)

$$
\begin{cases}
\mathrm{E}\left\{\dfrac{\partial}{\partial a}\bar{\varepsilon}(t;a,b)\right\}\Big|_{\substack{a=a_0\\ b=b_0}} = -\mathrm{E}\bar{\psi}_1^{\mathrm{T}}\Big|_{\substack{a=a_0\\ b=b_0}}, \\
\mathrm{E}\left\{\dfrac{\partial}{\partial b}\bar{\varepsilon}(t;a,b)\right\}\Big|_{\substack{a=a_0\\ b=b_0}} = -\mathrm{E}\bar{\psi}_2^{\mathrm{T}}\Big|_{\substack{a=a_0\\ b=b_0}}.
\end{cases} \tag{22}
$$

Under the conditions of $\|a-a_0\|<\delta$ and $\|b-b_0\|<\delta$ (δ is a positive infinitesimal constant), from (15)(19)(22),

$$
\begin{aligned}
\frac{\mathrm{d}V(a,b)}{\mathrm{d}\tau} &= -\bar{S}(a,b)f^{\mathrm{T}}(a)R_a^{-1}f(a) \\
&\quad - \bar{S}(a,b)f^{\mathrm{T}}(b)R_b^{-1}f(b).
\end{aligned} \tag{23}
$$

That is,

$$
\begin{cases}
\dfrac{\mathrm{d}V(a,b)}{\mathrm{d}\tau} \le 0, & (a^{\mathrm{T}},b^{\mathrm{T}})^{\mathrm{T}} \in D_A, \\
\dfrac{\mathrm{d}V(a,b)}{\mathrm{d}\tau} = 0, & (a^{\mathrm{T}},b^{\mathrm{T}})^{\mathrm{T}} \in D_c = \left\{(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}\right\}.
\end{cases}
$$

Thus the invariant set is D_c, D_A is the attraction domain of D_c. The estimated sequence $\left\{(\hat{a}^{\mathrm{T}},\hat{b}^{\mathrm{T}})^{\mathrm{T}}\right\}$ converges to $(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}$ (the true value and local stable stationary point of (16)) with probability one as t approaches infinity.

Providing the process noise Q_v=0, $\bar{K}(a,b)=0$ would be satisfied. From (15) and (19), (23) is satisfied globally by comparing (13)(14) with (20)(21). Thus, the estimated sequence $\left\{(\hat{a}^{\mathrm{T}},\hat{b}^{\mathrm{T}})^{\mathrm{T}}\right\}$ converges to $(a_0^{\mathrm{T}},b_0^{\mathrm{T}})^{\mathrm{T}}$ (the true value and global stable stationary point of (16)) with probability one as t approaches infinity.

Remark 2. The conditions of associated differential equation theory include C1-C6 originally[6]. Conditions C5 and C6 are satisfied obviously since $\gamma(t)=1/t$ in this paper.

IV. SIMULATION

Consider the following system with four parameters

$$
\begin{cases}
x(t+1) = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t) + w(t), \\
y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) + v(t).
\end{cases} \tag{24}
$$

Set $Q_v = \mathrm{I}_{2\times 2}$, R_v=0.01, a_{10}=−0.9, a_{20}=0.5, b_{10}=−1.88, b_{20}=−0.9. The simulation of parameter estimation for system (24) is shown in Figure 1. According to Figure 1, estimated values of a_1, a_2 and b_1 are convergent to the true values ultimately, estimated values of b_2 are convergent to the true value with small deviation. And, parameter estimation errors are computed as

$\delta_{a_1} = \|\hat{a}_1 - a_1\|/\|a_1\| = 0.114\%$, $\delta_{a_2} = \|\hat{a}_2 - a_2\|/\|a_2\| = 0.0702\%$,

$\delta_{b_1} = \|\hat{b}_1 - b_1\|/\|b_1\| = 0.165\%$, $\delta_{b_2} = \|\hat{b}_2 - b_2\|/\|b_2\| = 0.859\%$.

All stable parameter estimation errors are smaller than 1% from the above data. Thereby, the simulation results illuminate the effectiveness of EKF parameter estimation algorithm for system (24).

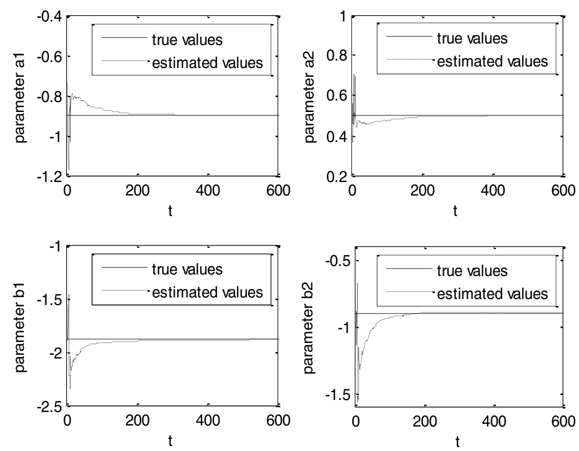

Figure 1. Estimation for a_1, a_2, b_1 and b_2.

V. CONCLUSION

EKF is a popular parameter estimation method for its easy applications. This paper concentrates on parameter estimation for a class of canonical dynamic systems by EKF algorithm. After extending the parameters to state variables, EKF can be used to estimate new state variables. The convergence of EKF parameter estimation for the canonical dynamic systems is studied by constructing associated differential equation. A convergence conclusion has been illustrated that estimated parameters are convergent to true values under some certain conditions.

ACKNOWLEDGMENT

This work is supported by Science Foundation of China University of Petroleum, Beijing (No.2462015YQ0515).

REFERENCES

[1] Z. Wang, Q. B. Jin and X. P. Liu, "Recursive least squares identification of hybrid Box-Jenkins model structure in open-loop and closed-loop," J. Frankl. Inst, vol. 353, pp. 265–278, January 2016.

[2] J. H. Li, “Parameter estimation for Hammerstein CARARMA systems based on the Newton iteration,” Appl. Math. Lett, vol. 26, pp. 91–96, January 2013.

[3] Z. N. Zhang, F. Ding, and X. G. Liu, “Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems,” Comput. Math. Appl, vol. 61, pp. 672-682, February 2011.

[4] F. Ding, “Combined state and least squares parameter estimation algorithms for dynamic systems,” Appl. Math. Model, vol. 38, pp. 403–412, January 2014.

[5] F. Ding, X. M. Liu and X. Y. Ma, “Kalman state filtering based least squares iterative parameter estimation for observer canonical state space systems using decomposition,” J. Comput. Appl. Math, vol. 301, pp. 135-143, August 2016.

[6] L. Ljung and T Söderström, Theory and Practice of Recursive Identification. MIT Press, 1983.

[7] E. A. Wan and R. Merwe, “The unscented Kalman filter for nonlinear estimation,” Proc. IEEE. Symp. Adaptive Systems for Signal Processing, Communication and Control(AS-SPCC), IEEE Press, 2000, pp. 153–158.

[8] L. Ljung, “Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems,” IEEE Trans. Automat. Contr., vol. 24, pp. 36–50, February 1979.

[9] D. Simon, Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches. Wiley-Interscience, 2006.

[10] Y. Wu and X. L. Luo, “Design of output data fusion soft sensor based on adaptive extended Kalman filtering algorithm,” J. Chem. Ind. Eng, vol. 61, pp. 2627–2635, October 2010.

[11] X. D. Sun and M. M. Xiong, “Extended Kalman filter for estimation of parameters in nonlinear state-space models of biochemical networks,” Plos One, vol. 3, pp. e3758, November 2008.