Received December 26, 2018, accepted January 14, 2019, date of publication January 24, 2019, date of current version February 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2894594

Robust Control of Uncertain Linear Systems Based on Reinforcement Learning Principles

DENGGUO XU[1,2], QINGLIN WANG[1], AND YUAN LI[1]
[1]School of Automation, Beijing Institute of Technology, Beijing 100081, China
[2]School of Mathematics and Statistics, Chuxiong Normal University, Chuxiong 675000, China

Corresponding author: Yuan Li (liyuan@bit.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61472037, Grant 61463002, and Grant 61375100.

ABSTRACT In this paper, a reinforcement learning (RL) approach is developed to solve the robust control for uncertain continuous-time linear systems. The objective is to find a feedback control law for the uncertain linear system using an online policy iteration algorithm. The robust control problem is solved by constructing an extended algebraic Riccati equation with properly defined weighting matrices for a general uncertain linear system. An online policy iteration algorithm is developed to solve the robust control problem based on RL principles without knowing the nominal system matrix. The convergence of the algorithm to the robust control solution for uncertain linear systems is proved. The simulation examples are given to demonstrate the effectiveness of the proposed algorithm. The results extend the design method of robust control to uncertain linear systems.

INDEX TERMS Reinforcement learning, uncertain linear system, robust control, algebraic Riccati equation.

I. INTRODUCTION

Over the past few decades, the problem of robust control design for uncertain systems has attracted considerable attention from researchers. Many significant results have been obtained on the topic, which can be applied not only in aircraft control and robot control [1], [2] but also in biological and physical sciences [3]. Many real models have uncertainties due to data measurement errors or disturbance. Several methods are available for robust control of uncertain systems, such as Kharitonov interval theory [4], structured singular value theory [5] and H_∞ control theory [6]. In this paper, we present a new robust controller design method based on reinforcement learning (RL) principles.

RL principles and adaptive dynamic programming (ADP) theory have been broadly applied to solve optimal control problems [7]–[10]. ADP consists of a class of RL methods that have shown their importance in a variety of applications, including feedback control of dynamical systems. In [11], for the first time, Werbos presented the idea of ADP that can approximate the (generally intractable) solution of Bellman's equation. In [12]–[14], RL techniques were first employed to seek solutions to the optimal regulator problem for discrete-time systems. In [15], an advantage updating algorithm for RL was presented. It does not require a model to be given or learned. An RL framework for continuous-time dynamical systems has been presented in [16], which derives algorithms for estimating value functions and improving policies with the use of function approximators. The application of RL in control theory is mainly related to optimal regulation and optimal tracking problems. For a linear system, optimal control or tracking is mainly achieved by online policy iteration or value iteration, which consists of the following two steps: policy evaluation and policy improvement [17]–[22]. RL application in nonlinear system control mainly combines integral RL and neural network approximation [23]–[25].

There are two advantages in using RL to solve feedback control problems. First, it can effectively solve the so-called "curse of dimensionality" in optimal control problems [23]. Second, the RL method can be used to solve the optimal control problem without knowledge of the system dynamics. In many practical applications, it is often difficult to fully know the system dynamics, and unknown model control problems can be solved using the RL algorithm.

On the other hand, robust controller design for uncertain systems has been studied by many scholars since the 1980s. A nonlinear controller that stabilizes an uncertain system was given in [26]. Linear controllers that guarantee stability were derived in [27]–[29]. These papers all require that the system satisfy the so-called "matching conditions". Using the algebraic Riccati equation (ARE), the design of linear robust controllers for uncertain linear systems was proposed in many studies. Petersen and Hollot [30] and

Schmitendorf [31] considered a special uncertain linear system in which the uncertainty matrix can be decomposed into a linear combination of *rank*-one matrices. ARE was used to obtain robust controllers for this special system. An optimal control design method for robust control of a general uncertain linear system was proposed in [32]. However, for an unmatched uncertain system, no direct design method was given. Using ARE and an optimal control method, a robust control law to stabilize an matched uncertain system exponentially was presented in [33]. An alternative method for the design of robust controllers was given in [34]. There, the main idea involves making the Lyapunov derivative negative via a one-dimensional parameter search. If the system satisfies the matching conditions, then the search is guaranteed to result in an answer. Later, for the matched system, a non-iterative design method for the robust controller was investigated by Jabbari and Schmitendorf [35]. The method proves to be effective since it does not require a numerical search procedure. However, precompensators are necessary for non-Hurwitz nominal systems. Tsay [36] and Dolphus and Schmitendorf [37] also investigated robust controllers for uncertain systems. However, their methods are not effective for general matched systems since strong conditions are imposed on the uncertainty in the input matrix of the system in the studies. It is noted that there are many more results on the robust stabilization of uncertain systems. The readers can refer to the literature [38]–[40] and the references cited in the articles. Among various methods for robust controller design, there are some results on optimal control-based methods. An optimal control approach to robust control design was presented in [41] and [42]. The robust control law is an optimal control law for a nominal system, obtained by selecting suitable weighting matrices. However, the results are based on time-invariable uncertainty in the system, and for the unmatched uncertain linear system, too many design parameters are required, while no effective parameter design method is proposed.

At present, the literature on robust control by RL is limited to nonlinear uncertain systems. A novel RL-based robust adaptive control algorithm was developed for a class of continuous-time uncertain nonlinear systems subject to input constraints [43]. Based on the neural network approximator, an online RL algorithm was proposed for a class of affine multiple input and multiple output nonlinear discrete-time systems with unknown functions and disturbances [44]. An online adaptive RL-based solution was developed for the infinite-horizon optimal control problem for continuous-time uncertain nonlinear systems in [45]. In addition, the robust control of uncertain nonlinear systems is also studied in other papers such as [46]–[50]. In all the papers listed above, neural networks are used to approximate the solutions of Hamilton Jacobi Bellman (HJB) equation. In fact, the infinite-time optimal control problem is mainly to solve the corresponding HJB equation [51]. The HJB equation is a nonlinear partial differential equation which is difficult to solve. For nonlinear system, neural network-based approximation is generally used to solve HJB equation. While in the linear system, the HJB equation reduces to ARE which does not need to use neural network to solve. There are some differences between the methods of solving robust control problems of linear systems and those of non-linear systems. However, to the best of our knowledge, almost no literature concerning RL-based robust control problems for uncertain linear systems has been presented.

The primary objective of this paper is to improve existing robust control methods for unmatched continuous-time linear systems with time-variable uncertainty. Using an extended ARE, an RL-based systemic method for the robust control of linear systems with unmatched uncertainties is proposed. According to the constructed ARE, an integral performance index is constructed. Based on the performance index, an integral RL algorithm is developed to solve the robust control for continuous-time linear systems with partially unknown nominal dynamics. To find a feedback control law for the uncertain linear system using an online policy iteration algorithm, the value function is prescribed by a quadratic form of the state and input of the system. The convergence of the algorithm to the robust control solution for uncertain linear systems is proved. Simulation examples are given to demonstrate the effectiveness of the proposed algorithm.

The main contributions of this paper include two aspects. First, we consider more general uncertain linear systems, with uncertainty entering both the system matrix and the input matrix. Based on the ARE, we propose a systemic method of robust control for uncertain linear systems. In [52], for linear systems with input matrix uncertainty, the author considered three cases. In this paper, we discuss the case that both the system matrix and the input matrix do not satisfy the matching conditions. The case discussed here is not included in [52]. Second, an online RL method is proposed to solve the robust control problem for uncertain linear systems. The corresponding algorithm is established, and its convergence is proved. Moreover, this is the first attempt to use the RL method to solve the robust control problem for uncertain linear systems. The advantage of using the RL algorithm is that there is no need to know the nominal system matrix.

The remainder of this paper is organized as follows. In Section II, we formulate the control objective and present basic results for the robust control problem. Then, in Section III, we propose an robust control design method for unmatched uncertain linear systems based on an extended ARE with properly defined weighting matrices. The RL algorithm and its convergence to the robust control solution are presented in Section IV. Specialization to matched uncertain linear systems is presented in Section V. To validate the theoretical results presented in this paper, three numerical examples are given in Section VI. Finally, the entire work is summarized and the future of this field is discussed in Section VII.

Notation: Throughout the paper, $\mathbb{R}$ denotes the set of real numbers; $A > 0$ and $A \geq 0$ denote that the matrix A is positive definite and positive semi-definite, respectively; $A > B$ and $A \geq B$ imply that $A - B > 0$ and $A - B \geq 0$, respectively; A^T denotes the transpose of the matrix A. The notation $\|.\|$ represents the Euclidean norm for vectors or the induced matrix norm for matrices, and I denotes the identity matrix of appropriate dimensions.

II. FORMATION OF THE PROBLEM AND PRELIMINARIES

Consider the uncertain continuous-time linear system described by

$$\dot{x}(t) = [A + \Delta A(s(t))]x(t) + [B + \Delta B(l(t))]u(t) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the state variable, $u(t) \in \mathbb{R}^m$ is the control input vector, and $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are nominal system and input matrices, respectively. $s(t) \in \mathcal{S}, l(t) \in \mathcal{L}$ are uncertain parameter vector functions that are Lebesgue measurable, where $\mathcal{S}$ and $\mathcal{L}$ are uncertain parameter sets. $\Delta A(s(t))$, $\Delta B(l(t))$ are the $n \times n, n \times m$ uncertain parameter matrices, which depend continuously on the uncertainty vectors $s(t)$ and $l(t)$, respectively.

For the sake of brevity, the argument of the time function is omitted in some of the following sections.

The objective of this paper is to find a feedback control law $u(t) = Kx(t)$ such that uncertain system (1) is asymptotically stable for all $s(t) \in \mathcal{S}$ and $l(t) \in \mathcal{L}$.

Definition 1: System (1) is said to satisfy the system matrix matched condition if for any $s(t) \in \mathcal{S}$, there exists an $m \times n$ matrix $\phi(s)$ such that

$$\Delta A(s) = B\phi(s) \quad (2)$$

where $\phi(s)$ is bounded.

Definition 2: System (1) is said to satisfy the input matrix matched condition if for any $l(t) \in \mathcal{L}$, there exists an $m \times m$ matrix $\bar{\phi}(l)$ such that

$$\Delta B(l) = B\bar{\phi}(l) \quad (3)$$

where $\bar{\phi}(l) \geq 0$.

Definition 3: System (1) is termed a matched uncertain linear system if it satisfies (2) and (3) for any $s(t) \in \mathcal{S}$ and $l(t) \in \mathcal{L}$.

Definition 4: System (1) is termed an unmatched uncertain linear system if it does not satisfy (2) or (3) for any $s(t) \in \mathcal{S}$ and $l(t) \in \mathcal{L}$.

The following lemma will be used later to prove the convergence of the algorithm proposed in this paper. For more details, see [53].

Lemma 1: Let $K_0 \in R^{m\times n}$ be any stabilizing feedback gain matrix for nominal system dynamics (A, B). $Q \geq 0, R > 0$ are weighting matrices of appropriate dimensions. Let P_i be the symmetric positive definite solution of the Lyapunov equation

$$(A - BK_i)^T P_i + P_i(A - BK_i) + Q + K_i^T RK_i = 0$$

where K_i, with $i = 1, 2, \ldots$, are defined recursively by

$$K_i = R^{-1}B^T P_{i-1}$$

Then, the following properties hold:

1. $A - BK_i$ is Hurwitz for $i = 1, 2, \ldots$,
2. $P^* \leq P_{i+1} \leq P_i$, and
3. $lim_{i\to\infty}K_i = K^*$, $lim_{i\to\infty}P_i = P^*$.

In [53], by iteratively solving the Lyapunov equation, which is linear in P_i, and updating K_i, the solution to the following ARE is numerically approximated

$$PA + A^T P + Q - PBRB^T P = 0.$$

III. UNMATCHED UNCERTAIN LINEAR SYSTEM

In this section, we consider the robust control problem of uncertain linear system (1), which does not necessary satisfy matched conditions (2) and (3). The robust control problem is solved by constructing an extended ARE with properly defined weighting matrices. A robust controller of uncertain linear systems can be obtained from the ARE solution.

By using the pseudo-inverse B^+ of B, we decompose the uncertain system matrix and the input matrix into a matched component and an unmatched component as follows:

$$\Delta A(s) = BB^+\Delta A(s) + (I - BB^+)\Delta A(s) \quad (4)$$

and

$$\Delta B(l) = BB^+\Delta B(l) + (I - BB^+)\Delta B(l) \quad (5)$$

where $B^+ = (B^T B)^{-1}B^T$.

Generally, the condition for the existence of pseudo-inverse of matrix B is that its column vectors are linearly independent [54]. In practical control systems, input matrix B is usually column full rank. Therefore, the existence condition of pseudo-inverse of input matrix B is generally satisfied. It should be noted here that the pseudo-inverse matrix B^+ satisfies $B^+B = I$ but does not satisfy $BB^+ = I$.

To obtain the solution of the robust control problem, the following assumptions are made.

Assumption 1: Assume that the nominal system dynamics (A, B) is controllable.

Assumption 2: There exists a positive semi-definite matrix F such that

$$\Delta A(s)^T (B^+)^T B^+\Delta A(s) \leq F \quad (6)$$

Assumption 3: There exists a positive semi-definite matrix H such that

$$\Delta A(s)^T \Delta A(s) \leq H \quad (7)$$

Denote $B_I = I - BB^+$. We construct an extended ARE as follows.

$$PA + A^T P + Q - P[BB^T + B_I B_I^T]P = 0 \quad (8)$$

where $Q = F + H + \beta^2 I$ and $\beta > 0$ is a design parameter.

The following theorem shows that robust controllers for uncertain linear systems can be obtained from solving the ARE (8).

Theorem 1: Assume that the matrix P is the symmetric positive definite solution of the ARE (8). For uncertain linear system (1) subjected to Assumption 1, Assumption 2 and

Assumption 3, one can choose the parameter β such that the following conditions hold

$$\beta^2 I - 2PB_I B_I^T P > 0, \ \Delta B(l) B^T \geq 0. \quad (9)$$

Then, the feedback control $u = Kx$ with $K = -B^T P$ can stabilize uncertain linear system (1) for all $s(t) \in \mathcal{S}$, $l(t) \in \mathcal{L}$.

Proof: Taking the time derivative of Lyapunov function $V(x) = x^T Px$ along uncertain linear system (1) with $u = Kx$, we can obtain

$$\begin{aligned}\frac{dV}{dt} &= x^T[(A + \Delta A(s))^T + K^T(B + \Delta B(l))^T]Px \\ &\quad + x^T P[(A + \Delta A(s)) + (B + \Delta B(l))K]x \\ &= x^T[\Delta A(s)^T P + P\Delta A(s)]x \\ &\quad + x^T[A^T P + PA + 2PBK + 2P\Delta B(l)K]x \quad (10)\end{aligned}$$

By using (4) and (5), we have

$$\begin{aligned}\frac{dV}{dt} &= x^T[A^T P + PA + \Delta A(s)^T B^{+T} B^T P \\ &\quad + \Delta A(s)^T B_I^T P]x \\ &\quad + x^T[PBB^+ \Delta A(s) + PB_I \Delta A(s)]x \\ &\quad + x^T[2PBK + 2P\Delta B(l)K]x\end{aligned}$$

It follows from (8) that

$$\begin{aligned}\frac{dV}{dt} &= -x^T Qx + x^T P[BB^T + B_I B_I^T]Px \\ &\quad + x^T \Delta A(s)^T B^{+T} B^T Px + x^T \Delta A(s)^T B_I^T Px \\ &\quad + x^T PBB^+ \Delta A(s)x + x^T PB_I \Delta A(s)x \\ &\quad + 2x^T PBKx + 2x^T P\Delta B(l)Kx\end{aligned}$$

By $K = -B^T P$, we have

$$\begin{aligned}\frac{dV}{dt} &= -x^T Qx + x^T K^T Kx + x^T PB_I B_I^T Px \\ &\quad - x^T \Delta A(s)^T B^{+T} Kx + x^T \Delta A(s)^T B_I^T Px \\ &\quad - x^T K^T B^+ \Delta A(s)x + x^T PB_I \Delta A(s)x \\ &\quad - 2x^T K^T Kx - 2x^T P\Delta B(l)B^T Px\end{aligned}$$

Let $L = -B_I^T P$; then,

$$\begin{aligned}\frac{dV}{dt} &= -x^T Qx - x^T K^T Kx + x^T L^T Lx \\ &\quad - 2x^T K^T B^+ \Delta A(s)x - 2\, x^T \Delta A(s)^T Lx \\ &\quad - 2x^T P\Delta B(l)B^T Px\end{aligned}$$

Since

$$\begin{aligned}&-x^T K^T Kx - 2x^T K^T B^+ \Delta A(s)x \\ &\quad = -x^T(K - B^+ \Delta A(s))^T(K - B^+ \Delta A(s))x \\ &\qquad + x^T(B^+ \Delta A(s))^T(B^+ \Delta A(s))x \\ &\quad \leq x^T(B^+ \Delta A(s))^T(B^+ \Delta A(s))x\end{aligned}$$

and

$$\begin{aligned}-2x^T L^T \Delta A(s)x &\leq x^T L^T Lx + \Delta A(s)^T \Delta A(s)x \\ &\leq x^T L^T Lx + x^T Hx\end{aligned}$$

we have

$$\begin{aligned}\frac{dV}{dt} &= -x^T Qx - x^T K^T Kx + x^T L^T Lx \\ &\quad - 2x^T K^T B^+ \Delta A(s)x - 2x^T \Delta A(s)^T Lx \\ &\quad - 2x^T P\Delta B(l)B^T Px \\ &\leq -x^T(F + H + \beta^2 I)x + x^T L^T Lx \\ &\quad + x^T Fx + x^T L^T Lx + x^T Hx \\ &\quad - 2x^T P\Delta B(l)B^T Px \\ &= -x^T(\beta^2 I - 2L^T L)x - 2x^T P\Delta B(l)B^T Px\end{aligned}$$

It follows from the conditions in (9) that

$$\frac{dV}{dt} < 0$$

Therefore, by the Lyapunov stability theory, the closed-loop unmatched uncertain linear system with $u(t) = Kx(t)$ is stable for all $s(t) \in \mathcal{S}$ and $l(t) \in \mathcal{L}$. In other words, $u(t) = Kx(t)$ is a stabilizing control law of unmatched uncertain linear system (1). The proof is complete.

Remark 1: In [52], for linear systems with time-invariance uncertainty, the author considered three cases. The first case is that both the system matrix and input matrix satisfy the matching conditions. The second case is that the system matrix does not satisfy the matching condition but the input matrix satisfies the matching conditions. The third case is that the input matrix does not satisfy the matching conditions but the system matrix has no uncertainty. This paper addresses systems with time varying uncertainties. Moreover, we discuss the case that both the system matrix and the input matrix do not satisfy the matching conditions. The case discussed here is not included in [52], so this paper extends the existing results.

Remark 2: The robust control design for unmatched uncertain linear systems is a challenging problem. In the existing literature on the robust control for unmatched linear systems, the design method is not direct or too many parameters in the design process need to be determined. This leads to problems that are not easy to solve. In Theorem 1, the conditions in (9) are easily satisfied. If the parameter β is selected such that it is large enough, then the first condition in (9) easily holds, as the term $2PB_I B_I^T P$ is a symmetric matrix. By properly selecting the nominal input matrix B, the second condition of (9) is also easily satisfied.

IV. RL ALGORITHM AND ITS CONVERGENCE

In this section, based on RL principles, an online iterative algorithm is derived to solve the robust control problem for unmatched uncertain linear system. The corresponding convergence analysis is presented.

A. RL ALGORITHM

According to the weighting matrices in (8), we construct a performance index as follows:

$$J = \int_t^\infty [x^T Qx + \bar{u}^T \bar{u}]dt \quad (11)$$

The performance index can be regarded as one of the optimal control problems of the system

$$\dot{x} = Ax + \bar{B}\bar{u} \tag{12}$$

where $\bar{B} = \begin{bmatrix} B & I - BB^+ \end{bmatrix}$ and $\bar{u} = \begin{bmatrix} u^T & v^T \end{bmatrix}^T$.

For any initial time t, the optimal value corresponding to (11) and (12) can be written as

$$\begin{aligned} V(x(t)) &= \int_t^\infty [x^T Qx + \bar{u}^T \bar{u}]dt \\ &= \int_t^{t+\Delta t} [x^T Qx + \bar{u}^T \bar{u}]dt + \int_{t+\Delta t}^\infty [x^T Qx + \bar{u}^T \bar{u}]dt \end{aligned}$$

that is,

$$V(x(t)) = \int_t^{t+\Delta t} [x^T Qx + \bar{u}^T \bar{u}]dt + V(x(t+\Delta t))$$

where $Q = F + H + \beta^2 I$.

Further, we can obtain

$$\begin{aligned} x(t)^T Px(t) = \int_t^{t+\Delta t} [x^T Qx + \bar{u}^T \bar{u}]dt \\ + x(t+\Delta t)^T Px(t+\Delta t) \end{aligned} \tag{13}$$

where P is the symmetric positive definite solution of the ARE (8). Based on this, the following algorithm can be used to compute a robust control law of uncertain linear system (1) with unmatched uncertainty.

Algorithm 1: Robust Control Algorithm For Unmatched Uncertain Linear System.

1. Choose the least upper bound matrices F and H satisfying (6) and (7), respectively, and a constant β. Compute $Q = F + H + \beta^2 I$;
2. Choose a stabilizing initial control policy $u = \bar{K}_0 x$ from the extended nominal linear system (12);
3. Policy evaluation: Given a control policy $\bar{K}_i$, solve P_i using the Bellman equation
$$\begin{aligned} x(t)^T P_i x(t) = \int_t^{t+\Delta t} [x^T Qx + x^T \bar{K}_i^T \bar{K}_i x]dt \\ + x(t+\Delta t)^T P_i x(t+\Delta t); \end{aligned} \tag{14}$$
4. Policy improvement: update the control input using
$$\bar{K}_{i+1} = -\bar{B}^T P_i; \tag{15}$$
5. Check whether the condition $\beta^2 I - 2PB_I B_I^T P > 0$ is satisfied. If the condition is not satisfied, return to the first step, and take a larger constant β.

The policy evaluation and improvement steps (14) and (15) are repeated until the policy improvement step no longer changes the present policy, that is, until $\|P_{i+1} - P_i\| \leq \varepsilon$ is satisfied, where ε is a small constant; thus, convergence to the optimal controller is achieved.

Remark 3: Although a robust controller for uncertain linear system (1) can be obtained by directly solving the ARE (8) offline, the nominal system matrix must be known, which is difficult to obtain in some cases. Algorithm 1 does not require knowledge of the nominal system matrix A. Compared with the existing robust control methods, it has advantages in terms of operability and practicability.

B. CONVERGENCE ANALYSIS

The convergence of the algorithm is derived from the following conclusions.

Lemma 2: Assuming that $(A + \bar{B}\bar{K}_i)$ is stable, solving for P_i in (14) is equivalent to finding the solution of the equation

$$(A + \bar{B}\bar{K}_i)^T P_i + P_i(A + \bar{B}\bar{K}_i) + Q + \bar{K}_i^T \bar{K}_i = 0 \tag{16}$$

Proof: Dividing both sides of (14) by Δt and taking the limit yields

$$\begin{aligned} 0 &= \lim_{\Delta t \to 0} \frac{x^T(t+\Delta t)P_i x(t+\Delta t) - x^T(t)P_i x(t)}{\Delta t} \\ &\quad + \lim_{\Delta t \to 0} \frac{\int_t^{t+\Delta t} x^T(Q + \bar{K}_i^T \bar{K}_i)xdt}{\Delta t} \\ &= \frac{dx^T(t)P_i x(t)}{dt} + \lim_{\Delta t \to 0} \frac{d}{d\Delta t} \int_t^{t+\Delta t} x^T(Q + \bar{K}_i^T \bar{K}_i)Xdt \\ &= x^T[(A + \bar{B}\bar{K}_i)^T P_i + P_i(A + \bar{B}\bar{K}_i) \\ &\quad + Q + \bar{K}_i^T \bar{K}_i]x \end{aligned}$$

Thus, (14) implies (16).

On the other hand, consider the stable system $\dot{x} = (A + \bar{B}\bar{K}_i)x$; taking the time derivative of Lyapunov function $V_i(x) = x^T P_i x$ along the closed-loop system yields

$$\frac{d}{dt}(x^T P_i x) = x^T(A + \bar{B}\bar{K}_i)^T P_i x + x^T P_i(A + \bar{B}\bar{K}_i)x$$

Integrating from t to $t + \Delta t$ on both sides yields

$$\begin{aligned} x(t+\Delta t)^T P_i x(t+\Delta t) - x(t)^T P_i x(t) \\ = -\int_t^{t+\Delta t} x^T(Q + \bar{K}_i^T \bar{K}_i)xd\tau, \end{aligned}$$

which results in (14). The proof is complete.

Lemma 3: Assume that the matrices F and H satisfy (6) and (7), respectively. For appropriately chosen parameter β that satisfies condition (9), the iteration (14) and (15) converges to the solution of the ARE (8). Thus, it converges to the solution of the robust control problem.

Proof: According to the the results of [53], the iteration (16) and (15) will converge to the solution of the corresponding ARE (8). It follows from Lemma 2 that the iteration (14) and (15) converges to the solution of the corresponding ARE because of the equivalence between (14) and (16) shown in Lemma 2. The proof is complete.

Remark 4: In Step 4 of Algorithm 1, solving P_i from (14) can be reduced to a least squares problem [18]. Through online reading of sufficient data along the system trajectory on the interval $[t, t + \Delta t]$, the matrix P_i can be computed using the least squares method.

Remark 5: Algorithm 1 is a online policy iteration algorithm based on reinforcement learning. In [18], this method is used to solve the optimal regulation problem of linear systems with unknown system dynamics. We develop this algorithm to solve robust control problems for uncertain continuous-time linear systems.

V. MATCHED UNCERTAIN LINEAR SYSTEM

In this section, we consider system (1) that satisfies the matched conditions (2) and (3), which is termed a matched uncertain system. Robust control of the uncertain linear system is obtained by constructing an ARE with properly chosen weighting matrices. An online reinforcement learning method is used to solve the robust control problem.

According to matched conditions (2) and (3), the matched uncertain linear system can be described by

$$\dot{x}(t) = [A + B\phi(s)]x(t) + [B + B\bar{\phi}(l)]u(t) \tag{17}$$

where $\phi(s), \bar{\phi}(l)$ are uncertain matrices of appropriate dimensions.

We construct an ARE as follows:

$$PA + A^T P + Q - PBB^T P = 0 \tag{18}$$

where $Q = F + I$ and F is an upper bound on uncertainty $\phi(s)^T\phi(s)$, i.e. $\phi^T(s)\phi(s) \le F$.

Theorem 2: Assume that $\bar{\phi}(l) \ge 0$ in system (17). Then, the feedback control $u = Kx$ with $K = -B^T P$ can stabilize uncertain linear system (17) for all $s(t) \in \mathcal{S}$, $l(t) \in \mathcal{L}$, where $P > 0$ is the solution of the ARE (18).

Proof: The proof is similar to that of theorem 1 and is omitted.

The linear system corresponding to the ARE (18) is

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{19}$$

and the performance index can be written as

$$V(x(t_0), u(.)) = \int_{t_0}^{\infty} [x^T Qx + u^T u]dt \tag{20}$$

where $Q = F + I > 0$.

Similar to algorithm 1, it is easy to prove the convergence of the following algorithm, which can be used to compute robust controllers for matched uncertain linear system (17).

Algorithm 2: Robust Control Algorithm For Matched Uncertain Linear System.

1. Choose F as an upper bound on uncertainty $\phi(s)^T\phi(s)$ and compute $Q = F + I$;
2. Choose a stabilizing initial control policy $u = K_0 x$ from the nominal linear system (19);
3. Policy evaluation: Given a control policy K_i, solve P_i using the Bellman equation

$$x(t)^T P_i x(t) = \int_t^{t+\Delta t} [x^T Qx + x^T K_i^T K_i x]dt + x(t+\Delta t)^T P_i x(t+\Delta t); \tag{21}$$

4. Policy improvement: update the control input using

$$K_{i+1} = -B^T P_i. \tag{22}$$

Remark 6: The convergence of the algorithm is completely analogous to algorithm 1, which is applicable to unmatched uncertain linear systems. From this perspective, matched linear systems are a special case of unmatched linear systems. Combining (4) and (5) with Definition 2 and Definition 3, the system (1) satisfies matched conditions (2) and (3) when $(I - BB^+)\Delta A(s) = 0$ and $(I - BB^+)\Delta B(l) = 0$. It is a matched system by letting $\phi(s) = B^+\Delta A(s)$ and $\bar{\phi}(l) = B^+\Delta B(l)$. This also shows that matched linear system is a special case of unmatched linear system. We propose a robust control method based on RL, which does not need to know the nominal system matrix of uncertain linear systems. This extends existing robust control methods.

VI. NUMERICAL EXAMPLES

In this section, three simulation examples are provided to demonstrate the feasibility of the theoretical results for robust control of uncertain linear systems.

Example 1: Consider the uncertain linear system

$$\dot{x}(t) = A(s(t))x(t) + B(l(t))u(t) \tag{23}$$

with

$$A(s) = \begin{bmatrix} -3 & -1 & 0 \\ 50 & 0 & s \\ 2 & 1+s & -1 \end{bmatrix}, \quad B(l) = \begin{bmatrix} 0 & 0 \\ 1+l & 0 \\ 0 & l \end{bmatrix}$$

where $-9 \le s(t) \le 3$ and $1 \le l(t) \le 3$ are uncertainties. The objective is to design gain K such that closed-loop matrix $A(s) + B(l)K$ is asymptotically stable for all $-9 \le s(t) \le 3$ and $1 \le l(t) \le 3$.

Let $A = \begin{bmatrix} -3 & -1 & 0 \\ 50 & 0 & 1 \\ 2 & 2 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}$. Obviously, nominal system dynamics (A, B) is controllable.

By simple computation, we have

$$\begin{aligned} \Delta A(s) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & s-1 \\ 0 & s-1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \frac{s-1}{2} \\ 0 & s-1 & 0 \end{bmatrix} \\ &\equiv B\phi(s) \end{aligned} \tag{24}$$

and

$$\begin{aligned} \Delta B(l) &= \begin{bmatrix} 0 & 0 \\ l-1 & 0 \\ 0 & l-1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{l-1}{2} & 0 \\ 0 & l-1 \end{bmatrix} \\ &\equiv B\bar{\phi}(l) \end{aligned} \tag{25}$$

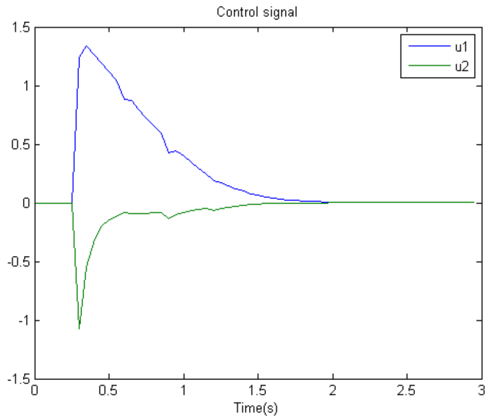

FIGURE 1. **Control signal.**

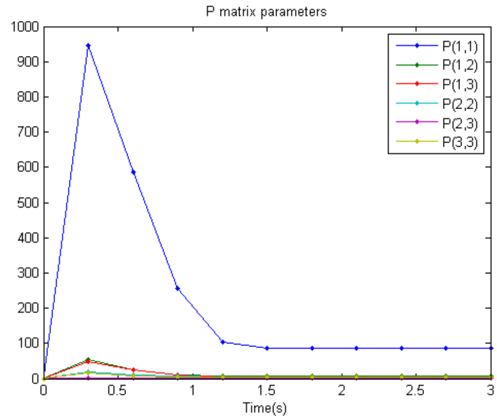

FIGURE 2. **P matrix iteration.**

which implies that the matched conditions are satisfied. It is clear that $\bar{\phi}(l) = \begin{bmatrix} \frac{l-1}{2} & 0 \\ 0 & l-1 \end{bmatrix} \geq 0$ for any $1 \leq l(t) \leq 3$.

$$\phi^T(s)\phi(s) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & (s-1)^2 & 0 \\ 0 & 0 & \frac{(s-1)^2}{4} \end{bmatrix} \leq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 25 \end{bmatrix} = F \tag{26}$$

The weight matrix Q in the ARE (18) is chosen as

$$Q = F + I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 101 & 0 \\ 0 & 0 & 26 \end{bmatrix} \tag{27}$$

The RL Algorithm 2 is used to obtain robust control for uncertain system (23). The initial control policy is chosen as $[0 \quad 0 \quad 0]$, as the nominal system is asymptotically stable. For the purpose of demonstrating the algorithm, the initial state of the nominal system is chosen as $x_0 = [0.1 \quad 0.2 \quad 0.1]$. Using MATLAB software, after 5 iterations, the control gain and P matrix parameters converge to the following optimal solutions:

$$K = \begin{bmatrix} -14.2312 & -9.4012 & -0.8424 \\ -3.0028 & -0.4212 & -4.2093 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 85.6724 & 7.1156 & 3.0028 \\ 7.1156 & 4.7006 & 0.4212 \\ 3.0028 & 0.4212 & 4.2093 \end{bmatrix}$$

Due to the 6 independent elements in the symmetric matrix P, 6 data samples are collected to perform batch least squares in each iteration. The evolution of the feedback control signal is presented in Fig. 1, where $u1$ and $u2$ are two control components. Fig. 2 shows the convergence process of the P matrix. Here, $P(i, j)$ is the element lying on the intersection of the ith row and the jth column in the symmetric matrix P, $i = 1, 2, 3$ and $j = 1, 2, 3$.

TABLE 1. **Eigenvalues for different *s* with *l* = 1.**

l=1,s	λ_1	λ_2	λ_3
l=1,s=-9	-22.8489	-2.0814 - 1.4468i	-2.0814 + 1.4468i
l=1,s=-7	-21.1450	-2.9333 - 1.7601i	-2.9333 + 1.7601i
l=1,s=-5	-19.5876	-3.7120 - 1.6630i	-3.7120 + 1.6630i
l=1,s=-3	-18.2843	-4.3637 - 1.2494i	-4.3637 + 1.2494i
l=1,s=-1	-17.4164	-4.7977 - 0.5057i	-4.7977 + 0.5057i
l=1,s=1	-17.1995	-4.4225	-5.3897
l=1,s=3	-17.7074	-4.6521 - 0.6450i	-4.6521 + 0.6450i

TABLE 2. **Eigenvalues for different *s* with *l* = 2.**

l=1,s	λ_1	λ_2	λ_3
l=2,s=-9	-32.4493	-4.0864-1.4071i	-4.0864 + 1.4071i
l=2,s=-7	-4.3491	-5.2498	-31.0233
l=2,s=-5	-3.6929	-7.1528	-29.7765
l=2,s=-3	-28.7761	-3.5011	-8.3450
l=2,s=-1	-28.1017	-3.3970	-9.1235
l=2,s=1	-27.8267	-3.3303	-9.4653
l=2,s=3	-27.9875	-3.2836	-9.3511

For comparison purposes, using MATLAB software to directly solve the ARE, we obtain the following optimal feedback gain and the P matrix. To avoid confusion, we use the following notations.

$$K_d = \begin{bmatrix} -14.2247 & -9.3959 & -0.8431 \\ -2.9985 & -0.4215 & -4.2089 \end{bmatrix}$$

$$P_d = \begin{bmatrix} 85.4828 & 7.1124 & 2.9985 \\ 7.1124 & 4.6980 & 0.4215 \\ 2.9985 & 0.4215 & 4.2089 \end{bmatrix}$$

Obviously, the results obtained by using the RL method are only marginal different from those obtained by the direct solution of the ARE.

The corresponding partial eigenvalues of the closed-loop uncertain linear system with $u = Kx$ for different s and l are listed in Table 1, Table 2 and Table 3. From Table 1, Table 2 and Table 3, we can see that the eigenvalues of

TABLE 3. Eigenvalues for different s with l = 3.

l=3,s	λ_1	λ_2	λ_3
l=3,s=-9	-41.9127	-3.0537	-9.2663
l=3,s=-7	-40.6478	-2.9968	-10.5881
l=3,s=-5	-39.5553	-2.9532	-11.7243
l=3,s=-3	-38.6787	-2.9158	-12.6382
l=3,s=-1	-38.0640	-2.8809	-13.2878
l=3,s=1	-37.7515	-2.8458	-13.6354
l=3,s=3	-37.7652	-2.8084	-13.6591

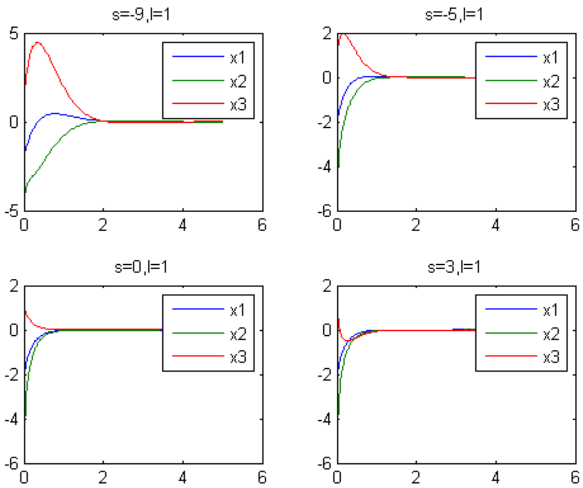

FIGURE 3. The state trajectories of the original system with l = 1.

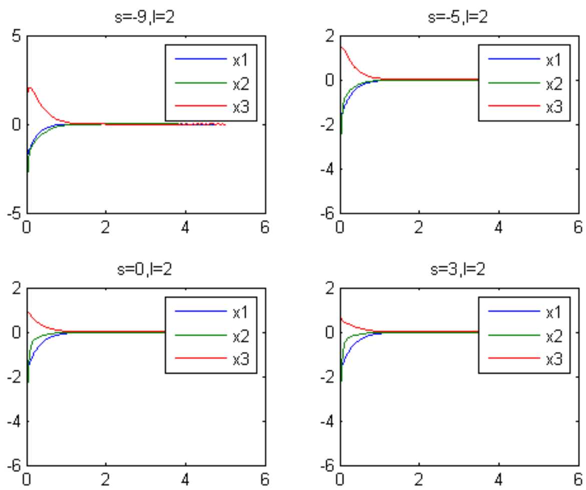

FIGURE 4. The state trajectories of the original system with l = 2.

the closed-loop system all have negative real parts. Thus, the uncertain linear system with robust control $u = Kx$ is asymptotically stable for all $-9 \le s(t) \le 3$ and $1 \le l(t) \le 3$. We chose the initial state of the original system (23) as $x_0 = [-2 \quad -5 \quad 1]$. For different values of the parameter s, Fig. 3, Fig. 4 and Fig. 5 show the orbits of the original system with $l = 1, 2$ and 3, respectively.

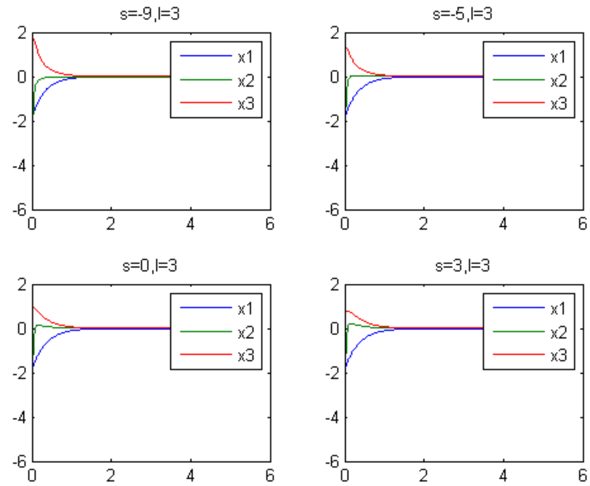

FIGURE 5. The state trajectories of the original system with l = 3.

Example 2: Consider the uncertain linear system with

$$A(s) = \begin{bmatrix} -2 & 0 \\ s & s-5 \end{bmatrix}, \quad B(l) = \begin{bmatrix} l+1 \\ l \end{bmatrix} \tag{28}$$

where $-1 \le s(t) \le 2$ and $0 \le l(t) \le 4$ are uncertainties. We would like to design gain K such that closed-loop matrix $A(s) + B(l)K$ is asymptotically stable for all $-1 \le s(t) \le 2$ and $0 \le l(t) \le 4$.

Let $A = \begin{bmatrix} -2 & 0 \\ 0 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Moreover,

$$\Delta A(s) = \begin{bmatrix} 0 & 0 \\ s & s \end{bmatrix}, \quad \Delta B(l) = \begin{bmatrix} l \\ l \end{bmatrix}$$

It is clear that the system matrix and the input matrix are all unmatched.

$$\begin{aligned} B^+ &= (B^T B)^{-1} B^T = \begin{bmatrix} 1 & 0 \end{bmatrix} \\ I - BB^+ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

We chose $\beta = 5$.

$$\begin{aligned} \Delta A(s)^T B^{+T} B^+ \Delta A(s) &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ &= F \\ \Delta A(s)^T \Delta A(s) &= \begin{bmatrix} s^2 & s^2 \\ s^2 & s^2 \end{bmatrix} \\ &\le \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} \\ &= H \end{aligned}$$

Therefore, the corresponding weight matrix Q in ARE (8) is obtained as follows.

$$Q = F + H + \beta^2 I = \begin{bmatrix} 29 & 4 \\ 4 & 29 \end{bmatrix}$$

By using the RL method, Algorithm 1 is implemented online to obtain robust control for unmatched uncertain linear system (28). The initial control policy is chosen as [0 0 0], as the extended nominal linear system $(A, \bar{B})$ is asymptotically stable. For the purpose of demonstrating the algorithm, the initial state of the extended nominal linear system is chosen as $x_0 = [3.5\ -5]$. Due to the 3 independent elements in the symmetric matrix P, 6 data samples are collected to perform batch least squares in each iteration. Using MATLAB software, after 4 iterations, the control gain and P matrix parameters converge to the following optimal solutions:

$$\bar{K} = \begin{bmatrix} -3.7365 & -0.3058 \\ 0 & 0 \\ -0.3058 & -2.3421 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 3.7365 & 0.3058 \\ 0.3058 & 2.3421 \end{bmatrix}$$

Thus, $K = \begin{bmatrix} -3.7365 & -0.3058 \end{bmatrix}$ is the feedback control gain of unmatched uncertain linear system (28). The feedback control signal evolution is presented in Fig. 6, where $u1$, $u2$ and $u3$ are three control components. Fig. 7 shows the convergence process of the P matrix. Here, $P(i, j)$ is the element lying on the intersection of the ith row and the jth column in the symmetric matrix P, $i = 1, 2$ and $j = 1, 2$.

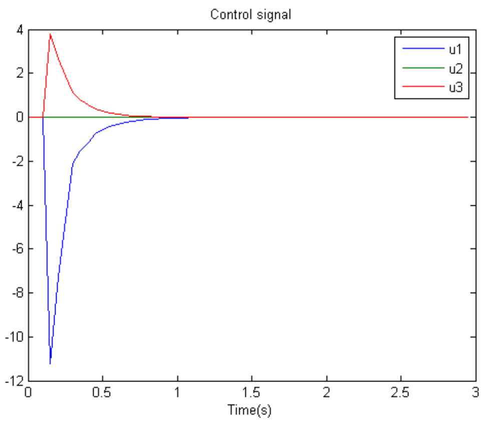

FIGURE 6. **Control signal.**

Using MATLAB software to directly solve the ARE, the following optimal feedback gain and the P matrix are obtained. We use the notations with subscript to distinguish them.

$$\bar{K}_d = \begin{bmatrix} -3.7364 & -0.3058 \\ 0 & 0 \\ -0.3058 & -2.3421 \end{bmatrix}$$

and

$$P_d = \begin{bmatrix} 3.7364 & 0.3058 \\ 0.3058 & 2.3421 \end{bmatrix}$$

By comparison, it is found that solving the robust control problem for an unmatched uncertain linear system using the RL method completely meets the error requirements.

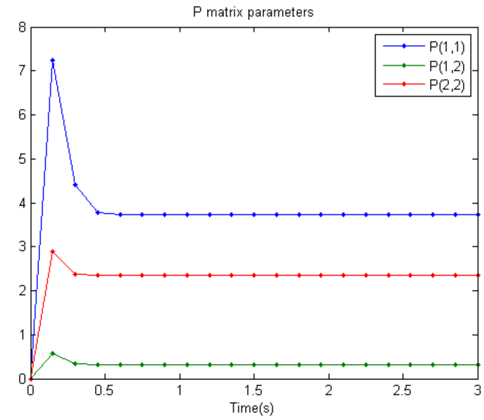

FIGURE 7. **P matrix iteration.**

TABLE 4. **Eigenvalues for different *s* with l = 0, 1.**

l,s	λ_1	λ_2
l=0,s=-1	-5.2998	-6.4367
l=0,s=0	-5.7365	-5.0000
l=0,s=1	-5.5376	-4.1989
l=0,s=2	-5.4910	-3.2455
l=1,s=-1	-10.2142	-5.5646
l=1,s=0	-9.9636	-4.8152
l=1,s=1	-9.7788	-4.0000
l=1,s=2	-9.6407	-3.1381

TABLE 5. **Eigenvalues for different *s* with *l* = 2, 3.**

l,s	λ_1	λ_2
l=2,s=-1	-14.2298	-5.5913
l=2,s=0	-14.0244	-4.7967
l=2,s=1	-13.8521	-3.9690
l=2,s=2	-13.7069	-3.1142
l=3,s=-1	-18.2624	-5.6010
l=3,s=0	-18.0739	-4.7895
l=3,s=1	-17.9074	-3.9560
l=3,s=2	-17.7598	-3.1036

TABLE 6. **Eigenvalues for different *s* with *l* = 4.**

l,s	λ_1	λ_2
l=4,s=-1	-22.2997 -5.6060	
l=4,s=0	-22.1200 -4.7857	
l=4,s=1	-21.9568 -3.9489	
l=4,s=2	-21.8081 -3.0976	

The corresponding partial eigenvalues of the closed-loop system for different s and l are listed in Table 4, Table 5 and Table 6. From the tables, it is clear that the eigenvalues of the uncertain closed-loop system all have a negative real part. Thus, unmatched uncertain linear system (28) with robust control $u = Kx$ is asymptotically stable for all $s \in [-1, 2]$

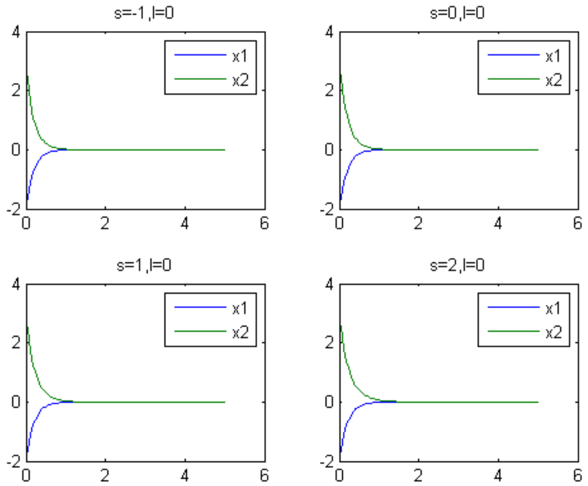

FIGURE 8. **The state trajectories of the original system with l = 0.**

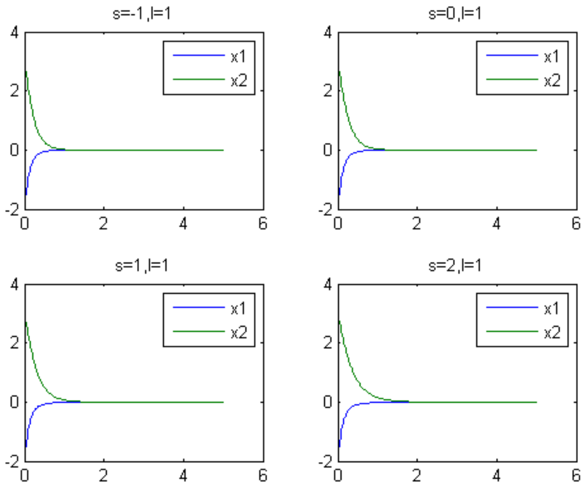

FIGURE 9. **The state trajectories of the original system with l = 1.**

and $l \in [0, 4]$. The initial state of the original system (28) is chosen as $x_0 = [-2\ 3]$. For different values of the parameter s, Fig. 8, Fig. 9 and Fig. 10 show the orbits of the original system with $l = 0$, 1 and 4, respectively.

Example 3: Consider the uncertain linear system with unstable nominal dynamics.

$$\dot{x}(t) = A(s(t))x(t) + B(l(t))u(t) \tag{29}$$

with

$$A(s) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & s-4 & 0 \\ 0 & s+1 & 2 \end{bmatrix}, \quad B(l) = \begin{bmatrix} 0 \\ l+1 \\ l+1 \end{bmatrix}$$

where $-3 \leq s(t) \leq 3$ and $0 \leq l(t) \leq 2$ are uncertainties. The goal is to design a robust controller based on an unstable nominal system.

Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & 0 \\ 0 & 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. Obviously, nominal system dynamics (A, B) is controllable and the matrix A has eigenvalue with positive real part.

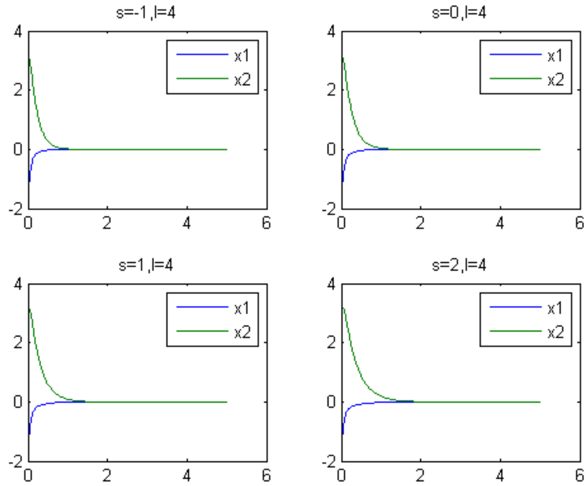

FIGURE 10. **The state trajectories of the original system with l = 4.**

By simple computation, we have

$$\begin{aligned} \Delta A(s) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & s-2 & 0 \\ 0 & s-2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & s-2 & 0 \end{bmatrix} \\ &\equiv B\phi(s) \end{aligned} \tag{30}$$

and

$$\Delta B(l) = \begin{bmatrix} 0 \\ l \\ l \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} l \equiv B\bar{\phi}(l) \tag{31}$$

which implies that the matched conditions are satisfied. It is clear that $\bar{\phi}(l) = l \geq 0$ for any $0 \leq l(t) \leq 2$.

$$\begin{aligned} \phi^T(s)\phi(s) &= \begin{bmatrix} 0 \\ s-2 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & s-2 & 0 \end{bmatrix} \\ &\leq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= F \end{aligned} \tag{32}$$

The weight matrix Q in the ARE (18) is chosen as

$$Q = F + I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 26 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{33}$$

The initial state of the nominal system is chosen as $x_0 = [1 \quad 2 \quad 1]$. The initial matrix P_0 is chosen as

$$P_0 = \begin{bmatrix} 0 & 0 & 2.5 \\ 0 & 0 & 1.7 \\ 2.5 & 1.7 & 1.5 \end{bmatrix}$$

The corresponding initially stabilizing control gain is $K_0 = [-2.5 \quad -1.7 \quad -3.2]$. The RL Algorithm 2 is used to obtain robust control for uncertain system (29). Using MATLAB

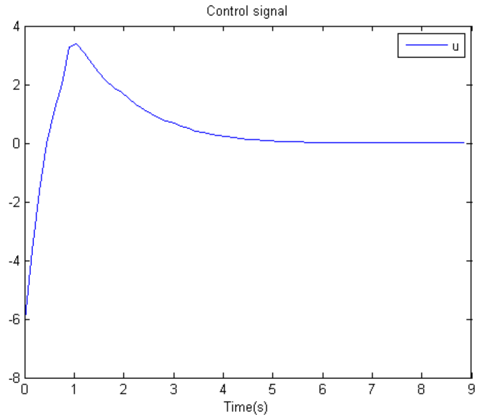

FIGURE 11. **Control signal.**

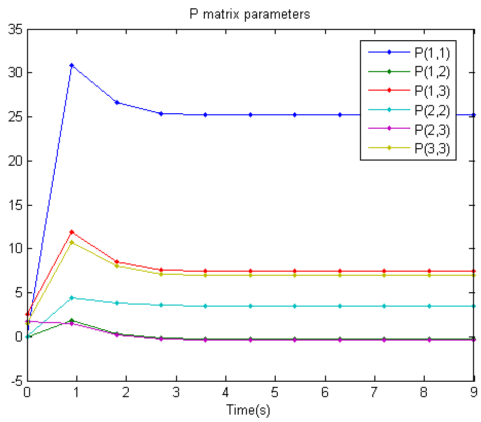

FIGURE 12. **P matrix iteration.**

software, after 4 iterations, the control gain and P matrix parameters converge to the following optimal solutions:

$$K = \begin{bmatrix} -7.1707 & -3.1336 & -6.6162 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 25.1930 & -0.2345 & 7.4052 \\ -0.2345 & 3.4994 & -0.3659 \\ 7.4052 & -0.3659 & 6.9820 \end{bmatrix}$$

The evolution of the control signal u is presented in Fig. 11. Fig. 12 shows the convergence process of the P matrix. Here, $P(i, j)$ is the element lying on the intersection of the ith row and the jth column in the symmetric matrix P, $i = 1, 2, 3$ and $j = 1, 2, 3$.

By solving ARE directly, the following results are obtained.

$$P = \begin{bmatrix} 25.1872 & -0.2346 & 7.4052 \\ -0.2346 & 3.4983 & -0.3672 \\ 7.4052 & -0.3672 & 6.9796 \end{bmatrix}$$

Obviously, the results obtained by using the RL method are only marginal different from those obtained by the direct solution of the ARE.

The initial state of the original system (29) is chosen as $x_0 = [-2\ -2\ 5]$. For different values of the parameter s,

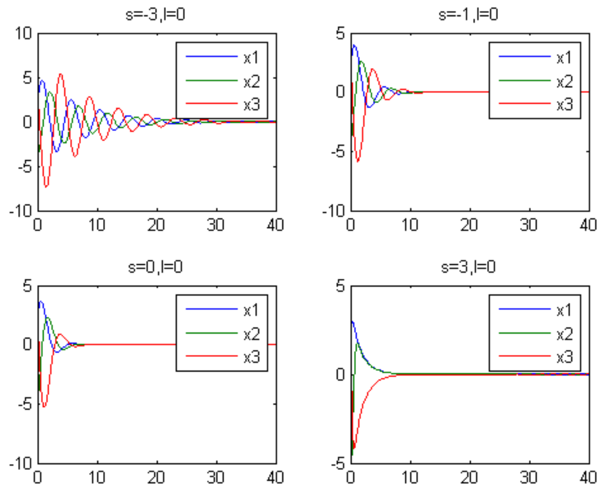

FIGURE 13. **The state trajectories of the original system with l = 0.**

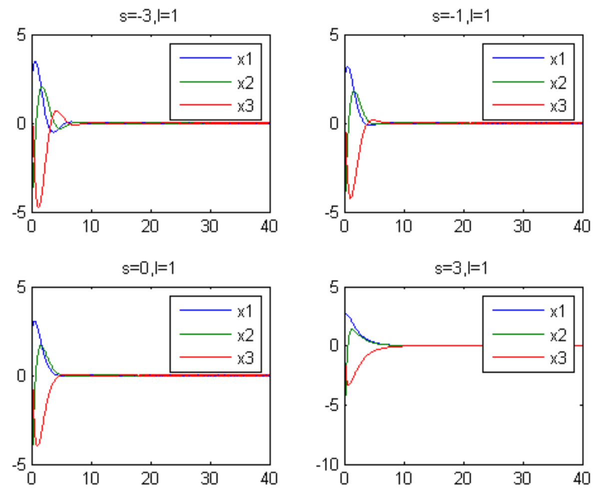

FIGURE 14. **The state trajectories of the original system with l = 1.**

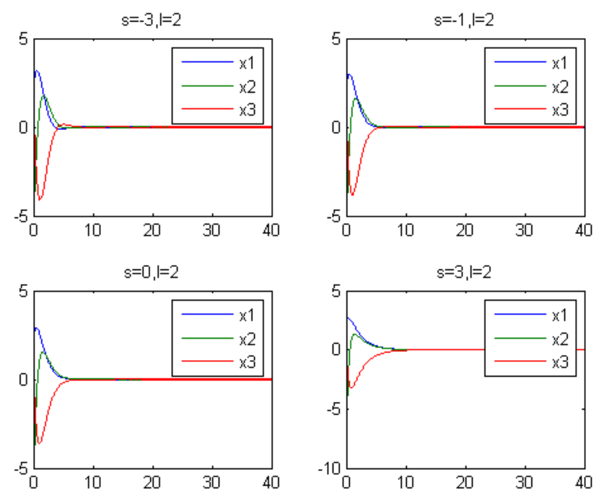

FIGURE 15. **The state trajectories of the original system with l = 2.**

Fig. 13, Fig. 14 and Fig. 15 show the orbits of the original system with $l = 0, 1$ and 2, respectively. As can be seen from the figures, the system orbit quickly converges to zero.

Thus, the uncertain linear system with robust control $u = Kx$ is asymptotically stable for all $-3 \leq s(t) \leq 3$ and $0 \leq l(t) \leq 2$.

VII. CONCLUSION

In this paper, a policy iteration algorithm is proposed to solve robust control problems for uncertain linear systems. The algorithm is based on online RL without the use of nominal system dynamics. Based on the corresponding nominal linear system, an online RL algorithm is established. By using equivalence between an ARE and the integral reinforcement relation, the convergence of the proposed algorithm is proved. Three numerical examples are given to show the correctness of the theoretical results. It can be seen that the RL method used to solve the robust control of uncertain linear systems is effective. This enriches the method for solving robust control problems for uncertain systems. Using the method in [55], the RL algorithm may be extended to solve the robust control of uncertain linear systems in the case of fully unknown nominal system dynamics. Moreover, the method used in this paper may be extended to the robust control problem of uncertain discrete-time systems.

ACKNOWLEDGMENT

The authors would like to thank the editors and reviewers for their helpful suggestions.

REFERENCES

[1] F. Liao, J. L. Wang, and G.-H. Yang, "Reliable robust flight tracking control: An LMI approach," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 1, pp. 76–89, Jan. 2002.
[2] A. A. G. Siqueira, M. H. Terra, and M. Bergerman, *Robust Control of Robots: Fault Tolerant Approaches*. London, U.K.: Springer, 2011.
[3] A. Belmiloudi, *Stabilization, Optimal and Robust Control: Theory and Applications in Biological and Physical Sciences*. London, U.K.: Springer, 2008.
[4] S. P. Bhattacharyya, H. Chapellat, and L. H. Keel, *Robust Control: The Parametric Approach*. Upper Saddle River, NJ, USA: Prentice-Hall, 1995.
[5] M. Djukanovic, M. Khammash, and V. Vittal, "Application of the structured singular value theory for robust stability and control analysis in multimachine power systems. I. Framework development," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1311–1316, Nov. 1998.
[6] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
[7] Y. Zhu, D. Zhao, and X. Li, "Using reinforcement learning techniques to solve continuous-time non-linear optimal tracking problem without system dynamics," *IET Control Theory Appl.*, vol. 12, no. 10, pp. 1339–1347, Jul. 2016.
[8] S. G. Khan, G. Herrmann, F. L. Lewis, T. Pipe, and C. Melhuish, "Reinforcement learning and optimal adaptive control: An overview and implementation examples," *Annu. Rev. Control*, vol. 36, no. 1, pp. 42–59, Apr. 2012.
[9] F. L. Lewis and K. G. Vamvoudakis, "Reinforcement learning for partially observable dynamic processes: Adaptive dynamic programming using measured output data," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 41, no. 1, pp. 14–25, Feb. 2011.
[10] X. Yang, D. Liu, and D. Wang, "Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints," *Int. J. Control*, vol. 87, no. 3, pp. 553–566, Mar. 2014.
[11] P. J. Werbos, "Elements of intelligence," *Cybernetica*, vol. 11, no. 3, p. 131, 1968.
[12] P. J. Werbos, "Approximate dynamic programming for real-time control and neural modeling," in *The Control Handbook*. New York, NY, USA: Van Nostrand, 1992, pp. 493–526.
[13] P. J. Webros, "A menu of designs for reinforcement learning over time," in *Neural Networks for Control*. Cambridge, MA, USA: MIT Press, 1990, pp. 67–95.
[14] P. J. Werbos, "Neural networks for control and system identification," in *Proc. 28th IEEE Conf. Decision Control*, Dec. 1989, pp. 260–265.
[15] L. C. Baird, "Reinforcement learning in continuous time: Advantage updating," in *Proc. IEEE World Congr. Comput. Intell.*, Jun./Jul. 1994, pp. 2448–2453.
[16] K. Doya, "Reinforcement learning in continuous time and space," *Neural Comput.*, vol. 12, no. 1, pp. 219–245, 2000.
[17] S. Ferrari, J. E. Steck, and R. Chandramohan, "Adaptive feedback control by constrained approximate dynamic programming," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 38, no. 4, pp. 982–987, Aug. 2008.
[18] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, no. 2, pp. 477–484, Feb. 2009.
[19] H. Modares and F. L. Lewis, "Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning," *IEEE Trans. Autom. Control*, vol. 59, no. 11, pp. 3051–3056, Nov. 2014.
[20] C. Qin, H. Zhang, and Y. Luo, "Online optimal tracking control of continuous-time linear systems with unknown dynamics by using adaptive dynamic programming," *Int. J. Control*, vol. 87, no. 5, pp. 1000–1009, Dec. 2014.
[21] W. Gao and Z.-P. Jiang, "Adaptive dynamic programming and adaptive optimal output regulation of linear systems," *IEEE Trans. Autom. Control*, vol. 61, no. 12, pp. 4164–4169, Dec. 2016.
[22] S. He, J. Song, Z. Ding, and F. Liu, "Online adaptive optimal control for continuous-time Markov jump linear systems using a novel policy iteration algorithm," *IET Control Theory Appl.*, vol. 9, no. 10, pp. 1536–1543, May 2015.
[23] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits Syst. Mag.*, vol. 9, no. 3, pp. 32–50, Aug. 2009.
[24] S. Bhasin, N. Sharma, P. Patre, and W. Dixon, "Asymptotic tracking by a reinforcement learning-based adaptive critic controller," *J. Control Theory Appl.*, vol. 9, no. 3, pp. 400–409, Jul. 2011.
[25] Y. Yang, D. Wunsch, and Y. Yin, "Hamiltonian-driven adaptive dynamic programming for continuous nonlinear dynamical systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 8, pp. 1929–1940, Aug. 2017.
[26] G. Leitmann, "Guaranteed asymptotic stability for some linear systems with bounded uncertainties," *J. Dyn. Syst. Meas. Control*, vol. 101, no. 3, pp. 212–216, Sep. 1979.
[27] J. S. Thorp and B. R. Barmish, "On guaranteed stability of uncertain linear systems via linear control," *J. Optim. Theory Appl.*, vol. 35, no. 4, pp. 559–579, Dec. 1981.
[28] B. R. Barmish, M. Corless, and G. Leitmann, "A new class of stabilizing controllers for uncertain dynamical systems," *SIAM J. Control Optim.*, vol. 21, no. 2, pp. 246–255, Mar. 1983.
[29] I. R. Petersen, "A stabilization algorithm for a class of uncertain linear systems," *Syst Control Lett.*, vol. 8, no. 4, pp. 351–357, Mar. 1987.
[30] I. R. Petersen and C. V. Hollot, "A Riccati equation approach to the stabilization of uncertain linear systems," *Automatica*, vol. 22, no. 4, pp. 397–411, Jul. 1986.
[31] W. E. Schmitendorf, "Designing stabilizing controllers for uncertain systems using the Riccati equation approach," *IEEE Trans. Autom. Control*, vol. AC-33, no. 4, pp. 376–379, Apr. 1988.
[32] S.-C. Tsay, I.-K. Fong, and T.-S. Kuo, "Robust linear quadratic optimal control for systems with linear uncertainties," *Int. J. Control*, vol. 53, no. 1, pp. 81–96, Jan. 1991.
[33] M.-L. Ni and H.-X. Wu, "A Riccati equation approach to the design of linear robust controllers," *Automatica*, vol. 29, no. 6, pp. 1603–1605, Nov. 1993.
[34] W. E. Schmitendorf, "Design methodology for robust stabilizing controllers," *J. Guid. Control Dyn.*, vol. 10, no. 3, pp. 250–254, May 1987.
[35] F. Jabbari and W. E. Schmitendorf, "A noniterative method for the design of linear robust controllers," *IEEE Trans. Autom. Control*, vol. 35, no. 8, pp. 954–957, Aug. 1990.
[36] S.-C. Tsay, "Robust control for linear uncertain systems via linear quadratic state feedback," *Syst. Control Lett.*, vol. 15, no. 3, pp. 199–205, Sep. 1990.
[37] R. M. Dolphus and W. E. Schmitendorf, "A non-iterative Riccati approach to robust control design," in *Proc. IEEE Amer. Control Conf.*, May 1990, pp. 916–918.

[38] P. P. Khargonekar, I. R. Petersen, and K. Zhou, "Robust stabilization of uncertain linear systems: Quadratic stabilizability and H_∞ control theory," *IEEE Trans. Autom. Control*, vol. 35, no. 3, pp. 356–361, Mar. 1990.
[39] N. S. Tripathy, I. N. Kar, and K. Paul, "Stabilization of uncertain discrete-time linear system with limited communication," *IEEE Trans. Autom. Control*, vol. 62, no. 9, pp. 4727–4733, Sep. 2017.
[40] L. Liu and L. Tang, "Partial state constraints-based control for nonlinear systems with backlash-like hysteresis," *IEEE Trans. Syst. Man Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2018.2841063.
[41] F. Lin and A. W. Olbrot, "An LQR approach to robust control of linear systems with uncertain parameters," in *Proc. 35th IEEE Conf.*, vol. 4, Dec. 1996, pp. 4158–4163.
[42] F. Lin, "An optimal control approach to robust control design," *Int. J. Control*, vol. 73, no. 3, pp. 177–186, 2000.
[43] D. Wang, H. He, and D. Liu, "Adaptive critic nonlinear robust control: A survey," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3429–3451, Oct. 2017.
[44] Y.-J. Liu, L. Tang, S.-C. Tong, C. L. P. Chen, and D.-J. Li, "Reinforcement learning design-based adaptive tracking control with less learning parameters for nonlinear discrete-time MIMO systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 1, pp. 165–176, Jan. 2015.
[45] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, Jan. 2013.
[46] Y.-J. Liu and S. Tong, "Adaptive NN tracking control of uncertain nonlinear discrete-time systems with nonaffine dead-zone input," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 497–505, Mar. 2015.
[47] H. Jiang, H. Zhang, Y. Cui, and G. Xiao, "Robust control scheme for a class of uncertain nonlinear systems with completely unknown dynamics using data-driven reinforcement learning method," *Neurocomputing*, vol. 273, pp. 68–77, Jan. 2018.
[48] Q. Zhang, D. Zhao, and D. Wang, "Event-based robust control for uncertain nonlinear systems using adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 37–50, Jan. 2018.
[49] X. Yang, H. He, and X. Zhong, "Adaptive dynamic programming for robust regulation and its application to power systems," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5722–5732, Jul. 2018.
[50] J. Song *et al.*, "A new iterative algorithm for solving H_∞ control problem of continuous-time Markovian jumping linear systems based on online implementation," *Int. J. Robust Nonlinear Control*, vol. 26, no. 17, pp. 3737–3754, Nov. 2016.
[51] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. Hoboken, NJ, USA: Wiley, 2012.
[52] F. Lin, *Robust Control Design: An Optimal Control Approach*. Hoboken, NJ, USA: Wiley, 2007.
[53] D. L. Kleinman, "On an iterative technique for Riccati equation computations," *IEEE Trans. Autom. Control*, vol. 13, no. 1, pp. 114–115, Feb. 1968.
[54] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*. New York, NY, USA: Springer, 2003.
[55] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

DENGGUO XU received the B.S. degree in mathematics from Yunnan Normal University, Kunming, China, in 2000, and the M.S. degree in mathematics from Yunnan University, Kunming, in 2006. He is currently pursuing the Ph.D. degree with the School of Automation, Beijing Institute of Technology, Beijing, China. He joined the School of Mathematics and Statistics, Chuxiong Normal University, Chuxiong, China, in 2006, where he is currently a Professor. His research interests include robust control, optimal control, neural network control, and fractional order system control theory.

QINGLIN WANG received the B.S. and M.S. degrees in control science and engineering from the Beijing Institute of Technology, Beijing, China, in 1983 and 1987, respectively, and the Ph.D. degree in control science and engineering from the Chinese Academy of Sciences, Beijing, in 1998. In 1990, he joined the School of Automation, Beijing Institute of Technology, where he is currently a Professor. His research interests include equilibrium state control theory, control of mobile robotics, and nonlinear control.

YUAN LI received the B.S. and M.S. degrees in electrical engineering from the Dalian Railway Institute, China, in 1999 and 2002, respectively, and the Ph.D. degree in control science and engineering from the Chinese Academy of Sciences, Beijing, China, in 2006. In 2006, he joined the Beijing Institute of Technology, where he is currently an Associate Professor. His research interests include vision measurement and control in robotics and Web intelligence.