MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Data-Driven Global Robust Optimal Output Regulation of Uncertain Partially Linear Systems

Odekunle, A.; Gao, W.; Wang, Y.

TR2019-080 August 10, 2019

Abstract

In this paper, a data-driven control approach is developed by reinforcement learning (RL) to solve the global robust optimal output regulation problem (GROORP) of partially linear systems with both static uncertainties and nonlinear dynamic uncertainties. By developing a proper feedforward controller, the GROORP is converted into a global robust optimal stabilization problem. A robust optimal feedback controller is designed which is able to stabilize the system in the presence of dynamic uncertainties. The closed-loop system is assured to be input-to-output stable regarding the static uncertainty as the external input. This robust optimal controller is numerically approximated via RL. Nonlinear small-gain theory is applied to show the input-to-output stability for the closed-loop system and thus solves the original GROORP. Simulation results validates the efficacy of the proposed methodology.

IEEE/CAA Journal of Automatica Sinica

Data-Driven Global Robust Optimal Output Regulation of Uncertain Partially Linear Systems

Adedapo Odekunle, *Student Member, IEEE,* Weinan Gao, *Member, IEEE* and Yebin Wang, *Senior Member, IEEE*

***Abstract*—In this paper, a data-driven control approach is developed by reinforcement learning (RL) to solve the global robust optimal output regulation problem (GROORP) of partially linear systems with both static uncertainties and nonlinear dynamic uncertainties. By developing a proper feedforward controller, the GROORP is converted into a global robust optimal stabilization problem. A robust optimal feedback controller is designed which is able to stabilize the system in the presence of dynamic uncertainties. The closed-loop system is assured to be input-to-output stable regarding the static uncertainty as the external input. This robust optimal controller is numerically approximated via RL. Nonlinear small-gain theory is applied to show the input-to-output stability for the closed-loop system and thus solves the original GROORP. Simulation results validates the efficacy of the proposed methodology.**

***Index Terms*—Reinforcement learning, small-gain theory, robust control, output regulation.**

I. Introduction

The output regulation problem aims at designing control strategies to achieve the rejection of a nonvanishing disturbance and forcing the output of dynamic systems to asymptotically track a desired reference. This problem has been tackled for linear systems since 1970s [1]. Due to its relevance to many real-world applications, the output regulation problem for nonlinear systems has also attracted considerable attention with focus on either local, semi-global or global stabilization [2], [3], [4], [5], [6].

In the most existing output regulation problems, both the nonvanishing disturbance and the reference are generated by an autonomous system, named exosystem—wherein the unmodeled disturbance is neglected. In [7], a more generalized case; H_∞ output regulation problem, where both the nonvanishing and unmodeled disturbances were considered. This H_∞ optimal control and output regulation techniques are efficiently combined for robust model-based control design. The H_∞ optimal control can be formulated a zero-sum differential games involving two players; the controller (player 1) and the unmodeled disturbance (player 2) which are the minimizing and the maximizing players, respectively [8]. In this setting, one can solve the output regulation problem for the system with unmodeled disturbances which are static uncertainties. Considering a nonlinear system with dynamic uncertainties, the notion of input-to-state stability and small-gain theory [9], [10] have been employed to solve the global robust output regulation problems [11]. However, we are not aware of any existing work on output regulation problems that taking both static and dynamic uncertainties into consideration. Also, it is noteworthy that most of the existing control strategies to output regulation problems are model-based, which means that an accurate knowledge of system's model is absolutely needed.

Reinforcement Learning (RL) is a non-model-based and data-driven approach which solves optimal control problems via online state and input information [12]. RL has been used to design optimal feedback controllers for both continuous-time and discrete-time systems for achieving stabilization; see [13], [14], [12], [15], [16], [17] wherein optimal cost and feedback controllers are computed using online data. In [18], a robust data-driven approach is proposed to solve control problems in linear and nonlinear systems with dynamic uncertainties. Reference [16] extends the solution to global optimal output regulation problems by incorporating dynamic uncertainties in the system. The exact knowledge of system dynamics and dynamic uncertainties are not required to design the robust optimal controllers.

This paper aims at proposing a novel data-driven solution to the global robust optimal output regulation problem (GROORP) for a class of partially linear composite systems. It is challenging since the system studied in this paper is with unknown dynamics, and both static and dynamic uncertainties. First, we convert the GROORP into a global robust optimal stabilization problem. Then, a data-driven approach is developed to compute the robust adaptive optimal controller and disturbance policy via online input and state information. In the presence of dynamic uncertainty, it is ensured to achieve the rejection of nonvanishing disturbance and forcing its trajectories to asymptotically track the desired reference. With both static and dynamic uncertainty, it is guaranteed that the closed-loop system is input-to-output stable regarding the static uncertainty as the external input. Optimality and global output regulation are both achieved for the class of partially linear systems.

The remainder of this paper is organized as follows. In Section II, we briefly review the linear optimal output regulation problem and linear optimal control theory. Considering static and nonlinear dynamic uncertainties, we formulate the GROORP for a class of partially linear systems in Section III. An offline solution on the basis of nonlinear small-gain theory is proposed therein. In Section IV, the RL technique is employed to design a robust optimal controller via online data. Simulation results on a partially linear system are provided in

A. Odekunle and W. Gao are with the Department of Electrical and Computer Engineering, Allen. E. Paulson College of Engineering and Computing, Georgia Southern University, Statesboro, GA, 30460, USA e-mail: ao02108@georgiasouthern.edu (Adedapo Odekunle), wgao@georgiasouthern.edu (Weinan Gao).

Y. Wang is with is with Mitsubishi Electric Research Laboratories, Cambridge, MA, 02139 USA e-mail: yebinwang@ieee.org.

Section V. Finally, concluding remarks are given in Section VI.

Notations. Throughout this paper, $\mathbb{R}_+$ (resp. $\mathbb{Z}_+$) denotes the set of nonnegative real numbers (resp. integers). $\mathbb{C}^-$ indicates the open left-half complex plane. $|\cdot|$ represents the Euclidean norm for vectors and the induced norm for matrices. A continuous function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ is of class $\mathcal{K}$ if it is strictly increasing and $\alpha(0) = 0$. It belongs to class $\mathcal{K}_\infty$ if, in addition, it is unbounded. A function $\beta : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is of class $\mathcal{KL}$ if for each fixed t, the function $\beta(\cdot, t)$ is of class $\mathcal{K}$ and, for each fixed s, the function $\beta(s, \cdot)$ is non-increasing and tends to 0 at infinity. Id stands for an identity function. $\otimes$ indicates the Kronecker product operator and $\text{vec}(A) = [a_1^T, a_2^T, \cdots, a_m^T]^T$, where $a_i \in \mathbb{R}^n$ are the columns of $A \in \mathbb{R}^{n\times m}$. For an arbitrary column vector $v \in \mathbb{R}^n$, $\text{vecv}(v) = [v_1^2, v_1v_2, \cdots, v_1v_n, v_2^2, v_2v_3, \cdots, v_{n-1}v_n, v_n^2]^T \in \mathbb{R}^{\frac{1}{2}n(n+1)}$. $\text{vecs}(P) = [p_{11}, 2p_{12}, \cdots, 2p_{1m}, p_{22}, 2p_{23}, \cdots, 2p_{m-1,m}, p_{mm}]^T \in \mathbb{R}^{\frac{1}{2}m(m+1)}$ for a symmetric matrix $P \in \mathbb{R}^{m\times m}$. $\lambda_M(P)$ (resp. $\lambda_m(P)$) denotes the maximum (resp. minimum) eigenvalue of a real symmetric matrix P. For any piecewise continuous function $u : \mathbb{R}_+ \to \mathbb{R}^m$, $\|u\|$ stands for $\sup_{t\geq 0} |u(t)|$. For two functions f and g, $(f \circ g)(\cdot) := f(g(\cdot))$.

II. Robust optimal output regulation of linear systems

Considering a class of linear systems with nonvanishing disturbance and reference signals generated by linear exosystems, the robust optimal output regulation problem (ROORP) is formulated by minimizing both static and dynamic optimization problems. Then, we recall the basics of robust control and policy iteration (PI) technique. An approach solving explicitly the regulator equation is presented as well.

A. Problem formulation

To begin with, consider the linear system

$$\dot{x} = Ax + B_1u + B_2\omega + Dv, \tag{1}$$
$$\dot{v} = Ev, \tag{2}$$
$$y = Cx, \tag{3}$$
$$y_d = -Fv, \tag{4}$$
$$e = y - y_d \tag{5}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ the control input, and $v \in \mathbb{R}^q$ the state of the exosystem (2). The exosystem generates both the nonvanishing disturbance $\eta = Dv$ and the reference $y_0 = -Fv$ for the output of the plant $y = Cx \in \mathbb{R}^r$. $e \in \mathbb{R}^r$ is the tracking error. $\omega \in \mathbb{R}^d$ is an unmodeled square integrable disturbance. $A \in \mathbb{R}^{n\times n}$, $B_1 \in \mathbb{R}^{n\times m}$, $B_2 \in \mathbb{R}^{n\times d}$, $C \in \mathbb{R}^{r\times n}$, $E \in \mathbb{R}^{q\times q}$, $F \in \mathbb{R}^{r\times q}$, and $D \in \mathbb{R}^{n\times q}$ are system matrices with (A, B_1) stabilizable. Throughout this paper, the following assumptions are made.

Assumption 2.1: The transmission zeros condition holds, i.e.,

$$\text{rank}\begin{bmatrix} A-\lambda I & B_1 \\ C & 0 \end{bmatrix} = n + r, \forall \lambda \in \sigma(E). \tag{6}$$

Assumption 2.2: All the eigenvalues of E are simple with zero real part.

Based on Assumptions 2.1-2.2, one can get the following technical result.

Theorem 2.1: Let the feedback gain $K \in \mathbb{R}^{m\times n}$ be such that $\sigma(A - B_1K) \in \mathbb{C}^-$. Then, if a controller is designed as $u = -K(x - Xv) + Uv$, where $X \in \mathbb{R}^{n\times q}$ and $U \in \mathbb{R}^{m\times q}$ solve the following equations:

$$\begin{aligned} XE &= AX + B_1U + D, \\ 0 &= CX + F, \end{aligned} \tag{7}$$

then the closed-loop linear system achieves disturbance rejection and asymptotic tracking.

Proof. Letting $\bar{x} = x - Xv, \bar{u} = u - Uv$ and using (7), we have

$$\begin{aligned} \dot{\bar{x}} &= Ax + B_1u + B_2\omega + Dv - XEv \\ &= Ax - B_1K(x - Xv) + (B_1U + D)v + B_2\omega - XEv \\ &= (A - B_1K)\bar{x} + B_2\omega. \end{aligned} \tag{8}$$

Since $\sigma(A - B_1K) \in \mathbb{C}^-$ and $w(t)$ is square integrable, we observe $\lim_{t\to\infty} \bar{x}(t) = 0$ and $\lim_{t\to\infty} \bar{u}(t) = 0$, which implies $\lim_{t\to\infty} e(t) = \lim_{t\to\infty} C\bar{x}(t) = 0$. The proof is completed.

Remark 2.1: (7) is called the linear regulator equation. Assumption 2.1 is made such that (7) is solvable for any matrices D, F [6].

Inspired by [19], [20], we tackle the robust optimal output regulation problem (ROORP) by solving a static optimization Problem 2.1 to find the optimal solution (X^*, U^*) to (7) and a dynamic optimization Problem 2.2 to find the optimal gains K^* and N^*.

Problem 2.1:

$$\begin{aligned} &\min_{(X,U)} \text{Tr}(X^T\bar{Q}X + U^T\bar{R}U), \\ &\text{subject to} \quad (7) \end{aligned} \tag{9}$$

where $\bar{Q} = \bar{Q}^T > 0, \bar{R} = \bar{R}^T > 0$.

One can write the error system of (1)-(5) as:

$$\dot{\bar{x}}^* = A\bar{x}^* + B_1\bar{u}^* + B_2\omega, \tag{10}$$
$$e = C\bar{x}^* \tag{11}$$

where $\bar{x}^* = x - X^*v, \bar{u}^* = u - U^*v$.

Problem 2.2:

$$\begin{aligned} &\min_{\bar{u}}\max_{\omega} \int_0^\infty [(\bar{x}^*)^TQ\bar{x}^* + (\bar{u}^*)^T\bar{u}^* - \gamma^{-2}\omega^T\omega]dt \\ &\text{subject to} \quad (10) - (11), \end{aligned}$$

where $Q = Q^T > 0$, and $\gamma \geq \gamma^* \geq 0$. The γ^* is named by H_∞ gain.

Remark 2.2: In order to solve the ROORP, we ought to design a control policy $u = -K^*(x - X^*v) + U^*v$ and a disturbance policy $\omega = N^*(x - X^*v)$ where optimal control gains K^* and N^*, and optimal regulator parameters X^* and U^* are achieved by solving optimization Problems 2.1 and 2.2. Theorem 2.1 ensures that the resultant closed-loop system achieves disturbance rejection and asymptotic tracking.

Remark 2.3: It is shown in [21, Remark 5] that the Problem 2.1 can be converted as a convex optimization problem with a quadratic cost and linear constraints. The solution to the Problem 2.1 is unique given positive definite matrices $\bar{Q}$ and $\bar{R}$. The motivation of introducing the Problem 2.1 is to optimize the steady-state behavior the system.

B. H_∞ control and Policy Iteration (PI)

By linear optimal control theory, we design an optimal feedback controller $\bar{u}^* = -K^*\bar{x}^*$ and a disturbance policy $\omega^* = N^*\bar{x}^*$ to minimize the cost of Problem 2.2. The optimal feedback control gain K^* and disturbance gain N^* are

$$K^* = {B_1}^T P^*, \tag{12}$$

$$N^* = \gamma^{-2}{B_2}^T P^*, \tag{13}$$

respectively, with $P^* = P^{*T} > 0$ the unique solution to the following game algebraic Riccati equation (GARE)

$$\begin{aligned}A^T P^* + &P^* A + Q \\ &-P^*(B_1 B_1^T - \gamma^{-2} B_2 B_2^T)P^* = 0.\end{aligned} \tag{14}$$

Remark 2.4: From (12) to (14), computing K^* and N^* does not depend on X^* or U^*. Problems 2.1 and 2.2 can be solved separately.

Lemma 2.1 ([22]): Let $K_0 \in \mathbb{R}^{m\times n}$ be any stabilizing control gain, and let $N_0 \in \mathbb{R}^{d\times n}$ be a zero matrix. $P_j = P_j^T > 0$ is the solution to the following Lyapunov equation

$$\begin{aligned}(A - &B_1K_j + B_2N_j)^T P_j + P_j(A - B_1K_j + B_2N_j) \\ &+ Q + K_j^T K_j - \gamma^2 N_j^T N_j = 0\end{aligned} \tag{15}$$

where K_j and N_j, with $j = 1, 2, \cdots$, are defined by

$$K_j = B_1^T P_{j-1} \tag{16}$$

$$N_j = \gamma^{-2} B_2^T P_{j-1} \tag{17}$$

Then, the following properties hold, for $j = 0, 1, 2, \cdots$,

1) $\sigma(A - B_1K_j) \in \mathbb{C}^-$,
2) $P^* \le P_{j+1} \le P_j$,
3) $\lim_{j\to\infty} K_j = K^*, \lim_{j\to\infty} N_j = N^*$ and $\lim_{j\to\infty} P_j = P^*$.

C. Solving regulator equations

Define a Sylvester map $\mathcal{S} : \mathbb{R}^{n\times q} \to \mathbb{R}^{n\times q}$ by

$$\mathcal{S}(X) = XE - AX, X \in \mathbb{R}^{n\times q}. \tag{18}$$

If we choose a $X_1 \in \mathbb{R}^{n\times q}$ such that $CX_1 + F = 0$, and $X_i \in \mathbb{R}^{n\times q}$, for $i = 2 \cdots h+1$, such that all the $\mathrm{vec}(X_i)$ form a basis of $\ker(I_q \otimes C)$, where h is the nullity of $I_q \otimes C$, then a pair $(X_\dagger^0, U_\dagger^0)$ is a solution to the regulator equation (7) if and only if there exist $\alpha_2^0, \alpha_3^0, \cdots, \alpha_{h+1}^0 \in \mathbb{R}$ such that

$$\mathcal{S}(X_\dagger^0) = B_1 U_\dagger^0 + D, \tag{19}$$

$$X_\dagger^0 = X_1 + \sum_{i=2}^{h+1} \alpha_i^0 X_i. \tag{20}$$

If the solution is not unique, we find all linearly independent vectors $\mathrm{vec}(\begin{bmatrix} X_\dagger^k \\ U_\dagger^k \end{bmatrix})$ by seeking sequences $\alpha_i^k \in \mathbb{R}$ such that, for $k = 1, 2, \cdots, H$ with $H = q(m - r)$,

$$X_\dagger^k = \sum_{i=2}^{h+1} \alpha_i^k X_i, BU_\dagger^k = \sum_{i=2}^{h+1} \alpha_i^k \mathcal{S}(X_i). \tag{21}$$

Then, the solution set of (7) is equivalent to

$$\begin{aligned}\mathbb{S} = &\{(X, U) | X = X_\dagger^0 + \sum_{k=1}^{H} \beta_k X_\dagger^k, U = U_\dagger^0 + \sum_{k=1}^{H} \beta_k U_\dagger^k, \\ &\forall \beta_1, \beta_2, \cdots, \beta_H \in \mathbb{R}\}.\end{aligned} \tag{22}$$

If we compute $\mathcal{S}(X_i)$ for $i = 0, 1, \cdots, h+1$ via online data, the solution set of the regulator equation (7) is obtained with unknown system matrices. The proposed method for solving the regulator equation paves the way for online robust optimal controller design in Section IV.

III. Global Robust Optimal Output Regulation of Partially Linear Systems

In this section, we formulate the GROORP of a class of partially composite linear systems. An offline solution to the GROORP is given by developing a global robust optimal controller.

A. GROORP formulation

Motivated by the class of partially linear systems in [23], we study a general class of perturbed partially linear systems:

$$\dot{\zeta} = g(\zeta, y, v), \tag{23}$$

$$\dot{x} = Ax + B_1[u + \Delta(\zeta, y, v)] + B_2\omega + Dv, \tag{24}$$

$$\dot{v} = Ev, \tag{25}$$

$$y = Cx, \tag{26}$$

$$y_d = -Fv, \tag{27}$$

$$e = y - y_d, \tag{28}$$

where $\zeta \in \mathbb{R}^p$ and $v \in \mathbb{R}^q$ represents the states of the dynamic uncertainty (23) and the exosystem (25), respectively. The functions $g(\zeta, y, v) : \mathbb{R}^p \times \mathbb{R}^r \times \mathbb{R}^q \to \mathbb{R}^p$, and $\Delta(\zeta, y, v) : \mathbb{R}^p \times \mathbb{R}^r \times \mathbb{R}^q \to \mathbb{R}^m$ are sufficiently smooth functions satisfying $g(0, 0, 0) = 0$ and $\Delta(0, 0, 0) = 0$. Suppose $A, B_1, B_2, D, g, \Delta$ are unknown with ζ unmeasurable.

Remark 3.1: The GROORP is solvable for the partially linear system (23)-(28) if a robust optimal controller is found by solving optimization Problems 2.1 and 2.2 such that for $v : [0, \infty) \to V$ with V being a prescribed compact set of $\mathbb{R}^q$, any initial conditions $\zeta(0), x(0)$, the trajectory of closed-loop system (23)-(28) exists and is bounded for any $t \ge 0$, and satisfies $\lim_{t\to\infty} e(t) = 0$ when $\omega \equiv 0$. And the system is input-to-output stable for nontrivial ω.

B. Offline solutions to GROORP

Let Σ_v be the class of piecewise functions from $[0,\infty)$ to V. Then two assumptions are made on the system (23)-(28):

Assumption 3.1: A sufficiently smooth function $\boldsymbol{\zeta}(v)$ with $\boldsymbol{\zeta}(0)=0$ exists satisfying the following equation for any $v\in\mathbb{R}^q$:

$$\begin{aligned}\frac{\partial\boldsymbol{\zeta}(v)}{\partial v}Ev &= g(\boldsymbol{\zeta}(v),y_d,v),\\ 0 &= \Delta(\boldsymbol{\zeta}(v),y_d,v).\end{aligned}\tag{29}$$

Under equations (7) and (29), we write the error system of (23)-(28) by letting $\bar{\zeta}=\zeta-\boldsymbol{\zeta}(v)$,

$$\dot{\bar{\zeta}}=\bar{g}(\bar{\zeta},e,v),\tag{30}$$

$$\dot{\bar{x}}=A\bar{x}+B_1[\bar{u}+\bar{\Delta}(\bar{\zeta},e,v)]+B_2\omega,\tag{31}$$

$$e=C\bar{x},\tag{32}$$

where

$$\begin{aligned}\bar{g}(\bar{\zeta},e,v)&=g(\zeta,y,v)-g(\boldsymbol{\zeta}(v),y_d,v),\\ \bar{\Delta}(\bar{\zeta},e,v)&=\Delta(\zeta,y,v)-\Delta(\boldsymbol{\zeta}(v),y_d,v).\end{aligned}$$

Two assumptions are made on the dynamic uncertainty, i.e., $\bar{\zeta}$-system, with e as the input and $\bar{\Delta}$ as the output.

Assumption 3.2: There exist a function σ_s of class $\mathcal{KL}$ and a function γ_s of class $\mathcal{K}$, both of which are independent of any $v\in\Sigma_v$ such that for any measurable locally essentially bounded e on $[0,T)$ with $0<T\leq+\infty$ and any $v\in\Sigma_v$, $\bar{\zeta}(t)$ right maximally defined on $[0,T')(0<T'\leq T)$ satisfies

$$|\bar{\zeta}(t)|\leq\sigma_s(|\bar{\zeta}(0)|,t)+\gamma_s(\|[e_{[0,t]}^T,\bar{\Delta}_{[0,t]}^T]^T\|),\forall t\in[0,T'),$$

where $e_{[0,t]}$ and $\bar{\Delta}_{[0,t]}$ are the truncated functions of e and $\bar{\Delta}$ over $[0,t]$, respectively.

Assumption 3.3: There exist a function σ_Δ of class $\mathcal{KL}$ and a function γ_Δ of class $\mathcal{K}$, both of which are independent of any $v\in\Sigma_v$ such that, for any initial state $\bar{\zeta}(0)$, any measurable locally essentially bounded e on $[0,T)$ with $0<T\leq+\infty$ and any $v\in\Sigma_v$, $\bar{\Delta}(t)$ right maximally defined on $[0,T')(0<T'\leq T)$ satisfies

$$|\bar{\Delta}(t)|\leq\sigma_\Delta(|\bar{\zeta}(0)|,t)+\gamma_\Delta(\|e_{[0,t]}\|),\forall t\in[0,T').\tag{33}$$

Remark 3.2: Assumptions 3.2 and 3.3 are made so that system (30) has strong unboundedness observability (SUO) [24] with zero-offset and input-to-output stability (IOS) [25] properties . Then by nonlinear small-gain theory, a controller exists to globally asymptotically stabilize the error system. Similar assumptions appear in [11].

Theorem 3.1: Under Assumptions 3.2 and 3.3, let symmetric matrices $Q\geq\gamma_x I_n, R=I_m$ with $\gamma_x>0$. If the gain function $\gamma_\Delta(s)$ satisfies the following inequality

$$\gamma_\Delta(s)\leq(Id+\rho_1)^{-1}\circ\gamma_e^{-1}\circ(Id+\rho_2)^{-1}(s),\forall s\geq 0\tag{34}$$

for $\gamma_e(s)=|C|\sqrt{1/\gamma_x}s$ and ρ_1,ρ_2 of class $\mathcal{K}_\infty$, then, for any exostate v, the error system (30)-(32) in closed-loop with the optimal control policy $\bar{u}=-K^*\bar{x}$ is globally asymptotically stable when $\omega\equiv 0$. Moreover, when ω is nontrivial, the closed-loop system is input-to-output stable regarding ω as an input [24].

Proof. The GARE can be rewritten as

$$\begin{aligned}&(A-B_1K^*)^TP^*+P^*(A-B_1K^*)+Q\\ &+P^*B_1B_1^TP^*+\gamma^{-2}P^*B_2B_2^TP^*=0\end{aligned}\tag{35}$$

Differentiating the Lyapunov function $V=\bar{x}^TP^*\bar{x}$ gives

$$\begin{aligned}\dot{V}=&\bar{x}^T[(A-B_1K^*)^TP^*+P^*(A-B_1K^*)]\bar{x}+2\bar{x}^TP^*B_1\bar{\Delta}\\ &+2\bar{x}^TP^*B_2\omega\\ =&-\bar{x}^T(Q+P^*B_1B_1^TP^*+\gamma^{-2}P^*B_2B_2^TP^*)\bar{x}\\ &+2\bar{x}^TP^*B_1\bar{\Delta}+2\bar{x}^TP^*B_2\omega\\ \leq&-\bar{x}^TQ\bar{x}-|\bar{\Delta}-B_1^TP^*\bar{x}|^2-|\gamma w-\gamma^{-1}B_2^TP^*\bar{x}|^2\\ &+|\bar{\Delta}|^2+\gamma^2|\omega|^2\\ \leq&-\bar{x}^TQ\bar{x}+|\bar{\Delta}|^2+\gamma^2|\omega|^2\\ \leq&-\gamma_x|\bar{x}|^2+|\bar{\Delta}|^2+\gamma^2|\omega|^2\end{aligned}\tag{36}$$

for any $t\geq 0$, we have

$$\begin{aligned}V(t)\leq&\exp\left(-\frac{\gamma_x}{\lambda_m(P^*)}t\right)V(0)+\frac{\lambda_m(P^*)}{\gamma_x}\|\bar{\Delta}\|^2\\ &+\frac{\gamma^2\lambda_m(P^*)}{\gamma_x}\|\omega\|^2.\end{aligned}\tag{37}$$

An immediate consequence of the previous inequality is

$$\begin{aligned}|\bar{x}(t)|\leq&\exp\left(-\frac{\gamma_x}{2\lambda_m(P^*)}t\right)\sqrt{\frac{\lambda_M(P^*)}{\lambda_m(P^*)}}|\bar{x}(0)|\\ &+\sqrt{\frac{1}{\gamma_x}}\|\bar{\Delta}\|+\gamma\sqrt{\frac{1}{\gamma_x}}\|\omega\|,\quad\forall t\geq 0,\end{aligned}\tag{38}$$

which implies that the $\bar{x}$-system with the pair $(\bar{\Delta},\omega)$ as the input is input-to-state stable [26]. One can write

$$|e(t)|\leq\sigma_e(|\bar{x}_0|,t)+\gamma_e\|\bar{\Delta}\|+\gamma\gamma_e\|\omega\|,\tag{39}$$

where

$$\sigma_e(|\bar{x}_0|,t)=|C|\exp\left(-\frac{\gamma_x}{2\lambda_m(P^*)}t\right)\sqrt{\frac{\lambda_M(P^*)}{\lambda_m(P^*)}}|\bar{x}_0|$$

is a function of $\mathcal{KL}$ and $\gamma_e=|C|\sqrt{1/\gamma_x}$, which guarantees that the $\bar{x}$-system with e as output has SUO property with zero-offset and IOS properties [24]. On the other hand, Assumptions 3.2 and 3.3 indicate that the $\bar{\zeta}$-system has SUO property with zero-offset and IOS properties with input-to-output gain function $\gamma_\Delta(s)$. By the nonlinear small-gain theory [24], under the following small-gain condition

$$(Id+\rho_2)\circ\gamma_e\circ(Id+\rho_1)\circ\gamma_\Delta(s)\leq s,\forall s\geq 0,\tag{40}$$

the error system (30)-(32) with $\bar{u}=-K^*\bar{x}$ is globally asymptotically stable at the origin if $\omega\equiv 0$. For a nontrivial square-integrable disturbance ω, one can achieve that the closed-loop system is input-to-output stable regarding ω as an external input. □.

C. Solvability of GROORP

Now, we are ready to design a robust optimal controller to solve the GROORP of the partially linear system (23)-(28).

Theorem 3.2: Under the conditions of Assumptions 2.1,2.2-3.3, if weight matrices are chosen $Q = Q^T \geq \gamma_x I_n, R = I_m$ such that small-gain condition (40) holds, then the GROORP of the partially linear system (23)-(28) is solvable by the robust optimal controller $u = -K^*(x - X^*v) + U^*v$.

Proof. By Theorem 3.1, the robust optimal feedback controller $\bar{u}^* = -K^*\bar{x}^*$ globally asymptotically stabilizes the error system (30)-(32) for any $v(t)$. Then, the trajectory of error system satisfies $\lim_{t\to\infty} \bar{\zeta}(t) = 0$ and $\lim_{t\to\infty} \bar{x}^*(t) = 0$ for $\omega \equiv 0$. We observe

$$\lim_{t\to\infty} e(t) = C\bar{x}^*(t) + (CX^* + F)v(t) = 0, \tag{41}$$

for any $x(0), \zeta(0)$. Also, it is checkable that the input-to-output stability of the closed-loop system still holds. The proof is completed.

IV. RL Online Learning

A novel online learning strategy is presented to solve X^*, U^* and online approximation of optimal values P^* and K^*. Similar as the robust adaptive dynamic programming strategy studied in [18], suppose Δ is available during the learning phase. Defining $\bar{x}_i = x - X_i v$ for $i = 0, 1, 2, \cdots, h+1$ with $X_0 = 0_{n\times q}$, we have

$$\begin{aligned}\dot{\bar{x}}_i =& Ax + B_1(u + \Delta) + B_2\omega + (D - X_iE)v \\ =& A_j\bar{x}_i + B_1(K_j\bar{x}_i + z) + B_2(\omega - N_j\bar{x}_i) \\ &+ (D - \mathcal{S}(X_i))v,\end{aligned} \tag{42}$$

where $A_j = A - B_1K_j + B_2N_j$, $z = u + \Delta$.

Then

$$\begin{aligned}&\bar{x}_i(t+\delta t)^T P_j \bar{x}_i(t+\delta t) - \bar{x}_i(t)^T P_j \bar{x}_i(t) \\ =& \int_t^{t+\delta t} \left[\bar{x}_i^T(A_j^TP_j + P_jA_j)\bar{x}_i + 2(z + K_j\bar{x}_i)^T B_1^T P_j \bar{x}_i \right. \\ &\left. + 2v^T(D - \mathcal{S}(X_i))^T P_j\bar{x}_i + 2(\omega - N_j\bar{x}_i)^T B_2^T P_j \bar{x}_i\right] d\tau \\ =& -\int_t^{t+\delta t} \bar{x}_i^T(Q + K_j^TRK_j - \gamma^2 N_j^TN_j)\bar{x}_i d\tau \\ &+ 2\int_t^{t+\delta t} (z + K_j\bar{x}_i)^T R K_{j+1}\bar{x}_i d\tau \\ &+ 2\int_t^{t+\delta t} v^T(D - \mathcal{S}(X_i))^T P_j \bar{x}_i d\tau \\ &+ 2\gamma^2\int_t^{t+\delta t} (\omega - N_j\bar{x}_i)^T N_{j+1}\bar{x}_i d\tau.\end{aligned} \tag{43}$$

For a large enough positive integer l and two vectors $a \in \mathbb{R}^{n_a}$, $b \in \mathbb{R}^{n_b}$, we define

$$\begin{aligned}&\Gamma_{ab} = [\int_{t_0}^{t_1} a\otimes b \mathrm{d}\tau, \int_{t_1}^{t_2} a\otimes b\mathrm{d}\tau, \cdots, \int_{t_{l-1}}^{t_l} a\otimes b\mathrm{d}\tau]^T, \\ &\delta_{\bar{x}_i\bar{x}_i} = [\mathrm{vecv}(\bar{x}_i(t_1)) - \mathrm{vecv}(\bar{x}_i(t_0)), \mathrm{vecv}(\bar{x}_i(t_2)) - \\ &\mathrm{vecv}(\bar{x}_i(t_1)), \cdots, \mathrm{vecv}(\bar{x}_i(t_l)) - \mathrm{vecv}(\bar{x}_i(t_{l-1}))]^T,\end{aligned}$$

where $t_0 < t_1 < \cdots < t_l$ are positive integers. (43) indicates the following equation.

$$\Psi_{ij}\begin{bmatrix}\mathrm{vecs}(P_j) \\ \mathrm{vec}(K_{j+1}) \\ \mathrm{vec}((D - \mathcal{S}(X_i))^TP_j) \\ \mathrm{vec}(N_{j+1})\end{bmatrix} = \Phi_{ij}, \tag{44}$$

where

$$\begin{aligned}\Psi_{ij} =& [\delta_{\bar{x}_i\bar{x}_i}, -2\Gamma_{\bar{x}_i\bar{x}_i}(I_n \otimes (K_j^TR) - 2\Gamma_{\bar{x}_iz}(I_n\otimes R), \\ &- 2\Gamma_{\bar{x}_iv}, -2\gamma^2(\Gamma_{\bar{x}_i\omega} - \Gamma_{\bar{x}_i\bar{x}_i}(I_n\otimes N_i^T))], \\ \Phi_{ij} =& -\Gamma_{\bar{x}_i\bar{x}_i}\mathrm{vec}(Q + (K_i^j)^TRK_i^j - \gamma^2 N_i^TN_i).\end{aligned}$$

Equation (44) is uniquely solved by the least squares method if the matrix Ψ_{ij} is of full column rank, i.e.,

$$\begin{bmatrix}\mathrm{vecs}(P_j) \\ \mathrm{vec}(K_i^{j+1}) \\ \mathrm{vec}((D - \mathcal{S}(X_i))^TP_j) \\ \mathrm{vec}(N_{j+1})\end{bmatrix} = (\Psi_{ij}^T\Psi_{ij})^{-1}\Psi_{ij}^T\Phi_{ij}. \tag{45}$$

Note that D is computable by (45) given $\mathcal{S}(X_0) = 0$. If we seek a sequence $\alpha_2^0, \alpha_3^0, \cdots, \alpha_{h+1}^0 \in \mathbb{R}$ and a matrix $U_\dagger^0 \in \mathbb{R}^{m\times q}$ such that

$$\mathcal{S}(X_1) + \sum_{i=2}^{h+1}\alpha_i^0\mathcal{S}(X_i) = P_j^{-1}K_{j+1}RU_\dagger^0 + D, \tag{46}$$

then $(X_\dagger^0, U_\dagger^0)$ is a solution to the regulator equation (7), where $X_\dagger^0 = X_1 + \sum_{i=2}^{h+1}\alpha_i^0X_i$.

If the solution to (7) is not unique, we find all linearly independent vectors $\mathrm{vec}(\begin{bmatrix}X_\dagger^k \\ U_\dagger^k\end{bmatrix})$ by seeking sequences $\alpha_2^k, \alpha_3^k, \cdots, \alpha_{h+1}^k \in \mathbb{R}$ such that for $k = 1, 2, \cdots, H$ with $H = q(m - r)$

$$X_\dagger^k = \sum_{i=2}^{h+1}\alpha_i^kX_i, \sum_{i=2}^{h+1}\alpha_i^k\mathcal{S}(X_i) = P_j^{-1}K_{j+1}RU_\dagger^k. \tag{47}$$

Then, we define a set:

$$\begin{aligned}\mathbb{S} =& \{(X, U)| X = X_\dagger^0 + \sum_{k=1}^H\beta_kX_\dagger^k, U = U_\dagger^0 + \sum_{k=1}^H\beta_kU_\dagger^k, \\ &\forall\beta_1, \beta_2, \cdots, \beta_H \in \mathbb{R}\}.\end{aligned} \tag{48}$$

Theorem 4.1: Given a stabilizing $K_0 \in \mathbb{R}^{m\times n}$, if Ψ_{ij} is in full column rank for $i = 0, 1, \cdots, h+1$, $j \in \mathbb{Z}_+$, the sequences $\{P_j\}_{j=0}^\infty$, $\{K_j\}_{j=1}^\infty$ obtained from solving (45) converge to P^* and K^*, respectively.

Proof. Given a stabilizing K_j, if $P_j = P_j^T$ is the solution of (15), K_{j+1} and N_{j+1} is determined by $K_{j+1} = R^{-1}B_1^TP_j$ and $N_{j+1} = \gamma^{-2}B_2^TP_j$, respectively. Let $T_j = (\mathcal{S}(X_i))^TP_j$. By (43), we know that P_j, K_{j+1} and T_j satisfy (45). On the other hand, let $P = P^T \in \mathbb{R}^{n\times n}$, $K \in \mathbb{R}^{m\times n}$, $N \in \mathbb{R}^{d\times n}$ and $T \in \mathbb{R}^{q\times n}$, such that

$$\Psi_{ij}\begin{bmatrix}\mathrm{vecs}(P) \\ \mathrm{vec}(K) \\ \mathrm{vec}(T) \\ \mathrm{vec}(N)\end{bmatrix} = \Phi_j.$$

Then, we have $P_j = P$, $K_{j+1} = K$, $N_{j+1} = N$, $T_j = T$. Moreover, P, K, N, T are unique when Ψ_{ij} is in full column rank. By Lemma 2.1, the convergence of P_j, K_j and N_j is proved.

Algorithm 1 RL Algorithm Algorithm for Solving GROORP

Select a K_0 such that $\sigma(A-B_1K_0) \in \mathbb{C}^-$ and a threshold $\epsilon > 0$. Choose $Q = Q^T \geq \gamma_x I_n$ such that the small-gain condition holds. Compute trails $X_0, X_1, \cdots, X_{h+1}$
Employ $u = -K_0x + \xi$ as the control input on $[t_0, t_l]$ with ξ an exploration noise.
$j \leftarrow 0$, $i \leftarrow 0$
repeat
  Solve P_j, K_{j+1}, N_{j+1} from (45).
  $j \leftarrow j+1$
until $|P_j - P_{j-1}| < \epsilon$
Obtain the approximated optimal control gains K^* and N^*, and approximated solution P^* to (14)
repeat
  Solve $\mathcal{S}(X_i)$ from (46)
  $i \leftarrow i+1$
until $i = h+2$
Obtain (X^*, U^*) by solving Problem 2.1
The robust optimal controller $u = -K_j(x - X^*v) + U^*v$ and the optimal disturbance policy $\omega = N_j(x - X^*v)$ are computed.

V. Example

Consider a partially linear system:

$$\begin{aligned}
\dot{\zeta} &= -\zeta^3 + \zeta e, \\
\dot{x} &= \begin{bmatrix} -1 & -2 \\ 0.5 & -2 \end{bmatrix} x + \begin{bmatrix} 2 \\ 2 \end{bmatrix} (u + v_1\zeta^2) + \begin{bmatrix} 1 \\ 4 \end{bmatrix} \omega, \\
\dot{v} &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} v, \\
e &= x_1 + v_2.
\end{aligned}$$

In this example, for any $v \in \mathbb{R}^2$, $\boldsymbol{\zeta}(v) = 0$ satisfies the Assumption 3.1. Taking $V_\zeta = \zeta^2/2$, the derivative of V along the trajectories of the dynamic uncertainty is given by

$$\begin{aligned}
\dot{V}_\zeta &= -\zeta^4 + \zeta^2 e \\
&= -0.5\zeta^4 - 0.5\zeta^4 + \zeta^2 e \\
&\leq -0.5\zeta^4, \quad \forall |\zeta| \geq \sqrt{\frac{|e|}{0.5}}
\end{aligned} \tag{49}$$

Given the fact that $\Delta = \zeta^2$, it is checkable that Assumptions 3.2 and 3.3 are satisfied with gain function

$$\gamma_\Delta(s) = \frac{s}{0.5}.$$

If $\gamma_e(s) < 0.5s$, the error system (30)-(31) is guaranteed globally asymptotically stable at the origin. In this paper, we choose $Q = 5I_2, \gamma = 11$, the initial stabilizing feedback control gain matrix as $K_0 = \begin{bmatrix} 0 & 0 \end{bmatrix}$, the initial disturbance control gain as $N_0 = \begin{bmatrix} 0 & 0 \end{bmatrix}$, the exploration noise as $\xi = \sin(2t)$, and the convergent criterion as $\epsilon = 10^{-8}$. For $i = 1, 2, 3$, matrices X_i are chosen

$$X_1 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}, X_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, X_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Moreover, the initial values of states are $x = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$ and $v = \begin{bmatrix} 3 & 0 \end{bmatrix}^T$. The online data is collected from $t = 0s$ to $t = 15s$. After that, we iteratively compute the optimal values and convergence is attained after 6 iterations. Figs. 1-3 depict the errors between P_j and P^*, between K_j and K^*, and N_j and N^*.

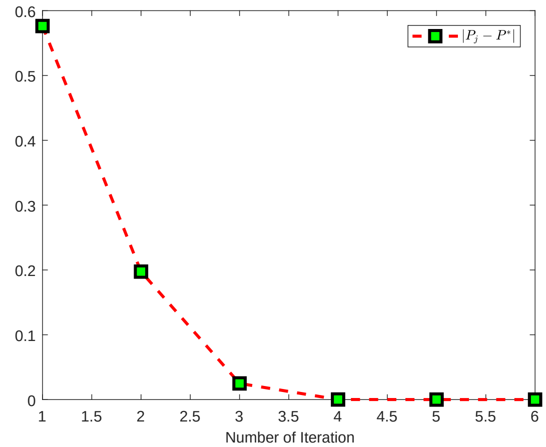

Fig. 1: Convergence of P_j to its optimal value P^* during the learning process

For $i = 0, 1, 2, 3$, we solve the linear map $\mathcal{S}(X)_i$ from online information. From (46) and (47), we get the set of unique solution of regulator equation, which is also the optimal solution (X^*, U^*):

$$X^* = \begin{bmatrix} 0.0000 & -1.0000 \\ 1.4999 & -1.0003 \end{bmatrix}, U^* = \begin{bmatrix} 0.9996 & -1.5002 \end{bmatrix}.$$

Then we get the robust optimal controller and optimal disturbance policy

$$\begin{aligned}
u &= -\begin{bmatrix} 0.8016 & 1.5239 \end{bmatrix} x + \begin{bmatrix} 3.2853 & -3.8262 \end{bmatrix} v, \\
\omega &= \begin{bmatrix} -0.0268 & 0.0553 \end{bmatrix} x + \begin{bmatrix} 0.0829 & -0.0285 \end{bmatrix} v,
\end{aligned} \tag{50}$$

respectively. The learned controller is implemented after $t = 15s$. Fig. 4 depicts that the output of the plant asymptotically tracks the reference. Figs. 5-8 depict the trajectories of the states, the control input, the disturbance and the dynamic uncertainty respectively.

In order to validate the effect of disturbances on the cost, we change the disturbance input by

$$\omega = \frac{1}{2}\left(\begin{bmatrix} -0.0268 & 0.0553 \end{bmatrix} x + \begin{bmatrix} 0.0829 & -0.0285 \end{bmatrix} v\right). \tag{51}$$

We record the cost

$$\int_0^{500} [(\bar{x}^*)^T Q \bar{x}^* + (\bar{u}^*)^T \bar{u}^* - \gamma^{-2}\omega^T\omega] dt$$

for different disturbances until $t = 500s$. This is reasonable since the cost does not change significantly after $t > 500s$ for a stabilized system. It is obtained that the cost under disturbance (51) has reduced by 21.7296 compared with the cost under (50).

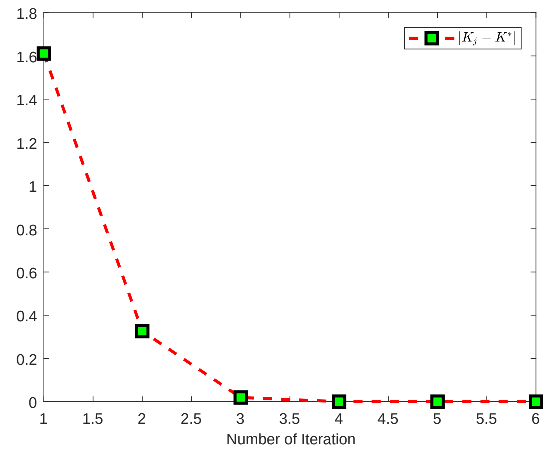

Fig. 2: Convergence of K_j to its optimal value K^* during the learning process

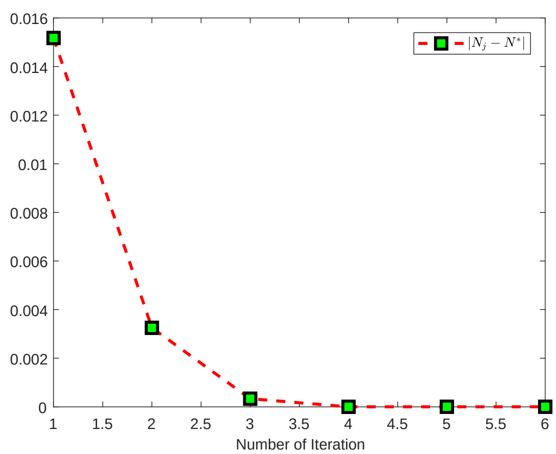

Fig. 3: Convergence of N_j to its optimal value N^* during the learning process

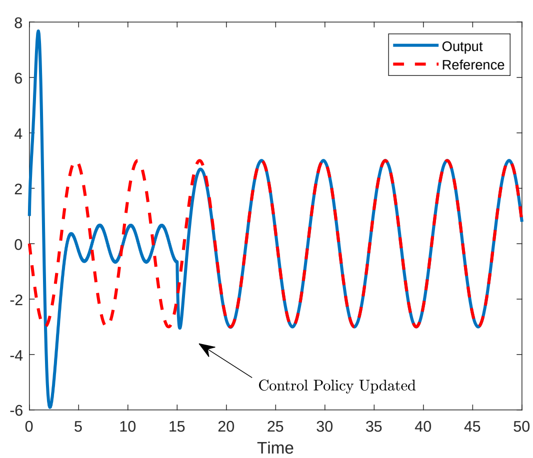

Fig. 4: Trajectories of the output and reference

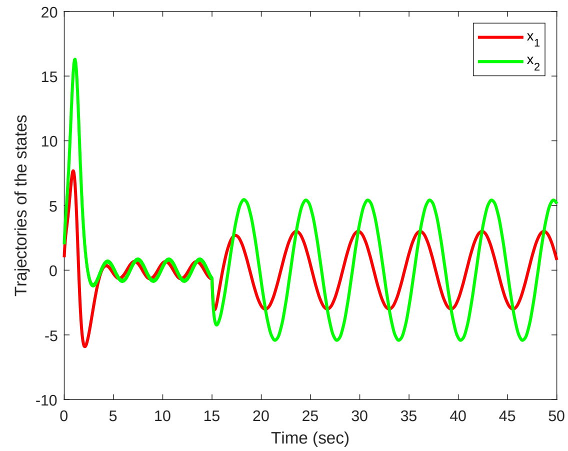

Fig. 5: Trajectories of the states

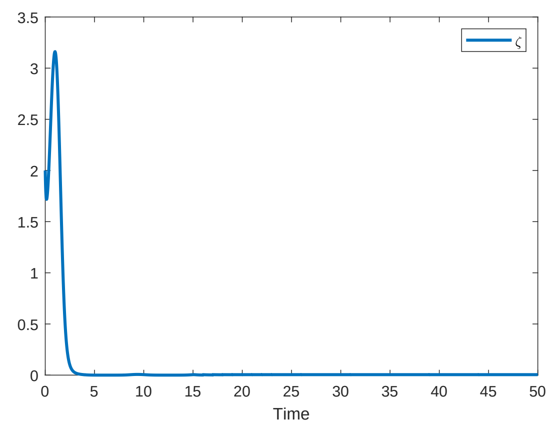

Fig. 6: Trajectory of the dynamic uncertainty $\zeta(t)$

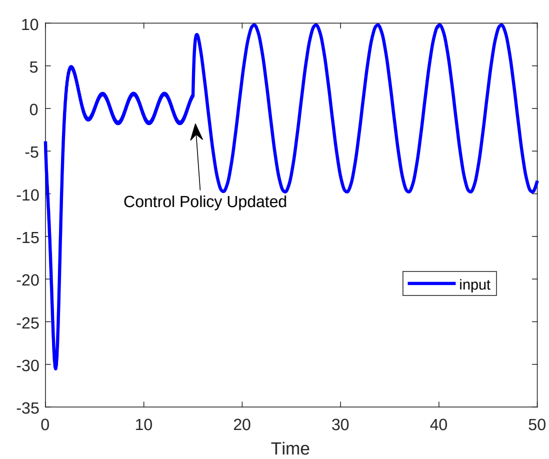

Fig. 7: Trajectory of control input $u(t)$

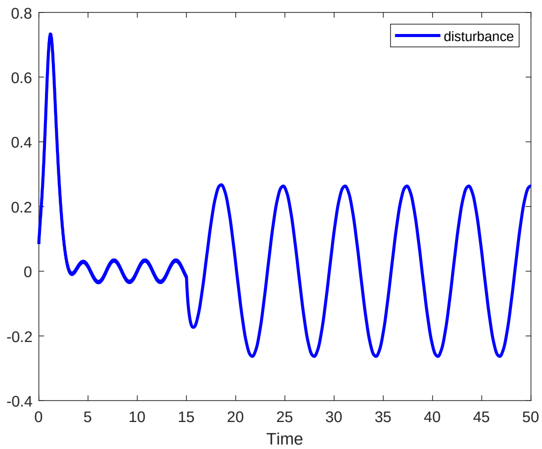

Fig. 8: Trajectory of the disturbance $\omega(t)$

VI. Conclusion

This paper proposes a novel control approach for global optimal output regulation of a class of partially linear systems with an exosystem and nonlinear dynamic uncertainties. By using reinforcement learning, a data-driven control strategy is proposed for designing robust adaptive optimal controllers and an optimal disturbance policy to achieve the rejection of nonvanishing disturbance and forcing the output to asymptotically track a desired output. The obtained simulation results ascertain the effectiveness of the proposed approach.

References

[1] B. Francis, "The linear multivariable regulator problem," *SIAM Journal on Control and Optimization*, vol. 15, no. 3, pp. 486–505, 1977.

[2] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 131–140, Feb 1990.

[3] J. Huang and W. J. Rugh, "On a nonlinear multivariable servomechanism problem," *Automatica*, vol. 26, no. 6, pp. 963 – 972, 1990.

[4] C. I. Byrnes, F. D. Priscoli, A. Isidori, and W. Kang, "Structurally stable output regulation of nonlinear systems," *Automatica*, vol. 33, no. 3, pp. 369 – 385, 1997.

[5] A. Serrani, A. Isidori, and L. Marconi, "Semiglobal nonlinear output regulation with adaptive internal model," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1178–1194, 2001.

[6] J. Huang, *Nonlinear Output Regulation: Theory and Applications*. Philadelphia, PA: SIAM, 2004.

[7] F. A. Yaghmaie, K. H. Movric, F. L. Lewis, R. Su, and M. Sebek, "H_∞ output regulation of linear heterogeneous multiagent systems over switching graphs," *Interational Journal of Robust and Nonlinear Control*, vol. 28, pp. 3852–3870, 2018.

[8] T. Başar and P. Bernhard, *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer Science & Business Media, 2008.

[9] T. Liu, D. J. Hill, and Z. P. Jiang, "Lyapunov formulation of ISS cyclic-small-gain in continuous-time dynamical networks," *Automatica*, vol. 47, no. 9, pp. 2088 – 2093, 2011.

[10] T. Liu and Z. Jiang, "A small-gain approach to robust event-triggered control of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 8, pp. 2072–2085, 2015.

[11] J. Huang and Z. Chen, "A general framework for tackling the output regulation problem," *IEEE Transactions on Automatic Control*, vol. 49, no. 12, pp. 2203–2218, 2004.

[12] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for feedback control," *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, 2009.

[13] 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.

[14] W. Gao, Y. Jiang, Z. P. Jiang, and T. Chai, "Adaptive and optimal output feedback control of linear systems: An adaptive dynamic programming approach," in *Proceedings of the 11th World Congress on Intelligent Control and Automation*, Shenyang, China, 2014, pp. 2085–2090.

[15] W. Gao, Z. P. Jiang, and K. Ozbay, "Adaptive optimal control of connected vehicles," in *Proceedings of the 10th International Workshop on Robot Motion and Control*, Poznan, Poland, 2015, pp. 288–293.

[16] W. Gao and Z. P. Jiang, "Global optimal output regulation of partially linear systems via robust adaptive dynamic programming," in *Proc. 1st Conference on Modelling. Identification and Control of Nonlinear Systems*, vol. 48, no. 11, Saint-Petersburg, Russia, 2015, pp. 742–747.

[17] B. Sun, M. He, Y. Wang, W. Gui, C. Yang, and Q. Zhu, "A data-driven optimal control approach for solution purification process," *Journal of Process Control*, vol. 68, pp. 171 – 185, 2018.

[18] Y. Jiang and Z. P. Jiang, "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 882–893, 2014.

[19] W. Gao and Z. P. Jiang, "Linear optimal tracking control: An adaptive dynamic programming approach," in *Proceedings of the American Control Conference*, Chicago, IL, 2015, pp. 4929–4934.

[20] A. J. Krener, "The construction of optimal linear and nonlinear regulators," in *Systems, Models and Feedback: Theory and Applications*, A. Isidori and T. J. Tarn, Eds. Birkhauser Boston, 1992, vol. 12, pp. 301–322.

[21] W. Gao and Z. P. Jiang, "Adaptive dynamic programming and adaptive optimal output regulation of linear systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4164–4169, 2016.

[22] H. Modares, F. L. Lewis, and Z. Jiang, "H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2550–2562, 2015.

[23] A. Saberi, P. Kokotovic, and S. Summers, "Global stabilization of partially linear composite systems," *SIAM Journal of Control and Optimization*, vol. 2, no. 6, pp. 1491–1503, 1990.

[24] Z. P. Jiang, A. R. Teel, and L. Praly, "Small-gain theorem for ISS systems and applications," *Mathematics of Control, Signals and Systems*, vol. 7, no. 2, pp. 95–120, 1994.

[25] E. D. Sontag, "Input to state stability: Basic concepts and results," in *Nonlinear and Optimal Control Theory*, P. Nistri and G. Stefani, Eds. Berlin: Springer-Verlag, 2007, pp. 163–220.

[26] ——, "Smooth stabilization implies coprime factorization," *IEEE Transactions on Automatic Control*, vol. 34, no. 4, pp. 435–443, 1989.