Reinforcement-Learning-Based Dual-Control Methodology for Complex Nonlinear Discrete-Time Systems With Application to Spark Engine EGR Operation

Peter Shih, Brian C. Kaul, S. Jagannathan, *Senior Member, IEEE*, and James A. Drallmeier

***Abstract*—A novel reinforcement-learning-based dual-control methodology adaptive neural network (NN) controller is developed to deliver a desired tracking performance for a class of complex feedback nonlinear discrete-time systems, which consists of a second-order nonlinear discrete-time system in nonstrict feedback form and an affine nonlinear discrete-time system, in the presence of bounded and unknown disturbances. For example, the exhaust gas recirculation (EGR) operation of a spark ignition (SI) engine is modeled by using such a complex nonlinear discrete-time system. A dual-controller approach is undertaken where primary adaptive critic NN controller is designed for the nonstrict feedback nonlinear discrete-time system whereas the secondary one for the affine nonlinear discrete-time system but the controllers together offer the desired performance. The primary adaptive critic NN controller includes an NN observer for estimating the states and output, an NN critic, and two action NNs for generating virtual control and actual control inputs for the nonstrict feedback nonlinear discrete-time system, whereas an additional critic NN and an action NN are included for the affine nonlinear discrete-time system by assuming the state availability. All NN weights adapt online towards minimization of a certain performance index, utilizing gradient–descent-based rule. Using Lyapunov theory, the uniformly ultimate boundedness (UUB) of the closed-loop tracking error, weight estimates, and observer estimates are shown. The adaptive critic NN controller performance is evaluated on an SI engine operating with high EGR levels where the controller objective is to reduce cyclic dispersion in heat release while minimizing fuel intake. Simulation and experimental results indicate that engine out emissions drop significantly at 20% EGR due to reduction in dispersion in heat release thus verifying the dual-control approach.**

***Index Terms*—Adaptive critic design, near-optimal control, nonstrict feedback nonlinear discrete-time system, output feedback control, separation principle.**

Manuscript received February 21, 2007; revised September 3, 2007 and December 29, 2007; accepted February 6, 2008. First published July 15, 2008; last published August 6, 2008 (projected). This work was supported by the National Science Foundation under Grants ECCS#0327877, and ECCS#0621924 and by the Intelligent Systems Center.

P. Shih and S. Jagannathan are with the Department of Electrical and Computer Engineering, University of Science and Technology (formerly University of Missouri-Rolla), Rolla, MO 65409 USA (e-mail: sarangap@mst.edu).

B. C. Kaul and J. A. Drallmeier are with the Department of Mechanical and Aerospace Engineering, University of Science and Technology (formerly University of Missouri-Rolla), Rolla, MO 65409 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2008.2000452

I. Introduction

ADAPTIVE control using neural network (NN) in general is now well understood for affine nonlinear discrete-time systems. Additionally, backstepping control of nonlinear discrete-time systems in strict feedback form in particular has been addressed in the literature [1]–[3]. The strict feedback nonlinear discrete-time system is expressed as [1]

$$x_i(k+1) = f_i\left(\overline{x}_i(k)\right) + g_i\left(\overline{x}_i(k)\right) x_{i+1}(k) \tag{1}$$

$$x_n(k+1) = f_n\left(\overline{x}_n(k)\right) + g_n\left(\overline{x}_n(k)\right) u(k) \tag{2}$$

where $x_i(k) \in \Re$ is the state, $u(k) \in \Re$ is the control input, $\overline{x}_i(k) = [x_1(k), \cdots, x_i(k)]^T \in \Re^i$, and $i = 1, \ldots, (n-1)$. For strict feedback nonlinear systems [1], the nonlinearities $f_i(\overline{x}_i(k))$ and $g_i(\overline{x}_i(k))$ depend only upon $x_1(k), \ldots, x_i(k)$, i.e., $\overline{x}_i(k)$. However, for a nonstrict feedback nonlinear system, where $f_i(\overline{x}_i(k))$ and $g_i(\overline{x}_i(k))$ depend on both $\overline{x}_i(k)$ and $x_{i+1}(k)$, there were no control design schemes available until recently [11]. Since available [1]–[3] methods when applied to a second-order nonlinear nonstrict feedback discrete-time systems will result in a noncausal controller (current control input depends on the future system states), control design for nonstrict feedback systems should be handled carefully [11]. On the other hand, though stable NN controllers [1]–[3], [10]–[13] are designed for nonlinear discrete-time systems, these control designs fail to render optimality since tracking error was utilized as the performance measure.

By contrast, the reinforcement-learning-based adaptive critic NN approach [4] has emerged as a promising tool to develop optimal (or suboptimal) NN controllers due to its potential to find approximate solutions to dynamic programming, where a *strategic* utility function, which is considered as the long-term system performance measure, can be optimized. Optimal NN controllers using reinforcement learning are considered as the third generation NN controllers and they are currently being pursued by many researchers [4]–[9], [15], [19], [21]. In supervised learning, an explicit signal is provided by the teacher to guide the learning process whereas in the case of reinforcement learning, the role of the teacher is more evaluative than instructional in nature. The critic NN monitors the system states and approximates the *strategic* utility function, and provides a better training signal to the action NN, which generates the near-optimal control action to the nonlinear system.

1045-9227/$25.00 © 2008 IEEE

There are many variants of adaptive critic NN controller architectures [4]–[9] using state feedback even though few results [6]–[9], [19] address the controller convergence. However, adaptive critic NN controller results were not available for the nonlinear discrete-time systems in nonstrict feedback form until recently [22]. By contrast, in [18], a novel NN controller using tracking error as the performance measure is introduced for nonlinear discrete-time systems in nonstrict feedback form. To the best of our knowledge, no known results are available for near-optimal control of complex nonlinear discrete-time systems using reinforcement learning such as the ones introduced in this paper.

In this paper, a novel adaptive critic NN-based feedback controller is developed to control a class of complex nonlinear discrete-time systems, which consists of a coupled nonstrict feedback plus an affine system, with bounded and unknown disturbances. A dual-control approach is undertaken where a primary adaptive critic NN controller is designed for nonlinear discrete-time systems in nonstrict feedback form and a secondary controller for the affine nonlinear discrete-time systems so that both controllers work together guaranteeing performance and stability.

For the case of a nonlinear discrete-time system in nonstrict feedback form, an adaptive NN backstepping is utilized for the controller design with two action NNs being used to generate the virtual and actual control inputs, respectively. The weights of the two action NNs are tuned by the critic NN signal to minimize the *strategic* utility function and their outputs. The critic NN approximates certain *strategic* utility function a variant of standard Bellman equation. The NN observer estimates the system states and output and the estimates are subsequently used in the controller. The proposed controller is *model-free* since the dynamics of the nonlinear discrete-time systems are not known beforehand. All the NN weights are tuned online. On the other hand, for the affine nonlinear discrete-time system, a separate critic and action NNs are utilized. The critic NN approximates the standard Bellmann equation and tunes the action NN so that the action NN generates a near-optimal signal to control the affine nonlinear discrete-time system. Both controllers work together and render guaranteed performance and closed-loop stability.

The main contributions of this paper can be summarized as follows.

1) The adaptive critic NN approach is extended to a complex nonlinear discrete-time system where a dual-control approach is undertaken while guaranteeing stability and performance.
2) Optimization of a long-term performance index is undertaken in contrast with traditional adaptive NN schemes [1], [2] where no optimization is performed.
3) Demonstration of the boundedness of the overall system is shown even in the presence of NN approximation errors and bounded unknown disturbances unlike in the existing adaptive critic design (ACD) works [7]–[9] where the convergence is presented under ideal circumstances. Stability proof is inferred even with an NN observer by relaxing the separation principle via novel weight-updating rules and by selecting the Lyapunov function consisting of the system estimation errors, tracking, and the NN weight estimation errors. A single critic NN is utilized to tune two action NNs.
4) A well-defined controller is presented by overcoming the problem of certain nonlinear function estimates becoming zero since a single NN is used to approximate both nonlinear functions $f_i(\overline{x}_i(k))$ and $g_i(\overline{x}_i(k))$ compared to [10].
5) The NN weights are tuned online instead of offline [5].
6) The assumption that $g_1(x_1(k), x_2(k))$ is bounded away from zero and its sign is known *a priori* is relaxed in contrast with [2].

The proposed primary controller is applied to control the spark ignition (SI) engine dynamics operating with high exhaust gas recirculation (EGR) levels, a practical complex nonlinear discrete-time system. The primary controller allows the engine, which is a nonstrict feedback nonlinear discrete-time system to operate in high EGR mode with fuel intake as the control input, where an inert gas displaces the stoichiometric ratio of fuel to air. The inert gas system is modeled as an affine nonlinear discrete-time system, and therefore, a separate secondary controller is designed. Both controllers enable the engine to operate in higher EGR mode compared to the uncontrolled case by reducing heat release dispersion while minimizing fuel intake. Consequently, the engine exhibits improved emissions and fuel efficiency compared to the uncontrolled case. Other controller designs can run an SI engine in lean mode [11]; however, engine catalysts cannot function efficiently with the lean exhaust chemistry.

EGR, on the other hand, allows for the efficient operation of standard three-way catalysts. Not only does it reduce precatalyst emissions, but it can improve fuel efficiency by reducing throttling losses. Therefore, the applicability of high EGR usage in the automotive engines is greater. Dilution with EGR also has wide practical applicability in diesel engines and in SI engines without three-way catalysts.

II. Complex Nonlinear Discrete-Time Systems

Consider the complex nonlinear discrete-time system, given in the following form:

$$x_1(k+1) = f_1(x_1(k), x_2(k), x_3(k)) + g_1(x_1(k), x_2(k), x_3(k))\, x_2(k) + d_1(k) \quad (3)$$

$$x_2(k+1) = f_2(x_1(k), x_2(k), x_3(k)) + g_2(x_1(k), x_2(k), x_3(k))\, u(k) + d_2(k) \quad (4)$$

$$x_3(k+1) = f_4(x_1(k), x_2(k), x_3(k)) + g_4(x_1(k), x_2(k), x_3(k))\, v(k) + d_3(k) \quad (5)$$

$$y(k+1) = f_3(x_1(k), x_2(k), x_3(k)) \quad (6)$$

where $x_i(k) \in \Re, i = 1, 2, 3$, are states, $u(k) \in \Re$ and $v(k) \in \Re$ are system inputs, and $d_1(k) \in \Re$, $d_2(k) \in \Re$, and $d_3(k) \in \Re$ are unknown but bounded disturbances. Bounds on the disturbances are given by $|d_1(k)| < d_{1m}$, $|d_2(k)| < d_{2m}$, and $|d_3(k)| < d_{3m}$ where d_{1m}, d_{2m}, and d_{2m} are unknown positive scalars. It is important to note that the output is a nonlinear function of states in contrast with available literature [12], [13] where the output is a linear function of the states. Finally, the output is considered measurable whereas the first two states

$x_1(k)$ and $x_2(k)$ are considered not available while $x_3(k)$ is assumed to be available for convenience. For the systems (3) and (4), not only should the systems' actual output converge to their target value, but the states should also converge to their respective desired values for the proposed application of engine control.

The controller development is presented separately for the two systems as the objectives are different even though the two controllers are designed for the same complex system. The first part uses (3), (4), and (6) to develop the primary controller. The second part uses (5) to develop the secondary controller. Stability for both systems is demonstrated together. The SI engine control with high levels is represented by (3)–(6), and therefore, the class of systems is of interest. One cannot directly apply the controller design from nonstrict feedback nonlinear discrete-time systems [11], [18] since the first system contains the third state unlike a typical nonstrict feedback system. Therefore, a dual approach is the best choice.

III. Primary Controller

To overcome the immeasurable states $x_1(k)$ and $x_2(k)$, an observer is used. It utilizes the current heat release output, $y(k)$, to estimate the future output $\hat{y}(k+1)$ and states $\hat{x}_1(k+1)$ and $\hat{x}_2(k+1)$. The design of the observer, which follows steps similar to [22] is discussed next.

A. Observer Design

For the observer design, the nominal values of the uncertainties are required since the nonlinearities as well as the input–output relationship is considered unknown. The nominal values of the unknown uncertainties can be obtained by a variety of ways. One of the ways is to apply Taylor series expansion without ignoring the higher order terms. Consider (3) and (4). We expand the individual nonlinear functions using Taylor series expansion into linear and higher order terms

$$f_1(\cdot) = f_{10} + \Delta f_1(\cdot) \tag{7}$$
$$f_2(\cdot) = f_{20} + \Delta f_2(\cdot) \tag{8}$$
$$g_1(\cdot) = g_{10} + \Delta g_1(\cdot) \tag{9}$$
$$g_2(\cdot) = g_{20} + \Delta g_2(\cdot) \tag{10}$$

where the first term in (7)–(10) are known nominal values and the second term are unknown higher order terms. The entire expansion for the terms in (7)–(10) is not necessary since it is not required for the controller design. Moreover, the higher order terms are not ignored. A two-layer feedforward NN with semirecurrent architecture and novel weight tuning are utilized to construct the output as

$$y(k+1) = w_1^T \phi\left(v_1^T z_1(k)\right) + \varepsilon\left(z_1(k)\right) \tag{11}$$

where $z_1(k) = [x_1(k), x_2(k), x_3(k), y(k), u(k)]^T \in R^5$ is the network input, $y(k+1)$ and $y(k)$ are the future and current outputs, $w_1 \in \Re^{n_1}$ and $v_1 \in \Re^{4\times n_1}$ denote the ideal output and constant hidden layer weight matrices, respectively, $u(k)$ is the control input, $\phi(v_1^T z_1(k))$ represents the hidden layer activation function, n_1 is the number of nodes in the hidden layer, and $\varepsilon(z_1(k)) \in \Re$ is the approximation error. For simplicity, the two equations can be represented as

$$\phi_1(k) = \phi\left(v_1^T z_1(k)\right) \tag{12}$$

and

$$\varepsilon_1(k) = \varepsilon\left(z_1(k)\right). \tag{13}$$

Rewrite (11) using (12) and (13) to obtain

$$y(k+1) = w_1^T \phi_1(k) + \varepsilon_1(k). \tag{14}$$

The states $x_1(k)$ and $x_2(k)$ are not measurable; therefore, $z_1(k)$ is not available either. Using the estimated and measured states and the output, $\hat{x}_1(k), \hat{x}_2(k), x_3(k)$, and $\hat{y}(k)$, respectively, instead of $x_1(k), x_2(k)$, and $y(k)$, the proposed observer is given as

$$\begin{aligned}\hat{y}(k+1) &= \hat{w}_1^T(k)\phi\left(v_1^T \hat{z}_1(k)\right) + l_1\tilde{y}(k)\\ &= \hat{w}_1^T(k)\hat{\phi}_1(k) + l_1\tilde{y}(k)\end{aligned} \tag{15}$$

where $\hat{z}_1(k) = [\hat{x}_1(k), \hat{x}_2(k), x_3(k), \hat{y}(k), u(k)]^T \in R^5$ is the input vector using estimated states, $\hat{y}(k+1)$ and $\hat{y}(k)$ are the estimated future and the current output, $\hat{w}_1(k)$ is the actual weight matrix, $u(k)$ is the estimated control input, $\hat{\phi}_1(k)$ is the hidden layer activation function, $l_1 \in R$ is the observer gain, and $\tilde{y}(k)$ is the heat release estimation error defined as

$$\tilde{y}(k) = \hat{y}(k) - y(k). \tag{16}$$

It is demonstrated in [14] that, if the hidden layer weights v_1 are chosen initially at random and kept constant, and the number of hidden layer nodes is sufficiently large, then the approximation error $\varepsilon(z_1(k))$ can be made arbitrarily small so that the bound $\|\varepsilon(z_1(k))\| \le \varepsilon_{1m}$ holds for all $z_1(k) \in S$ since the activation function forms a basis to the nonlinear function that the NN approximates. Now we choose, at our convenience, the observer structure as a function of output estimation errors and known quantities as

$$\hat{x}_1(k+1) = f_{10} - \hat{x}_2(k) + l_2\tilde{y}(k) \tag{17}$$
$$\hat{x}_2(k+1) = f_{20} + g_{20}u(k) + l_3\tilde{y}(k) \tag{18}$$

where $l_2 \in R$ and $l_3 \in R$ are design constants.

Define the state estimation and output errors as

$$\tilde{x}_i(k+1) = \hat{x}_i(k+1) - x_i(k+1), \qquad i \in \{1,2\} \tag{19}$$
$$\tilde{y}(k+1) = \hat{y}(k+1) - y(k+1). \tag{20}$$

Combining (17)–(20), to obtain the estimation and output error dynamics as

$$\begin{aligned}\tilde{x}_1(k+1) = {} & f_{10} - \hat{x}_2(k) + l_2\tilde{y}(k) - f_1(\cdot)\\ & - g_1(\cdot)x_2(k) - d_1(k)\end{aligned} \tag{21}$$
$$\begin{aligned}\tilde{x}_2(k+1) = {} & f_{20} + g_{20}u(k) + l_3\tilde{y}(k) - f_2(\cdot)\\ & - g_2(\cdot)u(k) - d_2(k)\end{aligned} \tag{22}$$

and

$$\tilde{y}(k+1) = \hat{w}_1^T(k)\hat{\phi}_1(k) + l_1\tilde{y}(k) - w_1^T\phi_1(k) - \varepsilon_1(k) \quad (23)$$

choose the weight tuning of the observer as

$$\hat{w}_1(k+1) = \hat{w}_1(k) - \alpha_1\hat{\phi}_1(k)\left(\hat{w}_1^T(k)\hat{\phi}_1(k) + l_4\tilde{y}(k)\right) \quad (24)$$

where $\alpha_1 \in R$ and $l_4 \in R$ are design constants. It will be shown in Section III-D that by using the above weight tuning, the separation principle is relaxed and the closed-loop signals will be bounded. Next, we present the theorem, where it is demonstrated that the state estimation and output estimation errors along with observer NN weight estimation errors are bounded. The following mild assumptions are required.

Assumption 1: The unknown smooth functions $f_2(\cdot)$ and $g_2(\cdot)$ are upper bounded within the compact set S as $f_{2\max} > |f_2(k)|$ and $g_{2\max} > |g_2(k)|$.

Remark: This assumption is a direct consequence of functions over the compact sets. This assumption is required for the NN universal approximation result to hold.

Theorem 1 (Observer Stability): Consider the system given by (3), (4), and (6), and the disturbance bounded by $|d_1(k)| < d_{1m}$ and $|d_2(k)| < d_{2m}$ where d_{1m} and d_{2m} are known positive scalars. Let the observer NN weight tuning be given by (24). The state estimation errors $\tilde{x}_1(k)$ and $\tilde{x}_2(k)$, output estimation errors $\tilde{y}(k)$, and NN weight estimate $\hat{w}_1(k)$ of the observer are uniformly ultimate boundedness (UUB), with the bounds specifically given by (B.11), with the controller design parameters selected as

$$0 < \alpha_1 \|\phi_1(k)\|^2 < 1 \quad (25)$$

$$|l_1| < \frac{1}{2} \quad (26)$$

$$|l_2| < \frac{\sqrt{3}}{3} \quad (27)$$

$$|l_3| < \frac{\sqrt{3}}{3} \quad (28)$$

$$|l_4| < \frac{\sqrt{3}}{3} \quad (29)$$

where α_1 is the NN adaptation gain and l_1, l_2, l_3, and l_4 are the observer design parameters.

Proof: Follow steps similar to [22]. See Appendix B for the proof.

Remark 1: In Theorem 1, state and output estimation errors and the NN weights of the observer are shown to be bounded. One can then design a controller by applying separation principle if the system under consideration is linear. Unfortunately, separation principle does not hold for nonlinear systems. Therefore, in Section III-D, the boundedness of the closed-loop system is demonstrated where the observer and controller signals are proven to be bounded without using separation principle. Next we discuss the design of the adaptive critic NN controller for the primary system and demonstrate that if the closed-loop system is bounded then the control inputs will be bounded.

B. Reinforcement Learning and Optimization

The purpose of the critic NN is to approximate the long-term performance index (or strategic utility function) of the nonlinear system through online weight adaptation. The critic signal estimates the future performance and tunes the two action NNs. The tuning will ultimately minimize the strategic utility function itself and the action NN outputs or control inputs to the system so that closed-loop stability is inferred.

The utility function $p(k) \in \Re$ is given by

$$p(k) = \begin{cases} 0, & \text{if } (|\tilde{y}(k)|) \le c \\ 1, & \text{otherwise} \end{cases} \quad (30)$$

where $c \in \Re$ is a user-defined threshold. The utility function $p(k)$ represents the current performance index. In other words, $p(k) = 0$ and $p(k) = 1$ refer to good and unsatisfactory tracking performance at the kth time step, respectively. The long-term *strategic* utility function $Q(k) \in \Re$ is defined as

$$\overline{Q}(k) = \beta^N p(k+1) + \beta^{N-1}p(k+2) + \cdots + \beta^{k+1}p(N) + \ldots \quad (31)$$

where $\beta \in \Re$, $0 < \beta < 1$, is the discount factor and N is the horizon index. The term $\overline{Q}(k)$ is viewed here as the long system performance measure for the controller since it is the sum of all future system performance indices. Equation (31) can also be expressed as $\overline{Q}(k) = \min_{u(k)}\{\alpha\overline{Q}(k-1) - \alpha^{N+1}p(k)\}$ after simple manipulation, which is similar to the standard Bellman equation.

C. Critic NN Design

We utilize the universal approximation property of NN to define the critic NN output and rewrite $\hat{\overline{Q}}(k)$ as

$$\hat{\overline{Q}}(k) = \hat{w}_2^T(k)\phi\left(v_2^T\hat{z}_2(k)\right) = \hat{w}_2^T(k)\hat{\phi}_2(k) \quad (32)$$

where $\hat{\overline{Q}}(k) \in \Re$ is the critic signal, $\hat{w}_2(k) \in \Re^{n_2}$ is the tunable weight matrix, $v_2 \in \Re^{3\times n_2}$ represents the constant input weight matrix selected initially at random, $\hat{\phi}_2(k) \in \Re^{n_2}$ is the activation function vector in the hidden layer, n_2 is the number of the nodes in the hidden layer, and $\hat{z}_2(k) = [\hat{x}_1(k), \hat{x}_2(k), x_3(k)]^T \in R^3$ is the input vector.

We define the prediction error as

$$e_c(k) = \hat{\overline{Q}}(k) - \beta\left(\hat{\overline{Q}}(k-1) - \beta^N p(k)\right) \quad (33)$$

where the subscript "c" stands for the "critic." We use a quadratic objective function to minimize

$$E_c(k) = \frac{1}{2}e_c^2(k). \quad (34)$$

The weight-update rule for the critic NN is based upon gradient adaptation, which is given by the general formula

$$\hat{w}_2(k+1) = \hat{w}_2(k) + \Delta\hat{w}_2(k) \quad (35)$$

where

$$\Delta\hat{w}_2(k) = \alpha_2\left[-\frac{\partial E_c(k)}{\partial\hat{w}_2(k)}\right] \quad (36)$$

or

$$\hat{w}_2(k+1) = \hat{w}_2(k) - \alpha_2 \hat{\phi}_2(k) \times \left(\hat{\overline{Q}}(k) + \beta^{N+1} p(k) - \beta \hat{\overline{Q}}(k-1)\right)^T \quad (37)$$

where $\alpha_2 \in \Re$ is the NN adaptation gain.

D. Action NN Design

In this section, the design of the virtual control input is discussed. Before we proceed, the following mild assumption is needed. Then, the system of nonlinear equations is rewritten.

Assumption 2: The unknown smooth function $g_2(\cdot)$ is bounded away from zero for all $x_1(k)$ and $x_2(k)$ within the compact set S. In other words, $0 < g_{2\min} < |g_2(\cdot)| < g_{2\max}$, $\forall x_1(k)$, and $x_2(k) \in S$, where $g_{2\min} \in \Re^+$ and $g_{2\max} \in \Re^+$. Without loss of generality, we will assume that $g_2(\cdot)$ is positive in this paper.

First, we simplify by rewriting the state equations with the following:

$$\Phi(\cdot) = f_1(x_1(k), x_2(k), x_3(k)) + g_1(x_1(k), x_2(k), x_3(k))\, x_2(k) + x_2(k). \quad (38)$$

Systems (3) and (4) can be rewritten as

$$x_1(k+1) = \Phi(\cdot) - x_2(k) + d_1(k) \quad (39)$$

$$x_2(k+1) = f_2(\cdot) + g_2(\cdot)u(k) + d_2(k). \quad (40)$$

1) Virtual Control Input Design: Our goal is to stabilize the system output $y(k)$ around a specified target point y_d by controlling the input. The secondary objective is to make $x_1(k)$ approach the desired trajectory $x_{1d}(k)$. At the same time, all signals in systems (3) and (4) must be UUB, all weights must be bounded, and a performance index must be minimized. Define the tracking error as

$$e_1(k) = x_1(k) - x_{1d}(k) \quad (41)$$

where $x_{1d}(k)$ is the desired trajectory. Using (39), (41) can be expressed as the following:

$$e_1(k+1) = x_1(k+1) - x_{1d}(k+1) = (\Phi(\cdot) - x_2(k) + d_1(k)) - x_{1d}(k+1). \quad (42)$$

By viewing $x_2(k)$ as a virtual control input, a desired virtual control signal can be designed as

$$x_{2d}(k) = \Phi(\cdot) - x_{1d}(k+1) + l_5 \hat{e}_1(k) \quad (43)$$

where l_5 is a gain constant. Since $\Phi(\cdot)$ is an unknown function, $x_{2d}(k)$ in (43) cannot be implemented in practice. We invoke the universal approximation property of NN to estimate this unknown function

$$\Phi(\cdot) = w_3^T \phi\left(v_3^T z_3(k)\right) + \varepsilon\left(z_3(k)\right) \quad (44)$$

where $z_3(k) = [x_1(k), x_2(k), x_3(k)]^T \in \Re^3$ is the input vector, $w_3^T \in \Re^{n_2}$ and $v_3^T \in \Re^{3\times n_3}$ are the ideal and constant input weight matrices, $\phi(v_3^T z_3(k)) \in \Re^{n_3}$ is the activation function vector in the hidden layer, n_3 is the number of the nodes in the hidden layer, and $\varepsilon(z_3(k))$ is the functional estimation error. It is demonstrated in [14] that, if the hidden layer weights v_1 are chosen initially at random and kept constant, and the number of hidden layer nodes is sufficiently large, then the approximation error $\varepsilon(z_3(k))$ can be made arbitrarily small so that the bound $\|\varepsilon(z_3(k))\| \le \varepsilon_{3m}$ holds for all $z_3(k) \in S$ in a compact set, since the activation function vector forms a basis to the nonlinear function that the NN approximates.

Rewriting (43) using (44), the virtual control signal can be rewritten as

$$x_{2d}(k) = w_3^T \phi\left(v_3^T z_3(k)\right) + \varepsilon\left(z_3(k)\right) - x_{1d}(k+1) + l_5 \hat{e}_1(k). \quad (45)$$

Replacing actual with estimated states, (45) becomes

$$\hat{x}_{2d}(k) = \hat{w}_3^T(k)\phi\left(v_3^T \hat{z}_3(k)\right) - x_{1d}(k+1) + l_5 \hat{e}_1(k) = \hat{w}_3^T(k)\hat{\phi}_3(k) - x_{1d}(k+1) + l_5 \hat{e}_1(k) \quad (46)$$

where $\hat{z}_3(k) = [\hat{x}_1(k), \hat{x}_2(k), x_3(k)]^T \in \Re^3$ is the input vector using estimated states, and $\hat{e}_1(k) = \hat{x}_1(k) - x_{1d}(k)$.

Define

$$e_2(k) = x_2(k) - \hat{x}_{2d}(k). \quad (47)$$

Equation (42) can be rewritten using (47) as

$$\begin{aligned} e_1(k+1) &= (\Phi(\cdot) - x_2(k) + d_1(k)) - x_{1d}(k+1) \\ &= \Phi(\cdot) - (e_2(k) + \hat{x}_{2d}(k)) + d_1(k) - x_{1d}(k+1) \\ &= \Phi(\cdot) - \hat{x}_{2d}(k) - e_2(k) - x_{1d}(k+1) + d_1(k). \end{aligned} \quad (48)$$

Combine (46) into (48), then (44) to get

$$\begin{aligned} e_1(k+1) &= \Phi(\cdot) - \left(\hat{w}_3^T(k)\hat{\phi}_3(k) - x_{1d}(k+1) + l_5 \hat{e}_1(k)\right) - e_2(k) - x_{1d}(k+1) + d_1(k) \\ &= \left(w_3^T \phi_3(k) + \varepsilon_3(k)\right) - \hat{w}_3^T(k)\hat{\phi}_3(k) - l_5 \hat{e}_1(k) - e_2(k) + d_1(k) \\ &= w_3^T \left(\hat{\phi}_3(k) - \tilde{\phi}_3(k)\right) - \hat{w}_3^T(k)\hat{\phi}_3(k) + \varepsilon_3(k) - l_5 \hat{e}_1(k) - e_2(k) + d_1(k) \\ &= w_3^T \left(\hat{\phi}_3(k) - \tilde{\phi}_3(k)\right) - \hat{w}_3^T(k)\hat{\phi}_3(k) + \varepsilon_3(k) - l_5 \hat{e}_1(k) - e_2(k) + d_1(k) \\ &= -\tilde{w}_3^T \hat{\phi}_3(k) - w_3^T \tilde{\phi}_3(k) + \varepsilon_3(k) - l_5 \hat{e}_1(k) - e_2(k) + d_1(k) \\ &= -\zeta_3(k) - w_3^T \tilde{\phi}_3(k) + \varepsilon_3(k) - l_5 \hat{e}_1(k) - e_2(k) + d_1(k) \end{aligned} \quad (49)$$

where

$$\zeta_3(k) = \tilde{w}_3^T(k)\hat{\phi}_3(k) = \hat{w}_3^T(k)\hat{\phi}_3(k) - w_3^T \hat{\phi}_3(k) \quad (50)$$

and

$$\tilde{\phi}_3(k) = \phi\left(v_3 \hat{z}_3(k)\right) - \phi\left(v_3 z_3(k)\right). \quad (51)$$

Let us define

$$e_{a1}(k) = \hat{w}_3^T(k)\hat{\phi}_3(k) + \left(\hat{\overline{Q}}(k) - Q_d(k)\right) \quad (52)$$

where $\hat{\overline{Q}}(k)$ is defined in (32), and the $a1$ subscript represents the error for the first action NN, $e_{a1}(k) \in \Re$. The desired

strategic utility function $Q_d(k)$ is "0" to indicate perfect tracking at all steps. Thus, (52) becomes

$$e_{a1}(k) = \hat{w}_3^T(k)\hat{\phi}_3(k) + \hat{\overline{Q}}(k). \tag{53}$$

The objective function to be minimized by the first action NN is given by

$$E_{a1}(k) = \frac{1}{2}e_{a1}^2(k). \tag{54}$$

The weight-update rule for the action NN is also a gradient-based adaptation, which is defined as

$$\hat{w}_3(k+1) = \hat{w}_3(k) + \Delta\hat{w}_3(k) \tag{55}$$

where

$$\Delta\hat{w}_3(k) = \alpha_3\left[-\frac{\partial E_{a1}(k)}{\partial \hat{w}_3(k)}\right] \tag{56}$$

$$\hat{w}_3(k+1) = \hat{w}_3(k) - \alpha_3\hat{\phi}_3(k)\left(\hat{\overline{Q}}(k) + \hat{w}_3^T(k)\hat{\phi}_3(k)\right) \tag{57}$$

with $\alpha_3 \in \Re$ is the NN adaptation gain.

2) Actual Control Design: Choose the following desired control input:

$$u_d(k) = \frac{1}{g_2(k)}\left(-f_2(k) + \hat{x}_{2d}(k+1) + l_6 e_2(k)\right). \tag{58}$$

Note that $u_d(k)$ is noncausal since it depends upon future value of $\hat{x}_{2d}(k+1)$. We solve this problem by using a semirecurrent NN since it can be a one-step predictor. The term $\hat{x}_{2d}(k+1)$ depends on state $x(k)$, virtual control input $\hat{x}_{2d}(k)$, desired trajectory $x_{1d}(k+2)$, and system errors $e_1(k)$ and $e_2(k)$. By taking the independent variables as the input to an NN, $\hat{x}_{2d}(k+1)$ can be approximated during control input selection. Consequently, in this paper, a feedforward NN with properly chosen weight tuning law rendering a semirecurrent or dynamic NN can be used to predict the future value. Alternatively, the value can be obtained by employing a filter [15]. The first layer of the second NN using the system errors, state estimates, and past value $\hat{x}_{2d}(k)$ as inputs generates $\hat{x}_{2d}(k+1)$, which in turn is used by the second layer to generate a suitable control input. The results in the simulation section show that the overall controller performance is satisfactory. On the other hand, one can use a single-layer dynamic NN to generate the future value of $\hat{x}_{2d}(k)$, which can be utilized as an input to a third control NN to generate a suitable control input. Here, these two single-layer NNs are combined into a single-multilayer NN.

Define input as $z_4(k) = [x_1(k), x_2(k), x_3(k), e_1(k), e_2(k), \hat{x}_{2d}(k), x_{1d}(k+2)]^T \in \Re^7$, then $u_d(k)$ can be approximated as

$$u_d(k) = w_4^T\phi\left(v_4^T z_4(k)\right) + \varepsilon\left(z_4(k)\right) = w_4^T\phi_4(k) + \varepsilon_4(k) \tag{59}$$

where $w_4 \in \Re^{n_4}$ and $v_4 \in \Re^{7\times n_4}$ denote the constant ideal output and hidden layer weight matrices, $\phi_4(k) \in \Re^{n_4}$ is the activation function vector, n_4 is the number of hidden-layer nodes, and $\varepsilon(z_4(k))$ is the estimation error. Again, we hold the input weights constant and adapt the output weights only. We also replace actual with estimated states to design the control input as

$$\hat{u}(k) = \hat{w}_4^T(k)\phi\left(v_4^T\hat{z}_4(k)\right) = \hat{w}_4^T(k)\hat{\phi}_4(k) \tag{60}$$

where $\hat{z}_4(k) = [\hat{x}_1(k), \hat{x}_2(k), x_3(k), \hat{e}_1(k), \hat{e}_2(k), \hat{x}_{2d}(k), x_{1d}(k+2)]^T \in \Re^7$, is the input vector. Rewriting (47) and substituting (58)–(60), we get

$$\begin{aligned}
e_2(k+1) &= x_2(k+1) - \hat{x}_{2d}(k+1)\\
&= \left(f_2(\cdot) + g_2(\cdot)\hat{w}_4^T(k)\hat{\phi}_4(k) + d_2(k)\right) - \hat{x}_{2d}(k+1)\\
&= f_2(\cdot) + g_2(\cdot)\left(\tilde{w}_4^T(k)\hat{\phi}_4(k) + w_4^T\phi_4(k) + w_4^T\tilde{\phi}_4(k)\right)\\
&\quad + d_2(k) - \hat{x}_{2d}(k+1)\\
&= f_2(\cdot) + g_2(\cdot)\left(w_4^T(k)\phi_4(k)\right)\\
&\quad + g_2(\cdot)\left(\zeta_4(k) + w_4^T\tilde{\phi}_4(k)\right) + d_2(k) - \hat{x}_{2d}(k+1)\\
&= f_2(\cdot) + g_2(\cdot)\left(u_d(k) - \varepsilon_4(k)\right)\\
&\quad + g_2(\cdot)\left(\zeta_4(k) + w_4^T\tilde{\phi}_4(k)\right) + d_2(k) - \hat{x}_{2d}(k+1)\\
&= l_6 e_2(k) - g_2(\cdot)\varepsilon_4(k) + g_2(\cdot)\zeta_4(k)\\
&\quad + g_2(\cdot)w_4^T\tilde{\phi}_4(k) + d_2(k)
\end{aligned} \tag{61}$$

where

$$\zeta_4(k) = \tilde{w}_4^T(k)\hat{\phi}_4(k) = \hat{w}_4^T(k)\hat{\phi}_4(k) - w_4^T\hat{\phi}_4(k) \tag{62}$$

and

$$\tilde{\phi}_4(k) = \hat{\phi}_4(k) - \phi_4(k). \tag{63}$$

Equations (49) and (61) represent the closed-loop error dynamics. Next we derive the weight-update law. Define

$$e_{a2}(k) = \hat{w}_4^T(k)\hat{\phi}_4(k) + \hat{\overline{Q}}(k) \tag{64}$$

where $e_{a2}(k) \in \Re$ is the error and the subscript $a2$ stands for the second action NN. Following the similar design, choose a quadratic objective function to minimize

$$E_{a2}(k) = \frac{1}{2}e_{a2}^2(k). \tag{65}$$

Define a gradient-based adaptation where the general form is given by

$$\hat{w}_4(k+1) = \hat{w}_4(k) + \Delta\hat{w}_4(k) \tag{66}$$

$$\Delta\hat{w}_4(k) = \alpha_4\left[-\frac{\partial E_{a2}(k)}{\partial \hat{w}_4(k)}\right] \tag{67}$$

or in other words

$$\hat{w}_4(k+1) = \hat{w}_4(k) - \alpha_4\hat{\phi}_4(k)\left(\hat{w}_4^T(k)\hat{\phi}_4(k) + \hat{\overline{Q}}(k)\right). \tag{68}$$

The proposed controller structure is shown in Fig. 1. Next, in Theorem 2, it is demonstrated that the closed-loop system is UUB. Before we proceed, the following assumptions are needed.

Assumption 3 (Bounded Ideal Weights): Let w_1, w_2, w_3, and w_4 be the unknown output layer target weights for the observer,

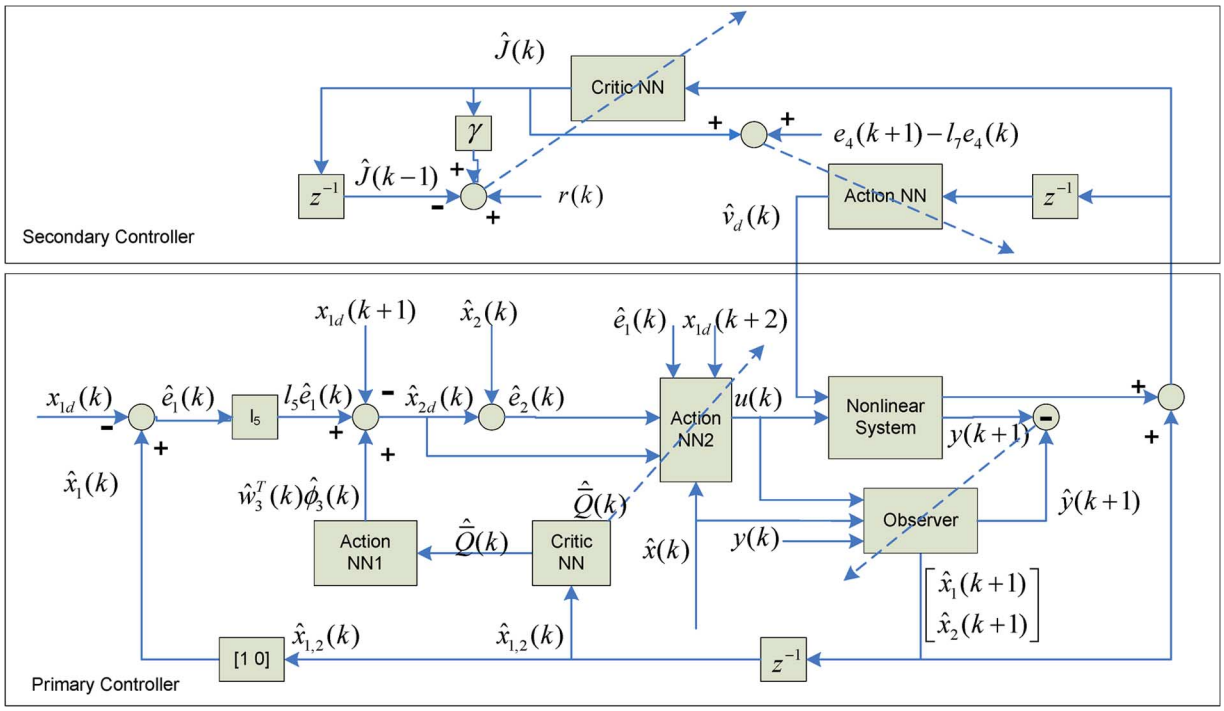

Fig. 1. Combined primary and secondary controller structure.

critic, and two action NNs, and assume that they are bounded above so that

$$\begin{aligned} \|w_1\| \le w_{1m} \quad \|w_2\| \le w_{2m} \\ \|w_3\| \le w_{3m} \quad \|w_4\| \le w_{4m} \end{aligned} \tag{69}$$

where $w_{om} \in R^+$, $w_{1m} \in R^+$, and $w_{2m} \in R^+$ represent the bounds on the unknown target weights, and where the Frobenius norm [15] is used.

Fact 1: The activation functions are bounded above by known positive values so that

$$\begin{aligned} \left\|\tilde{\phi}_1(\cdot)\right\| \le \tilde{\phi}_{1m} \quad \left\|\tilde{\phi}_2(\cdot)\right\| \le \tilde{\phi}_{2m} \\ \left\|\tilde{\phi}_3(\cdot)\right\| \le \tilde{\phi}_{3m} \quad \left\|\tilde{\phi}_4(\cdot)\right\| \le \tilde{\phi}_{4m} \end{aligned} \tag{70}$$

where $\hat{\phi}_{1m}, \tilde{\phi}_{1m} \in R^+$, $\hat{\phi}_{2m}, \tilde{\phi}_{2m} \in R^+$, $\hat{\phi}_{3m}, \tilde{\phi}_{3m} \in R^+$, and $\hat{\phi}_{4m}, \tilde{\phi}_{4m} \in R^+$ are the upper bounds.

Theorem 2 (Output Feedback Controller Stability): Consider the system given by (3) and (4) and the disturbance bounds d_{1m} and d_{2m} to be known constants. Let the observer, critic, virtual control, and control input NN weight tuning be given by (24), (37), (57), and (68), respectively. Let the virtual control input and control input be given by (46) and (60). The estimation errors and tracking errors $e_1(k)$ and $e_2(k)$ and weight estimates $\hat{w}_1(k)$, $\hat{w}_2(k)$, $\hat{w}_3(k)$, and $\hat{w}_4(k)$ are UUB, with the controller design parameters selected as

$$0 < \alpha_1 \|\phi_1(k)\|^2 < 1 \tag{71}$$

$$0 < \alpha_2 \|\phi_2(k)\|^2 < 1 \tag{72}$$

$$0 < \alpha_3 \|\phi_3(k)\|^2 < 1 \tag{73}$$

$$0 < \alpha_4 \|\phi_4(k)\|^2 < 1 \tag{74}$$

$$|l_1| < \frac{1}{2} \tag{75}$$

$$|l_2| < \frac{\sqrt{3}}{3} \tag{76}$$

$$|l_3| < \frac{\sqrt{3}}{3} \tag{77}$$

$$|l_4| < \frac{\sqrt{3}}{3} \tag{78}$$

$$|l_5| < \frac{1}{\sqrt{5}} \tag{79}$$

$$|l_6| < \frac{\sqrt{3}}{3} \tag{80}$$

$$0 < \beta < \frac{\sqrt{2}}{2} \tag{81}$$

where α_1, α_2, α_3, and α_4 are NN adaptation gains, l_1, l_2, l_3, l_4, l_5, and l_6 are controller gains, and β is employed to define the *strategic* utility function.

Proof: Follow steps similar to [22]. Proof is presented in Appendix B with Theorem 4. ■

Remark 2: A well-defined controller is developed in this paper since a single NN is utilized to approximate two nonlinear functions even though a second-order nonlinear nonstrict feedback discrete-time system is considered as the primary system. The causal problem encountered in the proposed work is due to the nonstrict feedback issue whereas the causal nature encountered in [23] is due to the nth-order strict feedback system. In [23], it was shown how a suitable coordinate transformation can be utilized to overcome the causal problem. Similar approach can be found for the proposed nonstrict feedback nonlinear discrete-time system as well. However, extension to an nth-order system is outside the scope of this work.

Remark 3: Generally, the separation principle used for linear systems does not hold for nonlinear systems, and hence it is relaxed in this paper for the controller design since the Lyapunov function is a quadratic function of system errors and weight estimation errors of the observer and controller NNs.

Remark 4: It is important to note that in this theorem persistency of excitation condition (PE) condition for the NN observer and NN controller and the linearity in the parameters assumption are not needed, in contrast with standard work in the discrete-time adaptive control, since the first difference does not require the PE condition to prove the boundedness of the weights. Even though the input to the hidden-layer weight matrix is not updated and only the hidden- to the output-layer weight matrix is tuned, the NN method relaxes the linear in the unknown parameter assumption. Additionally, the certainty equivalence principle is not used.

Remark 5: The NN weight tuning proposed in (24), (37), (57), and (68) renders a semirecurrent NN due to the proposed weight tuning law even though a feedforward NN is utilized. Here the NN outputs are not fed as delayed inputs to the network whereas the outputs of each layer are fed as delayed inputs to the same layer. This semirecurrent NN architecture renders a dynamic NN, which is capable of predicting the state one step ahead.

Remark 6: The need for an exact model of the nonlinear discrete-time in many existing ACD approaches [5], [6] is relaxed in our work. The action NNs will learn the unknown system dynamics through the feedback signals from the closed loop so that it can generate a near-optimal control input. The proposed actor–critic architecture will render a model-free approach.

Remark 7: It is important to note that the output-layer weights of the action and critic NN can be initialized at zero or random. This means that no explicit offline training phase is necessary and the updating of the NNs is performed in an online manner. This is in contrast with many ACD designs where some *a priori* training is needed. Additionally, the proposed methodology does not require stop/reset strategy utilized by certain adaptive critic schemes [6].

Remark 8: Compared to other adaptive critic or reinforcement learning schemes [5], [6], the proposed approach ensures closed-loop stability using the Lyapunov approach even though gradient-based adaptation is employed.

Remark 9: It is only possible to show the boundedness of all the closed-loop signals by using an extension of Lyapunov stability [15] due to the presence of approximation errors and bounded disturbances consistent with the literature. Consequently, a near-optimal solution can be demonstrated by using the update laws of the critic and action NNs.

Remark 10: Equations (71)–(80) relate to the selection of adaptation gains whereas (81) provides how the discount factor can be chosen in order to ensure stability and convergence. Such a relationship does not exist in the past adaptive critic literature where the discount factor and adaptation gains are selected by trail and error procedure.

Remark 11: With the proposed approach, the learning can be performed simultaneously both in the critic and action NNs in contrast with some of the available schemes where the learning is first accomplished by the critic NN and then by the action NN.

Corollary 1: The proposed adaptive critic NN controller and the weight-updating rules with parameter selection based on (71)–(81) cause the state $x_2(k)$ to approach the desired virtual control input $x_{2d}(k)$.

Proof: Combining (45) and (46), the difference between $\hat{x}_{2d}(k)$ and $x_{2d}(k)$ is given by

$$\hat{x}_{2d}(k) - x_{2d}(k) = \tilde{w}_3(k)\phi_3(k) - \varepsilon\left(z_3(k)\right) = \zeta_3(k) - \varepsilon_3(k) \tag{82}$$

where $\tilde{w}_3(k) \in \Re^{n_3}$ is the first action NN weight estimation error and $\zeta_3(k) \in \Re$ is defined in (50). Since both $\zeta_3(k) \in \Re$ and $\varepsilon_3(k)$ are bounded, $\hat{x}_{2d}(k)$ is bounded near $x_{2d}(k)$. In Theorem 1, we show that $e_2(k)$ is bounded, i.e., the state $x_2(k)$ is bounded to the virtual control signal $\hat{x}_{2d}(k)$. Thus, the state $x_2(k)$ is bounded to the desired virtual control signal $x_{2d}(k)$.

IV. Secondary Controller Design

To simplify the controller development, the third equation can be simplified as

$$x_3(k+1) = f_4\left(x(k)\right) + g_4\left(x(k)\right)v(k) + d_3(k) \tag{83}$$

where $x(k) = [x_1(k), x_2(k), x_3(k)]^T$. This equation can be represented as a standard affine nonlinear discrete-time system. The controller design for this system is different than the nonstrict feedback nonlinear discrete-time system given by (1) and (2). Therefore, the design of a novel reinforcement controller is introduced here by assuming that the third state is measurable. For maintaining dilution to a desired level, the third equation will be employed with EGR(k) as the control input and inert gas as an additional state. It is important to notice that the residual gas fraction is not known in advance though upper bounded is known. The objective of the secondary controller is to force the error between the actual state $x_3(k)$ to approach its target value $x_{3d}(k)$.

Define

$$x(k) = [x_1(k), x_2(k), x_3(k)]^T. \tag{84}$$

A. Critic NN Design

Let the long-term cost function be defined as

$$J(k) = \sum_{i=t_0}^{\infty} \gamma^i r(k+i) \tag{85}$$

where

$$r(k) = \left(x(k) - x_d(k)\right)^T Q\left(x(k) - x_d(k)\right) + v^T(k)Rv(k) \tag{86}$$

with R and Q are positive definite matrices and γ is the discount factor within the range of $0 \leq \gamma \leq 1$ selected by the designer. Invoke the universal approximation property of NN to estimate (85) as

$$J(k) = w_c^T \phi_c\left(v_c^T z_c(k)\right) + \varepsilon\left(z_c(k)\right) \tag{87}$$

where $\varepsilon(z_c(k))$ is the estimation error. Replace the states with estimated states

$$\hat{J}(k) = \hat{w}_c^T(k)\phi_c\left(v_c^T \hat{z}_c(k)\right) = \hat{w}_c^T(k)\phi_c(k) \tag{88}$$

where $\hat{w}_c \in \Re^{n_c}$ and $v_c \in \Re^{2\times n_c}$ denote the ideal output and constant hidden-layer weights, $\phi_c(k) \in \Re^{n_c}$ is the activation function vector, and n_c is the number of hidden-layer nodes.

Again, we hold the input weights constant and adapt the output weights only. Take $\hat{z}_c(k) = [\hat{x}_1(k), \hat{x}_2(k), x_3(k)]^T \in \Re^3$ as the input vector.

Define the prediction error as

$$\begin{aligned} e_c(k) &= \gamma \hat{J}(k) - \left[\hat{J}(k-1) - r(k)\right] \\ &= \gamma\zeta_c(k) + \gamma J(k) - \zeta_c(k-1) - J(k-1) \\ &\quad + r(k) - \varepsilon_c(k) + \varepsilon_c(k-1) \end{aligned} \tag{89}$$

where

$$\zeta_c(k) = \tilde{w}_c^T(k)\hat{\phi}_c(k) = \hat{w}_c^T(k)\hat{\phi}_c(k) - w_c^T\hat{\phi}_c(k). \tag{90}$$

Using a quadratic minimizing function

$$E_c(k) = \frac{1}{2}e_c^2(k) \tag{91}$$

Using a standard gradient-based adaptation method, the weight update is given by

$$\begin{aligned} \hat{w}_c(k+1) &= \hat{w}_c(k) + \alpha_c\left[-\frac{\partial E_c(k)}{\partial \hat{w}_c(k)}\right] \\ &= \hat{w}_c(k) - \alpha_c\gamma\phi_c(k)e_c(k) \\ &= \hat{w}_c(k) - \alpha_c\gamma\phi_c(k) \\ &\quad \times \left(\gamma\hat{J}(k) + r(k) - \hat{J}(k-1)\right). \end{aligned} \tag{92}$$

The control design is described next.

B. Action NN Design

The tracking error is defined as

$$\begin{aligned} e_4(k) &= x_3(k) - x_{3d}(k) \\ e_4(k+1) &= f_4(\cdot) + g_4(\cdot)v(k) + d_3(k) - x_d(k+1) \end{aligned} \tag{93}$$

where $x_{3d}(k)$ is the target bounded trajectory. Define the desired control signal as

$$v_d(k) = g_4^{-1}(\cdot)\left(-f_4(\cdot) + x_{3d}(k+1) + l_7e_4(k)\right) \tag{94}$$

where l_7 is a design parameter. Using the universal approximation property of NN and the approximate states

$$\hat{v}_d(k) = \hat{w}_a^T(k)\phi_a\left(v_a^T\hat{z}_a(k)\right) = \hat{w}_a^T(k)\hat{\phi}_a(k) \tag{95}$$

where $\hat{w}_a \in \Re^{n_a}$ and $v_a \in \Re^{2\times n_a}$ denote the ideal output- and constant hidden-layer weight matrices, $\phi_a(k) \in \Re^{n_a}$ is the activation function vector, n_a is the number of hidden layer nodes, and $\hat{z}_a(k) = [\hat{x}_1(k), \hat{x}_2(k), x_3(k)]^T \in \Re^3$ is the input vector. Again, we hold the input weights constant and adapt the output weights only. Rewrite (93) as

$$\begin{aligned} e_4(k+1) &= l_7e_4(k) + g_4(\cdot)\left(v(k) - v_d(k)\right) + d_3(k) \\ &= l_7e_4(k) + g_4(\cdot)\left(\tilde{w}_a^T(k)\phi_a(k) - \varepsilon_a(k)\right) + d_3(k) \\ &= l_7e_4(k) + g_4(\cdot)\zeta_a(k) + d_a(k) \end{aligned} \tag{96}$$

where

$$d_a(k) = -g_4(\cdot)\varepsilon_a(k) + d_3(k) \tag{97}$$

and

$$\zeta_a(k) = \tilde{w}_a^T(k)\hat{\phi}_a(k). \tag{98}$$

Define the control input cost function

$$\begin{aligned} e_a(k) &= \sqrt{g_4(\cdot)}\zeta_a(k) + \left(\sqrt{g_4(\cdot)}\right)^{-1}\left(\hat{J}(k) - J_d(k)\right) \\ &= \sqrt{g_4(\cdot)}\zeta_a(k) + \left(\sqrt{g_4(\cdot)}\right)^{-1}\hat{J}(k) \end{aligned} \tag{99}$$

where $J_d(k)$ is the desired long-term cost function and is equal to zero. Define a quadratic error to minimize

$$E_a(k) = \frac{1}{2}e_a^2(k). \tag{100}$$

Utilizing a gradient–decent minimization strategy

$$\begin{aligned} \hat{w}_a(k+1) &= \hat{w}_a(k) + \alpha_a\left[-\frac{\partial E_a(k)}{\partial \hat{w}_a(k)}\right] \\ &= \hat{w}_a(k) - \alpha_a\gamma\phi_a(k)\left(g_4(\cdot)\zeta_a(k) + J(k)\right)^T \\ &= \hat{w}_a(k) - \alpha_a\gamma\phi_a(k) \\ &\quad \times \left(e_4(k+1) - l_7e_4(k) - d_a(k) + \hat{J}(k)\right)^T. \end{aligned}$$

The disturbance and the approximation errors are taken as zero for the sake of weight tuning. Then, action NN weight tuning for the secondary controller can be expressed as

$$\begin{aligned} \hat{w}_a(k+1) &= \hat{w}_a(k) - \alpha_a\gamma\phi_a(k) \\ &\quad \times \left(e_4(k+1) - l_7e_4(k) + \hat{J}(k)\right)^T. \end{aligned} \tag{101}$$

The structure of the combined primary and secondary controller is shown in Fig. 1. Next the stability of the closed-loop system of the secondary controller is demonstrated. Before we proceed, the following mild assumptions are stated.

Assumption 4: Let w_a and w_c be the unknown output-layer target weights for the action and critic NNs, respectively, and assume that they are bounded such that $\|w_a\| \le w_{am}$ and $\|w_c\| \le w_{cm}$ where $w_{am} \in R^+$ and $w_{cm} \in R^+$ represent the bounds on the ideal weights.

Fact 2: The activation functions for the action and critic NNs of the secondary controller are bounded by known positive values, such that $\|\phi_a(k)\| \le \phi_{am}$ and $\|\phi_c(k)\| \le \phi_{cm}$, where $\phi_{am}, \phi_{cm} \in R^+$ denote the upper bound for the activation functions.

Assumption 5: The unknown smooth function $g_4(\cdot)$ is bounded away from zero for all $x_1(k)x_2(k), x_3(k)$ within the compact set S. In other words, $0 < g_{4\min} < |g_4(\cdot)| < g_{4\max}$, $\forall x_1(k), x_2(k), x_3(k) \in S$, where $g_{4\min} \in \Re^+$ and $g_{4\max} \in \Re^+$. Without loss of generality, we will assume that $g_4(\cdot)$ is positive in this paper. Additionally, the NN approximation errors $\varepsilon_a(z_a(k))$ and $\varepsilon_c(z_c(k))$ are bounded above in the compact set $S \subset R^n$ by ε_{am} and ε_{cm} [15].

Remark 12: The tracking and weight estimation errors will be expressed as a function of NN approximation errors and bounded disturbances. Even if the Assumption 5 is assumed, unless the NN weight tuning is properly selected, the entire system will not be stable.

Fact 2: With the Assumption 4, the term $d_a(k)$ in (97) is bounded over the compact set $S \subset R^n$ by $\|d_a(k)\| \le d_{am} = g_{4\max}\varepsilon_{am} + d_m$.

Theorem 3 (Secondary Controller Stability): Consider the secondary system given by (5) and the disturbance bound d_{3m} to be known constants. Let the virtual and control input NN weight tuning be given by (92) and (101), respectively. Let the control input be given by (95); the tracking error $e_4(k)$ and weight estimates $\hat{w}_a(k)$ and $\hat{w}_c(k)$ are UUB with the controller design parameters selected as

$$0 < \alpha_a \left\|\phi_a(k)\right\|^2 < 1 \tag{102}$$

$$0 < \alpha_c \left\|\phi_c(k)\right\|^2 < 1 \tag{103}$$

$$|l_7| < \frac{1}{2} \tag{104}$$

where α_a and α_c are NN adaptation gains and l_7 is the controller gain.

Proof: Proof is combined with Theorem 4. ■

Remark 13: Theorem 3 demonstrates the UUB of the secondary controller for the closed-loop system. The primary and secondary controller designs introduced here are analogous to dual-controller design approach in the literature [20].

Theorem 4 (Overall System Boundedness): Consider the system given by (3)–(6) and the disturbance bounds d_{1m}, d_{2m}, and d_{3m} to be known constants. Let the observer, critic, virtual control, and control input NN weight tuning be given by (24), (37), (57), and (68), respectively, for the primary adaptive controller whereas virtual and control input NN weight tuning be given by (92) and (101), respectively. Let the virtual control input and control input for the primary and secondary controllers be given by (46), (60), and (95), respectively, then the estimation and tracking errors $e(k) = [e_1(k), e_2(k), e_3(k), e_4(k)]^T$ and weight estimates $\hat{w}_1(k)$, $\hat{w}_2(k)$, $\hat{w}_3(k)$, $\hat{w}_4(k)$, $\hat{w}_a(k)$, and $\hat{w}_c(k)$ are UUB with the bounds given in (B.35) provided the design parameters are selected as (71)–(81) and (102)–(104).

Proof: See Appendix B.

V. Results and Analysis

EGR operation of an SI engine allows lower emissions and improved fuel efficiency. However, EGR operation destabilizes the engine due to the cyclic dispersion of heat release. The adaptive critic NN controller is designed to stabilize the SI engine operating at EGR conditions.

A. Daw Engine Model

SI engine dynamics can be expressed according to the Daw model as a class of nonlinear systems in a nonstrict feedback form [16]. At high EGR levels, the engine can be expressed as a combination of nonstrict feedback nonlinear discrete-time systems plus affine nonlinear discrete-time system given by [17]

$$\begin{aligned} x_1(k+1) &= AF(k) + F(k)x_1(k) - R\cdot F(k)CE(k)x_2(k) \\ &\quad + F(k)\left(r_{O_2}(k) + r_{N_2}(k)\right) + d_1(k) \end{aligned} \tag{105}$$

$$\begin{aligned} x_2(k+1) &= \left(1 - CE(k)\right)F(k)x_2(k) \\ &\quad + \left(MF(k) + u(k)\right) + d_2(k) \end{aligned} \tag{106}$$

$$\begin{aligned} x_3(k+1) &= F(k)\left(r_{CO_2}(k) + r_{H_2O}(k) + r_{N_2}(k)\right. \\ &\quad \left. + x_3(k) + EGR(k)\right) \end{aligned} \tag{107}$$

$$y(k) = x_2(k)CE(k) \tag{108}$$

$$\begin{aligned} \varphi(k) &= R\frac{x_2(k)}{x_1(k)} \\ &\quad \times \left[1 - \gamma\frac{x_3(k) + EGR(k)}{\left(x_3(k) + x_1(k) + x_3(k) + EGR(k)\right)}\right] \end{aligned} \tag{109}$$

$$CE(k) = \frac{CE_{\max}}{1 + 100^{-(\varphi(k) - \varphi_m)/(\varphi_u - \varphi_l)}} \tag{110}$$

$$\varphi_m = \frac{\varphi_u - \varphi_l}{2} \tag{111}$$

$$r_{H_2O}(k) = \gamma_{H_2O}x_2(k)CE(k) \tag{112}$$

$$r_{O_2}(k) = \gamma_{O_2}x_2(k)CE(k) \tag{113}$$

$$r_{N_2}(k) = \gamma_{N_2}R\cdot x_2(k)CE(k) \tag{114}$$

$$r_{CO_2}(k) = \gamma_{CO_2}x_2(k)CE(k) \tag{115}$$

where $x_1(k)$, $x_2(k)$, and $x_3(k)$ represent total mass of air, fuel, and inert gas, respectively, of the engine and $y_1(k)$ is the heat release at kth time instant. The value of combustion efficiency is denoted by $CE(k)$, which is bounded by $0 < CE_{\min} < CE(k) < CE_{\max}$ with maximum combustion efficiency being unity. Residual gas fraction $F(k)$ is bounded by $0 < F_{\min} < F(k) < F_{\max}$, which is typically unknown. $d_1(k)$ and $d_2(k)$ are unknown but bounded disturbances bounded by $|d_1(k)| < d_{1m}$ and $|d_2(k)| < d_{2m}$ with d_{1m} and d_{2m} being unknown positive scalars. The terms φ_m, φ_l, and φ_u are equivalence ratio system parameters. The terms $r_{H_2O}(k)$, $r_{O_2}(k)$, $r_{N_2}(k)$, and $r_{CO_2}(k)$ are the mass of water, oxygen, nitrogen, and carbon dioxide, respectively, whereas γ, γ_{H_2O}, γ_{O_2}, γ_{N_2}, and γ_{CO_2} are design constants, and constants associated with their respective chemicals.

Equations (105) and (106) represent the second-order nonstrict feedback nonlinear discrete-time systems whereas (107) can be controlled by the secondary controller. In this paper, for convenience, we assume that the secondary controller provides a bounded value to the primary system by setting the EGR at discrete levels. We set it to a constant that will simplify the controller implementation, as the third state is considered to be a fixed value. Note that this deterministic model accounts for stochastic effects by randomly fluctuating parameters such as injected air-fuel ratio or residual fraction. Other complex processes like temperature variation, turbulence, and fuel vaporization are not modeled but are assumed to add additional noise to the engine output. To implement the observer, replace the following from the Daw model into the general case:

$$\begin{aligned} f_1(\cdot) &= AF(k) + F(k)x_1(k) + F\left(r_{O_2}(k) + r_{N_2}(k)\right) \\ g_1(\cdot) &= -R\cdot F(k)CE(k) \\ f_2(\cdot) &= \left(1 - CE(k)\right)F(k)x_2(k) + MF(k) \\ g_2(\cdot) &= 1 \end{aligned} \tag{116}$$

and

$$\begin{aligned} f_{10} &= AF_0 + F_0\hat{x}_1(k) \\ g_{10} &= -R\cdot F_0CE_0 \\ f_{10} &= (1 - CE_0)F_0\hat{x}_2(k) + MF_0 \\ g_{10} &= 1. \end{aligned} \tag{117}$$

Note that we omitted the residuals in f_{10}, because they are not available. The error introduced by this is accounted for in

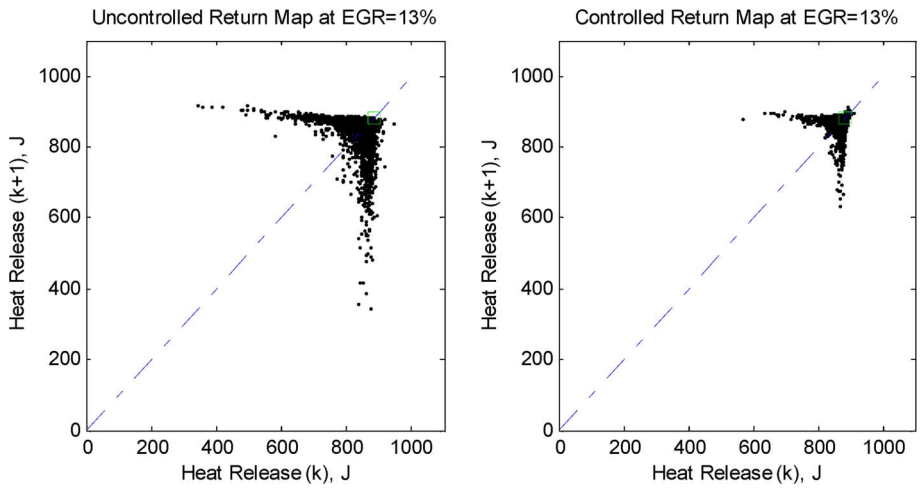

Fig. 2. Uncontrolled and controlled heat release return map at 13% EGR. Heat release at $k+1$ instance is plotted against heat release at k instance.

the air estimation error. To implement the controller, replace the following in place of $f_1(\cdot)$ and $g_1(\cdot)$:

$$\Phi(\cdot) = AF(k) + F(k)x_1(k) - R \cdot F(k)CE(k)x_2(k) + x_2(k) + F(k)\left(r_{O_2}(k) + r_{N_2}(k)\right). \quad (118)$$

To calculate the nominal values in (117), we run the engine at different EGR levels in an uncontrolled manner at the stoichiometric fuel to air ratio. That will give us the nominal fuel, air, and equivalence ratio—MF_0, AF_0, and φ_0. From those, combustion efficiency CE_0 is calculated.

B. Simulation Results

The controller is easily simulated in C in conjunction with the Daw model. The learning rates for the observer (71), critic (72), virtual control input (73), and control input (74) networks and for the secondary controller are taken as $0.01, 0.01, 0.01, 0.01, 0.01$, and 0.01, respectively. The gains l_1, l_2, l_3, l_4, l_5, l_6, and l_7 are selected as $0.05, 0.05, 0.04, 0.05, 0.2, 0.1$, and 0.1. The system constants $CE_{\max}$, φ_l, and φ_u are chosen as $1, 0.54$, and 0.58. The critic constants β and N are 0.4 and 4 for both controllers for all EGR levels. All NNs use 20 hidden neurons with hyperbolic tangent sigmoid activation functions in the hidden layer.

The simulation parameters selected were as follows. An equivalence ratio of one was maintained with stochastic variation of 1%, $R = 15.13$ for iso-octane, residual gas fraction $F = 0.09$, mass of nominal new air $= 0.52485$, mass of nominal new fuel $= 0.02428$, the standard deviation of mass of new fuel is 0.007, cylinder volume in moles $= 0.021$ to match the experimental constraint, molecular weight of fuel $= 114$, molecular weight of air $= 28.84$, $\phi_u = 0.665$, $\phi_l = 0.645$, and maximum combustion efficiency $= 1$. EGR was assumed to be an inert mixture with a molecular weight of 30.4.

The last two system variables, disturbances and stochastic effects, are modeled as follows. First, we assume a Gaussian distribution governs the two effects. We may inject disturbances to the two states in (105) and (106) due to $d_1(k)$ and $d_2(k)$, but a simpler method is to perturb the equivalence ratio (109). This simplification is sufficient because the states are not measurable; therefore, the disturbances are increasingly complex and immeasurable. Stochastic effects alter the output, and through the combustion efficiency (110) and finally the output (108), this single perturbation effectively models the last two system variables. The final model uses a Gaussian distribution noise injected into (109) centered around the target equivalence ratio and deviation of 0.007. The resulting simulation output matches to the output observed from the Ricardo engine. All simulations ran for uncontrolled 5000 cycles first, and then 5000 controlled cycles.

Fig. 2 shows two heat release return maps, one controlled and the other uncontrolled, for the set point at 13% EGR. Each figure shows the next time step versus the current time step heat release. Points centered along the 45° line represent heat release values that are equal to the next step heat release. Note the clustering of the points around the mean heat release of 850 J. The square represents the target heat release. At this set point, the heat release dispersion starts to affect the engine performance, indicated by the stray points away from the central cluster. There are no complete misfires, but the heat release variation can be clearly seen. Fig. 3 shows the time series of the heat release and control input at the same EGR level. The controller activates after several thousand cycles, indicated by the fluctuation of the control output. The controller converges quickly and to a stable operation point. The presence of spikes in the control output indicates a decline in heat release such as a misfire, translating into additional fuel control to counteract.

Figs. 4 and 5 depict another set point at 19% EGR. Similar features appear compared to the previous EGR level, except with higher frequency and amplitude of dispersion. Improvements shown reflect the assertion of the control action.

In order to quantify the performance of the controller, we compare the coefficient of variation (COV), which is the standard deviation normalized by dividing the mean of the heat release. As the COV decreases, the standard deviation decreases, which indicates that the engine heat release is more stable compared to higher COV. The controller performs better, and the return map consequently should approach the target value. Table I tabulates all of the data from the simulation. The COV of each set point decreased drastically (shown with a negative sign) as the controller operated. The performance exceeded the improvement due to the slight increase in the mean fuel input. Next, we show that experimental data supports the simulation data.

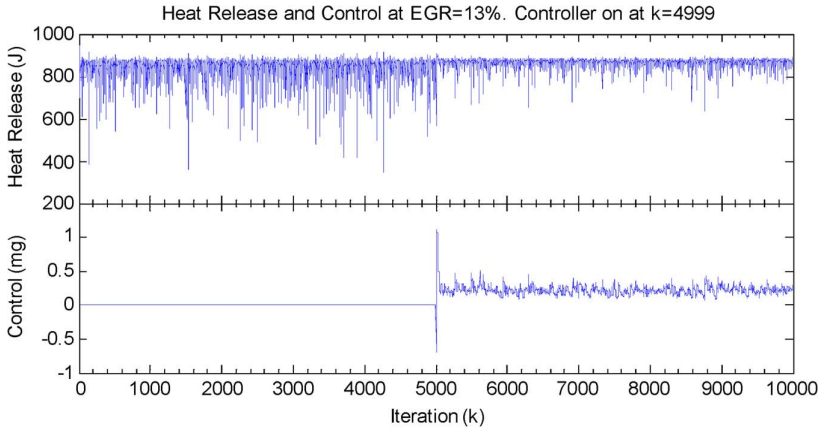

Fig. 3. Heat release versus iteration number at 13% EGR. Controller turns on at $k = 4000$. Note the almost instant learning convergence of the controller.

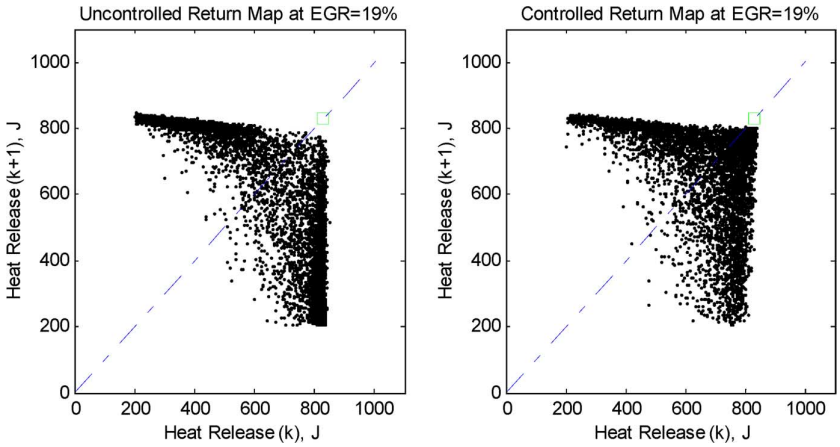

Fig. 4. Uncontrolled and controlled heat release return map at 19% EGR.

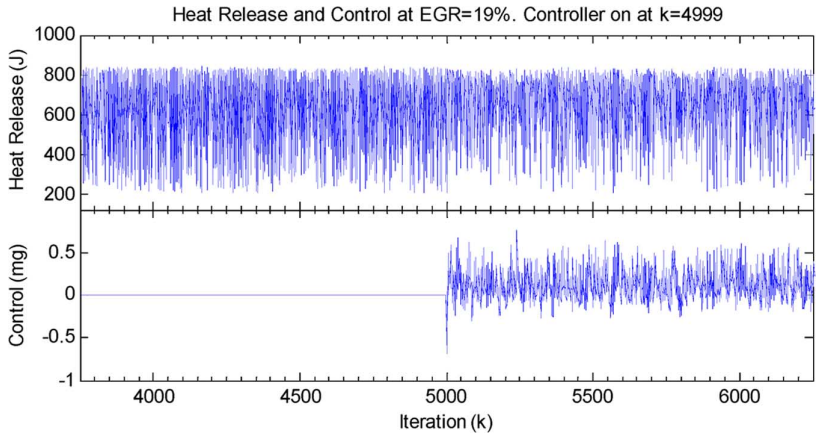

Fig. 5. Heat release and control input at 19% EGR.

C. Experimental Results Using Ricardo Engine

The experimental results are collected from a Ricardo hydra engine with a modern four-valve Ford Zetek head. It contains a single cylinder running at 1000 r/min with shaft encoders to signal each crank angle degree and start of cycle. There are 720° per engine cycle.

In the cylinder, a piezoelectric pressure transducer records pressure every crank angle degree. Combustion is considered to take place between 345° and 490°, for a total of 145 pressure measurements. The cylinder pressure is integrated along with volume during the 17.7-ms calculation window. All communications are completed at this time. The output of our controller controls the fuel input. This is controlled by a transistor–transistor logic (TTL) signal to a fuel injector driver circuit.

All signals communicate through a custom interface board using a microcontroller. The board interfaces with the PC

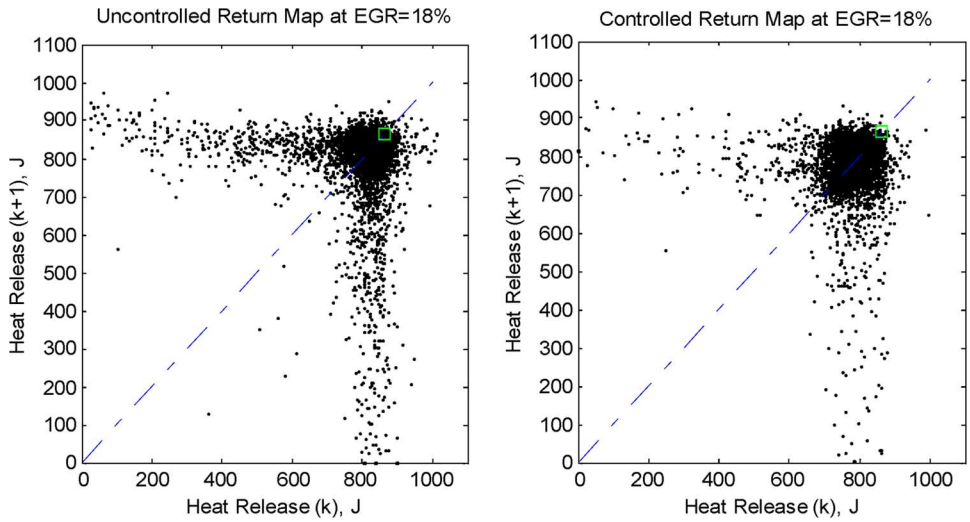

Fig. 6. Uncontrolled and controlled heat release return map at EGR $= 18\%$. Heat release at $k+1$ instance is plotted against heat release at k instance.

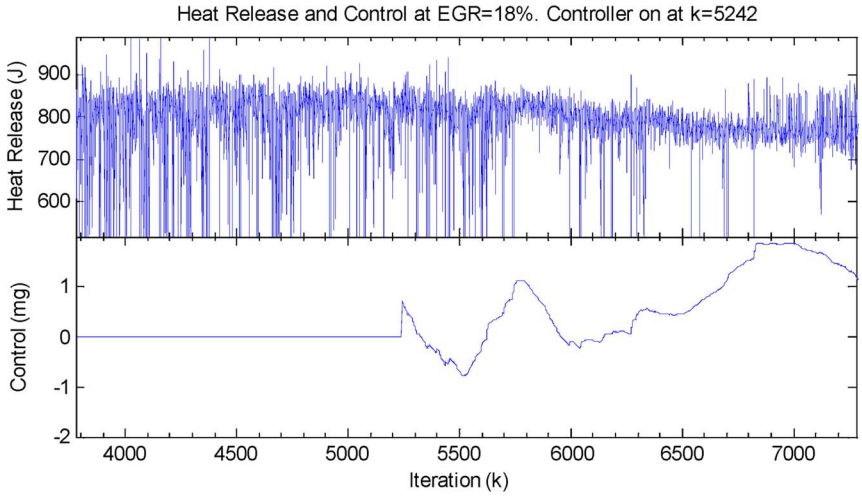

Fig. 7. Uncontrolled and controlled heat release return map at EGR $= 18\%$. Heat release at $k+1$ instance is plotted against heat release at k instance.

through a parallel port and with the engine hardware through an analog signal.

All constants given in the simulation section are used in the experiment. The first operation for an engine run is to measure the air flow and nominal fuel. The desired EGR set point equation is given by

$$\%\text{EGR} = 100 \times \left(\frac{m_{\text{EGR}}}{m_f + m_a + m_{\text{EGR}}} \right) \tag{119}$$

where m_{EGR} is the mass of inert gas introduced at each cycle, which is nitrogen in the lab and exhaust gas in production applications and m_f and m_a are mass of fuel and mass of air, respectively.

These values are loaded into the controller. Ambient pressure is used to reference the in-cylinder pressures when the exhaust valve is fully open and subtracted from the combustion pressure measurements. Uncontrolled and controlled data were collected at EGR percentages of 18, 20, and 23. The uncontrolled engine ran for 5000 cycles and then the controller is turned on for another 5000 cycles. Steady state was ensured prior to data collection by measuring stable exhaust temperatures.

Fig. 6 shows two heat release return maps, one controlled and the other uncontrolled, for the 18% EGR set point. The target heat release is at 870 J. At this EGR level, cyclic dispersion can clearly be seen, indicated by deviation of the points away from the main cluster on the 45° line. Fig. 7 shows the time series of the heat release and control input for the same set point.

Define the state and output tracking errors as

$$\begin{aligned} \hat{e}_1(k) &= \hat{x}_1(k) - x_{1d}(k) \\ \hat{e}_2(k) &= \hat{x}_2(k) - \hat{x}_{2d}(k) \end{aligned} \tag{120}$$

where $\hat{e}_1(k)$ and $\hat{e}_2(k)$ are state 1 and state 2 tracking errors, respectively. Fig. 8 shows the controller state tracking errors at a set point of 18% EGR. The range represents tracking error in percentage over and under the desired state trajectories. State 1 tracking error is considerably better than state 2 tracking. The second state tracks within 0.5%; therefore, both are performing

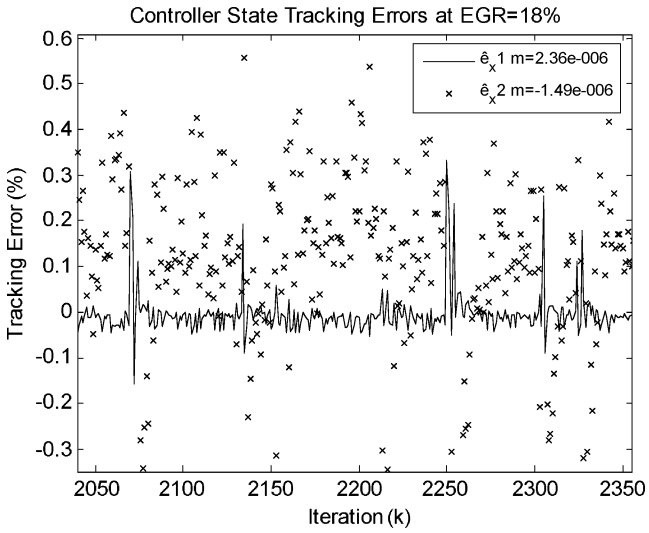

Fig. 8. State tracking errors.

TABLE I
COV AND FUEL DATA FOR EACH OF THE FOUR SET POINTS

EGR Fraction	Covariance (COV) Uncontrolled	Controlled	%COV Change	%Fuel Change
0.00	0.0058	0.0057	-0.75	0.00
0.13	0.0548	0.0384	-29.94	0.40
0.15	0.1387	0.0773	-44.30	0.71
0.19	0.3421	0.2383	-30.34	0.42

well. The spikes indicate unsuccessful tracking. Consequently, the observer and controller converged together to the desired states and estimated states, generating a stable error system.

Fig. 9 shows the return map of the heat release for 20% EGR. Note that as the equivalence ratio decreases, the return map spreads out and dispersion increases. Fig. 10 is the corresponding heat release and control time series. Misfires increase in frequency, as shown by the negative heat release spikes due to heat transfer from the cylinder to the environment without internal generation of useful work by combustion. Fig. 11 shows the increasing difficulty of the observer and controller to generate a low state tracking error compared to the previous case. As the engine operates in higher EGR modes, overall dispersion increases, thus degrading observer performance. Although the performance is reduced, the tracking error is well within satisfactory performance.

Fig. 12 illustrates a detailed view of 70 controlled cycles at 20% EGR. The controller generates decreasing control during cycles when the heat release is steady, indicated by cycles between 4805 and 4818 and between 4822 and 4836. However, during misfires or extreme dispersion in heat release, the controller attempts to compensate for the drop in heat release by pushing the control up, indicated by cycles 4819, 4847, etc. The controller compensates after a one cycle delay in the positive direction and attempts to recover the engine heat release towards the target point. It is difficult to determine success on cycles with no misfire, because no heat release plots are available for uncontrolled case during the same cycles when the controller is operating for comparison. Overall, the controller performs to general expectation.

Table II shows the improved COV when the controller is in operation compared to an uncontrolled engine along with the corresponding change in nominal fuel. At all EGR set points except 23%, the increase in fuel input is well within the tolerance of the equipment. On average, the COV decreases significantly by 25% compared to the uncontrolled case. For the 23% EGR, the drop is around 12%.

The COV and fuel change data indicate an improved performance compared to the previous controller where a simple tracking error is utilized without any optimization criteria [18]. The average drop in COV was 17% between uncontrolled and controlled cases, compared to 25% for the current controller. Although this seems to indicate an increase in performance, we must also consider the increase in average fuel input in conjunction. The previous controller increased the average fuel to 2.4%, which is well beyond the detection error and significant in the engine application. This controller, however, averages less than 1%, safely below the detection error. The controller fuel increases negligibly while performing better than the previous controller due to the incorporation of the performance index. Therefore, this controller outperforms other NN controller and at the same time exerts less impact on the fuel.

VI. CONCLUSION

The controller presented successfully controlled an SI engine to reduce cyclic dispersion under higher EGR conditions. The system is modeled as a combination of nonstrict feedback nonlinear discrete-time system and affine nonlinear discrete-time system. It converged upon a near-*optimal* solution through the use of a long-term strategic utility function even though the exact dynamics are not known beforehand. It was shown through simulation that the controller is stable under a variety of set points. In experimental results, the COV was reduced when the controller was turned on. At the same time, the average fuel input did not change significantly; therefore, the improvements are solely due to the effects of the controller. The output is stable, as predicted by the Lyapunov proof. There was also a significant reduction in unburned hydrocarbon between controlled and uncontrolled cases.

APPENDIX A

Tables III and IV present the improvement in emissions for several equivalence ratios. The improvement is better than what we have seen before [18] using another controller. NO_x is reduced by around 2%–7.4% from uncontrolled scenario. However CO_2 remains unchanged, whereas O_2 decreases by about 20%, as well as unburned hydrocarbons (uHC) decreasing with control nominally due to reduced cyclic dispersion.

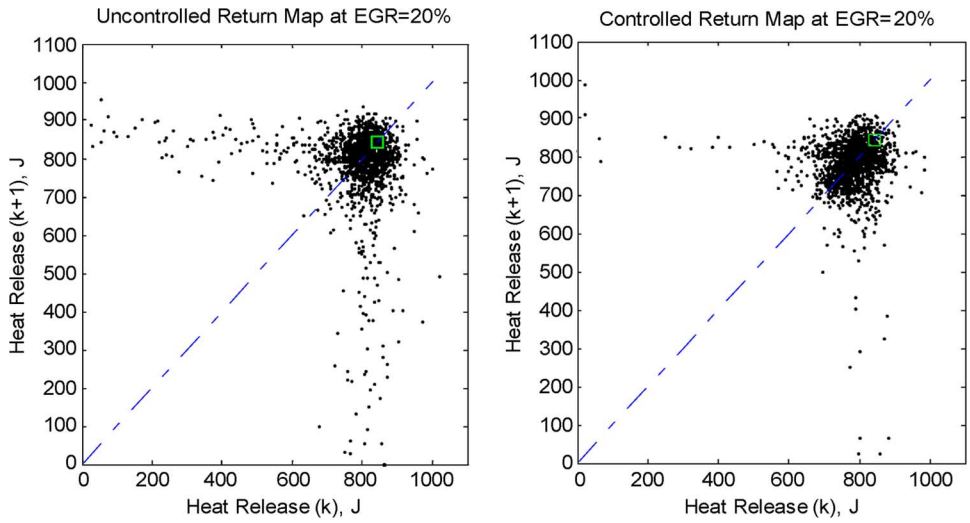

Fig. 9. Uncontrolled and controlled heat release return map at 20% EGR.

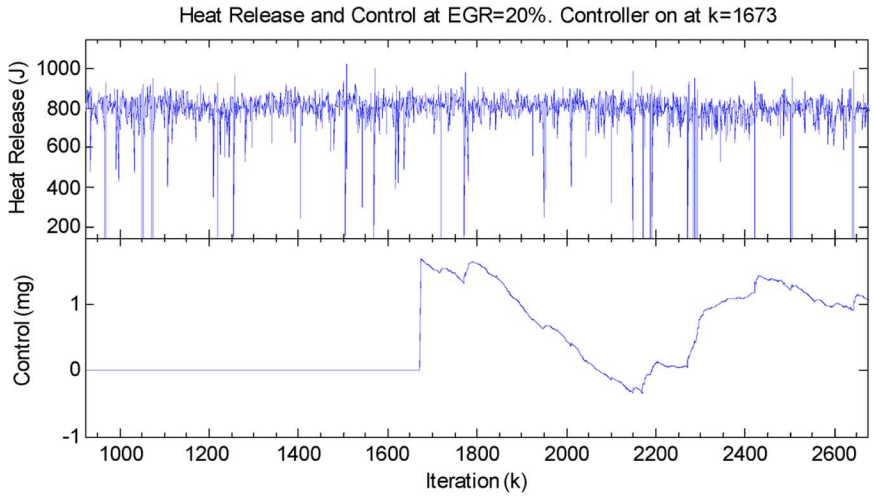

Fig. 10. Heat release and control input at 20% EGR.

APPENDIX B

Proof of Theorem 1: Define the Lyapunov function

$$J(k)=\sum_{i=1}^{4}J_i(k)$$
$$=\frac{\gamma_1}{\alpha_1}\tilde{w}_1^T(k)\tilde{w}_1(k)+\frac{\gamma_2}{3}\tilde{x}_1^2(k)+\frac{\gamma_3}{2}\tilde{x}_2^2(k)+\frac{\gamma_4}{3}\tilde{y}^2 \quad \text{(B.1)}$$

where $0<\gamma_i$, $i\in\{1,2,3,4\}$, are auxiliary constants. Take the first term, take the first difference, and substitute (24)

$$J_1(k)$$
$$=\frac{\gamma_1}{\alpha_1}\tilde{w}_1^T(k)\tilde{w}_1(k)$$
$$\frac{\alpha_1}{\gamma_1}\Delta J_1(k)$$
$$=\tilde{w}_1^T(k+1)\tilde{w}_1(k+1)-\tilde{w}_1^T(k)\tilde{w}_1(k)$$
$$=\left[\tilde{w}_1^T(k)-\alpha_1\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^T\hat{\phi}_1^T(k)\right]$$
$$*\left[\tilde{w}_1(k)-\alpha_1\hat{\phi}_1(k)\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)\right]$$
$$-\tilde{w}_1^T(k)\tilde{w}_1(k)$$
$$=\alpha_1^2\left\|\phi_1(k)\right\|^2\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2$$
$$-2\alpha_1\tilde{w}_1(k)\hat{\phi}_1(k)\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)$$
$$+\alpha_1\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2$$
$$-\alpha_1\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2$$
$$=-\alpha_1\left(1-\alpha_1\left\|\hat{\phi}_1(k)\right\|^2\right)\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2$$
$$+\alpha_1\left(\left(\zeta_1(k)+w_1^T\hat{\phi}_1+l_4\tilde{y}(k)\right)-\zeta_1(k)\right)^2-\alpha_1\zeta_1^2(k)$$
$$=-\alpha_1\left(1-\alpha_1\left\|\hat{\phi}_1(k)\right\|^2\right)\left(\hat{w}_1^T(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2$$
$$+\alpha_1\left(w_1^T\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2-\alpha_1\zeta_1^2(k). \quad \text{(B.2)}$$

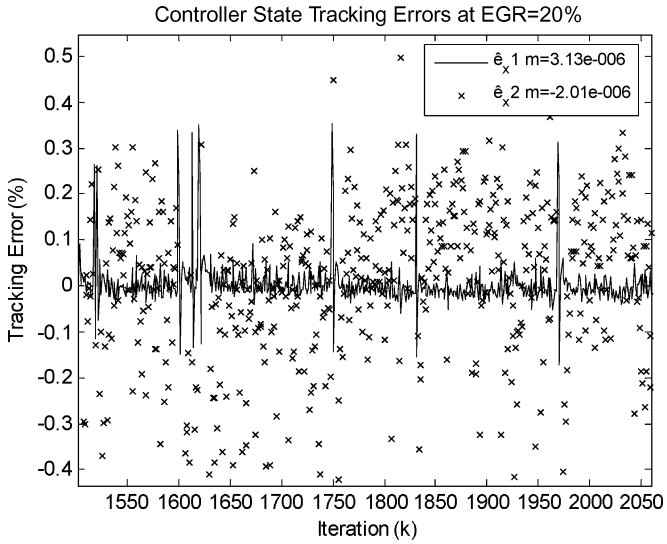

Fig. 11. State tracking errors.

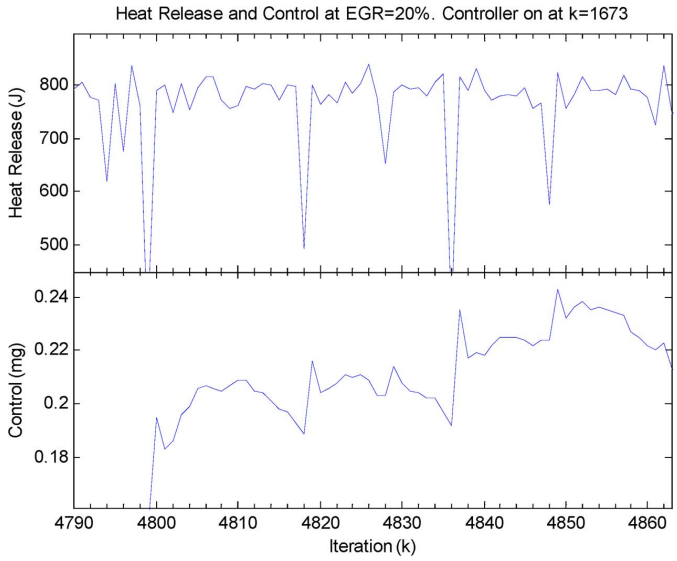

Fig. 12. Detailed view of 70 controlled cycles at 20% EGR.

TABLE II
COV AND FUEL DATA FOR EACH OF THE THREE SET POINTS

EGR	COV Uncontrolled	COV Controlled	%COV Change	%Fuel Change
0.18	0.2112	0.1511	-28.4	1.36
0.20	0.2139	0.1400	-34.6	0.77
0.23	0.5777	0.5066	-12.3	2.11

Invoke the Cauchy–Schwarz inequality, defined as

$$(a_1b_1+\ldots+a_nb_n)^2 \le \left(a_1^2+\ldots+a_n^2\right)\left(b_1^2+\ldots+b_n^2\right) \tag{B.3}$$

and simplify to get

$$\Delta J_1(k) \le -\gamma_1\left(1-\alpha_1\left\|\hat{\phi}_1(k)\right\|^2\right)\left(\hat{w}_1(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2 + 2\gamma_1\left(w_{1\mathrm{m}}\hat{\phi}_{1\mathrm{m}}\right)^2 + 2\gamma_1 l_4^2\tilde{y}^2(k) - \gamma_1\zeta_1^2(k). \tag{B.4}$$

Take the second term and substitute (21) to get

$$\Delta J_2(k) \le \gamma_2 l_2^2\tilde{y}^2(k) + \gamma_2\tilde{x}_2^2(k) + \gamma_2\left(w_{3\mathrm{m}}\phi_{3\mathrm{m}} + f_{10} + \varepsilon_{3\mathrm{m}} + d_{1\mathrm{m}}\right)^2 - \frac{\gamma_2}{3}\tilde{x}_1^2(k). \tag{B.5}$$

Take the third term in (B.1) and substitute (22), assuming bounded input $u_{\max} > |u(k)|$, to get

$$\Delta J_3(k) \le \gamma_3\left(f_{20} + (g_{20}+g_{2\max})u_{\max} + f_{2\max} + d_{2\mathrm{m}}\right)^2 + \gamma_3 l_3^2\tilde{y}^2(k) - \frac{\gamma_3}{2}\tilde{x}_2^2(k). \tag{B.6}$$

Take the fourth and final term in (B.1) and substitute (23) to obtain

$$\Delta J_4(k) \le \gamma_4\zeta_1^2(k) + \gamma_4 l_1^2\tilde{y}(k) + \gamma_4\left(w_{1\mathrm{m}}\tilde{\phi}_{1\mathrm{m}} + \varepsilon_{1\mathrm{m}}\right) - \frac{\gamma_4}{3}\tilde{y}^2(k). \tag{B.7}$$

Combine (B.2)–(B.7) and simplify to get the first difference of the Lyapunov function

$$\begin{aligned}\Delta J(k) \le & -\gamma_1\left(1-\alpha_1\left\|\hat{\phi}_1(k)\right\|^2\right)\left(\hat{w}_1(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2 \\ & -\left(\frac{\gamma_3}{2}-\gamma_2\right)\tilde{x}_2^2(k) - \frac{\gamma_2}{3}\tilde{x}_1^2(k) \\ & -\left(\frac{\gamma_4}{3} - 2\gamma_1 l_4^2 - \gamma_2 l_2^2 - \gamma_3 l_3^2 - \gamma_4 l_1^2\right)\tilde{y}^2(k) \\ & -(\gamma_1-\gamma_4)\zeta_1^2(k) + D_M^2\end{aligned} \tag{B.8}$$

where D_M^2 is defined as

$$\begin{aligned}D_M^2 = & 2\gamma_1(w_{1\mathrm{m}}\hat{\phi}_{1\mathrm{m}})^2 + \gamma_2(w_{3\mathrm{m}}\phi_{3\mathrm{m}} + f_{10} + \varepsilon_{3\mathrm{m}} + d_{1\mathrm{m}})^2 \\ & + \gamma_3\left(f_{20} + (g_{20}+g_{2\max})u_{\max} + f_{2\max} + d_{2\mathrm{m}}\right)^2 \\ & + \gamma_4(w_{1\mathrm{m}}\tilde{\phi}_{1\mathrm{m}} + \varepsilon_{1\mathrm{m}}).\end{aligned} \tag{B.9}$$

Select

$$\gamma_3 > 2\gamma_2 \quad \gamma_4 > 6\gamma_1 l_4^2 + 3\gamma_2 l_2^2 + 3\gamma_3 l_3^2 + 3\gamma_4 l_1^2 \quad \gamma_1 > \gamma_4. \tag{B.10}$$

This implies $\Delta J(k) < 0$ as long as (25)–(29) *and* the following hold:

$$|\tilde{x}_1(k)| > \frac{D_M}{\sqrt{\frac{\gamma_2}{3}}} \quad \text{or} \quad |\tilde{x}_2(k)| > \frac{D_M}{\sqrt{\frac{\gamma_3}{2}-\gamma_2}}$$

or

$$|\tilde{y}(k)| > \frac{D_M}{\sqrt{\frac{\gamma_4}{3} - 2\gamma_1 l_4^2 - \gamma_2 l_2^2 - \gamma_3 l_3^2 - \gamma_4 l_1^2}}$$

or

$$|\zeta_1(k)| > \frac{D_M}{\sqrt{\gamma_1-\gamma_4}}. \tag{B.11}$$

According to a standard Lyapunov extension theorem [15], this demonstrates that the estimation errors, the output error, and the NN observer weight estimation errors are UUB.

TABLE III
EMISSIONS DATA FOR SELECT EGR SET POINTS

	Uncontrolled			Controlled			Change(%)		
EGR	CO_2 (%)	CO (%)	O_2 (%)	CO_2 (%)	CO (%)	O_2 (%)	CO_2	CO	O_2
0.13	7.4	0.1	6.9	7.4	0.1	2.2	-0.4	0.0	-68.1
0.18	7.3	0.1	2.3	7.3	0.1	2.4	0.0	0.0	5.2
0.20	7.3	0.1	3.2	7.3	0.0	3.6	0.5	-10.0	10.8
0.23	7.3	0.1	5.4	7.3	0.1	5.8	-0.8	0.0	8.2

TABLE IV
UNBURNED HYDROCARBON EMISSION DATA

	Uncontrolled		Controlled		Change(%)	
EGR	NO_X (ppm)	uHC (ppm C_1)	NO_X (ppm)	uHC (ppm C_1)	NO_X	uHC
0.13	554.0	10619	478	10677	-13.7	0.5
0.18	82.0	13610	72	12605	-12.2	-7.4
0.20	51.0	14450	51	14108	0.0	-2.4
0.23	50.0	23928	45	22345	-5.0	-6.6

Proof of Theorem 4: Define the Lyapunov function

$$\begin{aligned}\overline{J}(k)&=\sum_{i=1}^{10}J_i(k)\\&=\frac{\gamma_1}{5}e_1^2(k)+\frac{\gamma_2}{3}e_2^2(k)+\sum_{j=3}^{6}\frac{\gamma_j}{\alpha_{j-2}}\tilde{w}_j^T(k)\tilde{w}_j(k)\\&\quad+\gamma_7\zeta_2^2(k-1)+\frac{\gamma_8}{3}\tilde{x}_1^2(k)+\frac{\gamma_9}{3}\tilde{x}_2^2(k)+\frac{\gamma_{10}}{3}\tilde{y}^2\end{aligned}\tag{B.12}$$

where $0<\gamma_i$, $i\in\{1,\ldots,6\}$, are auxiliary constants; the NN weights estimation errors $\tilde{w}_1^T(k+1)$, $\tilde{w}_2^T(k+1)$, $\tilde{w}_3^T(k+1)$, and $\tilde{w}_4^T(k+1)$ are defined in (24), (37), (57), and (68), by subtracting their respective ideal weights w_i, $i\in\{1,2,3,4\}$, on both sides; the observation errors $\tilde{x}_1(k+1)$ and $\tilde{x}_2(k+1)$ are defined in (21) and (22), respectively; the system errors $e_1(k+1)$ and $e_2(k+1)$ are defined in (49) and (61), respectively; and α_i, $i\in\{1,2,3,4\}$, are NN adaptation gains

$$\begin{aligned}J(k)&=\overline{J}(k)+\sum_{k=11}^{14}J_k\\&=\frac{\gamma_{11}}{3}\left\|e_4(k)\right\|^2+\frac{\gamma_{12}}{\alpha_4}tr\left(\tilde{w}_a^T(k)\tilde{w}_a(k)\right)\\&\quad+\frac{\gamma_{13}}{\alpha_c}tr\left(\tilde{w}_c^T(k)\tilde{w}_c(k)\right)+\gamma_{14}\left\|\zeta_c(k-1)\right\|^2\end{aligned}\tag{B.13}$$

where $0<\gamma_k$, $k\in\{11,\ldots,14\}$, are auxiliary constants; the NN weights estimation errors $\tilde{w}_a^T(k+1)$ and $\tilde{w}_c^T(k+1)$ are defined in (101) and (92), by subtracting their respective ideal weights w_i, $i\in\{a,c\}$, on both sides; the system error $e_4(k+1)$ is defined in (96); and α_i, $i\in\{a,c\}$, are NN adaptation gains.

The Lyapunov function (B.13) obviates the need for the CE condition. Take the first term from (B.12) and the first difference using (49) to get

$$\begin{aligned}J_1(k)&=\frac{\gamma_1}{5}e_1^2(k)\\\frac{5}{\gamma_1}\Delta J_1(k)&=e_1^2(k+1)-e_1^2(k)\\&=\Big(-\zeta_3(k)-w_3^T\tilde{\phi}_3(k)+\varepsilon_3(k)\\&\qquad-l_5\hat{e}_1(k)-e_2(k)+d_1(k)\Big)^2-e_1^2(k)\\&=\Big(-\zeta_3(k)-w_3^T\tilde{\phi}_3(k)+\varepsilon_3(k)-l_5\left(\tilde{x}_1(k)+e_1(k)\right)\\&\qquad-e_2(k)+d_1(k)\Big)^2-e_1^2(k).\end{aligned}\tag{B.14}$$

Invoke the Cauchy–Schwarz inequality defined as

$$(a_1b_1+\ldots+a_nb_n)^2\le\left(a_1^2+\ldots+a_n^2\right)\left(b_1^2+\ldots+b_n^2\right).\tag{B.15}$$

Simplify to get

$$\begin{aligned}\frac{1}{\gamma_1}\Delta J_1(k)&\le\Big(\zeta_3^2(k)+l_5^2\tilde{x}_1^2(k)+l_5^2e_1^2(k)+e_2^2(k)\\&\qquad+\left(\varepsilon_3(k)-w_3\tilde{\phi}_3(k)+d_1(k)\right)^2\Big)-\frac{1}{5}e_1^2(k)\\\Delta J_1(k)&\le\gamma_1\zeta_3^2(k)+\gamma_1l_5^2\tilde{x}_1^2(k)+\gamma_1l_5^2e_1^2(k)+\gamma_1e_2^2(k)\\&\quad+\gamma_1\left(\varepsilon_3(k)-w_3\tilde{\phi}_3(k)+d_1(k)\right)^2-\frac{\gamma_1}{5}e_1^2(k)\\&\le\gamma_1l_5^2\tilde{x}_1^2(k)+\gamma_1l_5^2e_1^2(k)+\gamma_1e_2^2(k)+\gamma_1\zeta_3^2(k)\\&\quad+\gamma_1(\varepsilon_{3\mathrm{m}}+w_{3\mathrm{m}}\tilde{\phi}_{3\mathrm{m}}+d_{1\mathrm{m}})^2-\frac{\gamma_1}{5}e_1^2(k).\end{aligned}\tag{B.16}$$

Take the second term from (B.12), substitute (61), invoke the Cauchy–Schwarz inequality, and simplify

$$\Delta J_2(k) \le 3l_6^2 e_2^2(k) + 3g_{2\max}^2\zeta_4^2(k) + \gamma_2(d_{2\mathrm{m}}+g_{2\max}\varepsilon_{4\mathrm{m}}+g_{2\max}w_{4\mathrm{m}}\tilde{\phi}_{4\mathrm{m}})^2 - e_2^2(k). \tag{B.17}$$

Take the third term, substitute (B.12), invoke the Cauchy–Schwarz inequality, and simplify

$$\Delta J_3(k) \le -\gamma_3\left(1-\alpha_1\left\|\hat{\phi}_1(k)\right\|^2\right)\left(\hat{w}_1(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2 + 2\gamma_3(w_{1\mathrm{m}}\hat{\phi}_{1\mathrm{m}})^2 + 2\gamma_3 l_4^2\tilde{y}^2(k) - \gamma_3\zeta_1^2(k). \tag{B.18}$$

Take the fourth term from (B.12), substitute (37), invoke the Cauchy–Schwarz inequality, and simplify

$$\begin{aligned}\Delta J_4(k) \le & -\gamma_4\left(1-\alpha_2\left\|\hat{\phi}_2(k)\right\|^2\right)\\ & \times\left(\hat{\overline{Q}}(k)+\beta^{N+1}p(k)-\beta\hat{\overline{Q}}(k-1)\right)^2\\ & -\gamma_4\zeta_2^2(k)+2\gamma_4\beta^2\zeta_2^2(k-1)\\ & +2\gamma_4\left(w_{2\mathrm{m}}\hat{\phi}_{2\mathrm{m}}(1+\beta)+\beta^{N+1}\right)^2.\end{aligned} \tag{B.19}$$

Take the fifth term from (B.12), substitute (57), invoke the Cauchy–Schwarz inequality, and simplify

$$\begin{aligned}\Delta J_5(k) \le & -\gamma_5\left(1-\alpha_3\left\|\hat{\phi}_3(k)\right\|^2\right)\\ & \times\left(\hat{\overline{Q}}(k)+\hat{w}_3^T(k)\hat{\phi}_3(k)\right)^2+2\gamma_5\zeta_2^2(k)\\ & +2\gamma_5(w_{2\mathrm{m}}\hat{\phi}_{2\mathrm{m}}+w_{3\mathrm{m}}\hat{\phi}_{3\mathrm{m}})^2-\gamma_5\zeta_3^2(k).\end{aligned} \tag{B.20}$$

Take the sixth term from (B.12), substitute (68), invoke the Cauchy–Schwarz inequality, and simplify

$$\begin{aligned}\Delta J_6(k) = & -\gamma_6\left(1-\alpha_4\left\|\hat{\phi}_4(k)\right\|^2\right)\\ & \times\left(\hat{w}_4^T(k)\hat{\phi}_4(k)+\hat{\overline{Q}}(k)\right)^2\\ & +2\gamma_6(w_{4\mathrm{m}}\hat{\phi}_{\mathrm{m}}+w_{2\mathrm{m}}\hat{\phi}_{2\mathrm{m}})^2+2\gamma_6\zeta_2^2(k)-\gamma_6\zeta_4^2(k).\end{aligned} \tag{B.21}$$

Take the seventh term from (B.12), set $\gamma_7 = 2\gamma_4\beta^2$

$$\Delta J_7(k) = 2\gamma_4\beta^2\zeta_2^2(k) - 2\gamma_4\beta^2\zeta_2^2(k-1). \tag{B.22}$$

Take the eighth term from (B.12), substitute (21), invoke the Cauchy–Schwarz inequality, and simplify

$$\Delta J_8(k) \le \gamma_8 l_2^2\tilde{y}^2(k) + \gamma_8\tilde{x}_2^2(k) + \gamma_8(w_{3\mathrm{m}}\phi_{3\mathrm{m}}+f_{10}+\varepsilon_{3\mathrm{m}}+d_{1\mathrm{m}})^2 - \frac{\gamma_8}{3}\tilde{x}_1^2(k). \tag{B.23}$$

Take the ninth term from (B.12), substitute (22), invoke the Cauchy–Schwarz inequality, and simplify

$$\Delta J_9(k) \le \gamma_9\left(f_{20}+(g_{20}+g_{2\max})w_{4\mathrm{m}}\hat{\phi}_{4\mathrm{m}}+f_{2\max}+d_{2\mathrm{m}}\right)^2 + \gamma_9(g_{20}+g_{2\max})\zeta_4(k) + \gamma_9 l_3^2\tilde{y}^2(k) - \frac{\gamma_9}{3}\tilde{x}_2^2(k). \tag{B.24}$$

Take the tenth term from (B.12), substitute (23), invoke the Cauchy–Schwarz inequality, and simplify

$$\Delta J_{10}(k) \le \gamma_{10}\zeta_1^2(k) + \gamma_{10}l_1^2\tilde{y}(k) + \gamma_{10}(w_{1\mathrm{m}}\tilde{\phi}_{1\mathrm{m}}+\varepsilon_{1\mathrm{m}}) - \frac{\gamma_{10}}{3}\tilde{y}^2(k). \tag{B.25}$$

Take the first term in the summation from (B.13) and replace (96)

$$\begin{aligned}\Delta J_{11}(k) &= \frac{\gamma_{11}}{3}\left(\|e_4(k+1)\|^2-\|e_4(k)\|^2\right)\\ &\le -\frac{\gamma_{11}}{3}\left(1-3l_7^2\right)\|e_4(k)\|^2+\gamma_{11}g_{4\max}^2\|\zeta_a(k)\|^2\\ &\quad+\gamma_{11}\|d_a(k)\|^2.\end{aligned} \tag{B.26}$$

Take the second term in the summation from (B.13) and replace (101)

$$\begin{aligned}\Delta J_{12}(k) = & \frac{\gamma_{12}}{\alpha_a}tr\left(\tilde{w}_a^T(k+1)\tilde{w}_a(k+1)-\tilde{w}_a^T(k)\tilde{w}_a(k)\right)\\ \le & \gamma_{12}\Bigg(-g_{4\min}\|\zeta_a(k)\|^2\\ & +\frac{1-\alpha_a\|\phi_a(k)\|^2 g_{4\min}}{g_{4\min}-\alpha_a\|\phi_a(k)\|^2 g_{4\max}^2}\\ & \times\|J(k)+d_a(k)\|^2\\ & -\left(g_{4\min}-\alpha_a\|\phi_a(k)\|^2 g_{4\max}^2\right)\\ & \times\left(\zeta_a(k)+\frac{I-\alpha_\alpha\|\phi_a(k)\|^2 g_{4\min}}{g_{4\min}-\alpha_a\|\phi_a(k)\|^2 g_{4\max}^2}\right)\Bigg).\end{aligned} \tag{B.27}$$

Define the following:

$$\gamma_{12} = \gamma'_{12}\gamma''_{12} = \gamma'_{12}\frac{1-\alpha_a\|\phi_a(k)\|^2 g_{4\min}}{g_{4\min}-\alpha_a\|\phi_a(k)\|^2 g_{4\max}^2} \le \frac{1}{2}\gamma'_{12}. \tag{B.28}$$

Rewrite (B.27) using (B.28)

$$\begin{aligned}\Delta J_{12} \le & -\gamma_{12}\left(g_{4\min}-\alpha_a\|\phi_a(k)\|^2 g_{4\max}^2\right)\\ & \times\left(\zeta_a(k)+\frac{I-\alpha_\alpha\|\phi_a(k)\|^2 g_{4\min}}{g_{4\min}-\alpha_a\|\phi_a(k)\|^2 g_{4\max}^2}\right)\\ & +\gamma'_{12}\|\zeta_c(k)\|^2+\gamma'_{12}\|d_a(k)\|^2-\gamma_{12}g_{4\min}\|\zeta_a(k)\|^2.\end{aligned} \tag{B.29}$$

Take the third term from (B.13) from the summation and replace (92)

$$\begin{aligned}\Delta J_{13} &= \frac{\gamma_{12}}{\alpha_c} tr\left(\tilde{w}_c^T(k+1)\tilde{w}_c(k+1) - \tilde{w}_c^T(k)\tilde{w}_c(k)\right)\\ &\le -\gamma_{13}\left(1-\alpha_c\gamma^2\left\|\phi_a(k)\right\|^2\right)e_c^2(k) - \gamma_{13}\gamma^2\zeta_c^2(k)\\ &\quad + \frac{\gamma_{13}}{4}\zeta_c^2(k-1) + \gamma_{13}Q_{\max}\left\|e(k)\right\|^2\\ &\quad + \frac{\gamma_{13}}{8}R_{\max}\left\|\zeta_a(k)\right\|^2 + \frac{\gamma_{13}}{8}R_{\max}\left\|w_a^T\phi_a(k)\right\|^2\\ &\quad + \gamma_{13}\varepsilon_{cm}^2. \end{aligned} \tag{B.30}$$

Take the fourth and final term from (B.13) from the summation and replace

$$\Delta J_{14} = \gamma_{14}\left(\left\|\zeta_c(k)\right\|^2 - \left\|\zeta_c(k-1)\right\|^2\right). \tag{B.31}$$

Combine (B.16)–(B.31) and simplify to get the first difference of the Lyapunov function

$$\begin{aligned}\Delta J \le & -\left(\frac{\gamma_1}{5}-\gamma_1 l_5^2\right)e_1^2(k) - \left(\frac{\gamma_2}{3}-\gamma_1-\gamma_2 l_6^2\right)e_2^2(k)\\ & -(\gamma_3-\gamma_{10})\zeta_1^2(k) - \left(\frac{\gamma_9}{3}-\gamma_8\right)\tilde{x}_2^2(k) - (\gamma_5-\gamma_1)\zeta_3^2(k)\\ & -\left(\gamma_6-\gamma_2 g_{2\max}^2-\gamma_9(g_{20}+g_{2\max})\right)\zeta_4^2(k)\\ & -\left(\frac{\gamma_8}{3}-\gamma_1 l_5^2\right)\tilde{x}_1^2(k) - (\gamma_4-2\gamma_5-2\gamma_6-2\gamma_4\beta^2)\zeta_2^2(k)\\ & -\left(\frac{\gamma_{10}}{3}-2\gamma_3 l_4^2-\gamma_8 l_2^2-\gamma_9 l_3^2-\gamma_{10}l_1^2\right)\tilde{y}^2(k)\\ & -\gamma_3\left(1-\alpha_1\left\|\hat{\phi}_1(k)\right\|^2\right)\left(\hat{w}_1(k)\hat{\phi}_1(k)+l_4\tilde{y}(k)\right)^2\\ & -\gamma_4\left(1-\alpha_2\left\|\hat{\phi}_2(k)\right\|^2\right)\\ & \times\left(\hat{\overline{Q}}(k)+\beta^{N+1}p(k)-\beta\hat{\overline{Q}}(k-1)\right)^2\\ & -\gamma_5\left(1-\alpha_3\left\|\hat{\phi}_3(k)\right\|^2\right)\left(\hat{\overline{Q}}(k)+\hat{w}_3^T(k)\hat{\phi}_3(k)\right)^2\\ & -\gamma_6\left(1-\alpha_4\left\|\hat{\phi}_4(k)\right\|^2\right)\left(\hat{w}_4^T(k)\hat{\phi}_4(k)+\hat{\overline{Q}}(k)\right)^2\\ & -\frac{\gamma_{11}}{3}\left(1-3l_7^2\right)\left\|e_4(k)\right\|^2+\gamma_{11}g_{4\max}^2\left\|\zeta_a(k)\right\|^2\\ & +\gamma_{11}\left\|d_a(k)\right\|^2-\gamma_{12}g_{4\min}\left\|\zeta_a(k)\right\|^2\\ & +\gamma_{12}\frac{1-\alpha_a\left\|\phi_a(k)\right\|^2 g_{4\min}}{g_{4\min}-\alpha_a\left\|\phi_a(k)\right\|^2 g_{4\max}^2}\left\|J(k)+d_a(k)\right\|^2\\ & -\gamma_{13}\left(1-\alpha_c\gamma^2\left\|\phi_a(k)\right\|^2\right)e_c^2(k)\\ & -\gamma_{12}\left(g_{4\min}-\alpha_a\left\|\phi_a(k)\right\|^2 g_{4\max}^2\right)\\ & \times\left(\zeta_a(k)+\frac{I-\alpha_\alpha\left\|\phi_a(k)\right\|^2 g_{4\min}}{g_{4\min}-\alpha_a\left\|\phi_a(k)\right\|^2 g_{4\max}^2}\right)\\ & -\gamma_{13}\gamma^2\zeta_c^2(k)+\frac{\gamma_{13}}{4}\zeta_c^2(k-1)+\gamma_{13}Q_{\max}\left\|e(k)\right\|^2\\ & +\frac{\gamma_{13}}{8}R_{\max}\left\|\zeta_a(k)\right\|^2+\gamma_{14}\left(\left\|\zeta_c(k)\right\|^2-\left\|\zeta_c(k-1)\right\|^2\right)\\ & +D_m^2 \end{aligned} \tag{B.32}$$

where

$$\begin{aligned}D_{\mathrm{m}}^2 =& \gamma_1(\varepsilon_{3\mathrm{m}}+w_{3\mathrm{m}}\tilde{\phi}_{3\mathrm{m}}+d_{1\mathrm{m}})^2\\ & +\gamma_2(d_{2\mathrm{m}}+g_{2\max}\varepsilon_{4\mathrm{m}}+g_{2\max}w_{4\mathrm{m}}\tilde{\phi}_{4\mathrm{m}})^2\\ & +2\gamma_3(w_{1\mathrm{m}}\hat{\phi}_{1\mathrm{m}})^3+2\gamma_4\left(w_{2\mathrm{m}}\hat{\phi}_{2\mathrm{m}}(1+\beta)+\beta^{N+1}\right)^2\\ & +2\gamma_5(w_{2\mathrm{m}}\hat{\phi}_{2\mathrm{m}}+w_{3\mathrm{m}}\hat{\phi}_{3\mathrm{m}})^2\\ & +2\gamma_6(w_{4\mathrm{m}}\hat{\phi}_{\mathrm{m}}+w_{2\mathrm{m}}\hat{\phi}_{2\mathrm{m}})^2\\ & +\gamma_8(w_{3\mathrm{m}}\phi_{3\mathrm{m}}+f_{10}+\varepsilon_{3\mathrm{m}}+d_{1\mathrm{m}})^2\\ & +\gamma_9\left(f_{20}+(g_{20}+g_{2\max})w_{4\mathrm{m}}\hat{\phi}_{4\mathrm{m}}+f_{2\max}+d_{2\mathrm{m}}\right)^2\\ & +\gamma_{10}(w_{1\mathrm{m}}\tilde{\phi}_{1\mathrm{m}}+\varepsilon_{1\mathrm{m}})^2+\frac{\gamma'_{12}n}{2}\left\|d_a(k)\right\|^2\\ & +\left(\frac{\gamma_{13}}{4}+\frac{\gamma'_{12}n}{2}\right)J_m^2+\frac{\gamma_{13}}{6}R_{\max}\left\|w_a^T\phi_a(k)\right\|^2\\ & +\gamma_{11}\left\|d_a(k)\right\|^2+\gamma_{13}\varepsilon_{cm}^2. \end{aligned} \tag{B.33}$$

Select

$$\begin{aligned}&\gamma_1 > 5\gamma_1 l_5^2 \quad \gamma_2 > 3\gamma_1 + 3\gamma_2 l_6^2 \quad \gamma_3 > \gamma_{10}\\ &\gamma_4 > 2\gamma_5 + 2\gamma_6 + 2\gamma_4\beta^2 \quad \gamma_5 > \gamma_1\\ &\gamma_6 > \gamma_2 g_{2\max}^2 + \gamma_9(g_{20}+g_{2\max}) \quad \gamma_7 = 2\gamma_4\beta^2\\ &\gamma_8 > 3\gamma_1 l_5^2 \quad \gamma_9 > 3\gamma_8\\ &\gamma_{10} > 6\gamma_3 l_4^2 + 3\gamma_8 l_2^2 + 3\gamma_9 l_3^2 + 3\gamma_{10} l_1^2. \end{aligned} \tag{B.34}$$

This implies $\Delta J(k) < 0$ as long as (71)–(81) and (102)–(104) hold *and* any *one* of the following holds:

$$\begin{aligned}&|e_1(k)| > \frac{D_m}{\sqrt{\frac{\gamma_1}{5}-\gamma_1 l_5^2}} \quad |e_2(k)| > \frac{D_m}{\sqrt{\frac{\gamma_2}{3}-\gamma_1-\gamma_2 l_6^2}}\\ &|\zeta_1(k)| > \frac{D_m}{\sqrt{\gamma_3-\gamma_{10}}}\\ &|\zeta_3(k)| > \frac{D_m}{\sqrt{\gamma_5-\gamma_1}} \quad |\zeta_2(k)| > \frac{D_m}{\sqrt{\gamma_4-2\gamma_5-2\gamma_6-2\gamma_4\beta^2}}\\ &|\zeta_4(k)| > \frac{D_m}{\sqrt{\gamma_6-\gamma_2 g_{2\max}^2-\gamma_9(g_{20}+g_{2\max})}}\\ &|\tilde{x}_1(k)| > \frac{D_m}{\sqrt{\frac{\gamma_8}{3}-\gamma_1 l_5^2}} \quad |\tilde{x}_2(k)| > \frac{D_m}{\sqrt{\frac{\gamma_9}{3}-\gamma_8}}\\ &|\tilde{y}(k)| > \frac{D_m}{\sqrt{\frac{\gamma_{10}}{3}-2\gamma_3 l_4^2-\gamma_8 l_2^2-\gamma_9 l_3^2-\gamma_{10}l_1^2}}\\ &\|e_4(k)\| > \frac{2\sqrt{3}D_m}{\sqrt{4\gamma_{11}(1-3l_7^2)-3\gamma_3 Q_{\max}}}\\ &\|\zeta_a(k)\| > \frac{2\sqrt{2}D_m}{\sqrt{8\gamma_{12}g_{4\min}-8\gamma_{11}g_{4\max}^2-\gamma_{13}R_{\max}}}\\ &\|\zeta_c(k)\| > \frac{D_m}{\sqrt{\gamma_{13}\gamma^2-\gamma'_{12}-\gamma_{14}}}. \end{aligned} \tag{B.35}$$

REFERENCES

[1] M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*. New York: Wiley, 1995.

[2] S. Jagannathan, "Control of a class of nonlinear systems using multilayered neural networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 5, pp. 1113–1120, Sep. 2001.
[3] F. C. Chen and H. K. Khalil, "Adaptive control of a class of nonlinear discrete-time systems using neural networks," *IEEE Trans. Autom. Control*, vol. 40, no. 5, pp. 791–801, May 1995.
[4] J. Si, in *NSF Workshop on Learning and Approximate Dynamic Programming*, Playacar, Mexico, 2002.
[5] P. J. Werbos, *Neurocontrol and Supervised Learning: An Overview and Evaluation*. New York: Van Nostrand Reinhold, 1992.
[6] J. J. Murray, C. Cox, G. G. Lendaris, and R. Saeks, "Adaptive dynamic programming," *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, vol. 32, no. 2, pp. 140–153, May 2002.
[7] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Balmont, MA: Athena Scientific, 1996.
[8] J. Si and Y. T. Wang, "On-line learning control by association and reinforcement," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 264–276, Mar. 2001.
[9] X. Lin and S. N. Balakrishnan, "Convergence analysis of adaptive critic based optimal control," in *Proc. Amer. Control Conf.*, 2000, vol. 12, pp. 264–276.
[10] F. L. Lewis, S. Jagannathan, and A. Yesilderek, *Neural Network Control of Robot Manipulators and Nonlinear Systems*. London, U.K.: Taylor & Francis, 1999.
[11] J. Vance, P. He, S. Jagannathan, and J. Drallmeier, "Neural network-based output feedback controller for lean operation of spark ignition engine," in *Amer. Control Conf.*, Portland, OR, 2006, pp. 8–13.
[12] N. Hovakimyan, F. Nardi, A. Calise, and N. Kim, "Adaptive output feedback control of uncertain nonlinear systems using single-hidden-layer neural networks," *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1420–1431, Nov. 2002.
[13] A. N. Atassi and H. K. Khalil, "A separation principle for the stabilization of a class of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 44, no. 9, pp. 1672–1687, Sep. 2003.
[14] B. Igelruk and Y. H. Pao, "Stochastic choice of basis functions in adaptive function approximation and the functional-link net," *IEEE Trans. Neural Netw.*, vol. 6, no. 6, pp. 1320–1329, Nov. 1995.
[15] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-Time Systems*. London, U.K.: Taylor & Francis, 2006.
[16] C. S. Daw, C. E. A. Finney, M. B. Kennel, and F. T. Connolly, "Observing and modeling nonlinear dynamics in an internal combustion engine," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 57, pp. 2811–2819, 1998.
[17] R. W. Sutton and J. A. Drallmeier, "Development of nonlinear cyclic dispersion in spark ignition engines under the influence of high levels of EGR," in *Proc. Central States Section Combustion Inst.*, Indianapolis, IN, 2000, pp. 175–180.
[18] J. B. Vance, A. Singh, B. Kaul, S. Jagannathan, and J. Drallmeier, "Neural network controller development and implementation for spark ignition engines with high EGR levels," *IEEE Tans. Neural Netw.*, vol. 18, no. 4, pp. 1083–1100, Jul. 2007.
[19] P. He and S. Jagannathan, "Reinforcement-based neuro-output feedback control of discrete-time systems with input constraints," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 35, no. 1, pp. 150–154, Feb. 2005.
[20] E. Yang, Y. Masuko, and T. Mita, "Dual-controller approach to three-dimensional autonomous formation control," *J. Guid. Control Dyn.*, vol. 27, no. 3, May–Jun. 2004.
[21] K. S. Hwang and H. J. Chao, "Adaptive reinforcement learning system for linearization control," *IEEE Trans. Ind. Electron.*, vol. 47, no. 5, pp. 1185–1188, Oct. 2000.
[22] P. Shih, B. Kaul, S. Jagannathan, and J. Drallmeier, "Near optimal output feedback control of nonlinear discrete-time systems in nonstrict feedback form with application to spark ignition engines," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2007, pp. 396–401.
[23] S. S. Ge, G. Y. Li, and T. H. Lee, "Adaptive NN control for a class of strict-feedback discrete-time nonlinear systems," *Automatica*, vol. 39, pp. 807–819, 2003.

Peter Shih was born in Taiwan R.O.C on August 5, 1980. He received the B.S. degree in biomedical engineering from Washington University in St. Louis, St. Louis, MO, in 2002 and the M.S. degree in computer engineering from University of Missouri—Rolla, Rolla, in May 2007.

Currently, he is a Software Engineer in Maryland.

Brian C. Kaul was born on December 9, 1978 in St. Louis County, MO. He received the B.S. (*summa cum laude*), M.S., and Ph.D. degrees in mechanical engineering from University of Missouri—Rolla, Rolla, in 2001, 2003, and 2008, respectively.

S. Jagannathan (S'89–M'89–SM'99) received the B.S. degree in electrical engineering from College of Engineering, Guindy at Anna University, Madras, India, in 1987, the M.S. degree in electrical engineering from the University of Saskatchewan, Saskatoon, SK, Canada, in 1989, and the Ph.D. degree in electrical engineering from the University of Texas, San Antonio, in 1994.

From 1986 to 1987, he was a Junior Engineer at Engineers India Limited, New Delhi, India; from 1990 to 1991, a Research Associate and Instructor at the University of Manitoba, Winnipeg, MB, Canada; and from 1994 to 1998, a Consultant at Systems and Controls Research Division, Caterpillar Inc., Peoria, IL. From 1998 to 2001, he was at the University of Texas, San Antonio, and since September 2001, he has been at the University of Missouri—Rolla, Rolla, where he is currently a Rutledge-Emerson Distinguished Professor and Site Director for the National Science Foundation (NSF) Industry/University Cooperative Research Center on Intelligent Maintenance Systems. He has coauthored more than 190 refereed conference and juried journal articles and several book chapters and three books entitled *Neural Network Control of Robot Manipulators and Nonlinear Systems*, (London, U.K.: Taylor & Francis, 1999), *Discrete-Time Neural Network Control of Nonlinear Discrete-Time Systems* (Boca Raton, FL: CRC Press, 2006), and *Wireless Ad Hoc and Sensor Networks: Performance, Protocols and Control* (Boca Raton, FL: CRC Press, 2007). He holds 17 patents with several pending. His research interests include adaptive and neural network control, computer/communication/sensor networks, prognostics, and autonomous systems/robotics.

James A. Drallmeier received the Ph.D. degree in mechanical engineering from the University of Illinois at Urbana-Champaign, Urbana, in 1989.

He then joined the faculty of the University of Missouri—Rolla, Rolla, where he is currently Professor of Mechanical Engineering. He operates the Spray Dynamics and Internal Combustion Engine Laboratories. His research interests lie in the fields of combustion, laser-based measurement systems, and internal combustion engines. Current research includes studying two-phase flows, particularly sprays and thin shear-driven films, and the dynamics of highly strained, dilute, intermittent combustion. He has been involved in developing and using laser-based diagnostic techniques for measuring spray and thin film dynamics over the past two decades. Additionally, he has been active in studying fuel systems and mixture preparation for advanced engine designs.