REGULAR PAPERS
Quaternion-Valued Echo State Networks .................................................. Y. Xia, C. Jahanchahi, and D. P. Mandic 663
Successive Overrelaxation for Laplacian Support Vector Machine .................................................. Z. Qi, Y. Tian, and Y. Shi 674
Adaptive Optimal Control of Highly Dissipative Nonlinear Spatially Distributed Processes With Neuro-Dynamic Programming ................................................................................ B. Luo, H.-N. Wu, and H.-X. Li 684
Convolutive Bounded Component Analysis Algorithms for Independent and Dependent Source Separation ............ H. A. Inan and A. T. Erdogan 697
Gaussian Kernel Width Optimization for Sparse Bayesian Learning .................................................. Y. Mohsenzadeh and H. Sheikhzadeh 709
Multiclass Semisupervised Learning Based Upon Kernel Spectral Clustering ............................................ S. Mehrkanoon, C. Alzate, R. Mall, R. Langone, and J. A. K. Suykens 720
A Universal Concept Based on Cellular Neural Networks for Ultrafast and Flexible Solving of Differential Equations ............................................................................................................ J. C. Chedjou and K. Kyamakya 749
Coupled Attribute Similarity Learning on Categorical Data .................................................. C. Wang, X. Dong, F. Zhou, L. Cao, and C.-H. Chi 781
Topology-Based Clustering Using Polar Self-Organizing Map .................................................. L. Xu, T. W. S. Chow, and E. W. M. Ma 798
Robust Consensus Tracking Control for Multiagent Systems With Initial State Shifts, Disturbances, and Switching Topologies ........................................................................................................ D. Meng, Y. Jia, and J. Du 809
$L_1$-Norm Low-Rank Matrix Factorization by Variational Bayesian Method .................................................. Q. Zhao, D. Meng, Z. Xu, W. Zuo, and Y. Yan 825
Nonlinear Model Predictive Control Based on Collective Neurodynamic Optimization .................................................. Z. Yan and J. Wang 840
Infinite Horizon Self-Learning Optimal Control of Nonaffine Discrete-Time Nonlinear Systems .................................................. Q. Wei, D. Liu, and X. Yang 866

BRIEF PAPERS
A Bootstrap Based Neyman–Pearson Test for Identifying Variable Importance .................................................. G. Ditzler, R. Polikar, and G. Rosen 880

ANNOUNCEMENTS
Call For Papers: WCCI 2016 ........................................................................................................ 887
Adaptive Optimal Control of Highly Dissipative Nonlinear Spatially Distributed Processes With Neuro-Dynamic Programming

Biao Luo, Huai-Ning Wu, and Han-Xiong Li, Fellow, IEEE

Abstract—Highly dissipative nonlinear partial differential equations (PDEs) are widely employed to describe the system dynamics of industrial spatially distributed processes (SDPs). In this paper, we consider the optimal control problem of the general highly dissipative SDPs, and propose an adaptive optimal control approach based on neuro-dynamic programming (NDP). Initially, Karhunen-Loève decomposition is employed to compute empirical eigenfunctions (EEFs) of the SDP based on the method of snapshots. These EEFs together with singular perturbation technique are then used to obtain a finite-dimensional slow subsystem of ordinary differential equations that accurately describes the dominant dynamics of the PDE system. Subsequently, the optimal control problem is reformulated on the basis of the slow subsystem, which is further converted to solve a Hamilton–Jacobi–Bellman (HJB) equation. HJB equation is a nonlinear PDE that has proven to be impossible to solve analytically. Thus, an adaptive optimal control method is developed via NDP that solves the HJB equation online using neural network (NN) for approximating the value function; and an online NN weight tuning law is proposed without requiring an initial stabilizing control policy. Moreover, by involving the NN estimation error, we prove that the original closed-loop PDE system with the adaptive optimal control policy is semiglobally uniformly ultimately bounded. Finally, the developed method is tested on a nonlinear diffusion-convection-reaction process and applied to a temperature cooling fin of high-speed aerospace vehicle, and the achieved results show its effectiveness.

Index Terms—Adaptive optimal control, empirical eigenfunction (EEF), highly dissipative partial differential equations (PDEs), neuro-dynamic programming (NDP), spatially distributed processes (SDPs).

I. INTRODUCTION

For many industrial systems, significant spatial variations are widely existed owing to the underlying physical phenomena, such as diffusion, convection, phase-dispersion, vibration, flow, and so on. These spatially distributed processes (SDPs) are naturally described by a set of nonlinear partial differential equations (PDEs) with homogeneous or mixed boundary conditions. Typical examples include the rapid thermal chemical vapor deposition [1] in microelectronics manufacturing, the temperature profile [2] of high-speed aerospace vehicle, catalytic packed-bed reactors [3], [4] used to convert methanol to formaldehyde, the Czochralski crystallization [5] of high-purity crystals, tokamak device [6] for nuclear fusion, an intangible doughnut-shaped bottle created by magnetic lines is used to confine the high-temperature plasma, wavy behavior [7] of excitable media in biology and chemistry, as well as several fluid dynamic systems [8], [9], and so on. Most of these SDPs are described with highly dissipative PDEs, which have an important feature that their long-term dynamic behavior is characterized by a finite number of degrees of freedom. This means that their dominant dynamics can be described by a few modes, which is a promising feature that makes it possible for extending the control theories and methods of ordinary differential equation (ODE) systems to the highly dissipative PDE systems.

One popular approach to the controller design of highly dissipative PDEs involves the application of model reduction techniques to derive an ODE system that describes the dynamics of the dominant (slow) modes of the PDE system, which are subsequently used as the basis for the synthesis of finite-dimensional controllers [10], [11]. Following this thought, many control approaches have been developed in the last few decades. Approximate inertial manifold [12] were employed to derive lower order ODE systems of the parabolic PDE systems [13], which were further used for the synthesis of finite-dimensional nonlinear feedback controllers that enforce stability and output tracking in the closed-loop system. In [14] and [15], adaptive model reduction technique based on Karhunen-Loève decomposition (KLD) was applied to derive the finite-dimensional ODE model of highly dissipative PDE systems, which was then used to design a nonlinear controller via state feedback [14] and output feedback [15], [16]. Optimal control of PDE systems refers to a class of methods that can be applied to synthesize a control policy, which results in best possible behavior with respect to the prescribed performance criterion. In [17], based on the empirical eigenfunctions (EEFs) computed via KLD, Bendersky and Christofides discretized the infinite-dimensional optimization problem of transport-reaction processes to a set of
low-dimensional nonlinear programming problems, which are then iteratively solved with successive quadratic programming (SQP); this method was further extended to highly dissipative PDE systems in [4], where an efficient nonlinear programming approach was presented. In [18], the PDE system was formulated as an ODE system via spatial discretization and then an optimal feedback control approach was proposed. By taking a fixed structure for the control law based on Ritz and extended Ritz approximation schemes, a mathematical programming problem was formulated and solved with SQP in [19]. In [2], a dual heuristic programming (DHP) approach was introduced to solve the temperature profile control problem of high-speed aerospace vehicle. Xu et al. [20] considered a class of bilinear parabolic PDE systems and suggested a sequential linear quadratic (LQ) regulator approach based on an iterative scheme, and in [6] this approach was successfully applied to the optimal tracking control of current profile in tokamaks. Luo and Wu [21] considered a class of parabolic PDE systems with nonlinear spatial differential operator, and developed neural network (NN)-based optimal control method by solving the Hamilton–Jacobi–Bellman (HJB) equation. However, most of these optimal control approaches are designed off-line and thus fragile to changing dynamics of a system during operation. To the best of our knowledge, the adaptive optimal control problem of general highly dissipative PDE systems, based on online design with neuro-dynamic programming (NDP), was rarely studied.

NDP is one class of reinforcement learning (RL) that have been widely investigated as a machine learning technique in artificial intelligence community [22]–[24]. NDP solves the dynamic programming problem forward-in-time using NN as value function approximator, and then avoids the so-called curse of dimensionality. In recent years, some NDP approaches have been introduced to synthesize controller for ODE systems [25]–[27]. For discrete-time (DT) optimal control problems, several NDP methods [28]–[33] were developed for solving the DT HJB equation in an offline iterative manner. Al-Tamimi et al. [28] suggested a heuristic dynamic programming (HDP) approach for solving the optimal control problem of nonlinear DT systems. In [29], the nonlinear DT HJB equation was converted to a sequence of linear generalized HJB equations. In [30], the infinite-time optimal tracking control problem with a new type of performance index was solved using the greedy HDP iteration algorithm. The optimal control problem for a class of nonlinear DT systems with control constraints [31] and time delays [34], respectively, were also solved with HDP methods. Fu et al. [35] investigated the adaptive learning and control for multiple-input-multiple-output system based on approximate dynamic programming (ADP). A finite-horizon optimal control problem was studied in [32] by introducing a $\varepsilon$-error bound; and finite-time problem with control constraint was considered by proposing a DHP scheme [36] with single NN. These works assumed that there are no NN reconstruction errors. To involve the effects of NN approximation errors, a neural HDP method in [37] was applied to learn state and output feedback adaptive critic control policy of nonlinear DT affine systems with disturbances. Dierks and Jagannathan [38] proposed a time-based ADP, which is an online control method without using value and policy iterations (PIs). Liu et al. [39]–[41] presented globalized DHP algorithms using three NNs for estimating system dynamics, cost function and its derivatives, and control policy, where model NN construction error was considered. For continuous-time (CT) optimal control problems, Beard et al. [42] applied Galerkin approximation to solve generalized HJB equation, and in [43] this method was extended to constrained control problem by introducing a nonquadratic performance functional and using NN to approximate cost function. Doya [44] discussed employing appropriate estimators for approximating value function such that the temporal difference error is minimized in RL approach. Murray et al. [45] suggested two (PI) algorithms that avoid the necessity of knowing the internal system dynamics. Vrabie et al. [46] extended their result and proposed a new PI algorithm to solve the LQ regulation problem online along a single state trajectory. A nonlinear version of this algorithm was presented in [47] using NN approximator. Vanvoudakis and Lewis [48] also gave a so-called synchronous PI algorithm, which tunes synchronously the weights of both actor and critic NNs for the optimal control problem nonlinear ODE systems. Luo et al. [49], [50] proposed off-policy RL approaches for optimal control and $H_\infty$ control design of CT systems. The off-policy RL is a kind of data-based method, which learns the control policy from the data generated with arbitrary policies rather than the evaluating policy. Liu et al. [51] proposed a novel online learning optimal control approach to deal with the decentralized stabilization problem for a class of CT nonlinear interconnected systems.

However, due to the infinite-dimensional nature of the PDE systems, it is still difficult to directly use these control design methods of ODE systems for the PDE systems. In this paper, we consider the optimal control problem of general highly dissipative PDE systems, and develop an adaptive optimal control method based on NDP.

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

1) An adaptive optimal control method based on NDP is proposed for nonlinear highly dissipative PDE systems. To the best our knowledge, NDP is rarely used for the control design of PDE systems and it still remains an open problem.

2) The developed adaptive optimal control method can solve the HJB equation online without requiring an initial stabilizing control law.

3) In the presence of the NN approximate error, we prove that the developed adaptive optimal control policy can ensure that the original closed-loop PDE system is semiglobally uniformly ultimately bounded (SGUUB), with singular perturbation (SP) theory.

4) To show the effectiveness of the developed adaptive optimal control method, simulation studies are conducted on a general nonlinear convection-diffusion-reaction process of four cases, and then applied to a complex temperature cooling fin of high-speed aerospace vehicle.
The rest of this paper is arranged as follows. The problem description is given in Section II. KLD and SP technique are used for model reduction in Section III. In Section IV, an adaptive optimal control method based on NDP is proposed and the stability of the closed-loop PDE system is proved. Simulation studies are conducted in Section V and a brief conclusion is given in Section VI.

Notation: \( \mathbb{R}, \mathbb{R}^n, \) and \( \mathbb{R}^{n \times m} \) are the set of real numbers, the \( n \)-dimensional Euclidean space and the set of all real matrices, respectively. \( \| \cdot \|, \| \cdot \|_1, \) and \( \langle \cdot, \cdot \rangle_{\mathbb{R}^n} \) denote the absolute value for scalars, Euclidean norm, and inner product for vectors, respectively. Let \( \mathbb{R}^n \) be the vector space of infinite sequences \( a \triangleq [a_1, \ldots, a_\infty]^T \) of real numbers equipped with the norm \( \|a\|_{\mathbb{R}^\infty} \triangleq \sqrt{\sum_1^\infty a_i^2} \), which a natural generalization of space \( \mathbb{R}^n \). The superscript \( T \) is used for the transpose and \( I \) denotes the identity matrix of appropriate dimension, and \( \nabla \triangleq \partial/\partial x \) denotes a gradient operator notation. \( \hat{\sigma}(\cdot) \) and \( \sigma(\cdot) \) denote the maximum singular value and the minimum singular value. For a symmetric matrix \( M, M > (\geq)0 \) means that it is a positive (semipositive) definite matrix. \( \|v\|^2_M \triangleq v^TMv \) for some real vector \( v \) and symmetric matrix \( M > (\geq)0 \) with appropriate dimensions. \( C^1(\Omega) \) is function space on \( \Omega \) with first derivatives are continuous. \( L_2([z_1, z_2], \mathbb{R}^n) \) is an infinite-dimensional Hilbert space of \( n \)-dimensional square integrable vector functions \( w(z) \in L_2([z_1, z_2], \mathbb{R}^n), z \in [z_1, z_2] \subset \mathbb{R} \) equipped with the inner product and norm: \( \langle w_1(z), w_2(z) \rangle \triangleq \int_{z_1}^{z_2} w_1(z)w_2(z)dz \) and \( \|w(z)\|^2_{L_2} \triangleq \langle w_1(z), w_1(z) \rangle^{1/2} \) for \( \forall w_1(z), w_2(z) \in L_2([z_1, z_2], \mathbb{R}^n) \).

II. PROBLEM DESCRIPTION

We consider a general class of CT SDPs, which are described by highly dissipative nonlinear PDEs with the following state-space representation:

\[
\begin{align*}
\frac{\partial y(z,t)}{\partial t} &= L(y, \frac{\partial y}{\partial z}, \frac{\partial^2 y}{\partial z^2}, \ldots, \frac{\partial^{n_0} y}{\partial z^{n_0}}) + \overline{B}(z)u(t) \\
y_h(t) &= \int_{z_1}^{z_2} H(z)y(z,t)dz
\end{align*}
\]

subjected to the mixed-type boundary conditions

\[
q \left( y, \frac{\partial y}{\partial z}, \frac{\partial^2 y}{\partial z^2}, \ldots, \frac{\partial^{n_0-1} y}{\partial z^{n_0-1}} \right) \bigg|_{z = z_1 \text{ or } z_2} = 0
\]

and the initial condition

\[
y(z,0) = y_0(z)
\]

where \( z \in [z_1, z_2] \subset \mathbb{R} \) is the spatial coordinate, \( t \in [0, \infty) \) is the temporal coordinate, \( y(z,t) = [y_1(z,t), \ldots, y_n(z,t)]^T \in \mathbb{R}^n \) is the state, and \( u(t) \in \mathbb{R}^p \) is the manipulated input. \( L \in \mathbb{R}^{n \times n} \) is a sufficiently smooth nonlinear vector function that involves a highly dissipative, possibly nonlinear, spatial differential operator of order \( n_0 + 1 \) (an even number), \( q \) is a sufficiently smooth nonlinear vector function, \( \overline{B}(z) \) and \( H(z) \) are known sufficiently smooth matrix functions of appropriate dimension, and \( y_0(z) \) is a smooth vector function representing the initial state profile.

Remark 1: The highly dissipative nonlinear PDE system (1)–(3) represents a large number of complex industrial SDPs. Representative examples include transport-reaction processes with significant diffusive and dispersive mechanisms that are naturally described by nonlinear parabolic PDEs (such as chemical reactors [52], [53], catalytic rod [4], heat transfer [2], rapid thermal chemical vapor deposition [1], crystal growth processes [5], FitzHugh–Nagumo equation [7], [15], etc.), and several fluid dynamic systems (such as Burgers’ equation [8] for gas dynamics and traffic flow, Kuramoto–Sivashinsky equation [9], Navier–Stokes equations [9], etc.).

In this paper, we consider the optimal control problem of PDE system (1)–(3) with the following infinite-horizon LQ cost functional:

\[
V(y_0) \triangleq \int_0^{+\infty} \left( \|y_h(t)\|^2 + \|u(t)\|^2 \right)dt
\]

where \( R > 0 \). The goal is to find a control policy \( u^*(t) \) such that the above cost functional is minimized, i.e.,

\[
u(t) \triangleq u^*(t) \triangleq \arg \min_u V(y_0(\cdot, t)).
\]

III. MODEL REDUCTION BASED ON KLD AND SP TECHNIQUE

It is known that [14]–[16], [54] the main feature of highly dissipative PDE systems is that they involve a spatial differential operator whose eigenspectrum can be partitioned into a finite-dimensional slow one and an infinite-dimensional stable fast complement. This means that their dominant dynamic behavior can be accurately described by finite-dimensional systems. However, for many real industrial SDPs with nonlinear spatial differential operators [1], [4], [20], [54]–[56], it is impossible to compute their analytic expressions of the eigenvalues and eigenfunctions, and thus, the direct use of basis functions are prohibited to derive finite-dimensional approximations of the PDE system. To overcome this difficulty, we initially compute a set of EEFs (dominant spatial patterns) of the PDE system using KLD based on the method of snapshots. These EEFs together with SP technique will be subsequently applied to obtain a slow subsystem that accurately describes the dominant dynamics of the PDE system.

A. EEFs Computation via KLD

KLD is a popular statistical pattern analysis method for seeking the dominant structures in an ensemble of a high-dimensional process, and obtaining low-dimensional approximate descriptions in many engineering fields. Given an ensemble of data, KLD yields a set of orthogonal EEFs for the representation of the ensemble, as well as a measure of the relative contribution of each EEF to the total energy (mean square fluctuation) of the ensemble. In this sense, the EEFs provide an optimal basis for the truncated series representation, which has a smaller mean square error than a representation by any other basis of the same dimension. In other words, the projection onto the first few EEFs capture most of the energy than any other projections. These properties make the
Algorithm 1 KLD Implementation Procedure

1. Step 1: For the real SDP, collect an ensemble of states \( \{y(z)\}, i = 1, \ldots, M \).
2. Step 2: Compute matrix \( C = (c_{ij})_{M \times M} \) with \( c_{ij} = \int \int \phi_i(z) \phi_j(z) dz \).
3. Step 3: Solve the eigenvalue problem \( CA_i = \lambda_i A_i \), and assume that \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M \).
4. Step 4: Normalize the eigenvector \( A_i \) to satisfy \( \langle A_i, A_i \rangle = \frac{1}{M} \). Then, EEFs can be computed with \( \phi_i = Y(z)A_i, \) where \( Y(z) = [y_1(z) \cdots y_M(z)] \).

EEFs a natural one to be considered when performing model reduction. KLD has already been applied to compute EEFs for PDE systems [4], [16], [57], [58]. Here, we omit the derivation of KLD, and directly give the implementation procedure in Algorithm 1.

B. Model Reduction With SP Technique

For convenience, we denote \( y(s, t) \triangleq y(z, t) \), and \( M(z) \triangleq M(z, s \in [z_1, z_2]) \) for some space-varying matrix function \( M(z) \). To simplify the notation, we consider the PDE system (1)–(3) with \( n = 1 \) without loss of generality. Assume that the PDE state \( y(z, t) \) can be represented as an infinite weighted sum of a complete set of orthogonal basis functions \( \{\phi_i(z)\} \)

\[
y(z, t) = \sum_{i=1}^{\infty} x_i(t) \phi_i(z) \tag{6}
\]

where \( x_i(t) \) is a time-varying coefficient named the mode of the PDE system. By taking inner product with \( \phi_i(z) \) \((i = 1, \ldots)\) on both sides of PDE system (1)–(3), we obtain the following infinite-dimensional ODE system:

\[
\begin{align*}
\dot{x}(t) &= f_s(x(t), x_f(t)) + Bu(t) \\
\dot{x}_f(t) &= L_f(x(t), x_f(t)) + B_f u(t) \\
y_0(t) &= H_f x_f(t) + J_f f_s(t) \\
x(0) &= x_0, x_f(0) = x_{f0}
\end{align*}
\tag{7}
\]

where

\[
x(t) = (y(t), \Phi_s(z)) \triangleq [x_1(t), \ldots, x_N(t)]^T \tag{8}
\]

and \( x_f(t) = (y(t), \Phi_f(z)) \triangleq [x_{N+1}(t), \ldots, x_{\infty}(t)]^T \in \mathbb{R}^{\infty}, f_s(x, x_f) \triangleq \langle \mathcal{L}_f(x, x_f), \Phi_f(z) \rangle, \mathcal{L}_f(x, x_f) \triangleq \langle \mathcal{L}(x, x_f), \Phi_f(z) \rangle, B \triangleq \langle \mathcal{B}(x, x_f), \Phi_f(z) \rangle, B_f \triangleq \langle \mathcal{B}_f(x, x_f), \Phi_f(z) \rangle, H_f \triangleq \int_{z_1}^{\infty} H(z) \Phi_f(z) dz, H_f \triangleq \int_{z_1}^{\infty} H(z) \Phi_f(z) dz, x_0 \triangleq \langle \phi_0(z), \Phi_s(z) \rangle, x_{f0} \triangleq \langle \phi_0(z), \Phi_f(z) \rangle \] with \( \Phi_s(z) = \{\phi_i(z), \ldots, \phi_N(z)\} \) and \( \Phi_f(z) = \{\phi_{N+1}(z), \ldots, \phi_{\infty}(z)\} \).

Owing to the highly dissipative nature of the PDE system (1)–(3), it is reasonable [55], [59] to assume that

\[
\mathcal{L}_f(x, x_f) = \frac{1}{\varepsilon} A_f x_f + f_f(x, x_f)
\]

where \( \varepsilon > 0 \) is a small positive parameter quantifying the separation between the slow (dominant) and fast (negligible) modes, \( A_f \) is a matrix that is stable (in the sense that the state of the system \( \dot{x}_f = A_f x_f \) tends exponentially to zero), and \( f_f(x, x_f) \) satisfies

\[
\|f_f(x, x_f)\| \leq k_1 \|x\| + k_2 \|x_f\| \tag{9}
\]

for \( \|x\| \leq \beta_1 \) and \( \|x_f\| \leq \beta_2 \) with \( k_1, k_2 > 0 \). Then, system (7) can be rewritten as the following standard singularly perturbed form:

\[
\begin{cases}
\dot{x}(t) = f_s(x(t), x_f(t)) + Bu(t) \\
\varepsilon \dot{x}_f(t) = A_f x_f(t) + \varepsilon f_f(x(t), x_f(t)) + \varepsilon B_f u(t) \\
y_0(t) = H_f x_f(t) + H_f x_f(t) \\
x(0) = x_0, x_f(0) = x_{f0}
\end{cases}
\tag{10}
\]

Introducing the fast time-scale \( \tau = t/\varepsilon \) and setting \( \varepsilon = 0 \), we obtain the following infinite-dimensional fast subsystem from:

\[
\frac{dx_f}{d\tau} = A_f x_f
\tag{11}
\]

which is exponentially stable [55], [59]. Then, setting \( \varepsilon = 0 \) in the system (10), we have that \( x_f = 0 \), and thus the following slow subsystem is obtained:

\[
\begin{cases}
\dot{x}(t) = f(x(t)) + Bu(t) \\
x(0) = x_0
\end{cases}
\tag{12}
\]

where \( f(x) = f_s(x, x_f) \).

Observe that the LQ cost functional (4) can be rewritten as

\[
V = \int_0^{+\infty} \left( \|x(t)\|_{Q_s}^2 + \|u(t)\|_{R_s}^2 \right) dt + \mathcal{R}(x, x_f) \tag{13}
\]

where \( \mathcal{R}(x, x_f) \triangleq \int_0^{+\infty} (2x^T Q_s x_f + x_f^T Q_f x_f) dt, \) \( Q_s \triangleq H_s^T H_s, Q_f \triangleq H_f^T H_f, \) \( H_f \triangleq H_s^T H_f, \) \( Q_f \triangleq H_f^T H_f. \) Using \( x_f = 0 \), we have that \( \mathcal{R}(x, x_f) = 0 \), then the LQ cost function (13) is written as

\[
V(x_0) = \int_0^{+\infty} \left( \|x(t)\|_{Q_s}^2 + \|u(t)\|_{R_s}^2 \right) dt. \tag{14}
\]

Now, the optimal control problem of the PDE system (1)–(3) can be reformulated as: find a state feedback control policy to minimize the LQ cost function (14) based on the slow subsystem (12).

In this paper, we use the EEFs for model reduction, i.e., the basis function \( \phi_i(z) \) is computed with KLD. The dimension of \( \Phi_s \) (i.e., \( N \)) is chosen such that it satisfies

\[
\sum_{i=1}^{N} \lambda_i \geq \frac{1}{\sum_{j=1}^{M} \lambda_j} \geq 1 - \zeta \tag{15}
\]

for a small positive real number \( \zeta > 0 \).

Remark 2: The use of EEFs for model reduction of nonlinear PDE systems has two merits compared with analytical eigenfunctions. The first is that EEFs can be computed for PDE systems with any form of nonlinear spatial differential operator, while analytical eigenfunctions are only available for PDE systems with a known linear spatial differential operator. Another important merit of EEFs is that they can be computed online on the basis of data collection of system states by using KLD, which does not require a mathematical system model. Thus, the highly complexity or unavailability of system dynamic model have no effects on the computation of EEFs.
IV. ADAPTIVE OPTIMAL CONTROL DESIGN WITH NDP

In this section, we design a NDP based adaptive optimal controller for the PDE system (1)-(3) on the basis of the finite-dimensional ODE model (12).

A. Adaptive Optimal Control Design

It is known [60] that the optimal control problem of the nonlinear finite-dimensional ODE system (12) with cost function (14) relies on the solution of the following HJB equation:

\[
(V^*)^T f + \|x\|_{Q_1}^2 - \frac{1}{4}(V^*)^T B R^{-1} B^T V^* = 0 \quad (16)
\]

where \( V^*(x) \in C^1(\Omega) \). Then, the optimal control policy \( u^* \) is given by

\[
u^*(x) = -\frac{1}{2} R^{-1} B^T \nabla V^*(x).
\]

Note that the HJB equation (16) is a nonlinear PDE that has proven to be impossible to solve analytically. To overcome this difficulty, we propose an adaptive optimal control method based on NDP, which can learn the solution of HJB equation online. NDP is implemented on an actor-critic structure, where actor and critic NNs are used for approximating control policy and value function, respectively.

From the well-known high-order Weierstrass approximation theorem [61], it follows that a continuous function can be represented by an infinite-dimensional linearly independent basis function set \( \{ \psi_k(x) \}_{k=1}^\infty \). For real practical applications, it is usually required to approximate the function in a compact set with a finite-dimensional function set. Thus, we consider the critic NN for approximating the solution \( V^*(x) \) of the HJB equation (16) on a compact set \( \Omega \). Let \( \Psi(x) \triangleq [\psi_1(x), \psi_2(x), \ldots, \psi_L(x)]^T \) be a vector of linearly independent activation functions, where \( L \) is the number of neurons in the NN hidden layer. Then, \( V^*(x) \) can be expressed with \( \Psi(x) \) as follows:

\[
V^*(x) = \theta^T \Psi(x) + \delta_1(x) \quad (18)
\]

where \( \theta^* \triangleq [\theta_1^*, \theta_2^*, \ldots, \theta_L^*]^T \) is the unknown ideal constant weight vector, and \( \delta_1(x) \) is the NN approximation error. The optimal control law (17) is rewritten as

\[
u^*(x) = -\frac{1}{2} R^{-1} B^T [\nabla \Psi(x)]^T \theta^* - \frac{1}{2} R^{-1} B^T \nabla \delta_1(x). \quad (19)
\]

For convenience, define a matrix \( D = BR^{-1}B^T \). It follows from (18) that the HJB equation (17) is represented as

\[
\theta^T [\nabla \Psi] f + \|x\|_{Q_1}^2 - \frac{1}{4} \|\theta^T [\nabla \Psi]\|_D^2 + \delta_2(x) = 0 \quad (20)
\]

where \( \delta_2(x) \) is the residual error of the HJB equation due to the critic NN error and given by

\[
\delta_2(x) \triangleq [\nabla \delta_1(x)]^T (f(x) + Bu^*(x)) - \frac{1}{4} \|\nabla \delta_1(x)\|_D^2. \quad (21)
\]

Although the ideal critic NN weight vector \( \theta^* \) provides the best approximate solution of the HJB equation (16), it is unknown. Actually, the output of the critic NN is

\[
\tilde{V}(x) = \tilde{\theta}^T \Psi(x) \quad (22)
\]

where \( \tilde{\theta} \triangleq [\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_L]^T \) is the estimate of \( \theta^* \). Thus, the actual output of the actor NN is

\[
\hat{u}(x) = -\frac{1}{2} R^{-1} B^T [\nabla \Psi(x)]^T \tilde{\theta}.
\]

It is observed from (20) and (21) that when the control policy is not stabilizing. This means that in this difficulty, we propose an adaptive optimal control method based on NDP, which can learn the solution of HJB equation online. NDP is implemented on an actor-critic structure, where actor and critic NNs are used for approximating control policy and value function, respectively.

B. Stability Analysis

Notice that the adaptive optimal control is designed based on slow subsystem. Thus, we should analyze the stability of the original closed-loop PDE system. Before starting, we introduce a definition and some reasonable assumptions.

Definition 1 [62]: Consider the system (12), the solution \( x(t) \) is SGUUB, if for any compact set \( \Omega \subset \mathbb{R}^N \) and \( x(t_0) = x_0 \in \Omega \), there exist positive constants \( \mu_1 \) and \( T(\mu_1, x_0) \) such that \( \|x(t)\| < \mu_1 \) for all \( t > t_0 + T \).
Assumption 1:
1) the ideal weight vector of the critic NN is bound, i.e., \( \| \theta^* \| \leq \theta_M \), where \( \theta_M \) is a positive number;
2) the critic NN approximation error \( \delta_1(x) \) and its derivative \( \nabla \delta_1(x) \) are bounded on a compact set containing \( \Omega \), i.e., \( |\delta_1| \leq \delta_M \) and \( |\nabla \delta_1| \leq \delta_{DM} \), where \( \delta_M \) and \( \delta_{DM} \) are positive numbers;
3) the critic NN activation functions and their gradients are bounded on a compact set containing \( \Omega \), i.e., \( |\psi| \leq \psi_M \) and \( |\nabla \psi| \leq \psi_{DM} \), where \( \psi_M \) and \( \psi_{DM} \) are positive numbers;
4) \( f(x) \) is bounded on a compact set containing \( \Omega \), i.e., \( \| f \| \leq f_M \), where \( f_M \) is a positive number. \( \square \)

Define the critic NN weight estimation error as \( \tilde{\theta}(t) = \tilde{\theta}(t) - \theta^* \), the value function estimation error as \( \tilde{V}(x) = \tilde{V}(x) - V^*(x) \), and the control estimation error as \( \tilde{u}(x) = \tilde{u}(x) - u^*(x) \), respectively. Then, we have the following result.

Theorem 1: Consider the optimal control problem of the PDE system (1)-(3) and the adaptive optimal control policy (23) with the tuning law (25) for critic NN weight vector. Let Assumption 1 hold. Then, there exist positive real constants \( \sigma_1, \sigma_2 \) and \( \varepsilon^* \), such that if \( \| x_0 \| \leq \sigma_1, \| x_f \|_{\mathbb{R}^N} \leq \sigma_2 \), and \( \varepsilon \in (0, \varepsilon^*) \), it follows that:
1) the critic NN weight estimation error \( \tilde{\theta} \) is SGUBB, and the state \( y(z, t) \) of the original closed-loop PDE system (1)-(3) is SGUBB in \( L_2 \)-norm, i.e., there exists a positive number \( \sigma_3 \) such that \( \| y(z, t) \|_{L_2} \leq \sigma_3 \); 
2) the state of slow subsystem \( x(t) \) converges to compact set: \( \Omega = \{ x | \| x \| \leq B_1 \} \), where \( B_1 \) is defined in (A23); 
3) the value function estimation error \( \tilde{V}(x) \), and the control estimation error \( \tilde{u}(x) \) are also SGUBB. \( \square \)

Proof: See the Appendix.

Remark 4: From Assumption 1 and the definition of bound \( B_1 \), it is noted that the compact set \( \Omega \) for NN approximation can be determined before control design. \( \square \)

Remark 5: The convergences of the \( \tilde{V}(x) \) and \( \tilde{u}(x) \) shows that, with the adaptive optimal control policy (23) and critic NN weight tuning law (25), \( \tilde{V} \), and \( \tilde{u} \) will approach \( V^* \) and \( u^* \), respectively. \( \square \)

Summarize briefly, the adaptive optimal control policy (23) and its weight tuning law (25) is based on NDP for solving the HJB equation online. It is noted that the control approach does not require an initial stabilizing control policy, and thus the initial critic NN weights can be given randomly. This is very important because an initial stabilizing control policy is often difficult to obtain for practical complex SDPs. Since the adaptive optimal controller is designed based on the slow subsystem, the SP theory is employed to prove the convergence of the original closed-loop PDE system in Theorem 1, where the NN estimation error is involved.

V. Simulation Studies
In this section, we first test the effectiveness of the developed adaptive optimal control on a general convection-diffusion-reaction process. Then, it is applied to a temperature cooling fin of high-speed aerospace vehicle.

A. Effectiveness Test on a General Convection-Diffusion-Reaction Process
Consider the following general nonlinear diffusion-convection-reaction process [4], [14], [56]:
\[
\begin{aligned}
\frac{\partial y(z, t)}{\partial t} &= (1 - a_{dc}) f_d(y, \tilde{\gamma} y/\tilde{\varepsilon} z, \tilde{\varepsilon}^2 y/\partial z^2) \\
&= a_d f_c(\tilde{\gamma} y/\partial z) + f_r(z, y) + b_T u(t) \\
\end{aligned}
\]
subjected to the Dirichlet boundary conditions
\[
y(0, t) = y(\pi, t) = 0
\]
and the initial condition
\[
y_0(t) = 0.3 \sin(3z)
\]
where \( f_c = \tilde{\gamma} y/\partial z \), \( f_d = \tilde{\gamma} y/\partial z \), \( k(y) \), \( f_r = \beta_T(z)(e^{-\gamma/(1+\gamma)} - e^{-\gamma}) - \beta_U y \), \( b_T(z) = \beta_U b(z) \), \( y \) is the PDE state, \( z \in [0, \pi] \), \( k(y) \) is the diffusion coefficient that may be constant or dependent on the state, \( \beta_T(z) \) is the heat transfer coefficient, \( \gamma \) is activation energy, and \( b(z) \) is the actuator distribution function. These parameters are given as: \( k(y) = 0.5 + 0.7/(1 + y) \), \( \beta_T(z) = 16(\cos(z) + 1) \), \( b_U = 1 \), \( \gamma = 4 \), \( b(z) = H(z - 0.2\pi) - H(z - 0.4\pi) \), where \( H(z) \) is the standard Heaviside function. The terms \( f_c, f_d, f_r \) are corresponding to convection, diffusion and reaction phenomena in the process, respectively. The coefficient \( a_{dc}, 0 \leq a_{dc} \leq 1 \) represents the weight of convection and diffusion in the process.

It is known [1]–[4], [6], [14], [20] that the diffusion term \( f_d \) is highly dissipative, which appears in parabolic PDE systems. Thus, their dominant system dynamics can be represented by a lower order slow subsystem, and an infinite-dimensional fast subsystem that is exponentially stable. In contrast, the convection term \( f_c \), which appears in hyperbolic PDE systems [63]–[66], does not have the slow-fast separation feature because their eigenmodes contain the same, or nearly the same amount of energy. This implies that for hyperbolic PDE systems, there does not exist an exponentially decay subsystem, and thus infinite number of modes is required to accurately describe their dynamic behavior. Then, controller design methods based on slow subsystem are usually difficult for hyperbolic systems. For convection-diffusion-reaction process (28)–(30), with the increase of the weight coefficient \( a_{dc} \), the weight of convection term \( f_c \) increasing and the weight of diffusion term \( f_d \) decreasing, and thus the control problem becomes more and more difficult for the developed adaptive optimal control method.

To test the efficiency of the developed method and show the influence of the weight coefficient \( a_{dc} \) between convection and diffusion, we conduct simulations on four cases respectively \( a_{dc} = 0.2, 0.5, 0.7 \) and 0.9. For four cases, we use the same parameters setting. The weighting matrix \( R \) in (4) is given as \( R = 1 \). Let the time and space sample interval be \( \Delta t = 0.05 s \) and \( \Delta z = 0.0628 \). Using different initial conditions and input signals, we collect 5000 snapshots for computing EEFs with KLD. The first three EEFs are employed to compute the state
of the slow subsystem. To use the developed adaptive optimal control method, we select critic NN activation function vector $\Psi(x)$ of size 18 (i.e., $L = 18$) as

$$
\Psi(x) = \left[ x_1^2 x_1 x_2 x_1 x_3 x_2^2 x_2 x_3 x_3^2 x_2^2 x_3^2 x_2^2 x_3^2 \right]^{T} \tag{31}
$$

and each element of initial NN weight vector $\hat{\theta}(0)$ is given as 1 for four cases (also can be generated randomly). Using the adaptive optimal control policy (23) and the NN weights tuning law (25) with $V = 0.5x^T x$ and $\alpha = 10$, closed-loop simulations are conducted. To show the effects of the neglected fast dynamics, we have computed the estimation error between the actual closed-loop PDE system state $y(z, t)$ and the approximated PDE state $\hat{y}(z, t) \triangleq \sum_{i=1}^{N} x_i(t) \phi_i(z)$ with

$$
y_{\text{error}}(z, t) \triangleq y(z, t) - \hat{y}(z, t). \tag{32}
$$

The simulation results of four cases are shown in Figs. 1–7. Fig. 1 shows the state profiles of open-loop PDE system and Fig. 2 shows the first three EEFs. Conducting closed-loop simulations on the PDE system with the developed adaptive optimal control method, Fig. 3 shows the first three representative critic NN weights. Figs. 4–7 show the control action, the state profiles of the closed-loop PDE system, the state trajectories of the slow subsystem, and PDE state estimation error, respectively.

Remark 6: Based on the well-known high-order Weierstrass approximation theorem [61], the solution of HJB equation is a continuous function that can be represented by any infinite-dimensional linearly independent basis function set $\{\psi_i(z)\}_{i=1}^{\infty}$. For the truncated case (22), the selections of finite-dimensional basis function set $\{\psi_i(z)\}_{i=1}^{L}$ and its size are often experience-based. Like all function approximation techniques, large size of the basis function set can improve the estimation performance at the price of highly computational effort. Thus, an appropriate selection of basis function set and its size are useful to balance the performance and computation. However, it is still difficult to develop a general optimal selection method for all systems. This is simply because the optimal selection is often different for different systems. In a word, for a specific system, prior experiences would be helpful for the selection of basis function set and its size.

Remark 7: It is noted from Theorem 1 that the state of closed-loop PDE system is SGUUB in $L_2$-norm. The error bound of the state mainly results from two aspects: 1) the
error between the original PDE state and the approximated PDE state via the slow subsystem modes χ and 2) the estimation error of critic NN for value function approximation of HJB equation. To reduce the error bound, four methods are potentially helpful: 1) increase the dimension of slow subsystem (12), i.e., N; 2) increase the number of neurons in the critic NN hidden layer, i.e., L; 3) construct NN activation functions well suited for special system; and 4) use more complicated value function approximate structure, such as nonlinear function approximation [67]. The first two methods will increase the computation load, and thus requires a tradeoff between accuracy and computation. The third and fourth methods are promising. Using system information collected in practice, it is possible to learn basis functions automatically for value function approximation. Nonlinear function approximation is still an open and difficult problem in RL community, because it may cause the system to become unstable. These issues still require a deep investigation and are left for future research. In fact, although there exists a theoretical error bound in PDE system state, it is found from the simulation results (Figs. 5 and 7) that both the PDE state and reconstruction error approach zero at a high accuracy for cases α_dC = 0.2, 0.5, and 0.7.

B. Application on Temperature Cooling Fin of High-Speed Aerospace Vehicle

In this section, the developed control method is applied to a complex temperature cooling fin of high-speed aerospace vehicle [2], the system dynamics of which is described with the following parabolic PDE:

\[
\rho C \frac{\partial T(l,t)}{\partial t} = k \frac{\partial^2 T(l,t)}{\partial l^2} - \frac{P h}{A} (T - T_{\infty}) + \frac{P \rho \sigma}{A} (T^4 - T_{\infty}^4) + \hat{B}(l)u(t)
\]  

(33)
subected to the boundary conditions

$$\frac{\partial T}{\partial l} \bigg|_{l=0} = 1, \quad \frac{\partial T}{\partial l} \bigg|_{l=L} = 0$$

(34)

and the initial condition

$$T(l, 0) = T_0(l)$$

(35)

where $T$ is the temperature of the fin, $l \in [0, L]$, $u(t) = [u_1(t) \ u_2(t) \ u_3(t)]^T$ is the control input vector, and $B(t) = [\hat{B}_1(t) \ \hat{B}_2(t) \ \hat{B}_3(t)]^T$ describe how control actions are distributed in spatial domain. The system parameters and their values are shown in Table I and the open-loop temperature profile is shown in Fig. 8(a). The control objective is to reach a desired constant temperature $T_d = 700 \ ^\circ C$, and achieve minimized cost.

To simplify the description of system (33)–(35), it is worthwhile to define the dimensionless temperature, desired temperature and spatial position variables as: $
abla \triangleq T/\overline{T}$, $\overline{y}_d \triangleq T_d/\overline{T}$, $z \triangleq l/L$, where $\overline{T} = 1000$ and it also can be a large number. Then, system (33) is rewritten as a dimensionless formulation

$$\frac{\partial \nabla}{\partial t} = \frac{k}{\rho C L^2} \frac{\partial^2 \nabla}{\partial z^2} - \frac{Ph}{\rho C A} \left( \nabla - \frac{T_{\infty}}{\bar{T}} \right) - \frac{P_{\varepsilon \sigma}}{\rho C T} (\nabla^4 - \nabla_d^4) + \frac{1}{\rho C T} \hat{B}(Lz)u(t).$$

(36)

Define state error $y \triangleq \nabla - \overline{y}_d$, and four coefficients: $a_1 \triangleq \frac{k}{\rho C L^2}$, $a_2 \triangleq \frac{Ph}{\rho C A}$, $a_3 \triangleq \frac{P_{\varepsilon \sigma}}{\rho C T}$, $a_4 \triangleq \frac{1}{\rho C T}$. Then, the system (33)(35) is briefly represented as

$$\frac{\partial y}{\partial t} = a_1 \frac{\partial^2 y}{\partial z^2} + a_2 y + a_3 T^4 (y + \overline{y}_d)^4 + u_d + B(z)u(t)$$

(37)

subjected to the boundary conditions

$$\frac{\partial y}{\partial z} \bigg|_{z=0} = \frac{L}{\bar{T}}, \quad \frac{\partial y}{\partial l} \bigg|_{l=1} = 0$$

(38)

and the initial condition

$$y(z, 0) = \overline{y}_0(z)$$

(39)

where $y_0(z) \triangleq \overline{y}_0(z) - \overline{y}_d$, $u_d \triangleq a_2 \left( \overline{y}_d - T_{\infty}/\bar{T} \right) - a_3 T^4 \overline{y}_d^4$, $B(z) \triangleq a_4 \hat{B}(Lz)$, and the initial state $y_0(z) = 0.5 \cos(\pi z)$, i.e., $T_0(l) = 0.5 T_c \cos(\pi l/L) + T_d$.

Then, the objective is to design control policy $u$ such that $y(z, t)$ of system (37)–(39) approaches zero, and achieve minimized LQ cost functional (4) with $R$ be an unit matrix, and objective output $y_d(t) = \int_0^L y(z, t)dz$.

Using KLD to compute EEFs, an ensemble of size 5000 (i.e., $M = 5000$) is collected from 50 independent simulations with different initial conditions and inputs. Fig. 8(b) shows the first three EEFs that are used for computing the state of slow subsystem. To solve the optimal control problem with the developed method, let $V = 0.5x^T x$, $\alpha = 10$ and the critic NN

<table>
<thead>
<tr>
<th>Table I: System Parameters and Their Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k = 19W/(m^3C)$</td>
</tr>
<tr>
<td>$A = 1m^2$</td>
</tr>
<tr>
<td>$P = 1.3716m$</td>
</tr>
<tr>
<td>$L = 1m$</td>
</tr>
<tr>
<td>$h = 20W/(m^3C)$</td>
</tr>
<tr>
<td>$T_{\infty} = 100^\circ C$</td>
</tr>
<tr>
<td>$T_{\infty} = 40^\circ C$</td>
</tr>
<tr>
<td>$\varepsilon = 0.965$</td>
</tr>
<tr>
<td>$\sigma = 5.669 \times 10^{-8}W/m^2K^4$</td>
</tr>
<tr>
<td>$\rho = 7685kg/m^3$</td>
</tr>
<tr>
<td>$C = 0.46kJ/(kg^\circ C)$</td>
</tr>
</tbody>
</table>

Fig. 8. Simulation results on the temperature cooling fin of high-speed aerospace vehicle. (a) Temperature profile of open-loop system. (b) First three EEFs (blue solid, green dash, and red dot lines represent EEFs $\phi_1(z) \sim \phi_3(z)$, respectively). (c) First three representative critic NN weights (blue solid, green dash, and red dot lines represent $\hat{B}_1(t) \sim \hat{B}_3(t)$, respectively). (d) Actual state trajectories of closed-loop slow subsystem (blue solid, green dash, and red dot lines represent $x_1(t) \sim x_3(t)$, respectively). (e) Control action (blue solid, green dash, and red dot lines represent $\hat{B}_1(t) \sim \hat{B}_3(t)$, respectively). (f) Temperature profile of closed-loop system. (g) State estimation error of the closed-loop PDE system. (h) Cost $J(t)$ with respect to time (blue solid and green dash lines are obtained by the adaptive optimal control and the existing method in [52], respectively).
activation function vector be selected as (30). Then, closed-loop simulation is conducted, and the results are demonstrated in Fig. 8. It is noted in Fig. 8(f) that the temperature of the cooling fin converges to the desired constant temperature profile \( T_d = 700 \, ^\circ C \).

We also conduct comparative simulation study between the developed adaptive optimal control and the existing optimal control approach based on LQ regulator [52]. Since the figures of closed-loop temperature profile, slow subsystem state trajectories and control trajectories of the two methods are similar, we omit these figures under the existing control approach of [52]. To compare the cost that generated from the two control methods, we define the cost with respect to time as \( J(t) \triangleq \int_t^{t+\infty} \left( \| y_f(\tau) \| + \| u(\tau) \| \right) d\tau \). Fig. 8(h) shows the curves of \( J(t) \), from which it is observed that the adaptive optimal control achieves a smaller cost than the optimal method in [52] as time increases.

VI. CONCLUSION

In this paper, we have developed a NDP-based adaptive optimal control method for SDPs described with highly dissipative nonlinear PDEs. The KLD is firstly employed to compute EEFs with the method of snapshots, based on which a finite-dimensional ODE model is derived via the SP technique. The ODE model is then used as the basis for optimal control design, which is converted to solve the HJB equation. Subsequently, we propose a NDP-based adaptive optimal control method to solve the HJB equation online, and derive a novel critic NN weights tuning law to overcome the difficulty of the requirement of an initial stabilizing control policy. Based on the Lyapunov theory, the stability of the original closed-loop PDE system is proved by involving the NN estimation error. Finally, we demonstrate the effectiveness of the developed adaptive optimal control method on a nonlinear diffusion-convection-reaction process and a temperature cooling fin of high-speed aerospace vehicle.

APPENDIX

A. Proof of Theorem 1

1) Under the adaptive optimal control policy (23) with the tuning law (25) of \( \hat{\theta} \), the closed-loop PDE system is written as

\[
\begin{aligned}
\frac{\partial y}{\partial t} = L y, \quad & \frac{\partial y}{\partial z}, \ldots, \frac{\partial^m y}{\partial z^m} - \frac{1}{2} B(z) R^{-1} B^T \nabla \Psi \nabla \theta \\
\hat{\theta} = -\alpha \mu \rho \phi (x) \partial \phi (x) + \alpha k \nabla \Psi (x) D \nabla \Psi (x).
\end{aligned}
\]

Using the model reduction procedure presented in Section III-B, the system (A1) is equivalently rewritten as

\[
\begin{aligned}
x_f = f_o (x, 0) - \frac{1}{2} R^{-1} B^T \nabla \Psi \nabla \theta + f_o (x, x_f) - f_s (x, 0) \\
\dot{x}_{f,f} = A_f x_f + \varepsilon f_f (x, x_f) - \frac{1}{2} \varepsilon B_f R^{-1} B^T \nabla \Psi \nabla \theta \\
\hat{\theta} = -\alpha \mu \rho \phi (x) + \alpha k \nabla \Psi (x) D \nabla \Psi (x).
\end{aligned}
\]

(A2)

Since \( L \) is a sufficiently smooth nonlinear vector function, it is obvious that there exists a constant \( k_3 > 0 \) such that

\[
\| f_s (x, x_f) - f_s (x, 0) \| \leq k_3 \| x_f \| \infty
\]

(A3)

for \( \| x \| \leq \sigma_1 \).

According to the exponential stability property of the \( x_f \)-subsystem (11), and the converse Lyapunov theorem [68], we have that there exist a Lyapunov function candidate \( V_f (x_f) \), and positive real numbers \( l_1, l_2, l_3, l_4 \) such that the following conditions hold:

\[
\begin{aligned}
l_1 \| x_f \|_{\infty}^2 & \leq V_f (x_f) \leq l_2 \| x_f \|_{\infty}^2 \\
\dot{V}_f (x_f) & \leq -\frac{1}{c} \| x_f \|_{\infty}^2 \\
\| \nabla V_f \|_{\infty} & \leq l_4 \| x_f \|_{\infty}
\end{aligned}
\]

(A4)

where \( \nabla V_f \equiv \partial V_f / \partial x_f \).

Before proving the stability of the closed-loop PDE system, we derive some useful conditions as follows:

\[
\begin{aligned}
\| \nabla V_f \| & = \| (\nabla \Psi)^T \phi \nabla \theta + \nabla \delta_1 \| \leq \sigma_{VD} M \quad (A5) \\
\| u^* \| & \leq \| -\frac{1}{c} R^{-1} B^T \nabla V_f \| \leq \sigma_{uM} \quad (A6) \\
\| f + B u^* \| & \leq \sigma_{BM} \quad (A7) \\
\| D \| & = \| B R^{-1} B^T \| \leq \sigma_{MD} \quad (A8)
\end{aligned}
\]

where \( \sigma_{RM} \triangleq \sigma (R), \sigma_{RM} \triangleq \sigma (R), \sigma_{uM} \triangleq \sigma (B), \sigma_{BM} \triangleq \sigma (B), \sigma_{VD} \triangleq \sigma (D_{\delta}), \sigma_{BM} \triangleq \sigma (D_{\phi}), \sigma_{uM} = \sigma_{RM} \sigma_{BM}/2, \sigma_{BM} = \sigma_{BM} \sigma_{BM} / 2, \sigma_{VD} = \sigma_{VD} / \sigma_{BM} \sigma_{BM} \text{ and } \sigma_{RM} = \sigma_{RM} \sigma_{BM} \).

Moreover, considering the fact \( \int_{\Omega_1} \nabla \Psi (z) R^2 (z) dz = B^T B + B_f B_f ^T \), we have

\[
\| B_f ^T \nabla V_f \| \leq \| (\nabla V_f)^T B_f B_f ^T \| \leq l_5 \| x_f \|_{\infty} \quad (A9)
\]

Now, let us consider the following Lyapunov function candidate:

\[
\bar{V}(x) = V^*(x) + V_f (x_f) + \nabla \bar{V}(x) + \bar{V}_f (x_f) \quad (A10)
\]

where \( V^*(x) \) is the solution of the HJB equation (16) and \( V_f (x_f) \) is the solution of the HJB equation (16). Differentiating \( \bar{V}^*(x) \), \( \bar{V}_f (x_f) \), \( \nabla \bar{V}(x) \) and \( \bar{V}_f (x_f) \) with respect to time along the trajectory of (A2), yield

\[
\begin{aligned}
\dot{\bar{V}}^* (x) & = (\nabla V^*)^T \left[ f_o (x, 0) + B \hat{\theta} + f_f (x, x_f) - f_s (x, 0) \right] \\
& = -\| x^* \|_{Q_o}^2 - \| u^* \|_{Q}^2 + (u^*)^T B^T (\nabla \Psi)^T \hat{\theta} - (u^*)^T B^T \nabla \delta_1 \\
& \quad + (\nabla V^*)^T \left[ f_o (x, x_f) - f_s (x, 0) \right] \\
& \leq -\sigma_1 \| x \|_{\infty}^2 + \sigma_2 \| x_f \|_{\infty} + \sigma_3 \| \hat{\theta} \| + \sigma_4 \quad (A11)
\end{aligned}
\]

where \( \sigma_1 \triangleq \sigma_{Q_o}, \sigma_2 \triangleq k_3 \sigma_{VD} M, \sigma_3 \triangleq \sigma_{uM} \sigma_{BM} \sigma_{BM}, \text{ and } \sigma_4 \triangleq \sigma_{uM} \sigma_{BM} \sigma_{BM} \). Also, using the \( \sigma_{Q_o} \) and \( \sigma_{Q} \) as the constants,

\[
\begin{aligned}
\dot{V}_f (x_f) & = \frac{1}{c} (\nabla V_f)^T A_f x_f + (\nabla V_f)^T f_f (x, x_f) \\
& \quad + (\nabla V_f)^T B_f u^* - \frac{1}{2} (\nabla V_f)^T B_f R^{-1} B^T (\nabla \Psi)^T \hat{\theta} \\
& \quad + \frac{1}{2} (\nabla V_f)^T B_f R^{-1} B^T \nabla \delta_1 \\
& \leq \left( -\frac{l_f}{c} + \sigma_5 \right) \| x_f \|_{\infty}^2 + \sigma_6 \| x \| \| x_f \|_{\infty} \| \hat{\theta} \| + \sigma_8 \| x_f \|_{\infty} \| \hat{\theta} \| \quad (A12)
\end{aligned}
\]
where \( \sigma_4 \triangleq l_{k2}, \sigma_6 \triangleq l_{k1}, \sigma_7 \triangleq l_{k3}u_{BM}^{-1}u_{BM}^{1/2}, \) and \( \sigma_8 \triangleq l_{k3}(u_{BM} + u_{BM}^{1/2})/2 \).

\[
\dot{V}(\chi) = (\nabla V)^T (f + B \dot{u})
= (\mu \dot{\theta} T \psi T D \theta + \frac{1}{2} \theta^T D \theta) + (\frac{1}{2} \sigma_9 \dot{\theta}^2 + \sigma_{10} \dot{\theta}^2 + \sigma_{11} \dot{\theta}^2 + \sigma_{12} \dot{\theta}^2)
\quad + \kappa \theta^T \psi D \nabla \psi.
\] (A14)

Thus, using (A11), (A12), (A13), and (A15), the time derivative of \( L(t) \) in (A10) satisfies

\[
L(t) = \dot{V}(\chi) + \dot{V}(\chi) + \dot{V}(\chi) + \dot{V}(\chi)
\quad + \nabla \dot{\theta}^T \psi D \nabla \theta
\] (A16)

Letting \( \epsilon^* \triangleq \min(\epsilon_1, \epsilon_2) \) where \( \epsilon_1 \triangleq 4 \sigma_1 l_{k1}/(4 \sigma_1 \sigma_5 + \sigma_5^2) \)
and \( \epsilon_2 \triangleq 4 \sigma_1 \sigma_{13} l_{k1}^2/(\sigma_2 \sigma_1 l_{k1} + \sigma_6^2 \sigma_{13} + 4 \sigma_1 \sigma_3 \sigma_{13}) \), then we have that, if \( \epsilon \in (0, \epsilon^*), \) then \( \Pi(\epsilon) > 0 \). Thus, we have that

\[
\| \nabla \dot{\theta} \| \geq (\sigma_2 + \sigma_8) \| \nabla \dot{\theta} \|.
\] (A20)

This means that \( x, \dot{x}, \) and \( \dot{\theta} \) are SGGUB. Due to the fact that \( \| (x, \dot{x}, \dot{\theta}) \|_2 \leq \| (x, \dot{x}, \dot{\theta}) \|_{\infty} \), the state \( (x, \dot{x}, \dot{\theta}) \) of the original closed-loop PDE system (1)-(3) is SGGUB in \( L_2 \)-norm. This complements the proof of the part 1) of Theorem 1.

2) Define

\[
B_4 \triangleq (\sigma_2 + \sigma_8)/\sigma_{14}.
\] (A23)

According to (A22) and the definition of \( \nu_{\chi,\theta} \), we have that

\[
\| x \|_2^2 + \| \dot{x} \|_{\infty}^2 + \| \dot{\theta} \|^2 < B_4^2,
\] then \( \| x \| \leq B_4 \). This completes the proof of the part 2) of Theorem 1.

3) It follows from the part 1) of Theorem 1 that \( \dot{\theta} \) is SGGUB, thus there exists an \( \nu_{\theta,\dot{\theta}} > 0 \) such that \( \| \dot{\theta} \| \leq \nu_{\theta,\dot{\theta}} \).

Considering

\[
\dot{V}(\chi) = \dot{\theta}^T \psi (\chi) - \delta_1
\]
and

\[
\ddot{u}(x) = -\frac{1}{2} R^{-1} B^T (\nabla \psi)^T \dot{\theta} + \frac{1}{2} R^{-1} B^T \nabla \delta_1.
\] we have that

\[
\| \ddot{V} \| \leq \| \dot{\theta} \| \| \psi \|_{\infty} + \| \delta_1 \| \leq \nu_{\theta,\dot{\theta}} \psi_{\infty} + \delta_1.
\]

\[
\| \ddot{u} \| \leq \frac{1}{2} R^{-1} B^T (\nabla \psi)^T \dot{\theta} + \frac{1}{2} R^{-1} B^T \nabla \delta_1
\]
\[
\leq \frac{1}{2} \nu_{\theta,\dot{\theta}} \psi_{\infty} \delta_1 + \delta_1.
\]

This implies that \( \ddot{V} \) and \( \ddot{u} \) are SGGUB. \( \square \)
REFERENCES


Biao Luo received the B.E. degree in measuring and control technology and instrumentations and the M.E. degree in control theory and control engineering from Xiangtan University, Xiangtan, China, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree in control science and engineering with Beihang University, Beijing, China.

He was a Research Assistant with the Department of System Engineering and Engineering Management, City University of Hong Kong, Hong Kong, and a Research Assistant with the Department of Mathematics and Science, Texas A&M University at Qatar, Doha, Qatar, in 2013. His current research interests include distributed parameter systems, optimal control, data-based control, fuzzy/neural modeling and control, hypersonic entry/reentry guidance, learning and control from big data, reinforcement learning, approximate dynamic programming, and evolutionary computation.

Mr. Luo was a recipient of the Excellent Master Dissertation Award of Hunan Province in 2011.

Hua-Ning Wu was born in Anhui, China, in 1972. He received the B.E. degree in automation from the Shandong Institute of Building Materials Industry, Jinan, China, and the Ph.D. degree in control theory and control engineering from Xi’an Jiaotong University, Xi’an, China, in 1992 and 1997, respectively.

He was a Post-Doctoral Researcher with the Department of Electronic Engineering, Beijing Institute of Technology, Beijing, China, from 1997 to 1999. In 1999, he joined the School of Automation Science and Electrical Engineering, Beihang University, Beijing. From 2005 to 2006, he was a Senior Research Associate with the Department of Manufacturing Engineering and Engineering Management (MEEM), City University of Hong Kong, Hong Kong, where he was a Research Fellow with the Department of MEEM from 2006 to 2008 and in 2010, and the Department of Systems Engineering and Engineering Management in 2011 and 2013. He is currently a Professor with Beihang University.

His current research interests include robust control, fault-tolerant control, distributed parameter systems, and fuzzy/neural modeling and control.

Dr. Wu serves as an Associate Editor of the IEEE *TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS: SYSTEMS*. He is a member of the Committee of Technical Process Failure Diagnosis and Safety, Chinese Association of Automation.

Han-Xiong Li (S’94–M’97–SM’00–F’11) received the B.E. degree in aerospace engineering from the National University of Defense Technology, Changsha, China, the M.E. degree in electrical engineering from the Delft University of Technology, Delft, The Netherlands, and the Ph.D. degree in electrical engineering from the University of Auckland, Auckland, New Zealand.

He is currently a Professor with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong. He has been involved in different fields, including military service, industry, and academia, over the last 30 years. He authored more than 140 SCI journal papers with an SCI h-index of 27. His current research interests include system intelligence and control, process design and control integration, and distributed parameter systems with applications to electronics packaging.

Dr. Li serves as an Associate Editor of the IEEE *TRANSACTIONS ON CYBERNETICS* and the IEEE *TRANSACTIONS ON INDUSTRIAL ELECTRONICS*. He was a recipient of the Distinguished Young Scholar (overseas) from the China National Science Foundation in 2004, the Chang Jiang Professor Award from the Ministry of Education, China, in 2006, and the National Professorship in the China Thousand Talents Program in 2010. He serves as a Distinguished Expert for the Hunan Government and the China Federation of Returned Overseas.