Incorporating Sustainability into the Conceptual Design of Chemical Process-Reaction Routes Selection
Kailiang Zheng, Helen H. Lou, Preeti Gangadharan, and Krishna RANCH
pp 9300-9309
Publication Date (Web): June 15, 2012 (Article)
DOI: 10.1021/ie3002952

Heuristic Dynamic Programming Algorithm for Optimal Control Design of Linear Continuous-Time Hyperbolic PDE Systems
Huai-Ning Wu and Biao Luo
pp 9310-9319
Publication Date (Web): June 4, 2012 (Article)
DOI: 10.1021/ie300897m

SEPARATIONS

A High Temperature Lithium Orthosilicate-Based Solid Absorbent for Post Combustion CO2 Capture
Robert Quinn, Ronald J. Kitzhofer, Jeffrey R. Hutson, and Timothy C. Golden
pp 9320-9327
Publication Date (Web): June 11, 2012 (Article)
DOI: 10.1021/ie300157m

Effect of pH, CO2, and High Glucose Concentrations on Polydimethylsiloxane Pervaporation Membranes for Ethanol Removal
Diana María Aguilar-Valencia, Miguel Ángel Gómez-García, and Javier Fontalvo
pp 9328-9334
Publication Date (Web): June 15, 2012 (Article)
DOI: 10.1021/ie3002765

Synergy of Electrochemical/O2 Process with Aluminum Electrodes in Industrial Wastewater Treatment
Carlos Barrera-Díaz, Lina A. Bernal-Martínez, Reyna Natividad, and Juan M. Peralta-Hernández
pp 9335-9342
Publication Date (Web): June 14, 2012 (Article)
DOI: 10.1021/ie3004144

New Approach for Modeling Hybrid Pressure Swing Adsorption–Distillation Processes
James A. Ritter, Fan Wu, and Armin D. Ebner
pp 9343-9355
Publication Date (Web): June 15, 2012 (Article)
DOI: 10.1021/ie300744n

GENERAL RESEARCH

Generation of Air–Water Two-Phase Flow Patterns by Altering the Helix Angle in Triple Helical Microchannels
Sambasiva Rao Gunneboyina and Animangsu Ghatak
pp 9356-9364
Publication Date (Web): June 27, 2012 (Article)
DOI: 10.1021/ie201249g
Heuristic Dynamic Programming Algorithm for Optimal Control Design of Linear Continuous-Time Hyperbolic PDE Systems

Huai-Ning Wu* and Biao Luo

ABSTRACT: This work considers the optimal control problem of linear continuous-time hyperbolic partial differential equation (PDE) systems with partially unknown system dynamics. To respect the infinite-dimensional nature of the hyperbolic PDE system, the problem can be reduced to finding a solution of the space-dependent Riccati differential equation (SDRDE), which requires the full system model. Therefore, a heuristic dynamic programming (HDP) algorithm is proposed to achieve online optimal control of the hyperbolic PDE system, which online collects data accrued along system trajectories and learns the solution of the SDRDE without requiring the internal system dynamics. The convergence of HDP algorithm is established by showing that the HDP algorithm generates a nondecreasing sequence which uniformly converges to the solution of the SDRDE. For implementation purposes, the HDP algorithm is realized by developing an approximate approach based on the method of weighted residuals. Finally, the application on a steam-jacketed tubular heat exchanger demonstrates the effectiveness of the developed control approach.

1. INTRODUCTION

Hyperbolic partial differential equations (PDEs) are an important class of PDEs that have been widely studied from a mathematical point of view. Moreover, hyperbolic PDEs can represent the dynamics of a large number of practical industrial processes, such as plug flow reactors, fixed-bed reactors, tubular heat exchangers, traffic flows, Saint-Venant equations for open channel hydraulic systems, and wave equations, and so forth. Due to their infinite-dimensional nature, it is very difficult directly using control design methods of ordinary differential equation (ODE) systems for these PDE systems. In addition, the control design for hyperbolic PDE systems is more complicated than that for highly dissipative ODE systems such as parabolic systems and Kuramoto-Sivashinsky equation (KSE). For highly dissipative PDE systems, since their dominant dynamic behavior is usually characterized by a finite (typically small) number of degrees of freedom, the controller design problem is often addressed following the reduce-then-design (RTD) framework (i.e., the PDE system is initially discretized into an approximate finite dimensional ODE model, and then the existing approaches for the ODE systems are used for controller synthesis). However, the RTD framework is not suitable for the control design of hyperbolic PDE systems, because the eigenvalues of their spatial differential operator cluster along vertical or nearly vertical asymptotes in the complex plane, such that they cannot be accurately represented by a finite-dimensional system via model reduction. In order to respect the distributed nature of hyperbolic PDE systems, the controller synthesis problem has to be addressed directly on the basis of the infinite-dimensional models themselves, i.e., following the design-then-reduce (DTR) framework. Over the past few years, some control approaches have been developed for hyperbolic PDE systems following the DTR framework, including the LQ optimal control method via spectral factorization and operator Riccati equation (ORE), boundary control method by using a strict Lyapunov function and backstepping, the sliding mode control method and model predictive control method on the basis of equivalent ODE realizations obtained by the method of characteristics, the nonlinear control method through a combination of PDE theory and geometric control techniques, the fuzzy control method via the T-S fuzzy PDE modeling, exponential stabilization with static output feedback control method, and robust control method via output feedback. However, most of these approaches are model-based and require full knowledge of the mathematical system models, which in most real cases are either unavailable or too costly to obtain. Furthermore, the modeling and identification for PDE systems are also very difficult. Thus, it is desirable to develop model-free control approaches.

Optimal control refers to a class of methods that can be used to synthesize a control policy which results in best possible behavior with respect to the prescribed performance criterion. It is well-known that the solution of optimal control problems can be obtained either by using Pontryagin’s minimum principle or by solving the Hamilton–Jacobi–Bellman (HJB) equation, which requires complete knowledge of the system dynamics. However, for many real processes, there often exist inevitable uncertainties in their constructed models. In this situation, the optimal controller derived with respect to the system’s assumed model will not perform optimally. Thus, adaptation of the controller parameters such that operation becomes optimal with respect to the behavior of the real process is highly desired. To this end, in recent decades, there has been a rapidly growing interest in the use of reinforcement learning.
learning (RL) techniques in adaptive optimal control design of ODE systems, which aims to learn optimal control policies in situations where the available training information is basically provided in terms of judging the success or failure of the observed behaviors. Heuristic dynamic programming (HDP) is one of the RL schemes based on value iteration, was used to solve optimal control problem of discrete-time (DT) ODE systems. For example, Al-Tamimi et al. used the HDP algorithm to solve the optimal control problem for nonlinear DT systems; Zhang et al. solved the infinite-time optimal tracking control problem with a new type of performance index by using the greedy HDP iteration algorithm; a neural HDP method was applied to learn state and output feedback adaptive critic control policy of nonlinear DT affine systems with disturbances. One important benefit of the HDP algorithm is that it can be implemented on an actor–critic structure without requiring the knowledge or identification/learning of the system dynamics. This makes HDP algorithm a promising model-free method for control design of real industrial processes. However, very limited works have focused on the application of the HDP method to PDE systems until present. Following the RTD framework, some control approaches without system models were developed. As such, the approximate dynamic programming (ADP)-based approaches were used to synthesize neural optimal control policies for DT chemical reactor process and DT temperature profile of a high-speed aerospace vehicle. Zhang et al. proposed a nonlinear observer for PDE process and then used it in nonlinear receding horizon control. But as discussed in the previous paragraph, the RTD framework is unsuitable for hyperbolic PDE systems. To the best of our knowledge, the theory of HDP algorithm for optimal control of linear continuous-time (CT) hyperbolic PDE systems from the DTR framework has not yet been established, which motivates the present study.

In this work, we consider the optimal control problem of a class of linear CT hyperbolic PDE systems with partially unknown system dynamics. This problem can be reduced to the problem description. The HDP algorithm is proposed as a new type of performance index by using the greedy HDP iteration algorithm; a neural HDP method was applied to learn state and output feedback adaptive critic control policy of nonlinear DT affine systems with disturbances. One important benefit of the HDP algorithm is that it can be implemented on an actor–critic structure without requiring the knowledge or identification/learning of the system dynamics. This makes HDP algorithm a promising model-free method for control design of real industrial processes. However, very limited works have focused on the application of the HDP method to PDE systems until present. Following the RTD framework, some control approaches without system models were developed. As such, the approximate dynamic programming (ADP)-based approaches were used to synthesize neural optimal control policies for DT chemical reactor process and DT temperature profile of a high-speed aerospace vehicle. Zhang et al. proposed a nonlinear observer for PDE process and then used it in nonlinear receding horizon control. But as discussed in the previous paragraph, the RTD framework is unsuitable for hyperbolic PDE systems. To the best of our knowledge, the theory of HDP algorithm for optimal control of linear continuous-time (CT) hyperbolic PDE systems from the DTR framework has not yet been established, which motivates the present study.

In this work, we consider the optimal control problem of a class of linear CT hyperbolic PDE systems with partially unknown system dynamics. This problem can be reduced to finding a solution of the space-dependent Riccati differential equation (SDRDE), but it requires the full system dynamics. Therefore, an HDP scheme is proposed for learning an optimal control policy from the DTR framework, and its convergence is rigorously proved. The HDP algorithm learns the solution of the SDRDE online, by measuring the system state without requiring the knowledge of the internal system dynamics. For implementation purposes, we develop an approximate approach based on the method of weighted residuals (MWR). Finally, the developed control method is applied to a steam-jacketed tubular heat exchanger to show its effectiveness. The main contributions of this work are briefly summarized as follows:

1. The HDP algorithm is proposed to synthesize the optimal control policy for linear hyperbolic PDE systems from the DTR framework. In detail, most of the existing RL-based methods were considered for finite-dimensional systems, while this work establishes the theory of the HDP scheme for linear infinite-dimensional systems. The convergence of the HDP algorithm is established rigorously by proving that it generates a nondecreasing sequence, which uniformly converges to the solution of SDRDE.

2. Most of the existing works focused on the DT ODE systems, while this work considers the CT hyperbolic PDE systems.

3. An online approximate approach based on the MWR is derived for implementation of the proposed HDP method.

The rest of the work is organized as follows: Section 2 presents the problem description. The HDP algorithm is proposed and its convergence is proven in Section 3. An MWR-based implementation method is developed in Section 4. Simulation studies are conducted, and a brief conclusion is drawn in Sections 5 and 6, respectively.

1. Notations. \( \mathbb{R} \), \( \mathbb{R}^n \) and \( \mathbb{R}^{\infty \times m} \) are the set of real numbers, the \( n \)-dimensional Euclidean space and the set of all real \( n \times m \) matrices, respectively. \( \| \cdot \| \) and \( \langle \cdot, \cdot \rangle_{\mathbb{R}^n} \) denote the Euclidean norm and inner product for vectors, respectively. Identity matrix, of appropriate dimensions, is denoted by \( I \). The superscript “\( -1 \)” is used for the transpose of a vector or a matrix. For a symmetric matrix \( M, M > (\geq, <, \leq) 0 \) means that it is positive definite (positive semidefinite, negative definite, negative semidefinite, respectively). The space-varying matrix function \( M(z), z \in [z_1, z_2] \) is positive definite (positive semidefinite, negative definite, negative semidefinite, respectively), if \( M(z) > (\geq, <, \leq) 0 \) for each \( z \in [z_1, z_2] \). The symbol \( * \) is used as an ellipsis in matrix expressions that are induced by symmetry, e.g.,

\[
[P(z)A(z) + *] = P(z)A(z) + A^T(z)P(z), \quad z \in [z_1, z_2]
\]

\( \mathcal{H}_n^+ = L^+_n([z_1, z_2]; \mathbb{R}^n) \) is an infinite-dimensional Hilbert space of \( n \)-dimensional square integrable vector functions \( \omega(z) \in \mathcal{H}_n^+, z \in [z_1, z_2] \subseteq \mathbb{R} \) equipped with the inner product and norm:

\[
\langle \omega_1(z), \omega_2(z) \rangle = \int_{z_1}^{z_2} \langle \omega_1(z), \omega_2(z) \rangle_{\mathbb{R}^n} \, dz \text{ and } \| \omega \|_1 = \langle \omega(z), \omega(z) \rangle_{\mathbb{R}^n}^{1/2},
\]

where \( \omega_1 \) and \( \omega_2 \) are any two elements of \( \mathcal{H}_n^+ \).

2. PROBLEM DESCRIPTION

Let us consider the following linear first-order CT hyperbolic PDE systems in one spatial dimension with a state-space description of the form:

\[
\frac{\partial y(z, t)}{\partial t} = A_1 \frac{\partial y(z, t)}{\partial z} + A_2(y(z, t)) + B(z)u(z, t)
\]

subject to the boundary condition,

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

and the initial condition,

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

where \( z \in [z_1, z_2] \subseteq \mathbb{R} \) and \( t \in [0, \infty) \) denote spatial position and time, respectively, \( y(z, t) = \{y_1(z, t) \ldots y_m(z, t)\}^T \in \mathbb{R}^m \) is the state, \( y_0(z) \in \mathbb{R}^m \) is the initial state, and \( u(z, t) = \{u_1(z, t) \ldots u_m(z, t)\}^T \in \mathbb{R}^m \) is the manipulated control input. \( A_1 \) is a real known diagonal \( n \times n \) matrix, \( A_2(z) \) and \( B(z) \) are real continuous space-varying matrix functions of appropriate dimensions. \( A_1(\partial y(z, t)/\partial z) + A_2(z) y(z, t) \) is referred to as internal system dynamics that is unknown in this work.

Remark 1. It is of practical interest to consider hyperbolic PDE systems with the unknown internal dynamics. For many real industrial processes, the accurate modeling and identification are impossible or too costly to conduct. One of their most prominent features is the presence of vast volume of data
accompanied by the lack of an effective process physical model that can support control. It is observed that the matrix $A_t$ in eq 1 is assumed to be diagonal. This is representative to some extent and is widely used in the literatures.\(^4,8,12,48\) Typical examples satisfying this assumption include plug flow reactors, fixed-bed reactors, and steam-jacketed tubular heat exchangers where the elements of $A_t$ are the fluid velocities.\(^{10,48,49}\)

We consider the following linear distributed state feedback control law:

$$u(z, t) = K(z)y(z, t)$$

where $K(z)$ is the control gain to be determined, which is a real continuous space-varying $m \times n$ matrix function defined on the interval $[z, \bar{z}]$. The substitution of the control law of eq 4 into eq 1 results in the following closed-loop PDE system:

$$\frac{\partial y(z, t)}{\partial t} = A_t\frac{\partial y(z, t)}{\partial z} + \bar{A}_t(z)y(z, t)$$

where

$$\bar{A}_t(z) \triangleq A_t(z) + B(z)K(z)$$

For convenience, we denote $y(\cdot, t) \triangleq y(z, t), z \in [z, \bar{z}]$, and $u(\cdot, t) \triangleq u(z, t), z \in [z, \bar{z}]$, and denote $M(\cdot) \triangleq M(z), z \in [z, \bar{z}]$ for some space-varying matrix function $M(z), z \in [z, \bar{z}]$. Define the following infinite-horizon linear quadratic (LQ) cost functional:

$$V_u(y(\cdot, t)) \triangleq \int_0^{+\infty} \left( y(\cdot, t), Q(\cdot)y(\cdot, t) \right) dt$$

where $Q(z) > 0$ and $R(z) > 0$.

The optimal control problem of the PDE system of eqs 1–3 here is to find a control law of eq 4 such that the LQ cost functional $V_u(y(\cdot, \cdot))$ is minimized, i.e.,

$$u(z, t) = u^*(z, t) \triangleq \arg \min_u V_u(y(\cdot, t))$$

The above LQ optimal control problem can be converted to solve a SDRDE, which is given in the following lemma.

**Lemma 1.** Consider the PDE system of eqs 1–3 with the state feedback control law of eq 4 and cost functional $V_u(y(\cdot, t))$. Suppose that $A_t < 0$. Let $P^*(z) > 0$ be a diagonal real continuous space-varying $n \times n$ matrix function defined on interval $[z, \bar{z}]$, and $P^*(\cdot) = 0$. If $P^*(z)$ satisfies the SDRDE

$$A_t\frac{dP^*(z)}{dz} = [P^*(z)A_t(z) + B(z)K(z)] + Q(z)$$

$$- P^*(z)B(z)R^{-1}(z)B^T(z)P^*(z), \quad P^*(\cdot) = 0$$

then $P^*(z)$ is the unique solution of the SDRDE 9 and the solution of the optimal control problem (i.e., eq 8) is as follows:

$$u^*(z, t) = K^*(z)y(z, t)$$

where

$$K^*(z) \triangleq -R^{-1}(z)B^T(z)P^*(z)$$

and the minimum cost is as follows:

$$V_u(y(\cdot, \cdot)) = \langle y(\cdot, \cdot), P^*(\cdot)y(\cdot, \cdot) \rangle$$

**Proof.** See ref 8.

**Remark 2.** It is worth mentioning that the SDRDE condition in Lemma 1, suggested by Aksikas et al.\(^8\) is only sufficient for the existence of the optimal control law. In the practical application, one can choose appropriate weighted matrices $Q(z)$ and $R(z)$ such that the solution of SDRDE 9 exists. Since this paper focuses on the development of online method when internal system dynamics is unknown, we leave theoretical analysis for the existence of solution of SDRDE 9 for future investigation.

Lemma 1 shows that the optimal control policy of eq 10 requires the solution of the SDRDE 9 for matrix $P^*(z)$. Aksikas et al.\(^8\) presented a model-based approach, which needs the full knowledge of the PDE model. Thus, it cannot be used for the case that internal system dynamics is unknown. Moreover, the modeling and identification procedures for the dynamics of PDE systems are often complicated. For the purpose of obtaining optimal controllers that minimize LQ cost functional $V_u(y(\cdot, \cdot))$ without making use of an internal system model, we propose an HDP algorithm for online solving the SDRDE 9. This means that as long as enough measuring information is available, the HDP algorithm can learn the optimal control policy online and avoid the modeling and identification procedures.

### 3. HDP ALGORITHM FOR OPTIMAL CONTROL DESIGN AND CONVERGENCE ANALYSIS

In this section, we will present the main results of this paper. In order to learn the optimal control policy (i.e., eq 10) of the PDE system, the HDP algorithm is proposed for learning SDRDE 9 online by measuring system state without requiring the model of internal system dynamics. Consequently, the convergence of the HDP algorithm is rigorously proved. We first give the procedure of the HDP algorithm as follows:

**Algorithm 1** (HDP algorithm).

**Step 1:** Let initial cost function $V^{(0)}(y(\cdot, t)) = 0$ and $i = 0$.

**Step 2:** Update $u^{(i)}(z, t)$ by

$$u^{(i)}(z, t) = \arg \min_u \left( \int_0^{\infty} \Xi(\tau)dt + V^{(i)}(y(\cdot, t + \Delta t)) \right)$$

where $\Xi(\tau) \triangleq \langle y(\cdot, \cdot), Q(\cdot)y(\cdot, \cdot) + u(\cdot, \cdot), R(\cdot)u(\cdot, \cdot) \rangle$.

**Step 3:** By using control policy $u^{(i)}(z, t)$, compute the state-value function $V^{(i+1)}(y(\cdot, t))$ with

$$V^{(i+1)}(y(\cdot, t)) = \min_u \left( \int_0^{\infty} \Xi(\tau)dt + V^{(i)}(y(\cdot, t + \Delta t)) \right)$$

$$= \int_0^{\infty} \Xi(\tau)dt + \int_0^{\infty} V^{(i)}(y(\cdot, t + \Delta t))dt$$

where $\Xi(\tau) \triangleq \langle y(\cdot, \cdot), Q(\cdot)y(\cdot, \cdot) + u^{(i)}(\cdot, \cdot), R(\cdot)u^{(i)}(\cdot, \cdot) \rangle$.

**Step 4:** Set $i = i + 1$. Go to Step 2 and continue.

**Remark 3.** Observe from Algorithm 1 that the HDP method does not require the knowledge of the internal dynamics of the PDE system of eqs 1–3, whose information is embedded in the online measurement of the system states $y(\cdot, t)$ and $y(\cdot, t + \Delta t)$, and evaluation of the cost $\int_0^{\infty} \Xi(\tau)dt$. The HDP algorithm involves two basis steps: policy improvement (Step 2) and policy evaluation (Step 3). The policy improvement is to help find a better control policy based on current state-value function. The policy evaluation step aims to compute the state-value function of an arbitrary control policy by using only one
Then, we parametrize follows from eqs 14 and 19 that,
\[ V(0)(y(\cdot, t)) = \min_u \int_t^{t + \Delta t} \Xi(t) \, dt \leq \int_t^{t + \Delta t} \hat{V}(0)(y(\cdot, t + \Delta t)) \]
\[ = \hat{V}(0)(y(\cdot, t)) \]
Second, suppose \( \hat{V}(i)(y(\cdot, t)) = \hat{V}(i)(y(\cdot, t)) \) for \( \forall y(\cdot, t) \), then it follows from eq 14 that,
\[ V(i+1)(y(\cdot, t)) = \min_u \int_t^{t + \Delta t} \Xi(t) \, dt \leq \int_t^{t + \Delta t} \hat{V}(i)(y(\cdot, t + \Delta t)) \]
\[ \leq \int_t^{t + \Delta t} \hat{V}(i)(y(\cdot, t)) \, dt + \hat{V}(i)(y(\cdot, t + \Delta t)) \]
\[ = \hat{V}(i)(y(\cdot, t)) \]

This completes the proof.

Theorem 1. Let \( P^0(z) \) be the solution of the SDRDE 9, \( P^0(z) \) = 0 for \( \forall z \in [z, z] \). If \( P^0(z)(i = 0,1,2,...) \) is the solution of eq 18, then,

1. \( P^0(z)(i = 0,1,2,...) \) converges uniformly to \( P^0(z) \) when \( i \to \infty \).

Proof. Let \( \{\mu^0\} \) be a sequence of control policies such that \( U^0(z,r) = U^0(z,r) \) on \( \tau \in [t + \Delta t], \). Assume that \( \hat{V}(0)(y(\cdot, t)) = 0 \) for \( \forall y(\cdot, t) \). Then, we have the following:

\[ \hat{V}(i)(y(\cdot, t)) = \int_t^{t + \Delta t} \hat{V}(i)(y(\cdot, t + \Delta t)) \]
\[ = \int_t^{t + \Delta t} \hat{V}(i)(y(\cdot, t)) \, dt \]
\[ \leq \int_t^{t + \Delta t} \hat{V}(i)(y(\cdot, t)) \, dt \]
\[ \leq \int_t^{t + \Delta t} \hat{V}(i)(y(\cdot, t)) \, dt + \hat{V}(i)(y(\cdot, t + \Delta t)) \]
\[ = \hat{V}(i)(y(\cdot, t)) \]

for \( \forall y(\cdot, t) \). Now, we will prove that \( \hat{V}(i)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \) for \( i = 0,1,2,... \) by induction. Because \( \hat{V}(0)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \), it follows from eq 14 that,

\[ V(i)(y(\cdot, t)) - \hat{V}(0)(y(\cdot, t)) = V(i)(y(\cdot, t)) - V(i+1)(y(\cdot, t)) \]
\[ = \int_t^{t + \Delta t} \Xi(t) \, dt \]
\[ \geq 0 \]

This means that \( \hat{V}(i)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \) is right for \( i = 0, \) i.e., \( \hat{V}(0)(y(\cdot, t)) \leq V(1)(y(\cdot, t)) \). Assume that \( \hat{V}(i)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \). Then subtracting eq 20 from eq 14 yields,

\[ V(i)(y(\cdot, t)) - \hat{V}(0)(y(\cdot, t)) = V(i)(y(\cdot, t)) - V(i+1)(y(\cdot, t)) \]
\[ \geq 0 \]

This implies that \( \hat{V}(i)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \) for \( i = 0,1,2,... \).

From Lemma 2, we have \( V(i)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \), thus,

\[ \hat{V}(i)(y(\cdot, t)) \leq V(i+1)(y(\cdot, t)) \leq V(i)(y(\cdot, t)) \]
\[ \leq V(i+1)(y(\cdot, t)) \]
\[ = \hat{V}(i)(y(\cdot, t)) \]

for \( \forall y(\cdot, t) \)
Thus, \( P^{(i+1)}(z) \geq P(z) \), \( i = 0, 1, 2, \ldots \), for all \( z \in \mathbb{R} \). This completes part (1) of Theorem 1.

(2) Let \( u(z, t) \) be an arbitrary stabilizing control law. Define a new sequence as follows:

\[
\hat{V}^{(i+1)}(y(t), t) \triangleq \int_t^{t+\Delta t} \Xi(r) \, dr + \hat{V}^{(i)}(y(t+\Delta t), t+\Delta t),
\]

where \( i = 0, 1, 2, \ldots \)

with \( \hat{V}^{(0)}(y(t), t) = 0 \) for all \( y(t) \). Similarly, we have the following:

\[
\hat{V}^{(i)}(y(t), t) \triangleq \int_t^{t+\Delta t} \Xi(r) \, dr + \hat{V}^{(i-1)}(y(t+\Delta t), t+\Delta t)
\]

The subtraction of eq (22) from eq (21) yields,

\[
\hat{V}^{(i+1)}(y(t), t) - \hat{V}^{(i)}(y(t), t) = \hat{V}^{(i)}(y(t, t+\Delta t)) - \hat{V}^{(i-1)}(y(t, t+\Delta t))
\]

\[
= \hat{V}^{(i)}(y(t, t+\Delta t)) - \hat{V}^{(i-1)}(y(t, t+2\Delta t))
\]

\[
\vdots
\]

\[
= \hat{V}^{(i)}(y(t, t+i\Delta t)) - \hat{V}^{(i-1)}(y(t, t+(i-1)\Delta t))
\]

\[
= \hat{V}^{(i)}(y(t, t+i\Delta t)) - \hat{V}^{(i-1)}(y(t, t+(i-1)\Delta t))
\]

\[
= \hat{V}^{(i)}(y(t, t+k\Delta t))
\]

\[
\leq \sum_{k=0}^{\infty} \hat{V}^{(i)}(y(t, t+k\Delta t))
\]

\[
= \sum_{k=0}^{\infty} \int_{t+(k+1)\Delta t}^{t+(k+1)\Delta t} \Xi(r) \, dr + \hat{V}^{(i)}(y(t, t+\Delta t))
\]

\[
= \int_t^{\infty} \Xi(r) \, dr
\]

From Lemma 2 and eq (23), we have \( V^{(i)}(y(t), t) \leq \hat{V}^{(i)}(y(t), t) \leq V^{(i)}(y(t), t) \) for all \( y(t) \). This implies that the sequence \( \{V^{(i)}(y(t), t)\} \) is upper bounded by \( V^{(i)}(y(t), t) \). Moreover, part (1) of Theorem 1 shows that \( \{V^{(i)}(y(t), t)\} \) is a nondecreasing sequence. Since a bounded monotone sequence always has a limit, we denote it as \( V^{(\infty)}(y(t), t) \). By eq (24), we get \( V^{(\infty)}(y(t), t) = \lim_{i \to \infty} \hat{V}^{(i)}(y(t), t) \).

Now, we show \( P^{(\infty)}(z) = P^{(\infty)}(z) \), i.e., \( P^{(\infty)}(z) \) satisfies the SDRDE. It follows from eq (17) that the control law associated to \( P^{(\infty)}(z) \) is as follows:

\[
u^{(\infty)}(z, t) = K^{(\infty)}(z)y(z, t)
\]

where \( K^{(\infty)}(z) \triangleq -R^{-1}(z)B^{T}(z)P^{(\infty)}(z) \). The closed-loop system under control law of eq (24) is as follows:

\[
\frac{dy(t, t)}{dt} = A_{i}y(t, t) + A_{i}^{\infty}(z)y(z, t)
\]

where,

\[
A_{i}^{\infty}(z) \triangleq A_{i}(z) + B(z)K^{(\infty)}(z)
\]

Differentiating \( (y(t, t), P^{(\infty)}(z)y(t, t)) \) with respect to time \( t \) along the state trajectories of the closed-loop system of eq (25), we get,

\[
\frac{d}{dt} \left( y(t, t), P^{(\infty)}(z)y(t, t) \right)
\]

\[
= \left( \frac{dy(t, t)}{dt}, P^{(\infty)}(z)y(t, t) \right)
\]

\[
+ \left( y(t, t), \frac{dP^{(\infty)}(z)}{dz} \frac{dy(t, t)}{dt} \right)
\]

\[
= \int_{\Xi} \left( \frac{dy(t, t)}{dz} \right)^{T} A_{i}^{T}P^{(\infty)}(z)y(z, t) dz
\]

\[
+ \int_{\Xi} y^{T}(z, t)P^{(\infty)}(z)A_{i} \frac{dy(t, t)}{dz} dz
\]

\[
+ \left( A_{i}^{T}(z)y(t, t), P^{(\infty)}(z)y(t, t) \right)
\]

From eq (26) and eq (27), we have

\[
\int_{\Xi} y^{T}(z, t)P^{(\infty)}(z)A_{i} \frac{dy(t, t)}{dz} dz = 0
\]

(see eq (15)), we have,

\[
y^{T}(z, t)P^{(\infty)}(z)A_{i}y(z, t)_{t=\Xi} = 0
\]

Since both \( P^{(\infty)}(z) \) and \( A_{i} \) are diagonal matrices, the equality \( P^{(\infty)}(z)A_{i}P^{(\infty)}(z) \) clearly holds. Then, it follows from eqs (27)–(29) that,
By eq 31, we have,
\[
\mathbb{A}_i \frac{dP^{(\infty)}(z)}{dz} = \left[ P^{(\infty)}(z)A_1(z) + * \right] + Q(z) - P^{(\infty)}(z)B(z)R^{-1}(z)B^T(z)P^{(\infty)}(z)
\]
which means that \( P^{(\infty)}(z) \) is the solution of SDRDE 9. By uniqueness, we have \( P^{(\infty)}(z) = P^*(z) \). The proof is completed.

It is seen from Theorem 1 that Algorithm 2 generates a nondecreasing sequence \{\( P^{(i)}(z) \)\}, which uniformly converges to the solution of SDRDE 9.

4. METHOD OF WEIGHTED RESIDUALS FOR HDP ALGORITHM

To compute the solution of the SDRDE 9 with the HDP algorithm, we need to solve the iterative eq 18 at each step. To this end, we derive a MWR based approximate approach on a set \( \mathcal{D} \subset \mathcal{H}^n \) such that \( y(t) \in \mathcal{D} \), to estimate the solution of eq 18.

Note that \( P^{(i)}(z) \) is a diagonal matrix, i.e.,
\[
P^{(i)}(z) = \text{diag}[p^{(i)}_1(z), \ldots, p^{(i)}_n(z)], \quad i = 0, 1, 2, \ldots
\]
Then eq 18 can be rewritten as follows:
\[
\sum_{j=1}^{n} \int_{T} \hat{p}^{(i+1)}(z) y^2_j(z, t) \, dz \\
= \sum_{j=1}^{n} \int_{T} \hat{p}^{(i)}(z) y^2_j(z, t + \Delta t) \, dz + \int_{T} \Xi^{(i)}(\tau) \, d\tau,
\]
i = 0, 1, 2, \ldots

(33)

From the well-known high-order Weierstrass approximation theorem,\(^{31}\) a continuous function can be uniformly approximated to any degree of accuracy by a set of complete linear independent basis functions. Let \( \psi_j(z) = [\psi_{j,1}(z) \ldots \psi_{j,n}(z)]^T \), \( j = 1, 2, \ldots, n \) be basis function vectors for approximating \( \hat{p}^{(i)}(z) \), where \( N_j \) is the number of basis functions in \( \psi_j(z) \). Then, the following expression is used to approximate \( \hat{p}^{(i)}(z) \):
\[
\hat{p}^{(i)}_{j}(z) = (w^{(i)}_{j})^T \psi_j(z) = \psi_j^T(z)w^{(i)}_{j}, \quad j = 1, 2, \ldots, n
\]
(34)

where \( w^{(i)}_j = [w^{(i)}_{j,1} \ldots w^{(i)}_{j,n}] \) is the weight vector. Thus, we have,
\[
\sum_{j=1}^{n} \int_{T} \hat{p}^{(i+1)}(z) y^2_j(z, t) \, dz \\
= \sum_{j=1}^{n} \int_{T} (w^{(i+1)}_j)^T \psi_j(z) y^2_j(z, t) \, dz \\
= (w^{(i+1)})^T \sum_{j=1}^{n} \int_{T} \psi_j(z) y^2_j(z, t) \, dz \\
= (w^{(i+1)})^T \theta(\mathbf{y}(t), t)
\]
(35)

where \( w^{(i+1)} = [w^{(i+1)}_1 \ldots w^{(i+1)}_n]^T \) and,
\[
\theta(\mathbf{y}(t), t) \triangleq \int_{T} \psi_j(z) y^2_j(z, t) \, dz \quad \text{and} \quad \int_{T} \psi_j(z) y^2_j(z, t) \, dz \int_{T} \psi_j(z) y^2_j(z, t) \, dz
\]
Define \( \Xi^{(i)}(\tau) \triangleq (\mathbf{y}(\tau), Q(\mathbf{y}(\tau))) + (\hat{u}^{(i)}(\tau), R(\hat{u}^{(i)}(\tau))) \), where,
Online collected, the control policy is approximately updated by eq 36, i.e.,
\[
\dot{\mathbf{w}}^{(i+1)} = \Pi^{(i)}(t, t + \Delta t) + \hat{\Xi}^{(i)}(t) \quad \text{for} \quad i = 0, \ldots, n
\]
\[
\Pi^{(i)}(t, t + \Delta t) = \int_{f_y(t)} P(y(t, t + \Delta t)) dy
\]
where \( P(y(t)) \) is chosen to be large enough, \( \hat{\Xi}^{(i)}(t) \) is the error vector, and \( \Pi^{(i)}(t, t + \Delta t) \) is the Jacobian matrix.

Remark 4. The convergence of the MWR in Algorithm 3 was provided in ref 52, which shows that if the number of basis function vectors \( \psi_j(z) \), \( j = 1, 2, \ldots, n \) is chosen to be large enough, then \( P^{(0)}(z) \) (each element of which is computed with eq 34) can approximate the solution \( P^{(0)}(z) \) of eq 18 in Algorithm 2 to any degree of accuracy. However, Theorem 1 shows that the HDP algorithm for online optimal control design of the PDE system of eqs 1–3, as follows:

Algorithm 3.

Step 1: Select basis function vectors \( \psi_j(z) \), \( j = 1, 2, \ldots, n \). Give an initial weight vector \( \mathbf{w}^{(0)} = 0 \) and let \( i = 0 \).

Step 2: Update the control policy \( \mathbf{w}^{(i)} \) with eq 36 at time instant \( t + \Delta t \).

Step 3: With the control policy \( \mathbf{w}^{(i)} \), collect \( N \) sample data sets along state trajectories of the closed-loop PDE system, and evaluate the associated matrices \( \Pi^{(i)} \) and \( \hat{\Xi}^{(i)} \) during time interval \( [t, t+\Delta t] \). Compute \( \mathbf{w}^{(i+1)} \) via eq 38 at time instant \( t = (i+1) \Delta t \).

Step 4: Set \( i = i+1 \). If \( \| \mathbf{w}^{(i+1)} - \mathbf{w}^{(i)} \|_\infty \leq \varepsilon \) (\( \varepsilon \) is a small positive real number), stop iteration and the weights remains invariable, else, go to Step 2 and continue.

5. APPLICATION TO STEAM-JACKET TUBULAR HEAT EXCHANGER

To illustrate the effectiveness of the developed HDP algorithm, we apply it to online learn the optimal control law for a steam-jacket tubular heat exchanger. The system model is given by the following:
\[
\frac{dT}{dt} = -\frac{\partial T}{\partial z} - \frac{hA}{\rho C_p}(T - T_0)
\]
subject to the boundary condition,
\[
T(0, t) = T_f, \quad t \in [0, +\infty)
\]
and initial condition,
\[
T(z, 0) = T_0(z), \quad z \in [0, L]
\]
In the above model, $T$ denotes the temperature in the tubular heat exchanger, $T_w$, $T_f$, and $T_0$ denote steam-jacket temperature, heat exchanger inlet constant temperature and initial temperature, respectively. In addition, $t$, $z$, and $L$ denote the independent time and space variables, and the length of the exchanger, respectively. By taking change of variables as follows:

$$y \triangleq T - T_f, \quad u \triangleq T_w - T_f, \quad y_0 \triangleq T_0 - T_f$$

and

$$a(z) \triangleq \frac{hA}{\rho C_p}$$

the equivalent representation of the model of eq 39 is obtained as follows:

$$\frac{\partial y}{\partial t} = -v \frac{\partial y}{\partial z} - a(z)y + a(z)u$$

subject to the boundary condition,

$$y(0, t) = 0, \quad t \in [0, +\infty)$$

and initial condition

$$y(z, 0) = y_0(z), \quad z \in [0, L]$$

The control objective is to determine the adjustment in the steam-jacket temperature $T_w$ (through a steam inlet valve), such that the exchanger temperature $T$ approaches the desired temperature profile $T_d(z)$ and achieves a minimum cost. Here, we are interested in the constant profile $T_d$. Since $T_d(0) = T_d$, then $T_d(z) = T_d$ for all $z \in [0, L]$. Thus, the desired profile for system of eq 41 is $y_d \triangleq T_d - T_f = 0$. According to eq 9, the optimal control of system of eq 41 is to solve the following SDRDE for $p^*(z)$:

$$-v \frac{\partial p^*(z)}{\partial z} = -a(z)p^*(z) - p^*(z)a(z) + q(z)$$

$$- p^*(z)a(z)r^{-1}(z)a(z)p^*(z)$$

$$p^*(z) > 0, \quad z \in [0, L], \quad p^*(L) = 0$$

The system parameters are set as follows: $v = 0.8$, $L = 1$, $a(z) = 4 - \exp(-2z/L)$, and $T_d = 340$ (then, $T_d = 340$). Let $y_0(z) = T_j \sin(5\pi z/L)$ (i.e., the initial temperature profile of exchanger is assumed to be $T_0(z) = T_f + T_d \sin(5\pi z/L)$). As mentioned in Remark 2, we need to select appropriate weighted matrices $Q(z) > 0$ and $R(z) > 0$. For simplicity, the weighted matrices in cost functional of eq 7 are chosen as $Q(z) = q(z) = 1$ and $R(z) = r(z) = 1$, $z \in [0, L]$ without loss of generality. We select 8 (i.e., $N = N_z = 8$) basis functions as $\psi_k(z) = \sin(0.8k(L - z)\pi/L)$, $k = 1,...,8$ for approximating $p^*(z)$. In Algorithm 3, select the value of stop criterion $\varepsilon = 10^{-5}$, the sampling step size in space $\delta = 0.02$ (i.e., $N_z = 50$) and in time $\delta t = 0.1(s)$. In each iterative step, after collecting 10 (i.e., $N = 10$) system state sets, eq 38 is used to update weights, that is, the weights is updated every one second (i.e., $\Delta t = 1(s)$).

Applying the developed HDP algorithm (i.e., Algorithm 3) for online learning the solution of SDRDE 42, Figure 1 shows the weights at each iterative step, where we observe that the weight vector converges to

$$[0.1335 \quad 0.0050 \quad 0.0198 \quad 0.0149 \quad -0.0029 \quad 0.0120 \quad -0.0030 \quad 0.0045]^T$$

with accuracy $\varepsilon$ at $i = 15$ iteration (i.e., at time instant $t = 15(s)$). Then, stop the HDP iteration and remain $\hat{p}(z)$ invariable from this instant on. Figure 2 gives the final profile of $\hat{p}(z)$ (denoted as $p^*(z)$). By using the convergent solution of the SDRDE for the closed-loop simulation, Figures 3 and 4 show the jacket’s temperature profile (i.e., control action) and
exchanger’s temperature profile (i.e., state of PDE system 39), respectively. From the simulation results, we find that the developed HDP algorithm is efficient for solving the optimal control problem of the steam-jacket tubular heat exchanger.

6. CONCLUSIONS

In this work, we have proposed an HDP algorithm from the DTR framework, for solving the optimal control problem of linear CT hyperbolic PDE systems. The HDP algorithm learns the solution of the SDRDE online without the requirement of the internal system dynamics. The convergence of the HDP algorithm is rigorously established by constructing a non-decreasing sequence that uniformly converges to the solution of SDRDE. To solve the iterative equation in the HDP algorithm, we have developed an approximate approach based on MWR. Finally, the effectiveness of the developed HDP algorithm is demonstrated by conducting simulation studies on a steam-jacketed tubular heat exchanger.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: whn@buaa.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by the National Basic Research Program of China (973 Program) (2012CB720003); in part by the National Natural Science Foundation of China under Grant Nos. 61074057, 61121003, and 91016004. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

■ REFERENCES


Figure 4. The evolution of the exchanger’s temperature profile.
(36) Li, X.; Yong, J. Optimal Control Theory for Infinite Dimensional Systems; Birkhäuser: Boston, 1995; p 338.