

# Boundary Control of Linear One-Dimensional Parabolic PDE using Neuro-Dynamic Programming

B. Talaei<sup>1</sup>, S. Jagannathan<sup>1</sup> and J. Singler<sup>2</sup>

**Abstract**—This paper develops a neuro-dynamic programming (NDP) based near optimal boundary control of distributed parameter systems (DPS) governed by linear one-dimensional parabolic partial differential equations (PDE) under Dirichlet boundary control condition. The structure of the optimal cost functional is defined as an extension of its definition from lumped parameter systems (LPS) but for the infinite dimensional state space. Subsequently, the Hamilton-Jacobi-Bellman (HJB) equation is formulated in the infinite dimensional domain without using any model reduction. Since solving the HJB equation for the exact optimal value functional is burdensome, a radial basis function network (RBF) is subsequently proposed to achieve a computationally feasible solution online and in a forward-in time manner. The optimal value functional is tuned by using conventional adaptive laws such that the HJB equation error is minimized and accordingly the optimal control policy is derived. Ultimate boundedness (UB) of the closed-loop system is verified by using the Lyapunov theory. The performance of proposed controller is successfully verified on an unstable diffusion reaction process.

## I. INTRODUCTION

Distributed parameter systems (DPS) are a major part of dynamical systems with wide range of applications [1]–[2]. As the name suggests, DPS arise in environments where system behavior changes continuously throughout the space. Similar to case of lumped parameter systems (LPS), the common attributes of a controller design for DPS should include: 1. a simple and reliable design; 2. feasible real-time implementation; 3. being robust to disturbances and modeling errors; and 4. closed-loop stability. However, the major challenge in the control of DPS when compared to a LPS is the infinite dimensional nature of state space modeled by partial differential equations (PDE). This characteristic makes control design qualities difficult to reach in contrast to finite set of ordinary differential equations (ODE) in LPS.

The successful results of linear optimal control of LPS motivated researchers to develop the operator theory for optimal control of DPS. The work was then extended to boundary control [3] where the design is performed in the infinite dimensional setting. However, a closed-form solution requires solving the operator Riccati equations which is significantly more time consuming in the infinite dimensional domain for DPS. Subsequently, various other approaches under the general category of *optimize-then-discretize* control were proposed. In particular, boundary

actuation received special attention in the previous decade due to *backstepping* approach [4] wherein the controller is developed through conventional calculus instead of operator theory. The backstepping approach further enhances the stability study through conventional calculus in the original PDE domain.

The advances in neuro-dynamic programming (NDP) on the other hand, and its successful forward-in-time implementation of optimal control policy for LPS [5], encouraged researchers to investigate application of this method to DPS [6]–[7]. However, due to the difficulty of designing the controller in infinite dimensional domain and with PDE dynamics, the DPS system was discretized into an approximate finite dimensional state space and subsequently the well-developed NDP algorithms [8] were utilized for this reduced order model. Design of a controller using a reduced order model is broadly categorized as *discretize-then-optimize* [9] method. The benefit of this design approach is the possibility of using either the state of the art NDP or other feedback suboptimal control designs. However, its primary limitation is the degraded performance due to a reduced order finite dimensional model of the system.

In contrast, in this paper, a new NDP based boundary control of DPS governed by linear parabolic PDE under Dirichlet boundary condition is introduced. Unlike previous NDP based control methods for DPS [6]–[7], no model reduction is utilized prior to the control design. The controller development is based on obtaining the Hamilton Jacobi Bellman (HJB) equation and the optimal boundary control policy in the infinite dimensional domain through calculus. The value functional is then approximated in order to design a suboptimal controller. The proposed suboptimal controller is the first to be designed according to the NDP for a DPS without any model reduction.

The development of the controller is as follows. First, the HJB equation is formulated in the infinite dimensional setting based on a novel definition of value functional. Subsequently, the optimal control policy is derived by using necessary conditions of optimality. Consequently, the value functional is estimated in a forward-in-time manner and the approximated optimal control policy is derived by using the estimation of value functional. Computationally efficient adaptive laws are proposed for online tuning of value functional weights. Eventually, Lyapunov analysis is utilized to demonstrate the closed-loop stability of the feedback system. Simulation results confirm the effectiveness of the proposed scheme on a diffusion-reaction process.

Throughout the paper,  $\|\cdot\|$  stands for Euclidean norm

<sup>1</sup>Department of Electrical and Computer Engineering, <sup>2</sup>Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, MO, 65409. E-mail: behzad.talaei@mst.edu, sarangap@mst.edu, singlerj@mst.edu. Research supported in part by NSF grant ECCS#1128281 and Intelligent Systems Center.

for vectors or Frobenius norm for matrices. We recall the inequality  $\|\cdot\|_2 \leq \|\cdot\|_F$  where  $\|\cdot\|_2$  and  $\|\cdot\|_F$  represent the induced 2 norm and the Frobenius norm, respectively. The  $\mathcal{L}_2$  norm is also defined as  $\|\cdot\|_{\mathcal{L}_2} = \left(\int_0^l \|\cdot\|^2 dz\right)^{\frac{1}{2}}$ .

The rest of the paper is organized as follows. In Section II the class of DPS under consideration is described and the state feedback optimal control approach is explained. Section III demonstrates the simulation results and Section IV provides the conclusion of the paper.

## II. ADAPTIVE OPTIMAL CONTROL OF DPS MODELED BY PARABOLIC PDE

In this paper, a DPS governed by a linear one-dimensional parabolic PDE dynamics expressed as

$$\frac{\partial x(z,t)}{\partial t} = \frac{\partial^2 x(z,t)}{\partial z^2} + \lambda(z)x(z,t), \quad (1)$$

is considered, where  $x(z,t) \in \mathcal{H}_2[0,l]$  is the system state with  $z \in [0,l]$  being the spatial variable,  $l > 0$ ,  $t$  representing time and  $\mathcal{H}_2$  being the Sobolev space of second order;  $\frac{\partial x}{\partial t}$  is the time derivative of  $x$ ,  $\frac{\partial^2 x}{\partial z^2}$  is its second spatial derivative and  $\lambda(\cdot) \in C_1([0,l])$  with  $C_1$  being the space of continuously differentiable functions. Considering the above PDE dynamics, in the following, the NDP controller is designed under Dirichlet boundary actuation condition.

Consider the linear DPS (1) with Dirichlet boundary control at  $z = l$  and general Robin boundary condition at  $z = 0$  as

$$x(l) = gu, \quad \frac{\partial x}{\partial z}\Big|_{z=0} = -hx(0). \quad (2)$$

where  $u$  is the control input and  $g, h \in \mathfrak{R}$  are known constants. The objective is to design a controller to minimize the following infinite horizon cost functional  $\bar{V}$  given by

$$\bar{V}(x,u) = \int_{t_0}^{\infty} \{Q(x) + ru^2\} dt, \quad (3)$$

where  $r$  is a positive constant and  $Q(x)$  is a positive definite function. If the system state is a finite dimensional  $n \times 1$  vector  $x_f$ , it is well-known that  $Q(x_f)$  could be defined in quadratic form as

$$Q_L(x_f) = x_f^T Q_L x_f = \sum_{i=1}^n \sum_{j=1}^n x_{f_i} Q_{L_{ij}} x_{f_j} \quad (4)$$

where  $Q_L$  is a positive definite  $n \times n$  kernel matrix,  $x_{f_i}$  is the  $i$ th element of vector  $x_f$  and  $Q_{L_{ij}}$  is the entry of matrix  $Q_L$  at the  $i$ th row and  $j$ th column. However, in the case of DPS, since there are infinitely many states  $[x(z_0), x(z_1), \dots, x(z_k), \dots]^T$   $z_i \in [0,l]$  [12] that are continuous in the spatial domain, the finite dimensional summations in (4) should be substituted by integrals. Therefore, intuitively, taking  $s, z \in [0,l]$  as continuous spatial variables for a surface kernel function  $q(s,z)$  which correspond to discrete variables  $i, j$  as rows and columns of matrix  $Q_L$  in (4),  $Q(x)$  for DPS

can be specified equivalent to (4) in the following surface integral form as

$$Q(x) = \int_0^l \int_0^l x(s)q(s,z)x(z)dsdz, \quad (5)$$

where  $q(\cdot, \cdot) \in C_2([0,l], [0,l])$ . Note that  $q(\cdot, \cdot)$  is a two dimensional continuous kernel function that has the same role of the kernel matrix  $Q_L$  in finite dimensional definition (4).

**Remark 1:** In order to further clarify the definition of  $Q(x)$  in (5) as extension of finite dimensional definition (4), take  $0 = z_0 \leq z_1 \leq \dots \leq z_{n-1} \leq z_n = l$  and  $0 = s_0 \leq s_1 \leq \dots \leq s_{n-1} \leq s_n = l$  as two partitions for  $[0,l]$ . Assigning  $x_f = [x(z_1), \dots, x(z_i), \dots, x(z_n)]^T$  as a  $n \times 1$  finite dimensional subset of  $x$  and defining  $\Delta z_i = z_i - z_{i-1}$  and  $\Delta s_i = s_i - s_{i-1}$  for  $1 \leq i \leq n$ , one can express  $Q(x)$  in (5) as

$$\begin{aligned} Q(x) &= \int_0^l \int_0^l x(s)q(s,z)x(z)dsdz \\ &\approx \sum_{i=1}^n \sum_{j=1}^n x_{f_i} q(s_i, z_j) x_{f_j} \Delta s_i \Delta z_j \end{aligned} \quad (6)$$

where the Riemann approximated definition of integrals is used. By defining  $Q_{L_{ij}} = q(s_i, z_j) \Delta s_i \Delta z_j$ , equation (6) can be viewed analogous to definition of  $Q_L$  in (4). This implies that definition of  $Q(x)$  in (5) reduces to (4) for conventional finite dimensional state spaces. In order to proceed, the following assumption is necessary:

**Assumption 1:** The function  $q(\cdot, \cdot)$  is symmetric, i.e.  $q(s,z) = q(z,s)$ , and  $Q(x)$  is *positive definite*, i.e.  $Q(x) \geq q_{min} \|x\|_{\mathcal{L}_2}^2$  with  $q_{min} \geq b_q$  where  $b_q$  is a positive constant that will be defined later in the paper. Note that  $Q(x)$  can be designed arbitrarily.

The optimal cost functional is represented by  $\bar{V}^*(x,t)$ . Similar to NDP control design of linear LPS with quadratic cost function [10], if  $p(s,z) \in C_1([0,l], [0,l])$  is a symmetric kernel function,  $\bar{V}^*(x,t)$  can be represented by a surface integral as

$$\bar{V}^*(x,t) = \frac{1}{2} \int_0^l V^*(x(z,t), z) dz. \quad (7)$$

where  $V^*(x(z,t), z) = \int_0^l x(s,t)p(s,z)x(z,t)ds$ . By taking the current time interval  $[t, t + \delta t)$ ,  $\bar{V}$  in (3) can be represented in the recursive form as

$$\bar{V}(x,u,t) = \int_t^{t+\delta t} \{Q(x) + ru^2\} dt + \bar{V}(x,u,t + \delta t), \quad (8)$$

where  $\bar{V}(x,u,t + \delta t)$  is the cost to go from time  $t + \delta t$  to  $\infty$ . Hence, the optimal value functional can be represented as

$$\bar{V}^*(x,t) = \min_u \left\{ \int_t^{t+\delta t} \{Q(x) + ru^2\} dt + \bar{V}(x,u,t + \delta t) \right\}. \quad (9)$$

Now invoking the principle of optimality, Equation (9) becomes

$$\bar{V}^*(x,t) = \min_u \left\{ \int_t^{t+\delta t} \{Q(x) + ru^2\} dt + \bar{V}^*(x,t+\delta t) \right\}. \quad (10)$$

It is assumed that  $\bar{V}^*(x,t)$  is Gâteaux analytic [13], i.e. its differential with respect to infinitesimal change of state  $x$  exists in the direction of system trajectory in the neighborhood of  $(x(z,t),t)$ . If an integral functional  $Y(x) = \int_0^l y(x) dz$  with  $y(x)$  being a function of  $x$  is Gâteaux analytic, according to calculus of variations [13],  $Y(x+\delta x)$  with  $\delta x$  being an infinitesimal change in  $x$  can be represented by its first order approximation as

$$Y(x+\delta x) \approx Y(x) + \int_0^l y_x \delta x dz \quad (11)$$

where  $y_x = \frac{\partial y}{\partial x}$ . Moreover, according to (7)

$$\frac{1}{2} \int_0^l V_{x(s,t)}^* \delta x(s,t) ds = \frac{1}{2} \int_0^l V_{x(z,t)}^* \delta x(z,t) dz. \quad (12)$$

Therefore, revisiting (10),  $\bar{V}^*(x,t+\delta t)$  can be expressed in its first order approximation form as

$$\begin{aligned} \bar{V}^*(x,t+\delta t) &\approx \bar{V}^*(x,t) + \frac{1}{2} \int_0^l V_{x(s,t)}^* \delta x(s,t) ds \\ &+ \frac{1}{2} \int_0^l V_{x(z,t)}^* \delta x(z,t) dz + \frac{\partial \bar{V}^*}{\partial t} \delta t = \bar{V}^*(x,t) \\ &+ \int_0^l V_{x(z,t)}^* \delta x(z,t) dz + \frac{\partial \bar{V}^*}{\partial t} \delta t, \end{aligned} \quad (13)$$

where  $\delta x$  is an infinitesimal variation in  $x$  as a consequence of  $\delta t$  change in time,  $V_{x(s,t)}^* = \frac{\partial V^*}{\partial x(s,t)}$  and  $V_{x(z,t)}^* = \frac{\partial V^*}{\partial x(z,t)}$  are partial derivatives of  $V^*$  with respect to  $x(s,t)$  and  $x(z,t)$ , respectively and  $\frac{\partial \bar{V}^*}{\partial t}$  is partial derivative of  $\bar{V}^*$  with respect to  $t$ . Substituting approximation (13) into equation (10) and canceling  $\bar{V}^*(x,t)$  on both sides yields

$$0 = \min_u \{Q(x) + ru^2 + \int_0^l V_{x(z,t)}^* \delta x(z,t) dz + \frac{\partial \bar{V}^*}{\partial t} \delta t\}. \quad (14)$$

Dividing through by  $\delta t$ , letting  $\delta t \rightarrow 0$  and substituting dynamics (1) yield

$$\begin{aligned} \min_u \{Q(x) + ru^2 + \int_0^l V_{x(z,t)}^* \frac{\partial x}{\partial t} dz\} &= \min_u \{Q(x) \\ &+ ru^2 + \int_0^l V_{x(z,t)}^* [\frac{\partial^2 x}{\partial z^2} + \lambda(z)x(z,t)] dz + \frac{\partial \bar{V}^*}{\partial t}\} = 0. \end{aligned} \quad (15)$$

Since due to infinite time horizon, the cost functional  $\bar{V}^*$  as defined in (7) is only state dependent and not explicitly dependent on time,  $\frac{\partial \bar{V}^*}{\partial t} = 0$  [10]. Therefore the Hamilton Jacobi Bellman (HJB) equation can be represented by

$$0 = \min_u \{Q(x) + ru^2 + \int_0^l V_{x(z,t)}^* [\frac{\partial^2 x}{\partial z^2} + \lambda(z)x(z,t)] dz\}. \quad (16)$$

In [11] a similar result but with using a different approach is derived for parabolic semi-linear PDE. Subsequently, the Hamiltonian is defined as

$$H = Q(x) + ru^2 + \int_0^l V_{x(z,t)}^* [\frac{\partial^2 x}{\partial z^2} + \lambda(z)x(z,t)] dz. \quad (17)$$

Accordingly, using integration by parts twice, Hamiltonian can be represented as

$$\begin{aligned} H &= Q(x) + ru^2 + \int_0^l V_x^* [\lambda(z)x] dz + V_x^*(l) \frac{\partial x}{\partial z} \Big|_{z=l} - V_x^*(0) \\ &\frac{\partial x}{\partial z} \Big|_{z=0} - \left\{ \frac{\partial V_x^*}{\partial z} \Big|_{z=l} g u - \frac{\partial V_x^*}{\partial z} \Big|_{z=0} x(0) - \int_0^l \frac{\partial^2 V_x^*}{\partial z^2} x dz \right\}, \end{aligned} \quad (18)$$

where  $\frac{\partial V_x^*}{\partial z}$  and  $\frac{\partial^2 V_x^*}{\partial z^2}$  are the first and second spatial derivative of  $V_x^*$  with respect to  $z$ . Moreover,  $x(z,t)$  and  $V_{x(z)}^*(x,z)$  are represented by  $x(z)$  and  $V_x^*(z)$  for brevity. This convention will be used throughout the paper. Based on necessary conditions of optimality, in order for the control input to be minimizing for Hamiltonian (18), the Fréchet derivative of this equation with respect to  $u$  should be zero [13]. Therefore,

$$\frac{\partial H}{\partial u} = -g \frac{\partial V_x^*}{\partial z} \Big|_{z=l} + 2ru = 0 \Rightarrow u^* = \frac{1}{2r} g \frac{\partial V_x^*}{\partial z} \Big|_{z=l}. \quad (19)$$

Substituting the optimal control in equation (18), the HJB equation for DPS (1) under Dirichlet boundary condition (2) can be represented in the form,

$$\begin{aligned} H^* = 0 &= Q(x) + \int_0^l V_x^* [\lambda(z)x(z)] dz + V_x^*(l) \frac{\partial x}{\partial z} \Big|_{z=l} \\ &- V_x^*(0) \frac{\partial x}{\partial z} \Big|_{z=0} + \int_0^l \frac{\partial^2 V_x^*}{\partial z^2} x dz \\ &+ \frac{\partial V_x^*}{\partial z} \Big|_{z=0} x(0) - \frac{1}{4r} g^2 \frac{\partial V_x^*}{\partial z} \Big|_{z=l}. \end{aligned} \quad (20)$$

Since solving partial integro-differential equation (PIDE) (20) for the exact  $V^*(z)$  is too difficult and time consuming, the objective is to find a suitable structure for estimation of  $V^*(z)$ . The continuous kernel function  $p(s,z)$  in (7) can be interpreted as an infinite dimensional array of unknown parameters to be approximated. It is well-known that radial basis function networks (RBF) can estimate an unknown continuous multi-variable function with a large enough yet finite set of basis functions [14]. The function  $p(s,z)$  in (7) can be represented in RBF approximated form as

$$p(s,z) = W_V^T \phi(s,z) + \varepsilon_p(s,z), \quad (21)$$

where  $\phi(s, z) : [0, l] \times [0, l] \rightarrow [C_2]^{n_p}$  is a vector of  $n_p$  radial basis functions  $\phi_j((s, z) - (s_{0j}, z_{0j}))$ ,  $1 \leq j \leq n_p$ ,  $W_V \in \mathfrak{R}^{n_p}$  and  $\varepsilon_p(s, z)$  is the estimation error. It is assumed that the uniform norm of estimation error and its first and second spatial derivatives, i.e.  $\frac{\partial \varepsilon_p}{\partial z}$ ,  $\frac{\partial^2 \varepsilon_p}{\partial z^2}$  are bounded with the bounds  $\varepsilon_{pM}$ ,  $\varepsilon_{pzM}$  and  $\varepsilon_{pzzM}$ , respectively. Note that the estimation error  $\varepsilon_p$  can be reduced by increasing the number of neurons.

Hence,  $V^*(x, z)$  can be represented in approximated form of

$$\begin{aligned} V^*(x, z) &= \int_0^l x(s) W_V^T \phi(s, z) x(z) ds + \varepsilon_V(x, z) \\ &= W_V^T \Phi(x, z) + \varepsilon_V(x, z), \end{aligned} \quad (22)$$

where  $\Phi(x, z)$  is a  $n_p \times 1$  vector defined as

$$\Phi(x, z) = \int_0^l x(s) \phi(s, z) x(z) ds, \quad (23)$$

and  $\varepsilon_V(x, z) = \int_0^l x(s) \varepsilon_p(s, z) x(z) ds$ . Subsequently, the optimal value functional can be expressed as

$$\bar{V}^*(x) = W_V^T \int_0^l \Phi(x, z) dz + \varepsilon_{\bar{V}}, \quad (24)$$

with  $\varepsilon_{\bar{V}} = \int_0^l \varepsilon_V dz$ . Consequently, the optimal control can be rewritten as

$$u^* = \frac{1}{2r} g \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=l} W_V + \varepsilon_u, \quad (25)$$

with  $\varepsilon_u = \frac{1}{2r} g \varepsilon_{V_x} \Big|_{z=l}$ . The HJB equation becomes

$$\begin{aligned} H^* = 0 &= Q(x) + \int_0^l W_V^T \Phi_x [\lambda(z) x(z)] dz + W_V^T \Phi_x(l) \\ &\frac{\partial x}{\partial z} \Big|_{z=l} - W_V^T \Phi_x(0) \frac{\partial x}{\partial z} \Big|_{z=0} + \int_0^l W_V^T \frac{\partial^2 \Phi_x}{\partial z^2} x dz + \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=0} \\ &W_V x(0) - \frac{1}{4r} W_V^T \frac{\partial \Phi_x}{\partial z} \Big|_{z=l} g^2 \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=l} W_V + \varepsilon_H, \end{aligned} \quad (26)$$

where  $\Phi_x(z)$  denotes  $\Phi_x(x, z)$  for brevity and  $\varepsilon_H$  is derived as

$$\begin{aligned} \varepsilon_H &= \int_0^l \varepsilon_{V_x}(z) [\lambda(z) x(z)] dz + \varepsilon_{V_x}(l) \frac{\partial x}{\partial z} \Big|_{z=l} \\ &- \varepsilon_{V_x}(0) \frac{\partial x}{\partial z} \Big|_{z=0} + \int_0^l \frac{\partial^2 \varepsilon_{V_x}}{\partial z^2} x dz + \frac{\partial \varepsilon_{V_x}}{\partial z} \Big|_{z=0} x(0) \\ &- \frac{1}{4r} g^2 \left( \frac{\partial \varepsilon_{V_x}}{\partial z} \Big|_{z=l} \right)^2 - \frac{1}{2r} W_V^T \frac{\partial \Phi_x}{\partial z} \Big|_{z=l} g^2 \frac{\partial \varepsilon_{V_x}}{\partial z} \Big|_{z=l}. \end{aligned} \quad (27)$$

If the value functional is estimated as

$$\hat{V}^*(x) = \int_0^l \hat{V}^*(z) dz = \hat{W}_V^T \int_0^l \Phi(x, z) dz, \quad (28)$$

the approximated HJB becomes

$$\begin{aligned} \hat{H} &= Q(x) + \int_0^l \hat{W}_V^T \Phi_x [\lambda(z) x(z)] dz + \hat{W}_V^T \Phi_x(l) \frac{\partial x}{\partial z} \Big|_{z=l} \\ &- \hat{W}_V^T \Phi_x(0) \frac{\partial x}{\partial z} \Big|_{z=0} + \int_0^l \hat{W}_V^T \frac{\partial^2 \Phi_x}{\partial z^2} x dz \\ &+ \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=0} \hat{W}_V x(0) - \frac{1}{4r} \hat{W}_V^T \frac{\partial \Phi_x}{\partial z} \Big|_{z=l} g^2 \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=l} \hat{W}_V, \end{aligned} \quad (29)$$

and therefore, the control input can be represented by

$$\hat{u} = \frac{1}{2r} g \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=l} \hat{W}_V. \quad (30)$$

The value functional weight tuning law for Dirichlet boundary control is defined as

$$\dot{\hat{W}}_V = -\alpha_1 \frac{\omega_D}{\zeta_D^2} \hat{H} - \alpha_2 \hat{W}_V - \alpha_3 \hat{W}_V \|\hat{W}_V\|^2 - \alpha_4 \|x\|_{\mathcal{L}_2}^2 \hat{W}_V, \quad (31)$$

where

$$\begin{aligned} \omega_D &= \int_0^l \Phi_x [\lambda(z) x(z)] dz + \int_0^l \frac{\partial^2 \Phi_x}{\partial z^2} x dz \\ &+ \Phi_x(l) \frac{\partial x}{\partial z} \Big|_{z=l} - \Phi_x(0) \frac{\partial x}{\partial z} \Big|_{z=0} + \frac{\partial \Phi_x}{\partial z} \Big|_{z=0} x(0) \\ &- \frac{1}{2r} \frac{\partial \Phi_x}{\partial z} \Big|_{z=l} g^2 \frac{\partial \Phi_x^T}{\partial z} \Big|_{z=l} \hat{W}_V, \end{aligned} \quad (32)$$

and  $\zeta_D = c_1 \|x\|_{\mathcal{L}_2}^2 + c_2 x(0)^2 + c_3 \left( \frac{\partial x}{\partial z} \Big|_{z=l} \right)^2 + c_4$  with  $c_1, \dots, c_4$  and  $\alpha_1, \dots, \alpha_4$  being appropriate positive design parameters.

**Remark 2:** The first term in update law (31) reduces the approximated Hamiltonian whereas the other terms are necessary to insure the closed loop stability as will be explained in the proof. Note that the convergence rate to the optimal control input depends on  $\alpha_1$ ,  $\omega_D$  and parameters  $c_1, \dots, c_4$ . Under the update law (31), and control input (30), it will be shown that system state vector  $x$  and the weights estimation error  $\hat{W}_V$  will be ultimately bounded (UB).

In the next theorem, the authors will illustrate the ultimately boundedness of the closed-loop system with the boundary control input (30) and under the value functional weight update law (31). Before proceeding, the following lemma is necessary:

**Lemma 1:** The following inequality holds

$$|\varepsilon_H| \leq c_{H1} \|x\|_{\mathcal{L}_2}^2 + c_{H2} x(0)^2 + c_{H3} \left( \frac{\partial x}{\partial z} \Big|_{z=l} \right)^2, \quad (33)$$

where  $c_{H1}, \dots, c_{H3}$  are positive constants. Moreover  $|\varepsilon_H| \leq \xi_{DM} \zeta_D$  with  $\zeta_D$  defined as in update law (31) and  $\xi_{DM}$  being a positive constant dependent on reconstruction error.

**Proof:** Omitted because of space constraints.

**Theorem:** Consider the DPS system with PDE dynamics (1) and boundary conditions (2). Under Assumption 1 with the boundary control input (30) and the infinite-horizon cost functional (3), the DPS system state  $x$  and estimation error  $\hat{W}_V$  will remain UB under the update law (31) where  $0 <$

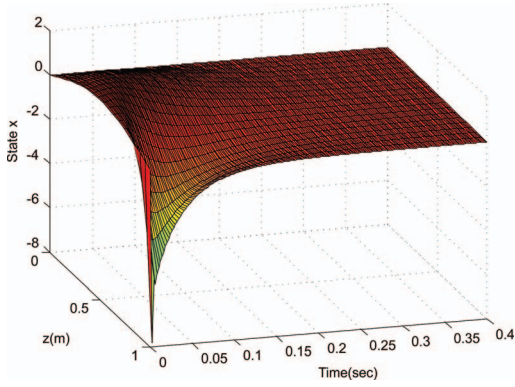


Fig. 1. Closed loop state evolution under Dirichlet boundary condition

$\alpha_1 < 1$ ,  $\alpha_2 > \alpha_3$  and  $\alpha_3 > \theta_{\alpha_3}^D$  with  $\theta_{\alpha_3}^D$  being a positive constant. Moreover, the actual control input will be bounded.

**Proof:** Refer to Appendix.

**Remark 3:** The presented control design is also applicable to more general linear PDE dynamics such as [4]

$$\begin{aligned} \frac{\partial x}{\partial t}(z,t) &= \frac{\partial^2 x}{\partial z^2}(z,t) + \lambda(z)x(z,t) + c(z)x(0,t) \\ &+ \int_0^\rho f(z,w)x(w,t)dw, \end{aligned} \quad (34)$$

where  $c(\cdot) \in C^1([0,l])$ ,  $f(\cdot, \cdot) \in C^1([0,l], [0,l])$ ,  $0 \leq \rho \leq l$  and  $w \in [0, \rho]$  is the integral variable. The basic PDE dynamics (1) are primarily chosen in this paper in order to avoid complicating illustrations. Moreover, it has been shown [4] that most widely applied DPS modeled by linear parabolic PDE are either in the form of or can be transformed into (1).

In the following section, numerical implementation of the controller will be illustrated through simulation on a reaction-diffusion process.

### III. SIMULATIONS

In this section, in order to verify the performance of proposed controller, a simulation example is provided. A typical linear reaction-diffusion system [4] with following linear parabolic PDE dynamics is considered:

$$\frac{\partial x}{\partial t} = \frac{\partial^2 x}{\partial z^2} + \beta_\alpha x \quad (35)$$

where  $\beta_\alpha > 0$ ,  $z \in [0, 1]$  and control input is only present at  $z = 1$ . The MATLAB `pdepe` function was used for simulation of dynamics in a real-time control setting with  $dz = 0.05$ . The boundary and initial conditions are expressed as

$$x(0) = 0, \quad x(1) = gu, \quad x(z, 0) = \sin(\pi z). \quad (36)$$

When  $u = 0$ , the open loop response of system can be obtained as

$$x(z,t) = 2 \sum_{n=1}^{\infty} e^{(\beta_\alpha - \pi^2 n^2)t} \sin(\pi n z) \int_0^1 \sin(\pi n z) x(z, 0) dz, \quad (37)$$

which is unstable for  $\beta_\alpha > \pi^2$ . For this example,  $\beta_\alpha$  was chosen to be 17 and  $g = 1$ . The sampling time for updating control input was  $t_s = 2msec$ . The chosen  $q(s, z)$  in cost functional (5) can be expressed as

$$q(s, z) = L_n(\text{distance}(s_0 + z_0 - 1 = 0, (s, z))) \quad (38)$$

with  $L_n(\cdot)$  being the Landau kernel, which is a continuous approximation for Dirac delta function, with  $n = 500$ , and the function  $\text{distance}(\cdot, \cdot)$  calculates the distance between the diameter  $s_0 + z_0 - 1 = 0$  and point  $(s, z) \in [0, 1] \times [0, 1]$ . The motivation behind choosing this  $q(s, z)$  is that it acts as an extension of the identity matrix for infinite dimensional cost functional (5). Thirty six radial basis functions were chosen as  $\phi_i s$  to approximate  $p(s, z)$  with the structure expressed as

$$\phi_i(s, z) = \frac{k}{1 + v \| (s, z) - (s_i, z_j) \|^2} \quad (39)$$

where  $v = 40$ ,  $k = 0.2$  and  $s_i s$  and  $z_j s$   $1 \leq i, j \leq 6$  were chosen from the set  $S = \{0.1, 0.25, 0.4, 0.55, 0.7, 1\}$ . However, since  $p(s, z)$  is symmetric, only 21 weights were needed to be updated and other weights could be found based on symmetry. The update law parameters were expressed as  $\alpha_1 = 0.1$ ,  $\alpha_2 = 0.2$ ,  $\alpha_3 = 0.1$ ,  $\alpha_4 = 0.01$ ,  $c_4 = 0.02$ ,  $c_1 = 2$ ,  $c_2 = c_3 = 0.01$  and  $\hat{W}_0$  was a random vector with positive entries. It can be easily verified that the necessary design conditions, i.e.  $\alpha_2 > \alpha_3$  and  $\alpha_3 > \theta_{\alpha_3}^D \simeq 0.08$  are satisfied for these parameters. Fig. 1 shows the good performance of controller in regulation of the DPS state. The smoothness of control input and fast convergence of HJB error are also shown in Fig. 2. In addition, as depicted in Fig. 2, increasing the update law parameter  $\alpha_1$  would speed up HJB error convergence rate. Fig. 3 shows the estimated  $\hat{p}(s, z)$  at the end of simulation. As shown, kernel  $\hat{p}(s, z)$  can be interpreted as a continuous two dimensional array which resembles a kernel matrix with infinitely many entries in LPS optimal control. Finally, Fig. 4 depicts the control input gain kernel  $K_u(s) = \hat{W}_V^T \frac{\partial \phi(s, z)}{\partial z} \Big|_{z=l}$ ,  $0 \leq s \leq l$  corresponding to estimated  $\hat{p}(s, z)$  in Fig 3. Qualitatively, Fig. 4 shows that feedback from the middle of spatial domain is significantly more important for system stabilization than places near the boundaries.

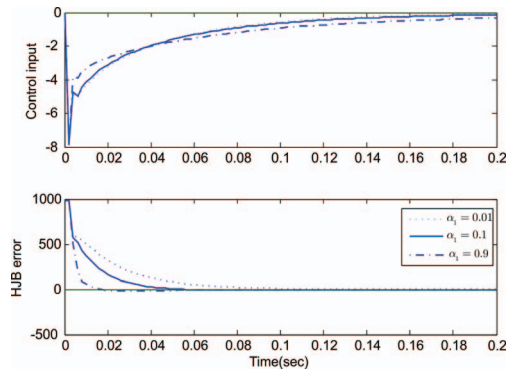


Fig. 2. Control input and HJB error for Dirichlet boundary controller

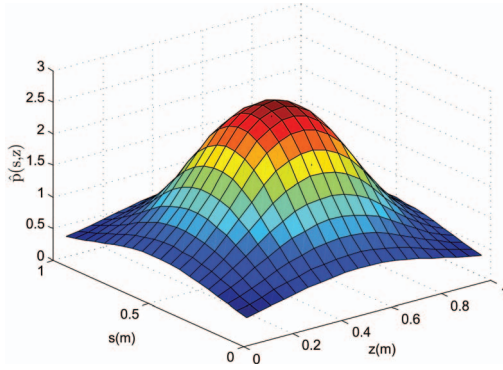


Fig. 3. Estimated  $\hat{p}(s, z)$  for Dirichlet boundary controller

#### IV. CONCLUSIONS

This paper developed an NDP-based near optimal boundary control scheme for DPS governed by one-dimensional PDE under Dirichlet actuation condition without any finite dimensional model approximation prior to control design. By defining the value functional as the extension of its definition in linear LPS optimal control design, the HJB equation was derived in original infinite dimensional state space. Based on defined structure for the value functional as a surface integral, a RBF was proposed to estimate its unknown parameters as a continuous two-variable kernel function. The update law for RBF unknown weights was defined to reduce the HJB error while insuring system stability. Ultimate boundedness of the closed-loop system was verified by using the standard Lyapunov theory. Since model reduction was not utilized in control development, the design is more reliable and can be applied to achieve accurate control and closed loop stability of the original infinite dimensional system. The performance of proposed control method was successfully verified on a diffusion reaction process.

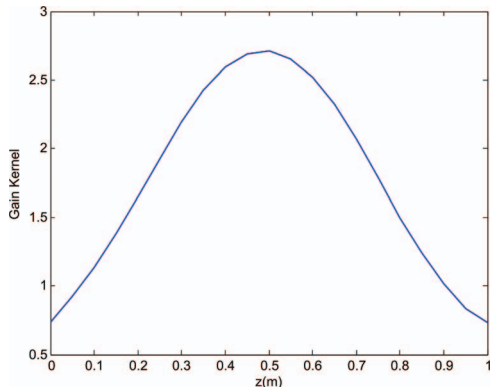


Fig. 4. Feedback gain kernel  $K_u(s)$  for Dirichlet boundary condition

#### V. APPENDIX

**Proof of Theorem:** Consider the following Lyapunov function

$$L = \underbrace{\frac{1}{2} \mu_a \tilde{W}_V^T \tilde{W}_V}_{L_a} + \underbrace{\int_0^l V^*(x) dz}_{L_b}, \quad (40)$$

where  $\mu_a$  is a positive tuning constant and  $\tilde{W}_V = \hat{W}_V - W_V$ . If  $\alpha_2 > \alpha_3$  and  $\alpha_3 > \theta_{\alpha_3}^D = \frac{\|\phi_{tz}\|_{\mathcal{L}_2}^4 g^4}{8c_1^2 r^2}$  define  $\eta = \frac{1}{2}(\alpha_2 - \alpha_3)$ . Therefore, if  $0 < \alpha_1 < 1$

$$\dot{L} \leq -\frac{q_{min}}{2} \|x\|_{\mathcal{L}_2}^2 - \mu_a \eta \|\tilde{W}_V\|^2 + \varepsilon. \quad (41)$$

where  $\varepsilon = \frac{\mu_a}{2} \xi_{DM}^2 + \frac{\mu_a \alpha_2}{2} W_{VM}^2 + \frac{81 \mu_a \alpha_3}{4} W_{VM}^4 + \frac{\mu_a \alpha_3}{2} W_{VM}^6$  with  $W_{VM}$  the bound for  $\|W_V\|$ . Therefore  $\dot{L}$  is always less than zero if  $\|x\|_{\mathcal{L}_2} > \sqrt{\frac{2\varepsilon}{q_{min}}}$  or  $\|\tilde{W}_V\| > \sqrt{\frac{\varepsilon}{\mu_a \eta}}$ . Consequently, the closed-loop system is UB.

Note that  $\varepsilon$  can be made small by reducing  $\xi_{DM}$  in Lemma 1 through increasing number of neurons and reducing  $\alpha_3$  through increasing the update parameter  $c_1$ . Moreover, the bound  $\|x\|_{\mathcal{L}_2} \in \sqrt{\frac{2\varepsilon}{q_{min}}}$  can be made smaller by increasing the design parameter  $q_{min}$  and  $\|\tilde{W}_V\| \in \sqrt{\frac{\varepsilon}{\mu_a \eta}}$  can be reduced by increasing  $\eta$  through the update parameter  $\alpha_2$ .

#### REFERENCES

- [1] A. A. Paranjape, J. Guan, S. J. Chung and Miroslav Krstic, "PDE boundary control for flexible articulated wings on a robotic aircraft," *IEEE Transactions on Robotics*, no. 3, pp. 625-640, 2013.
- [2] B. Talaei, H. Xu and S. Jagannathan, "Neural Network-based Near Optimal Constrained Control of Distributed Parameter Systems with Application to Diffusion-Reaction Processes," *IEEE MSC*, 2014.
- [3] I. Lasiecka, "Unified theory for abstract parabolic boundary problems—a semigroup approach," *Applied Mathematics and Optimization*, vol. 6, no. 1, pp. 287-333, 1980.
- [4] M. Krstic, and A. Smyshlyaev, "Boundary control of PDEs: A course on backstepping designs," SIAM, vol. 16, 2008.
- [5] D. Bertsekas, Dimitri P., "Dynamic programming and optimal control," Athena Scientific, vol. 1, no. 2, 1995.
- [6] Y. Luo, Q. Sun, H. Zhang and Lili Cui, "Adaptive critic design-based robust neural network control for nonlinear distributed parameter systems with unknown dynamics," *Neurocomputing* vol. 148, pp. 200-208, 2015.
- [7] B. Talaei, H. Xu and S. Jagannathan, "Near optimal boundary control of distributed parameter systems modeled as parabolic PDE by using finite difference neural network approximation," *IEEE Conference on Decision and Control*, 2014.
- [8] F. L. Lewis and D. Liu, editors, "Reinforcement Learning and Approximate Dynamic Programming for Feedback Control," John Wiley/IEEE Press, Computational Intelligence Series, 2012.
- [9] A. Alessandri, M. Gaggero and R. Zoppi, "Feedback optimal control of distributed parameter systems by using finite-dimensional approximation schemes," *IEEE Transaction on Neural Networks and Learning Systems*, vol. 23, no. 6, pp. 984-996, 2012.
- [10] F. L. Lewis, D. Vrabie and V. L. Syrmos, "Optimal control," John Wiley & Sons, 2012.
- [11] X. Li and J. Yong, *Optimal control theory for infinite dimensional systems*, Springer, 1995.
- [12] E. Kreyszig, "Introductory functional analysis with applications," Vol. 81, New York: Wiley, 1989.
- [13] D. R. Smith, "Variational methods in optimization," Courier Dover Publications, 1998.
- [14] M. D. Buhmann, "Radial basis functions: theory and implementations," Cambridge university press, vol. 5, 2003.