

Hamilton-Jacobi Equation for Optimal Control of Nonlinear Stochastic Distributed Parameter Systems Applied to Air Pollution Process

K.D. Do

Department of Mechanical Engineering
Curtin University, Perth, WA 6845, Australia

Copyright © 2014 K.D. Do. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper derives Hamilton-Jacobi equation (HJE) in Hilbert space for optimal control of stochastic distributed parameter systems (SDPSs) governed by partial differential equations (SPDEs) subject to both state-dependent and additive stochastic disturbances. First, nonlinear SDPSs are transformed to stochastic evolution systems (SESs), which are governed by stochastic ordinary differential equations (SODEs) in Hilbert space, using functional analysis. Second, the Hamilton-Jacobi equation (HJE), of which the solution results in an optimal control law, is derived. Third, a problem of optimal control of linear SDPSs, which include the air pollution process, with a quadratic cost functional is addressed as an application of the HJE. After, the control design is done, the SESs are transformed back to Euclidean space for implementation.

Keywords: Hamilton-Jacobi Equation, Hilbert space, stochastic distributed parameter system, stochastic evolution equation, optimal control, air pollution.

1 Introduction

Although optimal control of DPSs, i.e., systems governed by partial differential equations (PDEs), has been under development since 1960s, optimal control of nonlinear SDPSs subject to stochastic disturbances has been rarely addressed, [6], [40], [39], [36], [27], [25], [5], [10], [23]. The optimal control methods of the DPSs can be roughly classified into two main approaches.

The first approach, referred to as the modal control one, discretizes the PDEs to obtain lumped-parameter systems described in terms of modal coordinates, i.e.,

systems of ordinary differential equations (ODEs), to which the classical control design methods [1], [21], [22] can be applied. The modal control approach, see [30], [35], [18], can only control a certain number of modes of a DPS, and has difficulty in computing appropriate gain matrices.

The second approach applies the calculus of variations [8], [19] to derive a set of Euler-Lagrange (EL) equations in a form of the two-point boundary value (TPBV) problem, of which the solution results in optimal control inputs. Several techniques proposed to solve the TPBV problem include [26], [31] on hill-climbing algorithms, and [20], [32], [17], [3], [4] on an extension of the representer method to nonlinear systems based on linearization. However, the hill-climbing algorithms require a search in the whole control trajectories and the extended representer method requires many representers. To overcome the above issue, a non-climbing algorithm was proposed in [13] to solve a general TPBV problem.

On the other hand, the use of tools from functional analysis makes it possible to represent PDEs in Euclidean space as evolution systems, i.e., systems governed by ODEs in Hilbert space. Thus, we can utilize various ideas from control design methods well developed for systems governed by ODEs in Euclidean space. After nonlinear SDPSs are transformed to SESs in Hilbert space by using functional analysis, the goal of this paper is to present a derivation of the HJE in Hilbert space, which is relatively easy to follow, for optimal control of SDPSs subject to both state-dependent and additive stochastic disturbances. Several remarks on applications of the derived HJE are then made in relation to optimal control design procedure for SESs, stability analysis of SESs, and Itô's formula frequently used in control of stochastic systems. Finally, the HJE is applied to solve the optimal control problem of linear SDPSs, which include the air pollution process, with a quadratic cost functional.

2 Problem formulation

Let D be a open bounded set in Euclidean r -space \mathbb{R}^r with piecewise smooth boundary S , and let t denote time defined on an interval $T = [t_0, t_f]$ with $t_f > t_0 \geq 0$. In this paper, we consider the following SDPS in the Euclidean space:

$$\begin{aligned} \frac{\partial \mathbf{X}(\mathbf{x}, t)}{\partial t} &= \mathcal{N}\left(\mathbf{X}(\mathbf{x}, t), \frac{\partial \mathbf{X}(\mathbf{x}, t)}{\partial \mathbf{x}}, \dots, \frac{\partial^{r_1} \mathbf{X}(\mathbf{x}, t)}{\partial \mathbf{x}^{r_1}}, \mathbf{x}, t, \mathbf{u}_d(\mathbf{x}, t)\right) \\ &\quad + \mathcal{G}\left(\mathbf{X}(\mathbf{x}, t), \frac{\partial \mathbf{X}(\mathbf{x}, t)}{\partial \mathbf{x}}, \dots, \frac{\partial^{r_2} \mathbf{X}(\mathbf{x}, t)}{\partial \mathbf{x}^{r_2}}, \mathbf{x}, t\right) \dot{\mathbf{w}}(\mathbf{x}, t), \quad \mathbf{x} \in D \\ \mathbf{X}(\mathbf{x}, t_0) &= \mathbf{X}_0(\mathbf{x}), \quad \mathbf{x} \in D \\ \mathcal{N}_b\left(\mathbf{X}(\boldsymbol{\xi}, t), \frac{\partial \mathbf{X}(\boldsymbol{\xi}, t)}{\partial \mathbf{n}}, \dots, \frac{\partial^{r_3} \mathbf{X}(\boldsymbol{\xi}, t)}{\partial \mathbf{n}^{r_3}}, \mathbf{x}, t, \mathbf{u}_b(\boldsymbol{\xi}, t)\right), &\quad \boldsymbol{\xi} \in S, \end{aligned} \quad (1)$$

defined for $t \in T$, where $\mathbf{x} = \text{col}(x_1, \dots, x_r) \in D$ is the r -dimensional spatial coordinate vector; $\mathbf{X}(\mathbf{x}, t)$ is the n -dimensional vector function describing the system state; $\mathbf{w}(\mathbf{x}, t)$ is p -dimensional stochastic disturbance; the dot over $\mathbf{w}(\mathbf{x}, t)$ denotes the formal derivative with respect to time t ; $\mathcal{N}(\cdot)$ and $\mathcal{N}_b(\cdot)$ are the n -dimensional

vector functions; $\mathcal{G}(\cdot)$ are the $n \times p$ matrix function; $\mathbf{X}_0(\cdot)$ is the n -dimensional Gaussian random vector with zero mean and the covariance matrix $\mathbf{P}_0(\mathbf{x}, \mathbf{y})$; \mathbf{n} is the outward directed normal vector to the boundary S ; $\mathbf{u}_d(\cdot)$ is the m_d -dimensional distributed control input; and $\mathbf{u}_b(\cdot)$ is m_b -dimensional boundary control input. In (1), the following notations have been used

$$\begin{aligned} \frac{\partial \mathbf{X}^{r_i}(\cdot)}{\partial \mathbf{x}^{r_i}} &= \frac{\partial^{r_i} \mathbf{X}_d(\cdot)}{\partial x_1^{k_1} \dots \partial x_r^{k_r}}, \quad r_i = k_1 + \dots + k_r, \quad i = 1, 2, \\ \frac{\partial \mathbf{X}(\cdot)}{\partial \mathbf{n}} &= \sum_{i=1}^r \frac{\partial \mathbf{X}(\cdot)}{\partial x_i} \cos(\mathbf{n}, x_i) \end{aligned} \tag{2}$$

where $r_i \geq 0$, and k_i are nonnegative integers; and $\cos(\mathbf{n}, x_i)$ is the direction cosine of x_i . We assume that the p -dimensional stochastic disturbance vector $\mathbf{w}(\cdot)$ is a Wiener process with zero mean value and the covariance matrix function given by

$$\begin{aligned} \mathbb{E} \left\{ \langle \mathbf{h}(\mathbf{x}, t), \mathbf{w}(\mathbf{x}, t) \rangle_{\mathcal{L}_p^2(D)} \langle \mathbf{h}(\mathbf{y}, t), \mathbf{w}(\mathbf{y}, t) \rangle_{\mathcal{L}_p^2(D)} \right\} = \\ \int_D \langle \mathbf{h}(\mathbf{x}, t), \mathbf{Q}(\mathbf{x}, \mathbf{y}) \mathbf{h}(\mathbf{y}, s) \rangle_{\mathcal{L}_p^2(D)} d\mathbf{y} \min((t - t_0), (s - t_0)) \end{aligned} \tag{3}$$

for $t, s \in T$ and any p -dimensional square integrable function $\mathbf{h}(\mathbf{x}, t)$ on D , where $\mathbf{Q}(\mathbf{x}, \mathbf{y})$ is a $p \times p$ symmetric positive definite matrix function, and $\langle \cdot, \cdot \rangle_{\mathcal{L}_p^2(D)}$ is the inner product:

$$\langle \mathbf{h}(\mathbf{x}, t), \mathbf{h}(\mathbf{x}, s) \rangle_{\mathcal{L}_p^2(D)} = \int_D \mathbf{h}^T(\mathbf{x}, t) \mathbf{h}(\mathbf{x}, s) d\mathbf{x}. \tag{4}$$

Hence, $\dot{\mathbf{w}}(\mathbf{x}, t)$ can be regarded as the white Gaussian noise in t [9].

The **control objective** is to design the control inputs $\mathbf{u}_d(\cdot)$ and $\mathbf{u}_b(\cdot)$ so as to minimize the cost functional:

$$J(\mathbf{X}(\mathbf{x}, t_0), t_0) = \mathbb{E}_{(\mathbf{X}, t_0)} \left\{ J_f(\mathbf{X}(\mathbf{x}, t_f), t_f) + \int_{t_0}^{t_f} L(\mathbf{X}(\mathbf{x}, \tau), \tau, \mathbf{u}_d(\tau), \mathbf{u}_b(\tau)) d\tau \right\}, \tag{5}$$

where $\mathbb{E}_{(\mathbf{X}, t)}\{\cdot\}$ is the conditional expectation of $\{\cdot\}$ with respect to the σ -field generated by $\{(\mathbf{X}(\mathbf{x}, \tau), t_0 \leq \tau \leq t)\}$; and $J_f(\cdot)$ and $L(\cdot)$ are the positive and real-valued integrable functions with respect to time t .

3 Transformation of SDPS in Euclidean to SES in Hilbert space

In this section, functional analysis [41] is used to transform the SDPS (1) in the Euclidean space to a SES in the Hilbert space. As such, we first define $\mathcal{L}_n^2(D)$ and $\mathcal{H}_n^m(D)$ spaces by

$$\begin{aligned} \mathcal{L}_n^2(D) &= \left\{ \phi(\mathbf{x}, t) \mid \int_D \|\phi(\mathbf{x}, t)\|_n^2 d\mathbf{x} < \infty \right\} \\ \mathcal{H}_n^m(D) &= \left\{ \phi(\mathbf{x}, t) \mid \frac{\partial^{|\alpha|} \phi(\mathbf{x}, t)}{\partial x_1^{k_1} \dots \partial x_r^{k_r}} \in \mathcal{L}_n^2(D) \right\}, \end{aligned} \tag{6}$$

where $\|\cdot\|_n$ is the n -dimensional Euclidean norm; $|\alpha| = k_1 + \dots + k_r \leq m$ for $\alpha = (k_1, \dots, k_r)$ with k_1, \dots, k_r nonnegative integers. The spaces $\mathcal{L}_n^2(S)$ and $\mathcal{H}_n^m(S)$ are defined similarly by replacing D by S in (6). We now define the norms of $\mathcal{L}_n^2(D)$ and $\mathcal{H}_n^m(D)$ by

$$\begin{aligned} \|\phi(\mathbf{x}, t)\|_{\mathcal{L}_n^2(D)} &= \langle \phi(\mathbf{x}, t), \phi(\mathbf{x}, t) \rangle_{\mathcal{L}_n^2(D)} = \int_D \|\phi(\mathbf{x}, t)\|_n^2 d\mathbf{x}, \\ \|\phi(\mathbf{x}, t)\|_{\mathcal{H}_n^m(D)} &= \langle \phi(\mathbf{x}, t), \phi(\mathbf{x}, t) \rangle_{\mathcal{H}_n^m(D)} = \sum_{|\alpha| \leq m} \left\langle \frac{\partial^{|\alpha|} \phi(\mathbf{x}, t)}{\partial x_1^{k_1} \dots \partial x_r^{k_r}}, \frac{\partial^{|\alpha|} \phi(\mathbf{x}, t)}{\partial x_1^{k_1} \dots \partial x_r^{k_r}} \right\rangle_{\mathcal{L}_n^2(D)}. \end{aligned} \tag{7}$$

Thus, $\mathcal{L}_n^2(D)$ and $\mathcal{H}_n^m(D)$ are Hilbert spaces endowed with the norms $\|\phi(\mathbf{x}, t)\|_{\mathcal{L}_n^2(D)}$ and $\|\phi(\mathbf{x}, t)\|_{\mathcal{H}_n^m(D)}$, respectively, given in (7).

Considering $\mathbf{x} \in D$ as the parameter defined at very $t \in T$, we can regard $\mathbf{X}(\mathbf{x}, t)$ as $\mathbf{X}(t) \in \mathcal{U}_1 = \mathcal{H}_n^m(D)$ with $m \geq \max(r_i), i = 1, 2, 3$. Similarly, $\mathbf{w}(\mathbf{x}, t), \mathbf{u}_d(\mathbf{x}, t), \mathbf{X}(\boldsymbol{\xi}, t)$, and $\mathbf{u}_b(\boldsymbol{\xi}, t)$ are regarded as $\mathbf{w}(t) \in \mathcal{U}_2 = \mathcal{L}_n^2(D), \mathbf{u}_d(t) \in \mathcal{U}_3 = \mathcal{L}_n^2(D), \mathbf{X}(t) \in \mathcal{U}_4 = \mathcal{H}_n^m(S)$, and $\mathbf{u}_b(t) \in \mathcal{U}_5 = \mathcal{L}_n^2(S)$, respectively. Therefore, the SDPS (1) can be described as

$$\begin{aligned} d\mathbf{X}(t) &= \mathcal{N}(\mathbf{X}(t), t, \mathbf{u}_d(t))dt + \mathcal{G}(\mathbf{X}(t), t)d\mathbf{w}(t) \\ \mathbf{X}(t_0) &= \mathbf{X}_0 \\ \mathcal{N}_b(\mathbf{X}(t), t, \mathbf{u}_b(t)) &= 0, \end{aligned} \tag{8}$$

which is referred to as the SES in the Hilbert space. In (8), $\mathcal{N}(\cdot)$ maps $\mathcal{U}_1 \times T \times \mathcal{U}_3$ into $\mathcal{U}_6 = \mathcal{L}_n^2(D)$; $\mathcal{N}_b(\cdot)$ is the mapping defined on $\mathcal{U}_4 \times T \times \mathcal{U}_5$; and $\mathcal{G}(\cdot)$ defined on $\mathcal{U}_1 \times T$ is an element of $\mathcal{L}(\mathcal{U}_2, \mathcal{U}_6)$ showing the aggregate of bounded linear mapping from \mathcal{U}_2 into \mathcal{U}_6 .

The covariance matrix function of $\mathbf{w}(\mathbf{x}, t)$ is described from (3) by

$$\mathbb{E}\{\mathbf{w}(t) \circ \mathbf{w}(t)\} = \mathbf{Q} \min((t - t_0), (s - t_0)), \tag{9}$$

where \mathbf{Q} is a compact, positive, bounded, and trace class operator mapping \mathcal{U}_2 into \mathcal{U}_2 . If h_1, h_2 , and h are any elements of the Hilbert space \mathcal{H} , then $h_1 \circ h_2$ is an element of $\mathcal{L}(\mathcal{H}, \mathcal{H})$ defined by

$$(h_1 \circ h_2)h = h_1 \langle h_2, h \rangle_{\mathcal{H}}, \tag{10}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is an inner product in \mathcal{H} . The cost functional $J(\cdot)$ can be written from (5) as

$$J(\mathbf{X}(t_0), t_0) = \mathbb{E}_{\mathbf{X}, t_0} \left\{ J_f(\mathbf{X}(t_f), t_f) + \int_{t_0}^{t_f} L(\mathbf{X}(\tau), \tau, \mathbf{u}_d(\tau), \mathbf{u}_b(\tau))d\tau \right\}. \tag{11}$$

4 Hamilton-Jacobi equation in Hilbert space

In this section, we derive the HJE for the SES (8). As such, Let $W(\mathbf{X}(t), t)$ be the minimum of $J(\cdot)$ with respect to $\mathbf{u}_d(t) \in \mathcal{U}_3$ and $\mathbf{u}_b(t) \in \mathcal{U}_5$, i.e.,

$$W(\mathbf{X}(t), t) = \min_{\mathbf{u}_d(t) \in \mathcal{U}_3, \mathbf{u}_b(t) \in \mathcal{U}_5, t \leq \tau \leq t_f} [J(\mathbf{X}(t), t)]. \tag{12}$$

Let us assume that $W(\mathbf{X}(t), t)$ is continuously differentiable on T and continuously twice Fréchet differentiable on \mathcal{U}_1 . Applying Bellman’s optimality principle [2] to the integral term of (11) results in

$$\min_{\mathbf{u}_d(\tau) \in \mathcal{U}_3, \mathbf{u}_b(\tau) \in \mathcal{U}_5, t \leq \tau \leq t_f} [L_1 + L_2 - W(\mathbf{X}(t), t)] = 0, \tag{13}$$

where

$$\begin{aligned} L_1 &= \mathbb{E}_{\mathbf{X},t} \left\{ \int_t^{t+\sigma} J(\mathbf{X}(\tau), \tau, \mathbf{u}_d(\tau), \mathbf{u}_b(\tau)) d\tau \right\}, \\ L_2 &= \mathbb{E}_{\mathbf{X},t} \{ W(\mathbf{X}(t + \sigma), t + \sigma) \}, \end{aligned} \tag{14}$$

with $\sigma > 0$. Applying the mean-value theorem to the term L_1 yields

$$L_1 = \mathbb{E}_{\mathbf{X},t} \{ J(\mathbf{X}(t + \sigma\varepsilon), t + \sigma\varepsilon, \mathbf{u}_d(t + \sigma\varepsilon), \mathbf{u}_b(t + \sigma\varepsilon)) \} \sigma, \tag{15}$$

where $0 < \varepsilon < 1$. On the other hand, applying the Taylor expansion theorem [11] to the term L_2 results in

$$\begin{aligned} L_2 &= W(\mathbf{X}(t), t) + \sigma \frac{\partial W(\mathbf{X}(t), t)}{\partial t} + W_{\mathbf{X}}(\mathbf{X}(t), t) \mathbb{E}_{\mathbf{X},t} \{ \Delta \mathbf{X}(t) \} \\ &\quad + \frac{1}{2} \mathbb{E}_{\mathbf{X},t} \{ W_{\mathbf{X}\mathbf{X}}[\Delta \mathbf{X}(t), \Delta \mathbf{X}(t)] \} + \mathcal{O}(\|\Delta \mathbf{X}(t)\|_{\mathcal{U}_6}^3), \end{aligned} \tag{16}$$

where $W_{\mathbf{X}}(\cdot)$ and $W_{\mathbf{X}\mathbf{X}}(\cdot)$ are the first and the second order Fréchet derivatives on \mathcal{U}_1 , which are a linear mapping from \mathcal{U}_1 into the real-valued \mathbb{R} , and a bilinear mapping from $\mathcal{U}_1 \times \mathcal{U}_1$ into \mathbb{R} , respectively, and $\mathcal{O}(\cdot)$ is the same order infinitesimal, and

$$\Delta \mathbf{X}(t) = \mathbf{X}(t + \sigma) - \mathbf{X}(t). \tag{17}$$

We now calculate the term L_2 in (16). From (8), we have

$$\begin{aligned} \mathbb{E}_{\mathbf{X},t} \{ \Delta \mathbf{X}(t) \} &= \int_t^{t+\sigma} \mathcal{N}(\mathbf{X}(\tau), \tau, \mathbf{u}_d(\tau)) d\tau, \\ \mathbb{E}_{\mathbf{X},t} \{ \Phi[\Delta \mathbf{X}(t), \Delta \mathbf{X}(t)] \} &= \overline{\text{Tr}}(\Phi)(\mathcal{G}\sqrt{\mathbf{Q}})\sigma, \\ \overline{\text{Tr}}(\Phi)(\mathcal{G}\sqrt{\mathbf{Q}}) &= \sum_{i=1}^{\infty} \Phi \left[\mathcal{G}(\mathbf{X}(t), t) \sqrt{\lambda_i} e_i, \mathcal{G}(\mathbf{X}(t), t) \sqrt{\lambda_i} e_i \right]. \end{aligned} \tag{18}$$

where Φ is a bilinear mapping from $\mathcal{U}_1 \times \mathcal{U}_1$ into \mathbb{R} and $\{\lambda_i, e_i, i = 1, 2, \dots\}$ is an orthonormal set of eigenvalues and eigenfunctions of the operator \mathbf{Q} , i.e.,

$$\mathbf{Q}e_i = \lambda_i e_i. \tag{19}$$

Substituting (18) into (16) yields

$$\begin{aligned} L_2 &= W(\mathbf{X}(t), t) + \sigma \frac{\partial W(\mathbf{X}(t), t)}{\partial t} + W_{\mathbf{X}}(\mathbf{X}(t), t) \int_t^{t+\sigma} \mathcal{N}(\mathbf{X}(\tau), \tau, \mathbf{u}_d(\tau)) d\tau \\ &\quad + \frac{1}{2} \overline{\text{Tr}}(W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t)) (\mathcal{G}\sqrt{\mathbf{Q}})\sigma + \mathcal{O}(\|\Delta \mathbf{X}(t)\|_{\mathcal{U}_6}^3). \end{aligned} \tag{20}$$

On the other hand, from (8), (13), (14), and (15) we have

$$\mathcal{O}(\|\Delta \mathbf{X}(t)\|_{\mathcal{H}_6}^3) = \mathcal{O}(\sigma^3). \tag{21}$$

Substituting (15) and (20) with $\mathcal{O}(\|\Delta \mathbf{X}(t)\|_{\mathcal{H}_6}^3)$ satisfied (21) into (13) yields

$$\begin{aligned} -\frac{\partial W(\mathbf{X}(t), t)}{\partial t} &= \min_{\mathbf{u}_d(t+\sigma), \mathbf{u}_b(t+\sigma)} \left[\mathbb{E}_{\mathbf{X}, t} \{ J(\mathbf{X}(t + \sigma\varepsilon), t + \sigma\varepsilon, \mathbf{u}_d(t + \sigma\varepsilon), \mathbf{u}_b(t + \sigma\varepsilon)) \} \right. \\ &\quad \left. + \frac{1}{\sigma} W_{\mathbf{X}}(\mathbf{X}(t), t) \int_t^{t+\sigma} \mathcal{N}(\mathbf{X}(\tau), \tau, \mathbf{u}_d(\tau)) d\tau + \frac{1}{2} \overline{\text{Tr}}(W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t))(\mathcal{G}\sqrt{\mathbf{Q}}) \right]. \end{aligned} \tag{22}$$

From the Bochner theorem [41], we have

$$\lim_{\sigma \rightarrow 0} \frac{1}{\sigma} \int_t^{t+\sigma} \mathcal{N}(\mathbf{X}(\tau), \tau, \mathbf{u}_d(\tau)) d\tau = \mathcal{N}(\mathbf{X}(t), t, \mathbf{u}_d(t)). \tag{23}$$

Using this identity and letting $\sigma \rightarrow 0$ in (22) result in the following HJE in the Hilbert space:

$$-\frac{\partial W(\mathbf{X}(t), t)}{\partial t} = \min_{\mathbf{u}_d(t+\sigma), \mathbf{u}_b(t+\sigma)} \left[H(\mathbf{X}(t), t, \mathbf{u}_d(t), \mathbf{u}_b(t), W_{\mathbf{X}}(\mathbf{X}(t), t), W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t)) \right], \tag{24}$$

where $H(\cdot)$ is referred to as the system Hamiltonian, which is a real-valued functional given by

$$\begin{aligned} H(\cdot) &= J(\mathbf{X}(t), t, \mathbf{u}_d(t), \mathbf{u}_b(t)) + \langle W_{\mathbf{X}}(\mathbf{X}(t), t), \mathcal{N}(\mathbf{X}(t), t, \mathbf{u}_d(t)) \rangle_{\mathcal{H}_6} \\ &\quad + \frac{1}{2} \text{Tr}[\mathcal{G}(\mathbf{X}(t), t)\mathbf{Q}\mathcal{G}^*(\mathbf{X}(t), t)W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t)], \end{aligned} \tag{25}$$

with $\text{Tr}[\cdot]$ and the asterisk being the trace and the adjoint of the corresponding operator, respectively.

Remark 4.1 *If the functional space is replaced by a space of finite dimension, then $\mathcal{N}(\cdot)$ and $\mathcal{G}(\cdot)$ become a vector and a matrix in the finite dimensional space, respectively, and the HJE (24) is reduced to the HJE derived for stochastic lumped-parameter systems (SLPSs) in Euclidean space [24]. Moreover, if the SDPS (8) is linear subject to additive stochastic noise, the HJE (24) is reduced to the results in [39].*

Remark 4.2 *The optimal controls $\mathbf{u}_d(t)$ and $\mathbf{u}_b(t)$ can be found by performing the following three-step procedure:*

1. *Minimize the Hamiltonian $H(\cdot)$ to obtain $\mathbf{u}_d(\cdot)$ and $\mathbf{u}_b(\cdot)$ as functions of $\mathbf{X}(t)$, t , and $W(t)$, i.e., $\mathbf{u}_d(\cdot) = \boldsymbol{\varpi}_d(\mathbf{X}(t), t, W(t))$ and $\mathbf{u}_b(\cdot) = \boldsymbol{\varpi}_b(\mathbf{X}(t), t, W(t))$;*
2. *Substitute the above $\boldsymbol{\varpi}_d(\cdot)$ and $\boldsymbol{\varpi}_b(\cdot)$ into (24), then solve for $W(\cdot)$, which must satisfy the aforementioned conditions such as the boundary defined in the third equation of (8) and twice Fréchet differentiable;*

3. Substitute the found $W(\cdot)$ into the expressions of $\varpi_d(\cdot)$ and $\varpi_b(\cdot)$ to obtain the controls $\mathbf{u}_d(\cdot)$ and $\mathbf{u}_b(\cdot)$.

In general, the first two steps of the above procedure are formidable for a predefined cost functional $J(\cdot)$ in (11) and a general nonlinear SES in (8). This formidable task was tackled by approximating methods on either dealing with the above procedure directly or converting the HJE (24) to a two-point boundary value problem, see [26, 31] on hill-climbing algorithms; [3, 17, 20, 32] on an extension of the representer method [3] to nonlinear systems based on linearization; and [13] on a non-climbing algorithm. Moreover, no formal proof of convergence of the approximating solutions by the above approximating methods to the true ones even though there only are several works on existence and uniqueness of (weak/viscosity) solutions of HJEs to certain SLPSs and SDPSs [7, 29, 38].

Remark 4.3 If the controls $\mathbf{u}_d(t)$ and $\mathbf{u}_b(t)$ in the SES (8) are set to zero or are designed such that they are functions of $\mathbf{X}(t)$ and t , the HJE (24) can be used for stability analysis of the SES (8) without controls or of the resulting closed loop system. This was used in [12] for stability analysis of linear DPSs. Moreover, if the SES (8) does not contain the controls $\mathbf{u}_d(t)$ and $\mathbf{u}_b(t)$, and the cost functional $J(\cdot)$ is chosen as

$$J(\cdot) = -\frac{dW(\mathbf{X}(t), t)}{dt} + \langle W_{\mathbf{X}}(\mathbf{X}(t), t), \mathcal{G}(\mathbf{X}(t), t)\mathbf{w}(t) \rangle_{\mathcal{W}_6}, \quad (26)$$

the HJE (24) reduces to Itô's formula [34].

5 Application to optimal control of an air pollution process

5.1 Mathematical model of an air pollution process

Let D be a open bounded set in Euclidean three-dimensional space \mathbb{R}^3 with piecewise smooth boundary S , let t denote time defined on an interval $T = [t_0, t_f]$ with $t_f > t_0 \geq 0$, and let $\mathbf{r} = \text{col}(x, y, z)$. An air pollution process can be modeled by a system of advection-diffusion equations, which are a set of the following partial differential equations in three dimensional space as follows [37]:

$$\begin{aligned} \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial t} = & -\frac{\partial(V^x \mathbf{c}(\mathbf{r}, t))}{\partial x} - \frac{\partial(V^y \mathbf{c}(\mathbf{r}, t))}{\partial y} - \frac{\partial(V^z \mathbf{c}(\mathbf{r}, t))}{\partial z} + \mathbf{f}(\mathbf{r}, t) + \frac{\partial}{\partial x} \left(\mathbf{K}^x \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial x} \right) \\ & + \frac{\partial}{\partial y} \left(\mathbf{K}^y \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial y} \right) + \frac{\partial}{\partial z} \left(\mathbf{K}^z \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial z} \right) - \mathbf{K} \mathbf{c}(\mathbf{r}, t) + \mathbf{B}_d(\mathbf{r}, t) \mathbf{u}_d, \quad \forall \mathbf{r} \in D, \\ \bar{\mathbf{K}}^x \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial x} \mathbf{l}_x + \bar{\mathbf{K}}^y \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial y} \mathbf{l}_y + \bar{\mathbf{K}}^z \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial z} \mathbf{l}_z + \bar{\mathbf{K}}_b \mathbf{c}(\mathbf{r}, t) + \mathbf{B}_b(\mathbf{r}, t) \mathbf{u}_b = & 0, \quad \forall \mathbf{r} \in S, \\ \mathbf{c}(\mathbf{r}, t_0) = & \mathbf{c}_0(\mathbf{r}) \end{aligned} \quad (27)$$

where $\mathbf{c}(\mathbf{r}, t) = \text{col}(c_1(\mathbf{r}, t), \dots, c_n(\mathbf{r}, t))$ is the n -dimensional vector of pollutant concentrations in location \mathbf{r} and at time t measured in parts per million (p.p.m); V^x , V^y and V^z are wind velocities in the x , y and z directions, respectively (m/s); \mathbf{K}^x , \mathbf{K}^y and \mathbf{K}^z are turbulent diffusivities which are simply called diffusion coefficient matrices in the x , y and z directions, respectively (m^2/s); $\bar{\mathbf{K}}^x$, $\bar{\mathbf{K}}^y$ and $\bar{\mathbf{K}}^z$ are the mean parts of \mathbf{K}^x , \mathbf{K}^y and \mathbf{K}^z , respectively; \mathbf{K} is the deposition coefficient matrix; $\bar{\mathbf{K}}_b$ is a constant parameter matrix; $\mathbf{f}(\mathbf{r}, t)$ represents the emission source; l_x , l_y and l_z are direction cosines of the outward normal to the boundary surface S ; $\mathbf{B}_d(\mathbf{r}, t)$ and $\mathbf{B}_b(\mathbf{r}, t)$ are $n \times m_d$ and $n \times m_b$ matrix functions; \mathbf{u}_d and \mathbf{u}_b are m_d -dimensional and m_b -dimensional control input vectors. It is noted that the controls \mathbf{u}_d and \mathbf{u}_b can be either distributed or discrete over the domain D and the boundary S , see Subsection 5.6.

Now the wind velocities (V^x, V^y, V^z), and the turbulent diffusivities ($\mathbf{K}^x, \mathbf{K}^y, \mathbf{K}^z$) are decomposed of their mean and stochastic parts as follows

$$\begin{aligned} V^x &= \bar{V}^x + W^{Vx}, \quad V^y = \bar{V}^y + W^{Vy}, \quad V^z = \bar{V}^z + W^{Vz}, \\ \mathbf{K}^x &= \bar{\mathbf{K}}^x + \mathbf{W}^{\mathbf{K}x}, \quad \mathbf{K}^y = \bar{\mathbf{K}}^y + \mathbf{W}^{\mathbf{K}y}, \quad \mathbf{K}^z = \bar{\mathbf{K}}^z + \mathbf{W}^{\mathbf{K}z}, \\ \mathbf{K} &= \bar{\mathbf{K}} + \mathbf{W}^{\mathbf{K}}, \end{aligned} \tag{28}$$

where $\bar{\bullet}$ and W^\bullet denote the mean and stochastic parts of \bullet , respectively. Substituting (28) into (27) results in

$$\begin{aligned} \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial t} &= -\frac{\partial(\bar{V}^x \mathbf{c}(\mathbf{r}, t))}{\partial x} - \frac{\partial(\bar{V}^y \mathbf{c}(\mathbf{r}, t))}{\partial y} - \frac{\partial(\bar{V}^z \mathbf{c}(\mathbf{r}, t))}{\partial z} + \frac{\partial}{\partial x} \left(\bar{\mathbf{K}}^x \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial x} \right) \\ &+ \frac{\partial}{\partial y} \left(\bar{\mathbf{K}}^y \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial y} \right) + \frac{\partial}{\partial z} \left(\bar{\mathbf{K}}^z \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial z} \right) - \bar{\mathbf{K}} \mathbf{c}(\mathbf{r}, t) + \mathbf{B}_d(\mathbf{r}, t) \mathbf{u}_d \\ &- \frac{\partial(W^{Vx} \mathbf{c}(\mathbf{r}, t))}{\partial x} - \frac{\partial(W^{Vy} \mathbf{c}(\mathbf{r}, t))}{\partial y} - \frac{\partial(W^{Vz} \mathbf{c}(\mathbf{r}, t))}{\partial z} + \frac{\partial}{\partial x} \left(\mathbf{W}^{\mathbf{K}x} \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial x} \right) \\ &+ \frac{\partial}{\partial y} \left(\mathbf{W}^{\mathbf{K}y} \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial y} \right) + \frac{\partial}{\partial z} \left(\mathbf{W}^{\mathbf{K}z} \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial z} \right) - \mathbf{W}^{\mathbf{K}} \mathbf{c}(\mathbf{r}, t) + \mathbf{f}(\mathbf{r}, t) \quad \forall \mathbf{r} \in D, \\ \bar{\mathbf{K}}^x \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial x} l_x + \bar{\mathbf{K}}^y \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial y} l_y + \bar{\mathbf{K}}^z \frac{\partial \mathbf{c}(\mathbf{r}, t)}{\partial z} l_z + \bar{\mathbf{K}}_b \mathbf{c}(\mathbf{r}, t) + \mathbf{B}_b(\mathbf{r}, t) \mathbf{u}_b &= 0, \quad \forall \mathbf{r} \in S, \\ \mathbf{c}(\mathbf{r}, t_0) &= \mathbf{c}_0(\mathbf{r}). \end{aligned} \tag{29}$$

It is now seen that the air pollution process has been rewritten as a system of stochastic PDEs. In the next subsection, we will present an application of the control design in the previous sections to a fairly general linear SDPS, which covers the stochastic air pollution process (29). The control objective is to design the controls \mathbf{u}_d and \mathbf{u}_b so as to minimize the pollutant concentration $\mathbf{c}(\mathbf{r}, t)$ while putting appropriate weights on the controls \mathbf{u}_d and \mathbf{u}_b .

5.2 Linear stochastic distributed parameter system

Consider the following linear SDPS subject to both state-dependent and additive stochastic disturbances:

$$\begin{aligned} \frac{\partial \mathbf{X}(\mathbf{x}, t)}{\partial t} &= \mathbf{A}_x \mathbf{X}(\mathbf{x}, t) + \mathbf{B}_d(\mathbf{x}, t) \mathbf{u}_d(\mathbf{x}, t) + \mathbf{G}_1(\mathbf{x}, t) \mathbf{X}(\mathbf{x}, t) \dot{w}_1(\mathbf{x}, t) \\ &\quad + \mathbf{G}_0(\mathbf{x}, t) \dot{w}_0(\mathbf{x}, t), \quad \mathbf{x} \in D \\ \mathbf{X}(\mathbf{x}, t_0) &= \mathbf{X}_0(\mathbf{x}), \quad \mathbf{x} \in D \\ \beta_\xi \mathbf{X}(\xi, t) &= \mathbf{B}_b(\xi, t) \mathbf{u}_b(\xi, t), \quad \xi \in S \end{aligned} \tag{30}$$

where

$$\begin{aligned} \mathbf{A}_x[\bullet] &= \sum_{i,j=1}^n \mathbf{A}_{ij}(\mathbf{x}, t) \frac{\partial^2[\bullet]}{\partial x_i \partial x_j} + \sum_{i=1}^n \mathbf{B}_i(\mathbf{x}, t) \frac{\partial[\bullet]}{\partial x_i} + \mathbf{C}(\mathbf{x}, t)[\bullet], \\ \beta_\xi[\bullet] &= \sum_{j=1}^n \mathbf{A}_j(\xi, t) \frac{\partial[\bullet]}{\partial x_j} + \mathbf{F}(\xi, t)[\bullet], \end{aligned} \tag{31}$$

where the $\mathbf{A}_{ij}(\mathbf{x}, t)$, $\mathbf{B}_i(\mathbf{x}, t)$, $\mathbf{C}(\mathbf{x}, t)$, $\mathbf{F}(\xi, t)$, $\mathbf{G}_1(\mathbf{x}, t)$ are $r \times r$ symmetric matrices defined on $D \times T$ and $\mathbf{A}_{ij}(\mathbf{x}, t) = \mathbf{A}_{ji}(\mathbf{x}, t)$, and $\mathbf{A}_j(\xi, t)$ is given by

$$\mathbf{A}_j(\xi, t) = \sum_{i=1}^n \mathbf{A}_{ij}(\xi, t) \cos(\mathbf{n}, x_i), \tag{32}$$

with \mathbf{n} being the outward normal to the boundary S at the point $\xi \in S$, and (\mathbf{n}, x_i) being the angle between the outward normal \mathbf{n} and the x_i -axis. In (30), $\mathbf{G}_0(\mathbf{x}, t)$ is an n -dimensional vector function; $\mathbf{G}_1(\mathbf{x}, t)$ is a $n \times n$ matrix function; $\mathbf{B}_d(\mathbf{x}, t)$ and $\mathbf{B}_b(\xi, t)$ are $n \times m_d$ and $n \times m_b$ matrix functions; $\mathbf{w}(\mathbf{x}, t) = \text{col}(w_1(\mathbf{x}, t), w_0(\mathbf{x}, t))$ is mutually independent Wiener process, of which intensity $\mathbf{Q}(\mathbf{x}, \mathbf{y})$ of the covariance matrix function defined by (3) is given by

$$\mathbf{Q}(\mathbf{x}, \mathbf{y}) = \text{diag}(Q_1(\mathbf{x}, \mathbf{y}), Q_0(\mathbf{x}, \mathbf{y})). \tag{33}$$

Moreover, the functionals $J_f(\cdot)$ and $L(\cdot)$ in (5) are specified in a quadratic form as

$$\begin{aligned} J_f(\mathbf{X}(\mathbf{x}, t_f), t_f) &= \frac{1}{2} \int_D \mathbf{X}^T(\mathbf{x}, t_f) \Omega_f(\mathbf{x}, \mathbf{y}) \mathbf{X}(\mathbf{y}, t_f) d\mathbf{y}, \\ L(\mathbf{X}(\mathbf{x}, t), t, \mathbf{u}_d(\mathbf{x}, t), \mathbf{u}_b(\xi, t)) &= \frac{1}{2} \int_D \mathbf{X}^T(\mathbf{x}, t) \Omega(\mathbf{x}, \mathbf{y}) \mathbf{X}(\mathbf{y}, t) d\mathbf{y} \\ &\quad + \frac{1}{2} \int_D \mathbf{u}_d^T(\mathbf{x}, t) \mathbf{R}_d(\mathbf{x}, \mathbf{y}) \mathbf{u}_d(\mathbf{y}, t) d\mathbf{y} + \frac{1}{2} \int_S \mathbf{u}_b^T(\xi, t) \mathbf{R}_b(\xi, \eta) \mathbf{u}_b(\eta, t) dS_\eta, \end{aligned} \tag{34}$$

where $\Omega_f(\mathbf{x}, \mathbf{y})$ and $\Omega(\mathbf{x}, \mathbf{y})$ are $n \times n$ positive definite matrix functions; and $\mathbf{R}_d(\mathbf{x}, \mathbf{y})$ and $\mathbf{R}_b(\xi, \eta)$ are $m_d \times m_d$ and $m_b \times m_b$ positive definite matrix functions, respectively. It is clearly seen that the linear SDPS (30) includes the stochastic air pollution process (29).

5.3 Transformation to SES in Hilbert space

Applying the procedure in the previous section, we can transform the linear SDPS (30) to the following SES in Hilbert space:

$$\begin{aligned} d\mathbf{X}(t) &= \mathbf{A}(t)\mathbf{X}(t) + \mathbf{B}_d(t)\mathbf{u}_d(t) + \mathbf{G}_1(t)\mathbf{X}(t)\dot{w}_1(t) + \mathbf{G}_0(t)\dot{w}_0(t), \\ \mathbf{X}(t_0) &= \mathbf{X}_0, \\ \beta\mathbf{X}(t) &= \mathbf{B}_b(t)\mathbf{u}_b(t), \end{aligned} \quad (35)$$

and the functionals $J_f(\cdot)$ and $L(\cdot)$ from (34) are transformed to

$$\begin{aligned} J_f(\mathbf{X}(t_f), t_f) &= \frac{1}{2}\langle \mathbf{X}(t_f), \boldsymbol{\Omega}_f \mathbf{X}(t_f) \rangle_{\mathcal{L}_n^2(D)}, \\ L(\mathbf{X}(t), t, \mathbf{u}_d(t), \mathbf{u}_b(t)) &= \frac{1}{2}\langle \mathbf{X}(t), \boldsymbol{\Omega} \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} + \frac{1}{2}\langle \mathbf{u}_d(t), \mathbf{R}_d \mathbf{u}_d(t) \rangle_{\mathcal{L}_{m_d}^2(D)} \\ &\quad + \frac{1}{2}\langle \mathbf{u}_b(t), \mathbf{R}_b \mathbf{u}_b(t) \rangle_{\mathcal{L}_{m_b}^2(S)} \end{aligned} \quad (36)$$

Therefore, the Hamiltonian $H(\cdot)$ of the system (35) is given from (25) by

$$\begin{aligned} H(\cdot) &= \frac{1}{2}\langle \mathbf{X}(t), \boldsymbol{\Omega} \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} + \frac{1}{2}\langle \mathbf{u}_d(t), \mathbf{R}_d \mathbf{u}_d(t) \rangle_{\mathcal{L}_{m_d}^2(D)} + \frac{1}{2}\langle \mathbf{u}_b(t), \mathbf{R}_b \mathbf{u}_b(t) \rangle_{\mathcal{L}_{m_b}^2(S)} \\ &\quad + \langle W_{\mathbf{X}}(\mathbf{X}(t), t), \mathbf{A}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} \\ &\quad + \frac{1}{2}\text{Tr}[(\mathbf{G}_1(t)\mathbf{X}(t)Q_1\mathbf{X}^*(t)\mathbf{G}_1^*(t)W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t) + \mathbf{G}_0(t)Q_0\mathbf{G}_0^*(t))W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t)]. \end{aligned} \quad (37)$$

5.4 Control design

To design the controls $\mathbf{u}_d(t)$ and $\mathbf{u}_b(t)$, we need to solve the HJE (24) with $H(\cdot)$ given in (37). Let us assume that $W(\mathbf{X}(t), t)$ is of the following form

$$W(\mathbf{X}(t), t) = \frac{1}{2}\langle \mathbf{X}(t), \mathbf{P}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} + \frac{1}{2}P_0(t) \quad (38)$$

where $\mathbf{P}(t) \in \mathcal{L}(\mathcal{U}_1 \times \mathcal{U}_1)$ is a self-adjoint positive definite trace class operator with the kernel $\mathbf{P}(\mathbf{x}, \mathbf{y}, t)$ and $P_0(t)$ is a positive and real-valued function of t . From (38), the first and second Fréchet derivatives $W_{\mathbf{X}}(\mathbf{X}, t)$ and $W_{\mathbf{X}\mathbf{X}}(\mathbf{X}, t)$ are given by

$$\begin{aligned} W_{\mathbf{X}}(\mathbf{X}(t), t) &= \mathbf{P}(t)\mathbf{X}(t), \\ W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t) &= \mathbf{P}(t). \end{aligned} \quad (39)$$

Substituting (39) into (37) results in

$$\begin{aligned} H(\cdot) &= \frac{1}{2}\langle \mathbf{X}(t), \boldsymbol{\Omega} \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} + \frac{1}{2}\langle \mathbf{u}_d(t), \mathbf{R}_d \mathbf{u}_d(t) \rangle_{\mathcal{L}_{m_d}^2(D)} + \frac{1}{2}\langle \mathbf{u}_b(t), \mathbf{R}_b \mathbf{u}_b(t) \rangle_{\mathcal{L}_{m_b}^2(S)} \\ &\quad + \langle \mathbf{P}(t)\mathbf{X}(t), \mathbf{A}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} \\ &\quad + \frac{1}{2}\text{Tr}[(\mathbf{G}_1(t)\mathbf{X}(t)Q_1\mathbf{X}^*(t)\mathbf{G}_1^*(t)W_{\mathbf{X}\mathbf{X}}(\mathbf{X}(t), t) + \mathbf{G}_0(t)Q_0\mathbf{G}_0^*(t))\mathbf{P}(t)]. \end{aligned} \quad (40)$$

Applying Green's formula, see Lemma 3.1 in [14] to $\langle \mathbf{P}(t)\mathbf{X}(t), \mathbf{A}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)}$ yields

$$\begin{aligned} \langle \mathbf{P}(t)\mathbf{X}(t), \mathbf{A}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} &= \langle \mathbf{A}^*(t)\mathbf{P}(t)\mathbf{X}(t), \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} \\ &+ \langle \boldsymbol{\beta}(t)\mathbf{X}(t), \mathbf{P}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(S)} - \langle \boldsymbol{\beta}^*(t)\mathbf{P}(t)\mathbf{X}(t), \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(S)}, \end{aligned} \tag{41}$$

where $\mathbf{A}^*(t)$ and $\boldsymbol{\beta}^*(t)$ are given by

$$\begin{aligned} \mathbf{A}^*(t)[\bullet] &= \mathbf{A}_x^*[\bullet] = \sum_{i,j=1}^n \frac{\partial^2(\mathbf{A}_{ij}(\mathbf{x}, t)[\bullet])}{\partial x_i \partial x_j} - \sum_{i=1}^n \frac{\partial(\mathbf{B}_i(\mathbf{x}, t)[\bullet])}{\partial x_i} + \mathbf{C}(\mathbf{x}, t)[\bullet], \\ \boldsymbol{\beta}^*(t)[\bullet] &= \boldsymbol{\beta}_\xi^*[\bullet] = \sum_{j=1}^n \mathbf{A}_j(\boldsymbol{\xi}, t) \frac{\partial[\bullet]}{\partial x_j} - \sum_{i=1}^n \left[\mathbf{B}_i(\boldsymbol{\xi}, t) - \sum_{j=1}^n \frac{\partial \mathbf{A}_{ij}(\boldsymbol{\xi}, t)}{\partial x_j} \right] \\ &\quad \times \cos(\mathbf{n}_\xi, x_i)[\bullet] + \mathbf{F}(\boldsymbol{\xi}, t)[\bullet]. \end{aligned} \tag{42}$$

Substituting (40) with the use of (41) into the HJE (24) yields the following optimal controls

$$\begin{aligned} \mathbf{u}_d(t) &= -\mathbf{R}_d^{-1} \mathbf{B}_d^*(t)\mathbf{P}(t)\mathbf{X}(t), \\ \mathbf{u}_b(t) &= -\mathbf{R}_b^{-1} \mathbf{B}_b^*(t)\mathbf{P}(t)\mathbf{X}(t), \end{aligned} \tag{43}$$

and $\mathbf{P}(t)$ and $P_0(t)$ must satisfy the following equations:

$$\begin{aligned} \frac{d\mathbf{P}(t)}{dt} &= -\mathbf{A}^*(t)\mathbf{P}(t) - (\mathbf{A}^*(t)\mathbf{P}(t))^* - \boldsymbol{\Omega} - \mathbf{G}_1^*(t)\mathbf{P}(t)\mathbf{Q}_1\mathbf{G}_1(t) \\ &\quad + \mathbf{P}(t)\mathbf{B}_d(t)\mathbf{R}_d^{-1}\mathbf{B}_d^*(t)\mathbf{P}(t) + \mathbf{P}_b(t)\mathbf{B}_b(t)\mathbf{R}_b^{-1}\mathbf{B}_b^*(t)\mathbf{P}(t), \\ \mathbf{P}(t_f) &= \boldsymbol{\Omega}_f, \boldsymbol{\beta}^*\mathbf{P}(t) = 0, \\ \frac{dP_0(t)}{dt} &= -\text{Tr}(\mathbf{G}_0^*(t)\mathbf{P}(t)\mathbf{Q}_0\mathbf{G}_0(t)), P_0(t_f) = 0, \end{aligned} \tag{44}$$

where $(\mathbf{A}^*(t)\mathbf{P}(t))^*$ and $\mathbf{P}_b(t)$ are the adjoint of $\mathbf{A}^*(t)\mathbf{P}(t)$ and an element of $\mathcal{L}(\mathcal{L}_n^2(S), \mathcal{L}_n^2(S))$, respectively, and satisfy:

$$\begin{aligned} \langle \mathbf{X}(t), \mathbf{A}^*(t)\mathbf{P}(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)} &= \langle (\mathbf{A}^*(t)\mathbf{P}(t))^* \mathbf{X}(t), \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)}, \\ \langle \mathbf{P}(t)\mathbf{X}(t), \mathbf{X}(t) \rangle_{\mathcal{L}_n^2(S)} &= \langle \mathbf{X}(t), \mathbf{P}_b(t)\mathbf{X}(t) \rangle_{\mathcal{L}_n^2(D)}. \end{aligned} \tag{45}$$

5.5 Transformation of SES in Hilbert space to PDE in Euclidean space

For implementation, we now need to transform the controls $\mathbf{u}_d(t)$ and $\mathbf{u}_b(t)$ in (43) and $\mathbf{P}(t)$ and $P_0(t)$ in (44) to Euclidean space. For symmetric operators, the following two equations hold from the Schwartz kernel theorem [28]:

$$\begin{aligned} \mathbf{P}(t)\phi(t) &= \int_D \mathbf{P}(\mathbf{x}, \mathbf{y}, t)\phi(\mathbf{y}, t)d\mathbf{y}, \\ \mathbf{P}^{-1}(t)\phi(t) &= \int_D \bar{\mathbf{P}}(\mathbf{x}, \mathbf{y}, t)\phi(\mathbf{y}, t)d\mathbf{y}, \end{aligned} \tag{46}$$

where $\phi(t)$ is an infinitely differentiable function with compact support in D , $\mathbf{P}(\mathbf{x}, \mathbf{y}, t)$ and $\bar{\mathbf{P}}(\mathbf{x}, \mathbf{y}, t)$ are kernels of $\mathbf{P}(t)$ and $\mathbf{P}^{-1}(t)$, respectively. With (46), the controls $\mathbf{u}_d(t)$ and $\mathbf{u}_b(t)$ in (43) are transformed to Euclidean space as follows:

$$\begin{aligned} \mathbf{u}_d(\mathbf{x}, t) &= -\int_{D^2} \bar{\mathbf{R}}_d(\mathbf{x}, \mathbf{y}) \mathbf{B}_d^T(\mathbf{y}, t) \mathbf{P}(\mathbf{y}, \mathbf{y}', t) \mathbf{X}(\mathbf{y}', t) d\mathbf{y} d\mathbf{y}' \\ \mathbf{u}_b(\boldsymbol{\xi}, t) &= -\int_S \int_D \bar{\mathbf{R}}_b(\boldsymbol{\xi}, \boldsymbol{\eta}) \mathbf{B}_b^T(\boldsymbol{\eta}, t) \mathbf{P}(\boldsymbol{\eta}, \mathbf{x}, t) \mathbf{X}(\mathbf{x}, t) d\mathbf{x} dS_\eta. \end{aligned} \tag{47}$$

Moreover, $\mathbf{P}(t)$ and $P_0(t)$ in (44) are transformed to Euclidean space as follows:

$$\begin{aligned} \frac{\partial \mathbf{P}(\mathbf{x}, \mathbf{y}, t)}{dt} &= -\mathbf{A}_x^* \mathbf{P}(\mathbf{x}, \mathbf{y}, t) - (\mathbf{A}_y^* \mathbf{P}(\mathbf{x}, \mathbf{y}, t))^T - \boldsymbol{\Omega}(\mathbf{x}, \mathbf{y}) \\ &\quad - \int_D \mathbf{G}_1^T(\mathbf{x}, t) \mathbf{P}(\mathbf{x}, \mathbf{x}', t) Q_1(\mathbf{x}', \mathbf{y}) \mathbf{G}_1(\mathbf{y}, t) d\mathbf{x}' \\ &\quad + \int_{D^2} \mathbf{P}(\mathbf{x}, \mathbf{x}', t) \mathbf{B}_d(\mathbf{x}', t) \bar{\mathbf{R}}_d(\mathbf{x}', \mathbf{y}') \mathbf{B}_d^T(\mathbf{y}', t) \mathbf{P}(\mathbf{y}', \mathbf{y}, t) d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_{S^2} \mathbf{P}(\mathbf{x}, \boldsymbol{\xi}, t) \mathbf{B}_b(\boldsymbol{\xi}, t) \bar{\mathbf{R}}_b(\boldsymbol{\xi}, \boldsymbol{\eta}) \mathbf{B}_b^T(\boldsymbol{\eta}, t) \mathbf{P}(\boldsymbol{\eta}, \mathbf{y}, t) dS_\xi dS_\eta, \end{aligned} \tag{48}$$

$$\mathbf{P}(\mathbf{x}, \mathbf{y}, t_f) = \boldsymbol{\Omega}_f(\mathbf{x}, \mathbf{y}), \quad \beta_\xi^* \mathbf{P}(\mathbf{x}, \boldsymbol{\xi}, t) = 0,$$

$$\frac{dP_0(t)}{dt} = -\int_{D^2} \mathbf{G}_0^T(\mathbf{x}, t) \mathbf{P}(\mathbf{x}, \mathbf{y}, t) Q_0(\mathbf{x}, \mathbf{y}) \mathbf{G}_0(\mathbf{y}, t) d\mathbf{x} d\mathbf{y}, \quad P_0(t_f) = 0.$$

5.6 Finite number of controls

We assume that there are M_d controllers at fixed points $\mathbf{x}_1, \dots, \mathbf{x}_{M_d}$ of the domain D and M_b controllers at fixed points $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_{M_b}$ of the boundary S . Thus, by setting

$$\begin{aligned} \mathbf{B}_d(\mathbf{x}, t) \mathbf{u}_d(\mathbf{x}, t) &= \sum_{i=1}^{M_d} \mathbf{B}_d(\mathbf{x}_i, t) \mathbf{u}_d(\mathbf{x}_i, t) \mathbf{I} \delta(\mathbf{x} - \mathbf{x}_i), \\ \mathbf{B}_b(\boldsymbol{\xi}, t) \mathbf{u}_b(\boldsymbol{\xi}, t) &= \sum_{i=1}^{M_b} \mathbf{B}_b(\boldsymbol{\xi}_i, t) \mathbf{u}_b(\boldsymbol{\xi}_i, t) \mathbf{I} \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_i), \\ \mathbf{u}_d^T(\mathbf{x}, t) \mathbf{R}_d(\mathbf{x}, \mathbf{y}) \mathbf{u}_d(\mathbf{y}, t) &= \sum_{i,j=1}^{M_d} (\mathbf{u}_d^T(\mathbf{x}_i, t) \mathbf{R}_d(\mathbf{x}_i, \mathbf{y}_j) \mathbf{u}_d(\mathbf{y}_j, t)) \mathbf{I} \delta(\mathbf{x} - \mathbf{x}_i) \mathbf{I} \delta(\mathbf{y} - \mathbf{y}_j), \\ \mathbf{u}_b^T(\boldsymbol{\xi}, t) \mathbf{R}_b(\boldsymbol{\xi}, \boldsymbol{\eta}) \mathbf{u}_b(\boldsymbol{\eta}, t) &= \sum_{i,j=1}^{M_b} (\mathbf{u}_b^T(\boldsymbol{\xi}_i, t) \mathbf{R}_b(\boldsymbol{\xi}_i, \boldsymbol{\eta}_j) \mathbf{u}_b(\boldsymbol{\eta}_j, t)) \mathbf{I} \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_i) \mathbf{I} \delta(\boldsymbol{\eta} - \boldsymbol{\eta}_j), \end{aligned} \tag{49}$$

where $\delta(\cdot)$ is the Dirac delta function, we can use the same control design procedure in the previous subsections to obtain the following optimal controls:

$$\begin{aligned} \mathbf{u}_d(\mathbf{x}_i, t) &= -\sum_{j=1}^{M_d} \int_D \bar{\mathbf{R}}_d(\mathbf{x}_i, \mathbf{x}_j) \mathbf{B}_d^T(\mathbf{x}_j, t) \mathbf{P}(\mathbf{y}, \mathbf{x}_j, t) \mathbf{X}(\mathbf{y}, t) d\mathbf{y} \\ \mathbf{u}_b(\boldsymbol{\xi}_i, t) &= -\sum_{j=1}^{M_b} \int_D \bar{\mathbf{R}}_b(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) \mathbf{B}_b^T(\boldsymbol{\xi}_j, t) \mathbf{P}(\boldsymbol{\xi}_j, \mathbf{x}, t) \mathbf{X}(\mathbf{x}, t) d\mathbf{x}. \end{aligned} \tag{50}$$

Moreover, $\mathbf{P}(\mathbf{x}, \mathbf{y}, t)$ is given by:

$$\begin{aligned} \frac{\partial \mathbf{P}(\mathbf{x}, \mathbf{y}, t)}{dt} = & -\mathbf{A}_x^* \mathbf{P}(\mathbf{x}, \mathbf{y}, t) - (\mathbf{A}_y^* \mathbf{P}(\mathbf{x}, \mathbf{y}, t))^T - \Omega(\mathbf{x}, \mathbf{y}) \\ & - \int_D \mathbf{G}_1^T(\mathbf{x}, t) \mathbf{P}(\mathbf{x}, \mathbf{x}', t) \mathbf{Q}_1(\mathbf{x}', \mathbf{y}) \mathbf{G}_1(\mathbf{y}, t) d\mathbf{x}' \\ & + \sum_{i,j=1}^{M_d} \mathbf{P}(\mathbf{x}, \mathbf{x}'_i, t) \mathbf{B}_d(\mathbf{x}'_i, t) \bar{\mathbf{R}}_d(\mathbf{x}'_i, \mathbf{y}'_j) \mathbf{B}_d^T(\mathbf{y}'_j, t) \mathbf{P}(\mathbf{y}'_j, \mathbf{y}, t) \\ & + \sum_{i,j=1}^{M_b} \mathbf{P}(\mathbf{x}, \boldsymbol{\xi}_i, t) \mathbf{B}_b(\boldsymbol{\xi}_i, t) \bar{\mathbf{R}}_b(\boldsymbol{\xi}_i, \boldsymbol{\eta}_j) \mathbf{B}_b^T(\boldsymbol{\eta}_j, t) \mathbf{P}(\boldsymbol{\eta}_j, \mathbf{y}, t), \end{aligned} \quad (51)$$

$$\mathbf{P}(\mathbf{x}, \mathbf{y}, t_f) = \Omega_f(\mathbf{x}, \mathbf{y}),$$

$$\beta_{\boldsymbol{\xi}}^* \mathbf{P}(\mathbf{x}, \boldsymbol{\xi}, t) = 0,$$

and $P_0(t)$ is the same as given in (48).

6 Conclusions

This paper has presented a procedure to transform the nonlinear SDPSs in the Euclidean space to the SESs in the Hilbert space. Then, the HJE for the SESs was derived for solving the optimal control problem of nonlinear SDPSs under both state-dependent and additive stochastic disturbances. The HJE was applied to solve the optimal control problem of linear SDPSs, which include the stochastic air pollution process. The designed control laws were then transformed back to Euclidean space for implementation using the Schwartz kernel theorem. The proposed optimal control design can be applied to further improve performance of controlling other practical DPSs such as marine riser systems in [15], [16], and [33].

References

- [1] B. D. O. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*, Dover Publications, 1990.
- [2] R. Bellman, *Dynamic programming*, Princeton University Press, 1957.
- [3] A. Bennett, *Inverse modeling of the ocean and atmosphere*, Cambridge University Press, Cambridge, 2002.
- [4] A. F. Bennett, B. S. Chua, and L. M. Leslie, *Generalized inversion of a global numerical weather prediction model*, *Meteorology and Atmospheric Physics* **60** (1996), 165–178.
- [5] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and control of infinite dimensional systems*, Birkhauser, Boston, MA, 2007.

- [6] A. G. Butkovskiy and A. Y. Lerner, *The optimal control of systems with distributed parameters*, Automation remote Control **21** (1960), 472–477.
- [7] P. Cannarsa and G. D. Prato, *Some results on non-linear optimal control problems and Hamilton-Jacobi equations in infinite dimensions*, Journal of Functional Analysis **90** (1990), no. 1, 27–47.
- [8] R. Courant and D. Hilbert, *Methods of mathematical physics, vol. 1*, Wiley Interscience, New York, 1953.
- [9] R. F. Curtain and P. L. Falb, *Ito's lemma in infinite dimensions*, Journal of Mathematical Analysis and Applications **31** (1970), 434–448.
- [10] R. F. Curtain and H. J. Zwart, *An introduction to infinite-dimensional linear systems theory*, Springer-Verlag, New York, 1995.
- [11] J. Dieudonne, *Foundations of modern analysis*, Academic Press, New York, 1960.
- [12] K. D. Do, *Optimal control allocations for inverse optimal control of linear distributed parameter systems*, Journal of Systems Science and Complexity, Submitted (2013).
- [13] ———, *Optimal control of second-order nonlinear distributed parameter systems*, Systems & Control Letters, submitted (2013).
- [14] ———, *Inverse optimal filtering of linear distributed parameter systems*, Applied Mathematical Sciences **7** (2013, <http://dx.doi.org/10.12988/ams.2013.37361>), no. 119, 5901–5925.
- [15] K. D. Do and J. Pan, *Boundary control of transverse motion of marine risers with actuator dynamics*, Journal of Sound and Vibration **318** (2008), no. 4-5, 768–791.
- [16] ———, *Boundary control of three-dimensional inextensible marine risers*, Journal of Sound and Vibration **327** (2009), no. 3-5, 299–321.
- [17] G. Evensen, *Data assimilation*, Springer, Berlin, 2009.
- [18] W. Gawronski, *Dynamics and control of structures a modal approach*, Springer-Verlag, New York, 1998.
- [19] I. M. Gelfand and S. V. Fomin, *Calculus of variations*, Prentice-Hall, Englewood Cliffs, N.J, 1963.
- [20] M. A. Iglesias and C. Dawson, *The representer method for state and parameter estimation in single-phase darcy flow*, Computer Methods in Applied Mechanics and Engineering **196** (2007), 4577–4596.
- [21] H. Khalil, *Nonlinear systems*, Prentice Hall, 2002.

- [22] M. Krstic, I. Kanellakopoulos, and P.V. Kokotovic, *Nonlinear and adaptive control design*, Wiley, New York, 1995.
- [23] M. Krstic and A. Smyshlyaev, *Boundary control of pdes: A course on backstepping designs*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008.
- [24] H. J. Kushner, *Stochastic stability and control*, Academic Press, New York, 1967.
- [25] I. Lasiecka and R. Triggiani, *Control theory for partial differential equations: Continuous and approximation theories*, Cambridge University Press, Cambridge, UK, 2000.
- [26] M. D. Levine, *Trajectory optimization using the newton-raphson method*, *Automatica* **3** (1966), no. 3-4, 203–217.
- [27] J. Lions and S. K. Mitter, *Optimal control of systems governed by partial differential equations*, Springer-Verlag, Berlin, 1971.
- [28] P. L. Lions, *Optimal control of systems governed by partial differential equations*, Springer-Verlag, Berlin, 1971.
- [29] ———, *Generalized solutions of Hamilton-Jacobi Equations*, Pitman Lecture Notes, London, 1982.
- [30] L. Meirovitch, *Principles and techniques of vibrations*, Prentice-Hall, 1997.
- [31] S. K. Mitter, *Successive approximation methods for the solution of optimal control problems*, *Automatica* **3** (1966), no. 3-4, 135–149.
- [32] H. E. Ngodock, S. R. Smith, and G. A. Jacobs, *Cycling the representer algorithm for variational data assimilation with a nonlinear reduced gravity ocean model*, *Ocean Modelling* **19** (2007), no. 3-4, 101–111.
- [33] T. L. Nguyen, K. D. Do, and J. Pan, *Boundary control of two-dimensional marine risers with bending couplings*, *Journal of Sound and Vibration* **332** (2013), no. 16, 3605–3622.
- [34] E. Pardoux, *Equations aux derivees partielles stochastiques nonlineaires monotones*, Thesis, Universite Paris XI, 1975.
- [35] S. Ravindran, *A reduced-order approach for optimal control of fluids using proper orthogonal decomposition*, *International Journal for Numerical Methods in Fluids* **34** (2000), no. 5, 425–448.
- [36] A. C. Robinson, *A survey of optimal control of distributed-parameter systems*, *Automatica* **7** (1971), no. 3, 371–388.

- [37] J. H. Seinfeld and S. N. Pandis, *Atmospheric chemistry and physics: From air pollution to climate change*, John Wiley & Sons, New York, 1998.
- [38] P. E. Souganidis, *Existence of viscosity solutions of hamilton-jacobi equations*, Journal of Differential Equations **56** (1985), no. 3, 345–390.
- [39] S. G. Tzafestas and J. M. Nightingale, *Optimal control of a class of linear stochastic distributed-parameter systems*, Proceedings of IEE **115** (1968), no. 8, 1213–1220.
- [40] P. K. C. Wang, *Control of distributed parameter systems*, In Advances in Control Systems; Theory and Applications (Editor C. T. LEONDES), Academic Press, New York **1** (1964), 75–1972.
- [41] K. Yoshida, *Functionl analysis*, Springer-Verlag, Berlin, 1968.

Received: January 15, 2014