POLLUTION CONTROL FOR SWITCHING DIFFUSION MODELS: APPROXIMATION METHODS AND NUMERICAL RESULTS

CAOJIN ZHANG AND GEORGE YIN*
Department of Mathematics
Wayne State University
Detroit, MI 48202, USA

QING ZHANG
Department of Mathematics
University of Georgia
Athens, GA 30602, USA

LE YI WANG
Department of Electrical and Computer Engineering
Wayne State University
Detroit, MI 48202, USA

(Communicated by Xiaoying Han)

ABSTRACT. This work focuses on optimal pollution controls. The main effort is devoted to obtaining approximation methods for optimal pollution control. To take into consideration of random environment and other random factors, the control system is formulated as a controlled switching diffusion. Markov chain approximation techniques are used to design the computational schemes. Convergence of the algorithms are obtained. To demonstrate, numerical experimental results are presented. A particular feature is that computation using real data sets is provided.

1. Introduction. Professor Peter E. Kloeden has made fundamental contributions to stochastic differential equations, numerical methods, and applications. The book by Kloeden and Platen [8] has been used training a generation of mathematicians and scientists. An updated treatments can be found in Han and Kloeden [3] and references therein. Devoted to numerical solutions of pollution management problems, on the happy occasion of celebrating his 70th birthday, we dedicate this paper to Peter.

Environment pollution was identified as responsible for 9 million premature deaths in 2015, or 16% of all deaths worldwide. The Lancet Commission on Pollution and Health estimated that in 2016, ambient air pollution alone costs the
global economy 5.7 trillion dollars, or 4.4 percent of global GDP. Consequently, it has been widely recognized that pollution management is of vital importance for environmental health, social welfare, and world economy. A major issue of pollution management is concerned with the tradeoff of pollution accumulation and consumption.

Following the seminal paper of Keeler et al. [6], much work has been devoted to the study of optimal control of the related dynamic systems or economic systems. In [5], Kawaguchi and Morimoto treated a pollution accumulation problem of maximizing the long-run average welfare using a controlled diffusion model. Assume that an economy consumes some goods and also generates certain amount of pollution. The pollution stock is gradually degraded and its instantaneous growth rate encounters randomness that can be approximately represented by a random disturbance with mean zero and constant variance. The social welfare is then defined as a suitable combination of the utility of the consumption and the disutility of pollution. The basic optimal control problem is to find the optimal consumption strategies for the society in the long-run average sense.

Departing from the existing frameworks, we consider a general and yet more realistic model. Real-world economic systems are highly complex. It is difficult to capture the underlying structural and dynamic changes by merely the traditional diffusion models. To accommodate high level uncertainty, we introduce a discrete-event component. The new model is of hybrid in nature in the sense that it involves both continuous dynamics and discrete events. The continuous and discrete events coexist and interact. Some of the recent formulations and related work can be found in Yin and Zhu [16]; see also Jasso-Fuentes and Yin [4]. In this paper, we use a continuous-time Markov chain to represent the discrete event process. In our model, the Markov chain serves as a modulating force. The systems that we consider in this paper are controlled switching diffusions.

Within this formulation, stochastic control techniques can be employed to find optimal controls and establish their essential properties. At present, closed-form solutions to controlled switching diffusions are difficult to find. In order to implement the control policy in practice, one has to resort to numerical solutions. This paper is devoted to computational methods and numerical solutions for the class of controlled switching diffusion arising from pollution control problems. However, computable solutions are highly appreciated.

Our contributions of this paper can be summarized as follows.

(i) We propose a new model to address issues arising in environment management. Switching processes are incorporated into the existing diffusion models to capture random structural and dynamic changes in pollution dynamics that were not considered in the existing literature.

(ii) We develop new feasible numerical methods for the resulting stochastic control problem. By employing the key idea of Markov chain approximation techniques, we construct discrete-time controlled Markov chains, and show that the controlled chains are consistent with the original continuous-time controlled switching diffusions. Furthermore, we prove that our numerical schemes are convergent by means of the weak convergence method.

(iii) We verify our theoretical results by a number of numerical case studies.

(iv) We further validate our results by using real-world data. One of the interesting aspects is that this paper might be one of the first to use numerical stochastic control techniques to deal with real data problem arising from pollution con-
Problem formulation. Suppose that \( d, d_1, d_2 \) are positive integers, and that \( \mathcal{M} = \{1, 2, \ldots, m\} \) is a finite set. Let \( C(t) = (c_1(t), \ldots, c_d(t))' \) denote the consumption rate of goods, \( G \in \mathbb{R}^{d \times d_1} \) denote the generation of polluting materials when consumption plan implemented, and \( X(t) = (x_1(t), \ldots, x_d(t))' \) be the amount of polluting materials. For each \( i \in \mathcal{M} \), suppose that \( \Xi(i) = \begin{pmatrix} \xi_1(i) \\ \vdots \\ \xi_d(i) \end{pmatrix} \in \mathbb{R}^{d \times d} \), which represents the self growth rate and interaction between the pollution materials, where \( \xi_j(i) \) are row vectors for \( j = 1, \cdots, d \). \( \sigma(x, i) \) is an \( \mathbb{R}^{d \times d_2} \) matrix, \( \alpha(t) \) is a continuous-time Markov chain with state space \( \mathcal{M} \), and \( W(\cdot) \) is an \( \mathbb{R}^{d_2} \)-dimensional vector-valued standard Brownian motion independent of \( \alpha(t) \).

The amount of pollution materials \( X(t) \) satisfies the system equation

\[
dX(t) = [GC(t) + \Xi(\alpha(t))X(t)]dt + \sigma(X(t), \alpha(t))dW(t), \quad X(0) = x, \quad \alpha(0) = \alpha. \tag{1}
\]

The corresponding objective function is given by

\[
J(x, \alpha, C(\cdot)) = E \int_0^\infty e^{-\lambda t} U(X(t), \alpha(t), C(t))dt, \tag{2}
\]

where \( U(\cdot) \) is a utility function satisfying the following conditions:

(H1) Given \( x \) and \( i \), suppose that \( U \) is a smooth function w.r.t. the control \( C \) in that \( U(x, i, \cdot) \in C^2[0, \infty) \).

(H2) Given \( i \) and \( C \), the function \( U \) satisfies: For some \( \kappa \geq 1 \),

- \( U(x, i, C) \geq 0 \) and \( U \) is convex.
- There exists a constant \( D_0 > 0 \) such that \( \frac{1}{\lambda_0} |x| - D_0 \leq U(x, i, C) \leq D_0(1 + |x|^\kappa) \)
- \( |U(x, i, C) - U(y, i, C)| \leq D_0 |x - y|(1 + |x|^\kappa - 1 + |y|^\kappa - 1) \) for \( x, y \in \mathbb{R}^d \).

(H3) For each \( i \in \mathbb{Z}_+ \), \( \sigma(\cdot, i) \) is continuous. Moreover, \( \sigma(x, i)\sigma'(x, i) > 0 \) for each \( i \in \mathbb{Z}_+ \) and each \( x \) where \( \sigma'(x, i) \) denotes the transpose of \( \sigma(x, i) \).

We say that the consumption rate is admissible if it is \( \mathcal{F}_t \)-progressively measurable, where \( \mathcal{F}_t = \{ (X(s), \alpha(s)) : s \leq t \} \) such that \( 0 \leq |C(t)| \leq K_0 \) for some \( K_0 > 0 \). Assume that the pollution level \( \alpha(\cdot) \) is a continuous-time Markov chain with a finite...
state space $M = \{1, \ldots, m\}$. As another generalization of Morimoto [5], we assume that $\sigma$ to be dependent of $(X(t), \alpha(t))$.

**Example 2.1.** A particular, one-dimensional example is the following pollution control model. Let $c(t) \in \mathbb{R}$, $t \geq 0$, denote the consumption rate. Here $c(t)$ represents how fast the underlying economy consumes some goods and in the meantime generates pollution. The rate $c(t)$ is real valued also called the flow of pollution and is the control process. Let $X(t)$ be a real-valued process denoting the stock of pollution at time $t$ and let $\alpha(t)$ be a finite-state continuous-time Markov chain representing the environment mode. The pollution decay rate is given by $\rho(\alpha(t))$, a function of $\alpha(t)$. These variables satisfy the following stochastic differentiable equation

$$dX(t) = [c(t) - \rho(\alpha(t))X(t)]dt + \sigma_0(X(t), \alpha(t))dW(t), \quad X(0) = x,$$

where $\sigma_0$ is a function of $(x, \alpha)$ and $W(t)$ is a standard Brownian motion.

In this paper, the environment mode process $\alpha(t)$ represents the states of macroeconomic factors such as the level of resources available for environment protection, government regulatory policies on pollution control, etc. We assume that $\alpha(t)$ takes values in $M = \{1, 2, \ldots, m\}$.

The overall objective is the consumption of goods at a level for maintaining good economic activities and in the meantime keeping pollution in check. To quantify, let the social utility function be $\tilde{U}(c)$ and the social disutility of pollution be $D(x)$. The goal is to choose $c(t)$ over time to maximize the expected welfare

$$J(x, \alpha, c(\cdot)) = E \int_0^\infty e^{-\lambda t}[\tilde{U}(c(t)) - D(X(t))]dt,$$

where $\lambda > 0$ is a discount factor. For example, we can take $\tilde{U}(c) = c^\beta$ with $0 < \beta < 1$ and $D(x) = x^2$.

3. **HJB equations.** Consider the objective (2) with constraint (1). Denote the corresponding value function by

$$V(x, \alpha) = \max_C J(x, \alpha, C),$$

where $x \in \mathbb{R}^d$, $\alpha \in \{1, 2, \ldots\}$. We obtain the following system of dynamic programming equations

$$\max_C [LV(x, \alpha) - \lambda V(x, \alpha) + U(x, \alpha, C)] = 0$$

(6)

where the operator $L$ is defined as:

$$Lv(x, \alpha) = \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d a_{ij}x_ix_j \frac{\partial^2 v(x, \alpha)}{\partial x_i \partial x_j} + \sum_{i=1}^d (G_i C + \xi_i(\alpha) x) \frac{\partial v(x, \alpha)}{\partial x_i} + QV(x, \cdot)(\alpha).$$

(7)

Assume that for each $\alpha$, $v(\cdot, \alpha)$ has continuous partial derivatives up to the second order, that $Q = (q_{ij})$ is the generator of the Markov chain $\alpha(\cdot)$, and that

$$Qv(x, \cdot) = \sum_j q_{ij} v(x, j)$$

$$a_{ij} = (\sigma(x, \alpha) \sigma'(x, \alpha))_{ij}, G_i \text{ is the } i\text{-th row of matrix } G, \xi_i(\alpha) \text{ is the } i\text{-th row of } \Xi(\alpha), x \text{ is a } d\text{-dimension vector with } x = (x_1, x_2, \ldots, x_d)', \text{ and } z' \text{ denotes the transpose of } z \in \mathbb{R}^d.$$


4. Numerical method. To solve equation (6), a closed-form solution is difficult or virtually impossible to obtain due to the Markov switching. We need to deal with a system of equations instead of a single equation. As a viable alternative, we aim to obtain numerical solutions.

First, using the finite difference method with step size $h > 0$, we approximate the derivatives as:

$$
\frac{\partial V}{\partial x_i} = \begin{cases} 
\frac{V(x + e_i h, \alpha) - V(x, \alpha)}{h} & \text{if } G_i C - \xi_i(\alpha)x > 0 \\
-\frac{V(x, \alpha) - V(x - e_i h, \alpha)}{h} & \text{if } G_i C - \xi_i(\alpha)x < 0, 
\end{cases}
$$

(8)

$$
\frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{2V(x, \alpha) + V(x + e_i h + e_j h, \alpha) + V(x - e_i h - e_j h, \alpha)}{2h^2} - \frac{V(x + e_i h, \alpha) + V(x - e_i h, \alpha) + V(x + e_j h, \alpha) + V(x - e_j h, \alpha)}{2h^2}, 
$$

(9)

if $a_{ij}x_i x_j > 0$,

$$
\frac{\partial^2 V}{\partial x_i \partial x_j} = \frac{2V(x, \alpha) + V(x + e_i h - e_j h, \alpha) + V(x - e_i h + e_j h, \alpha)}{2h^2} + \frac{V(x + e_i h, \alpha) + V(x - e_i h, \alpha) + V(x + e_j h, \alpha) + V(x - e_j h, \alpha)}{2h^2}, 
$$

if $a_{ij}x_i x_j < 0$,  

(10)

where $\{e_i\}_{i=1}^d$ is the standard basis in $\mathbb{R}^d$. It follows that we can rewrite (6) as an iterative algorithm:

$$
V^n_{i+1} h, \alpha) = \max C \left\{ \begin{array}{c}
\{ \sum_{i=1}^d V^n_h(x + e_i h, \alpha)A_i^+ + V^n_h(x - e_i h, \alpha)A_i^- \\
\times \sum_{i=1}^d \sum_{j \neq i} \left[ V^n_h(x + e_i h + e_j h, \alpha)B^+_{i,j} + V^n_h(x - e_i h - e_j h, \alpha)B^+_{i,j} \\
+ \sum_{k \neq \alpha} q_{i,k} V^n(x, k) + U(x, C, \alpha) \right] \}
\end{array} \right\},
$$

(11)

with

$$
D = \sum_{i=1}^d \frac{G_i C + \xi_i(\alpha)x}{h} + \frac{a_{i,i} x_i^2}{2h^2} - \sum_{i \neq j} \frac{a_{i,j} x_i x_j}{h^2},
$$

$$
A_i^+ = \frac{a_{i,i} x_i^2}{2h^2} - \sum_{i \neq j} \frac{a_{i,j} x_i x_j}{2h^2} + \frac{(G_i C + \xi_i(\alpha)x)^+}{h},
$$

$$
A_i^- = \frac{a_{i,i} x_i^2}{2h^2} - \sum_{i \neq j} \frac{a_{i,j} x_i x_j}{2h^2} + \frac{(G_i C + \xi_i(\alpha)x)^-}{h},
$$

$$
B_{i,j}^+ = \frac{(a_{i,j} x_i x_j)^+}{2h^2},
$$

$$
B_{i,j}^- = \frac{(a_{i,j} x_i x_j)^-}{2h^2}.
$$

NUMERICAL METHODS FOR POLLUTION CONTROL 3671
Theorem 4.1. As $n \to \infty$, $V_h^{n+1}(x, \alpha)$ converges to a limit, denoted by $V_h(x, \alpha)$, such that $V_h = TV_h$.

It can be proved that the operator $T$ is a contraction mapping. We refer the reader to Barles and Jakobsen [1] and Lu et al. [12] for details. Then by the Banach fixed-point theorem, we conclude that there exists a unique $V_h$ satisfying $V_h = TV_h$.

5. Convergence of Markov chain approximation. By Theorem 4.1, we need only demonstrate that this limit gives a good approximation to the value function of the original problem (5). In this section, we first that show the limit in theorem 4.1 coincides with the value function for a constructed controlled Markov chain, denoted by $V_{MC}^h$. Then we demonstrate that $V_{MC}^h$ approximates the value function $V$ in equation (5).

5.1. Markov chain approximation. We propose to simulate the controlled stochastic process (1) by a controlled Markov chain and study underlying relation between our numerical approximation and the optimization problem under the constructed Markov chain.

To begin, we consider a discrete Markov chain with state space $S \times M$, where $S = \{x : x = ke_ih, i = 1, \ldots, d, k = 0, 1, 2, \ldots\}$, which corresponds to the nodes in our discrete domain with step size $h$, $M$ the state space of the finite state continuous-time Markov chain $\alpha(\cdot)$. We keep track of the path of such controlled Markov chain by $\{(\eta_i^h, \alpha_i^h) : i = 0, 1, 2, \ldots\}$ and denote the temporal difference as $\Delta t$. Furthermore, we let $C_n^h$ be a sequence of control actions, which is adapted to the smallest $\sigma$-algebra generated by $\{(\eta_i^h, \alpha_i^h), (\eta_i^h, \alpha_i^h), (\eta_i^h, \alpha_i^h), C_0^h, \ldots, C_{n-1}^h\}$.

To retain the local behavior of the original process (similar to the consideration of local consistency defined in Kushner and Dupuis [10, p.71]), we define the transition probabilities of the Markov chain $\{(\eta_i^h, \alpha_i^h)\}$ given control action $C$ by:

$$
P((x, \alpha), (x + e_i h, \alpha)|C) = \frac{A_i^+}{D},$$

$$P((x, \alpha), (x - e_i h, \alpha)|C) = \frac{A_i^-}{D},$$

$$P((x, \alpha), (x + e_i h \pm e_j h, \alpha)|C) = \frac{B_{ij}^+}{D},$$

$$P((x, \alpha), (x - e_i h \pm e_j h, \alpha)|C) = \frac{B_{ij}^-}{D},$$

$$P((x, \alpha), (x, k)|C) = \frac{q_{\alpha,k}}{D} \quad \text{for } k \neq \alpha,$$

$$P((x, \alpha), (y, k)|C) = 0, \quad \text{for any } y \in S, \text{ otherwise.}$$

We proceed to verify that our constructed approximating Markov chain is locally like the switching diffusion (1). Due to the presence of $\alpha(t)$, we need to modify the local consistency definition in [10]. The definition is given as follows.

Definition 5.1. The sequence $\{(\eta_i^h, \alpha_i^h)\}$, with temporal difference $\Delta t(x, i, C)$, is said to be locally consistent with equation (1) if

$$E_{x,i,n,C}(\Delta \eta_n^h) = [GC - \Xi(x)]\Delta t(x, i, C) + o(\Delta t(x, i, C)).$$
and use continuous-time interpolations where

\[ \eta^h(t) = \eta^h_{|n}, \quad \alpha^h(t) = \alpha^h_{|n}, \quad C^h(t) = C^h_{|n}, \quad \zeta^h(t) = n \quad \text{for} \quad t \in [\theta_n^h, \theta_{n+1}^h), \]

where \( \theta_n^h = \sum_{i=0}^{n-1} \Delta \eta^h_i \). Using the interpolation of the controlled Markov chain \( \{(\eta^h(\cdot), \alpha^h(\cdot), C^h(\cdot))\} \), we consider an optimal control problem with the objective function as (2). The corresponding problem is formulated as:

\[
J_{MC}^h(\eta^h(\cdot), \alpha^h(\cdot), c^h(\cdot)) = \mathbb{E} \sum_{i=0}^{\infty} e^{-\lambda_i} U(\eta^h_i, \alpha^h_i, C^h_i) \Delta t(\eta^h_i, \alpha^h_i, C^h_i),
\]

where

\[
V_{MC}^h(x, \alpha) = \max_{C} J_{MC}^h(\eta^h(\cdot), \alpha^h(\cdot), C^h(\cdot)).
\]

which can be regarded as a Markov decision process (MDP) Khalvati et al. [7]. Through dynamic programming principle, we can obtain the iterative form of the value function as

\[
V_{MC}^h(x, \alpha) = \max_{C} \left\{ \sum_{(y, j)} P((x, i), (y, j)|C) V_{MC}^h(y, j) + U(x, \alpha, C) \right\}.
\]

Note that the operator on the right-hand side agrees with that in (11). By the uniqueness of fixed point, we conclude that for any \( x > 0 \),

\[ V_h(x, \alpha) = V_{MC}^h(x, \alpha). \]

This shows that the value function obtained by our numerical algorithm can be considered as the value function of a related problem by solving an Markov decision process based on a simulated process (controlled Markov chain).

5.2. Weak convergence to controlled switching diffusions. This section demonstrates that appropriately interpolated process converges weakly to the controlled switching diffusion. First, the process \( \eta_n \) may be written as

\[
\eta^h_n = x + \sum_{i=0}^{n-1} \Delta \eta^h_i = x + \sum_{i=0}^{n-1} E_i \Delta \eta^h_i + \sum_{i=0}^{n-1} [\Delta \eta^h_i - E_i \Delta \eta^h_i].
\]

We denote

\[
M^h_n = \sum_{i=0}^{n-1} \Delta M^h_i = \sum_{i=0}^{n-1} [\Delta \eta^h_i - E_i \Delta \eta^h_i],
\]


which is a martingale w.r.t. the smallest $\sigma$-algebra generated by \{\(\eta_i^h, \alpha_i^h, C_i^h\), \(i \leq n\)\}. Denote the piecewise constant interpolation of \(M_n^h\) as \(M^h(t)\). That is, \(M^h(t) = M_n^h\) for \(t \in [t_n^h, t_{n+1}^h)\).

Using the local consistence property, we can interpret its continuous-time interpolation as a stochastic differential equation with driving force \(W^h(\cdot)\),

\[
\eta^h(t) = \int_0^t [G\sigma(c) + \Xi(\alpha^h(s))\eta^h(s)]ds + \int_0^t \sigma(\eta^h(s), \alpha^h(s))dW^h(s) + o^h(t), \tag{18}
\]

where

\[
W^h(t) = \int_0^t [\sigma(\eta^h(s), \alpha^h(s))]^{-1}dM^h(s), \tag{19}
\]

and \(o^h(t) \to 0\) in probability as \(h \to 0\).

To proceed, it is more convenient that we use a relaxed control representation [10, Section 9.5], denoted by \(m^h(\cdot)\). Using the relaxed control to represent the control actions \(C^h(\cdot)\), leads to the use of a nice compactness property in the analysis. To proceed, we denote the set of all admissible controls by \(\mathcal{C}\). We map the set \(\mathcal{C}\) into a larger space with \(m^h(\cdot, \cdot)\) defined on \(B(\mathcal{C} \times [0, \infty])\), where \(B\) denotes the Borel sets, which satisfies:

\[
m^h(\mathcal{C}, [0, t]) = t, \quad m^h(D, [0, t]) \text{ is } \mathcal{F}_t \text{ measurable for any } D \in B(\mathcal{C}), \tag{20}
\]

where \(\mathcal{F}_t\) is the smallest \(\sigma\)-algebra generated by \{\(\eta^h(s), \alpha^h(s), s \leq t\)\}. Moreover, we can further rewrite the process \(\eta^h(\cdot)\) as

\[
\eta^h(t) = \int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha^h(s))\eta^h(s)]m^h_s(dc)ds + \int_0^t \sigma(\eta^h(s), \alpha^h(s))dW^h(s), \tag{21}
\]

where we have used \(m^h_s(dc)\), the “derivative” (see [10, p.267] and also Song et al. [13]).

**Lemma 5.3.** The sequence \{\(\eta^h(\cdot), \alpha^h(\cdot), m^h(\cdot), W^h(\cdot)\)\} is tight.

**Proof.** We define \(F_t^h\) as the smallest \(\sigma\) algebra generated by

\[
\{\eta^h(s), \alpha^h(s), m^h(s), W^h(s), s \leq t\}.
\]

Denote the corresponding conditional expectation on \(F_t^h\) by \(E_{t}^h\). For any \(\varepsilon > 0\), and \(t, s > 0\) satisfying \(s \leq \varepsilon\), it can be shown that (see, for example, using similar techniques in [14, pp. 218-220]) for some \(\gamma(h, \varepsilon) > 0\),

\[
E_{t}^h|\alpha^h(t + s) - \alpha^h(t)|^2 \leq E_{t}^h\gamma(h, \varepsilon)
\]

such that

\[
\lim_{\varepsilon \to 0} \lim_{h \to 0} E\gamma(h, \varepsilon) = 0.
\]

The process \(\alpha^h(\cdot)\) is thus tight by Kushner [9, Theorem 3, p.47].

Using the representation \(W^h(\cdot)\) in (19), assumption (H3), and the martingale properties of \(M_n^h\) defined in (17), By the local consistence, we have for any \(\varepsilon > 0\),
any \( t, s > 0 \) with \( 0 \leq s \leq \varepsilon \), there is a \( \gamma(h, \varepsilon) \) satisfying
\[
E_h^t |W^h(t + s) - W^h(s)|^2 \\
\leq E_h^t \left( \sum_{i=0}^\infty \sigma^{-1}(\eta_i^h, \alpha_i^h)(\Delta \eta_i^h - \Delta \eta_i^h) \right)^2 \\
\times \sum_{j=0}^\infty \sigma^{-1}(\eta_j^h, \alpha_j^h)(\Delta \eta_j^h - \Delta \eta_j^h) \\
\leq E_h^t \gamma(h, \varepsilon) + o(h),
\]
and
\[
\lim_{\varepsilon \to 0} \limsup_{h \to 0} E\gamma(h, \varepsilon) = 0.
\]
Thus we obtain the tightness of \( \{\eta^h(\cdot)\} \). Likewise, we can show \( \eta^h(\cdot) \) is tight.
Next, note that the measures \( m^h(\cdot) \) live in a compact set, thus it is tight. Putting the estimates together, the desired result of the lemma follows.

**Lemma 5.4.** The weak limit \( \{\eta^h(\cdot), \alpha^h(\cdot), m^h(\cdot), W^h(\cdot)\} \) satisfies equation (1).

**Proof.** By the tightness and the Prohorov theorem, there is a convergent subsequence of \( \{\eta^h(\cdot), \alpha^h(\cdot), m^h(\cdot), W^h(\cdot)\} \). For notational simplicity, we still index the subsequence by \( h \) with limit denoted by \( \{\eta(\cdot), \alpha(\cdot), m(\cdot), \hat{W}(\cdot)\} \).

We first show that \( W^h(\cdot) \) converges weakly to \( \hat{W}(\cdot) \), which is a standard Brownian motion. For \( \varepsilon > 0 \) and any process \( y(\cdot) \), define the process \( y_\varepsilon(\cdot) \) by \( y_\varepsilon(t) = y(\varepsilon t), t \in [\varepsilon, \varepsilon + \varepsilon] \). We note that for any \( t, s > 0 \),
\[
W^h(t + s) - W^h(t) = \int_s^{t + s} \sigma^{-1}(\eta^h(v), j)I(\alpha^h(v) = j) dM^h(v) \\
= \sum_{j=1}^m \int_s^{t + s} \sigma^{-1}(\eta^h(v), \alpha^h(v)) dM^h(v) + \epsilon^h(t),
\]
where \( E|\epsilon^h(t)| \to 0 \) as \( \varepsilon \to 0 \) uniform in \( t \) in any bounded time interval.

To characterize \( \hat{W}(\cdot) \), let \( t > 0, \delta > 0, \kappa_1, \kappa_2, \{t_k : k \leq \kappa_1\} \) be given such that \( t_k \leq t \leq t + \delta \) for all \( k \leq \kappa_1 \), \( \phi_j(\cdot) \) for \( j \leq \kappa_2 \) is real-valued and continuous functions on \( U \times [0, \infty) \) with compact support for all \( j \leq \kappa_2 \). Define
\[
(\phi_j, m)_t \overset{\text{def}}{=} \int_0^t \int_C \phi_j(r, s) m(dr)ds = \int_0^t \int_C \phi_j(r, s) m_s(dr)ds.
\]
Let \( H(\cdot) \) be a real-valued and continuous function of its arguments with compact support. Note that \( W^h(\cdot) \) is an \( F^h_t \)-martingale. Thus we obtain
\[
EH(\eta^h(t_k), \alpha^h(t_k), W^h(t_k), (\phi_j, m^h)_{t_k}, j \leq \kappa_2, k \leq \kappa_1)[W^h(t + s) - W^h(t)] = 0.
\]
Using the Skorohod representation and the dominant convergence theorem, passing to the limits as \( h \to 0 \), we have
\[
EH(X(t_k), \alpha(t_k), \hat{W}(t_k), (\phi_j, m)_{t_k}, j \leq \kappa_2, k \leq \kappa_1) \\
\times (\hat{W}(t + s) - \hat{W}(t)) = 0.
\]
Since \( \hat{W}(\cdot) \) has continuous sample paths, \( \hat{W}(\cdot) \) is a continuous \( \mathcal{F}_t \)-martingale. Note that
\[
E[W^h(t + s)W^h(t + s) - W^h(t)W^h(t)] = E[(W^h(t + s) - W^h(t))'(W^h(t + s) - W^h(t))].
\]
By the weak convergence and the Skorohod representation theorem, we have
\[
EH(X(t_k), \alpha(t_k), \widehat{W}(t_k), (\phi_j, m)_{i_k}, j \leq \kappa_2, k \leq \kappa_1) \\
\times (\widehat{W}(t + s) \widehat{W}(t + s) - \widehat{W}(t) \widehat{W}(t) - sI) = 0.
\]
The quadratic variation of the martingale \( \widehat{W}(t) \) is \( t \), which implies \( \widehat{W}(\cdot) \) is an \( \mathcal{F}_t \)-Wiener process. Then, by the tightness of \( \{ \eta^k(\cdot), \alpha^k(\cdot) \} \), we can further get from (21),
\[
\eta^h(t) = \int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha^h(s))\eta^h(s)]m^h_s(\text{d}c)\text{d}s + \int_0^t \sigma(\eta^h_s, \alpha^h_s)\text{d}W^h(s) + e^h_s(t),
\]
where
\[
\lim_{\varepsilon \to 0} \lim_{h \to 0} \sup E|e^h_s(t)| \to 0,
\]
uniformly on any finite time interval.

Letting \( h \to 0 \), the convergence with probability one (using Skorohod representation) yields
\[
E \left| \int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha^h(s))\eta^h(s)]m^h_s(\text{d}c)\text{d}s \\
- \int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha(s))X(s)]m_s(\text{d}c)\text{d}s \right. \to 0,
\]
uniformly in \( t \). In addition, \( \{ m^h(\cdot) \} \) converges in the “compact weak” topology. For any bounded and continuous function \( \phi(\cdot) \) with compact support,
\[
(\phi, m^h)_{\infty} = \int_0^\infty \int_U \phi(r, s)m^h_s(\text{d}r)\text{d}s \\
\to \int_0^\infty \int_U \phi(r, s)m_s(\text{d}r)\text{d}s = (\phi, m)_{\infty}.
\]
The weak convergence and the Skorohod representation imply that
\[
\int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha^h(s))\eta^h(s)]m^h_s(\text{d}r)\text{d}s \\
\to \int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha(s))X(s)]m_s(\text{d}r)\text{d}s,
\]
uniformly in \( t \) on any bounded interval with probability one (using Skorohod representation). Because \( \eta^h(\cdot) \) and \( \alpha^h(\cdot) \) are piecewise constant functions, it follows from the probability one convergence,
\[
\int_0^t \sigma(\eta^h_s, \alpha^h_s)\text{d}W^h(s) \\
\to \int_0^t \sigma(\eta_s, \alpha_s)\text{d}\widehat{W}(s).
\]
Combining the estimates obtained thus far,
\[
X(t) = x + \int_0^t \int_{\mathcal{C}} [Gc + \Xi(\alpha(s))X(s)]m_s(\text{d}c)\text{d}s \\
+ \int_0^t \sigma(\eta_s, \alpha_s)\text{d}\widehat{W}(s) + e(t),
\]
where \( \lim_{\varepsilon \to 0} E|e(\varepsilon(t)| = 0 \) uniformly on any bounded interval. Letting \( \varepsilon \to 0 \) in the above leads to the desired result. \( \square \)
5.3. **Convergence of objective and value functions.** It remains to demonstrate that $V_{MC}^h(x, \alpha)$ approximates $V(x, \alpha)$, the value function of the original problem.

**Lemma 5.5.** $J_{MC}^h(\eta^h(\cdot), \alpha^h(\cdot), C^h(\cdot)) \rightarrow J(X(\cdot), \alpha(\cdot), C(\cdot))$.

*Proof.* In view of the weak convergence and the Skorohod representation theorem, the desired result follows by noting the continuity and hence the boundedness of the objective function. □

**Theorem 5.6.** $\lim \inf h V_{MC}^h(x, \alpha) = V(x, \alpha) = \lim \sup h V_{MC}^h(x, \alpha)$

*Proof.* First, it is readily seen that

$$\lim \sup h V_{MC}^h(x, \alpha) \leq V(x, \alpha)$$

follows directly from the weak convergence in Theorem 5.4. For any $\varepsilon > 0$, we need to prove

$$\lim \inf h V_{MC}^h(x, \alpha) > V(x, \alpha) - \varepsilon,$$

first, suppose $C_{op}(\cdot)$ is the optimal control for with trajectory $X^\varepsilon(\cdot)$ and relaxed control representation $m_{op}(\cdot)$. By the chattering lemma [10], there exists an ordinary control $C_{op}^b(\cdot)$, which is constant on $[k\delta, (k+1)\delta]$, and only takes finite values such that

$$J(x, \alpha, C_{op}^b) > J(x, \alpha, C_{op}) - \varepsilon = V(x, \alpha) - \varepsilon.$$

Then we consider optimal control problem of the process $X^{\varepsilon, \Delta t}(t) = X^\varepsilon(k\Delta t)$ for $t \in [k\Delta t, (k+1)\Delta t]$, where $k = 0, 1, 2, \ldots$, which can be considered as a piecewise constant approximation of the process $X^\varepsilon(\cdot)$. Denote the optimal control as $\hat{C}(\cdot)$, with trajectory $\hat{X}^\varepsilon(\cdot)$ and relaxed control $\hat{m}(\cdot)$. It follows:

$$J(x, \alpha, \hat{C}) > J(x, \alpha, C_{op}^b) > V(x, \alpha) - \varepsilon.$$

By the chattering lemma again, there exists $\Delta t > 0$ and an ordinary control $C^{\Delta t}(\cdot)$, which takes finite values and which is constant on $[k\Delta t, (k+1)\Delta t]$, for $k = 0, 1, 2, \ldots$ such that

$$J(x, \alpha, C^{\Delta t}) > J(x, \alpha, \hat{C}) - \varepsilon.$$

It follows $C^{\Delta t}(\cdot)$ is admissible with respect to the controlled Markov chain defined in Section 5.1. Using lemma 5.5 we got

$$\lim \inf h V^h(x, \alpha) > J^h(x, \alpha, C^{\Delta t}) \rightarrow J(x, \alpha, C^{\Delta t}),$$

$$\lim \inf h V^h(x, \alpha) > V(x, \alpha) - 2\varepsilon.$$

The proof is complete. □

6. **Numerical examples.** In this section, we use several examples to illustrate our numerical approximation methods.

**Example 6.1.** We first consider a one-dimension case given in (3) and (4). In particular, we take $D(x) = 100x$, $\tilde{U}(c) = 2\sqrt{c}$, and $\sigma_0(x, i) = \sigma(i)x$. Suppose that the Markov chain has two states, denoted by 1 and 2. The generator of the Markov chain is given by

$$Q = \begin{pmatrix} -2.2 & 2.2 \\ 3.1 & -3.1 \end{pmatrix}.$$
Let $\lambda = 0.05$ with pollution generation/decay rates $\rho(1) = 0.7$, $\rho(2) = 0.3$, the noise “volatility” $\sigma(1) = 0.1$, $\sigma(2) = 2.3$, and consumption rate $c(t) \in [0, 3]$. We use our numerical scheme to obtain the optimal control actions and corresponding value functions. The results are depicted in Figures 1 and 2, respectively. As it can be seen that the optimal controls are of the threshold type. In state 1, if the accumulation of pollution is less than $x = 63$, we can perform full consumption rate. If the stock exceeds $x = 63$, we should stop consuming. In state 2, the threshold value is $x = 80$. This value is greater than that of state 1. It is due to the greater decay rate $\rho(2)$ in state 2. We obtain the following result.

Example 6.2. If we change the social disutility function to $D(x) = 100x^2$, a quadratic function, and keep the rest of settings as the previous example, the corresponding control actions and value functions are demonstrated in Fig. 3. The thresholds are smaller than those in Example 6.1. The reason is that a quadratic disutility function has a much larger rate of increasing with respect to $x$, the stock of pollution. With control actions, although it will bring the benefits from utility function $U(\cdot)$, the disutility part increases more dramatically with the accumulation of pollution. As a tradeoff, the optimal decision is not to implement the control when $x > 41$ in state 1 and $x > 69$ in state 2.

Example 6.3. Next, we try to find the optimal strategy from a real-data scenario. We use real air pollution data provided in De Vito et al. [2] to calibrate our regime-switching model. In particular, it records the accumulation of 13 materials that
generate air pollution. The data collection period is from 03/10/2004, 18:00 to 04/04/2005, 14:00. Through statistical analysis, we identify that the distributions of the hourly change rate of the log NOx concentration and log NO\textsubscript{2} concentration have three underlying states, respectively. Their histograms are shown in Fig. 4. For experimental purpose, we split the data set into two parts. The first part contains the data from 03/10/2004, 18:00 to 11/15/2004, 17:00 and is used for model calibration and for derivation of the optimal strategy, denoted as the training set \( D_{\text{train}} \). The second part is from 11/15/2004, 18:00 to 04/04/2005, 14:00, denoted as the testing set \( D_{\text{test}} \). We assume \( D_{\text{train}} \) is historical data and \( D_{\text{test}} \) is data collected in future which is not observed by the model. We calibrate the model using \( D_{\text{train}} \) and apply the model implied consumption strategy on \( D_{\text{test}} \) to test the model’s generalization ability on new samples.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure3}
\caption{The control actions and value functions in two states}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figure4}
\caption{The histograms of distributions}
\end{figure}
In this example, we analyze the accumulation process of NO\textsubscript{2}. Using an Expectation Maximization (EM) based algorithm combined with a Bayesian Inference method as Zhang et al. [18], we fit the hourly observations of NO\textsubscript{2} in \(D_{\text{train}}\) into a one-dimensional regime-switching diffusion model (1) with 3 underlying states.

We obtain \(\rho(0) = 0.0078, \sigma^2(0) = 0.043\) for the 0th hidden state, \(\rho(1) = -8.609E, E\sigma^2(1) = 0.566\) for the first hidden state, and \(\rho(2) = 8.422, \sigma^2(2) = 0.640\) for the second hidden state. Also, the estimated generator is

\[
Q = \begin{pmatrix}
-1.087 & 0.992 & 0.095 \\
2.41 & -24 & 21.59 \\
23.90 & 0 & -23.90 \\
\end{pmatrix}.
\]

Given the utility function:

\[
U(c) - D(x) = 2\sqrt{e} - 3\ln x,
\]

with the control set \(c \in [0, 1]\). We consider optimal consumption plan as to the maximize the utility function under the pollution accumulation process. Using our numerical method, we demonstrate the value functions and corresponding optimal controls in Figure 5, respectively.

Using our numerical method, the optimal controls are obtained as

- when \(\alpha = 0\),
  \[
  c(x, 0) = \begin{cases} 
  1 & \text{if } \ln x < -4.7 \text{ or } \ln x \geq 1.1, \\
  0 & \text{otherwise}; 
  \end{cases}
  \]

- when \(\alpha = 1\),
  \[
  c(x, 1) = \begin{cases} 
  1 & \text{if } \ln x < 0 \text{ or } \ln x \geq 0.1, \\
  0 & \text{otherwise}; 
  \end{cases}
  \]

- when \(\alpha = 2\),
  \[
  c(x, 2) = \begin{cases} 
  1 & \text{if } \ln x < -4.8 \text{ or } \ln x \geq -0.1, \\
  0 & \text{otherwise}. 
  \end{cases}
  \]

We fix our model and test its performance on testing set \(D_{\text{test}}\) and keep track of the empirical accumulation utility

\[
\frac{1}{|D_{\text{test}}|} \sum_{x \in D_{\text{test}}} e^{-\lambda t_i} [U(c) - D(x)] \Delta t.
\]

We demonstrate the control actions with respect to time span on Fig. 6. The accumulative loss is \(-55.89\).

**Example 6.4.** Next, we consider the NO\textsubscript{x} accumulation process. Similar as before, using \(D_{\text{train}}\), we calibrate the model to get parameter estimates as: \(\rho(0) = 0.0095, \sigma^2(0) = 0.121, \rho(1) = -8.698, \sigma^2(1) = 0.70, \rho(2) = 8.432, \sigma^2(2) = 0.963\). The estimated generator matrix is

\[
Q = \begin{pmatrix}
-1.069 & 0.979 & 0.09 \\
2.22 & -24 & 21.78 \\
23.79 & 0.105 & -23.895 \\
\end{pmatrix}.
\]

Given the same utility function, we obtain the optimal controls as
• when $\alpha = 0$,

$$c(x) = \begin{cases} 
1 & \text{if } \ln x < -4.5 \text{ or } \ln x \geq 2.4, \\
0 & \text{otherwise}; 
\end{cases}$$
when $\alpha = 1$,
\[
c(x) = \begin{cases} 
1 & \text{if } \ln x < -1.2 \text{ or } \ln x \geq 5.1, \\
0 & \text{otherwise}; 
\end{cases}
\]

when $\alpha = 2$,
\[
c(x) = \begin{cases} 
1 & \text{if } \ln x < -4.8 \text{ or } \ln x \geq 0.4, \\
0 & \text{otherwise}. 
\end{cases}
\]

The control actions on the testing set is demonstrated in Fig. 7 with accumulative loss $-69.59$.

Example 6.5. We consider a two-dimensional system. As can be seen that the processes NOx and NO$_2$ are highly correlated. It is natural to integrate them as a 2-dimension correlated process of (1) with a full covariance matrix. Consider $D_{\text{train}}$ for NO$_2$ and NO$_x$: $\{x_1^1, \ldots, x_N^1\}$, $\{x_1^2, \ldots, x_N^2\}$, respectively. It is known that combustion of natural gas and coal are two important industrial sources to produce pollution materials NO$_x$ and NO$_2$. We consider control variable $C = (c_1, c_2)$, corresponding to the consumption of natural gas and coal, respectively. We assume that the consumption of per unit natural gas generates 0.75 for NO$_x$ and 0.45 for NO$_2$, and the consumption of per unit coal generates 2.13 NO$_x$ and 1.32 NO$_2$. It follows that
\[
G = \begin{pmatrix} 
0.75 & 0.45 \\
2.13 & 1.32 
\end{pmatrix}.
\]

We further assume the pollution materials can be described by
\[
\begin{align*}
\ln x_1^1 &= G_1 C + \Xi_1,1(\alpha) \Delta t \ln x_{i-1}^1 + \Xi_1,2(\alpha) \Delta t \ln x_{i-1}^2 + \sigma_1,1(\alpha) \sqrt{\Delta t} \epsilon_1 + \sigma_1,2(\alpha) \sqrt{\Delta t} \epsilon_2 \\
\ln x_1^2 &= G_2 C + \Xi_2,1(\alpha) \Delta t \ln x_{i-1}^1 + \Xi_2,2(\alpha) \Delta t \ln x_{i-1}^2 + \sigma_2,1(\alpha) \sqrt{\Delta t} \epsilon_1 + \sigma_2,2(\alpha) \sqrt{\Delta t} \epsilon_2,
\end{align*}
\]
which is known as the system of mixture regression model, where $\Delta t$ is the length of each time interval, and $\epsilon_1, \epsilon_2 \sim N(0, 1)$ are two independent standard Gaussian noises.
Using the EM based algorithm, we calibrate the mixture regression model (24) with data from $D_{\text{train}}$ and obtain parameter estimates as

$$
\Xi(1) = \begin{pmatrix} 0.004 & 0.002 \\ 0.0001 & 0.006 \end{pmatrix},
$$

$$
\Xi(2) = \begin{pmatrix} 6.5 & 2.1 \\ 3.4 & 5.9 \end{pmatrix},
$$

$$
\Xi(3) = \begin{pmatrix} -8.7 & -4.4 \\ -2.3 & -8.5 \end{pmatrix},
$$

$$
a(1) = \sigma(1)\sigma'(1) = \begin{pmatrix} 0.11 & 0.06 \\ 0.06 & 0.036 \end{pmatrix},
$$

$$
a(2) = \sigma(2)\sigma'(2) = \begin{pmatrix} 13.7 & 12.8 \\ 12.8 & 13.06 \end{pmatrix},
$$

$$
a(3) = \sigma(3)\sigma'(3) = \begin{pmatrix} 0.69 & 0.29 \\ 0.29 & 0.19 \end{pmatrix},
$$

and the generator is given by

$$
Q = \begin{pmatrix} -1.2 & 1.1 & 0.1 \\ 2.0 & -24.0 & 22.0 \\ 23.8 & 1.1 & -24.9 \end{pmatrix}.
$$

Based on these relations, we obtain the optimal control policy given in Fig 8 (b), (d), and (f). Consider the utility function given by

$$
U(x_1, x_2, c_1, c_2) = 2\sqrt{c_1} + 2\sqrt{c_2} - 3\ln x_1 - 3\ln x_2,
$$

where $x_1, x_2$ represent the state variables of NOx and NO$_2$ concentration, respectively, and $c_1, c_2 \in [0, 1]$ are the corresponding control variables. We solve the maximization problem and obtain 4 optimal control actions: $(c_1, c_2) = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$ indexed as 0,1,2,3. We then obtain the value functions and optimal actions in

Next, we examine the performance of our optimal strategy on the testing data. Using a K-Nearest-Neighbor (KNN) Larose [11] algorithm, we estimate the optimal control actions on the testing data set and calculate the accumulative utility: $-114.25$.

**Example 6.6.** Consider a multi-dimensional case. In this example, we aim to obtain optimal control $C = (c_1, c_2)$ for a state space of 4 pollution materials NO$_x$, NO$_2$, CO, and O$_3$, with

$$
G = \begin{pmatrix} 0.75 & 0.45 \\ 2.13 & 1.32 \\ 3.42 & 2.54 \\ 1.68 & 2.01 \end{pmatrix}.
$$

To overcome the curse of dimensionality, we use the principle component analysis to identify the most significant contributing sources of pollution. This involves a
“learning” map $H$ to project the covariate space into a $\mathbb{R}^2$ space. We use $D_{\text{train}}$ to learn the map: $W$ where $H' = \begin{pmatrix} 0.718 & 0.696 & -0.008 & 0.007 \\ -0.056 & 0.048 & -0.989 & -0.131 \end{pmatrix}$.

Then we use the dimension reduction by $\tilde{X} = XH$ to obtain a covariate matrix $R_{6000 \times 2}$ and fit it into a 2-dimension regime switching diffusion model as (24). We obtain the parameter estimates as:

$$\Xi(1) = \begin{pmatrix} -0.002 & 0 \\ 0 & -0.002 \end{pmatrix},$$

$$\Xi(2) = \begin{pmatrix} -12.69 & 0 \\ 0 & 0.767 \end{pmatrix},$$

$$\Xi(3) = \begin{pmatrix} 4.94 & 0 \\ 0 & -0.272 \end{pmatrix}.$$
\[ a(1) = \sigma(1)\sigma'(1) = \begin{pmatrix} 0.106 & -0.09 \\ -0.09 & 0.107 \end{pmatrix}, \]
\[ a(2) = \sigma(2)\sigma'(2) = \begin{pmatrix} 0.91 & -0.36 \\ -0.36 & 2.05 \end{pmatrix}, \]
\[ a(3) = \sigma(3)\sigma'(3) = \begin{pmatrix} 36.32 & 6.14 \\ 6.14 & 7.56 \end{pmatrix}, \]
and generator
\[
Q = \begin{pmatrix} -4.2 & 3.1 & 0.1 \\ 8.0 & -30.0 & 22.0 \\ 23.8 & 4.8 & -28.6 \end{pmatrix}.
\]

We consider the objective function
\[
U(x_1, \ldots, x_4, c_1, c_2) = 2\sqrt{c_1} + 2\sqrt{c_2} - 3 \sum_{i=1}^{4} \ln x_i.
\]

To proceed, we apply our numerical method to derive the optimal controls \( C^*(\tilde{X}) = (c_1, c_2) \) with respect to the state space \( \tilde{X} \). There are 4 optimal control actions: \((c_1, c_2) = \{(0,0), (1,0), (0,1), (1,1)\}\) indexed as 0,1,2,3. We demonstrate the value functions and optimal actions with respect to the state space \( \tilde{X} \) as in Figure 9.

We compute the optimal controls on \( D_{test} \), by implementing the data transformation \( \tilde{X} = XH \) and exploiting our optimal controls \( C^*(\tilde{X}) \). The optimal control on \( D_{test} \) is demonstrated in Fig. 10. The accumulative utility is \(-415.37\), which is much better than the loss \(-1503.33\) by using a random policy.

7. Further remarks. This paper is devoted to numerical solutions of a pollution control problem. While this study is limited to a simplified model structure, more general model structures can be considered such as nonlinear models in the continuous state \( x \) system. In addition, to balance the future and current generations, more advanced criteria [4] can be treated. These will be pursued in our future studies.

REFERENCES

[1] G. Barles and E. R. Jakobsen, On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations, ESAIM: Mathematical Modelling and Numerical Analysis, 36 (2002), 33–54.
[2] S. De Vito, E. Massera, M. Piga, L. Martinotto and G. Di Francia, On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, Sensors and Actuators B: Chemical, 129 (2008), 750–757.
[3] X. Han and P. E. Kloeden, Random Ordinary Differential Equations and Their Numerical Solution, Springer, Singapore, 2017.
[4] H. Jasso-Fuentes and G. Yin, Advanced Criteria for controlled Markov-Modulated Diffusions in an Infinite Horizon: Overtaking, Bias, and Blackwell Optimality, Science Press, Beijing, China, 2013.
[5] K. Kawaguchi and H. Morimoto, Long-run average welfare in a pollution accumulation model, J. Econom. Dyn. Control, 31 (2007), 703–720.
[6] E. Keefer, M. Spence and R. Zeckhauser, The optimal control of pollution, J. Economic Theory, 4 (1972), 19–34.
[7] K. Khalvati, S. A. Park, J. Dreher and R. Rao, A probabilistic model of social decision making based on reward maximization, Advances in Neural Information Processing Systems, (2016), 2901–2909.
Figure 9. Value functions and optimal controls for a 4 dimension case.

Figure 10. Control actions on testing set based on optimal strategy

[8] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1992.

[9] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.
[10] H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer Science & Business Media, 1992.

[11] D. Larose, K-nearest neighbor algorithm, *Discovering Knowledge in Data: An Introduction to Data Mining*, (2005), 90–106.

[12] X. Lu, G. Yin, Q. Zhang, C. Zhang and X. Guo, Building up an illiquid stock position subject to expected fund availability: Optimal controls and numerical methods, *Appl. Math. Optim.*, 76 (2017), 501–533.

[13] Q. Song, G. Yin and Z. Zhang, Numerical method for controlled regime-switching diffusions and regime-switching jump diffusions, *Automatica*, 42 (2006), 1147–1157.

[14] G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Two-Time-Scale Approach*, Second Edition, Springer, New York, 2013.

[15] G. Yin, Q. Zhang and G. Badowski, Discrete-time singularly perturbed Markov chains: Aggregation, occupation measures, and switching diffusion limit, *Adv. in Appl. Probab.*, 35 (2003), 449–476.

[16] G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer, New York, 2010.

[17] Q. Zhang and G. Yin, On nearly optimal controls of hybrid LQG problems, *IEEE Trans. Automat. Control*, 44 (1999), 2271–2282.

[18] Q. Zhang, C. Zhang, G. Yin, Optimal stopping of two-time scale Markovian systems: Analysis, numerical methods, and applications, *Nonlinear Analysis: Hybrid Systems*, 26 (2017), 151–167.

Received March 2018; revised July 2018.

E-mail address: czhang@wayne.edu
E-mail address: gyin@wayne.edu
E-mail address: qz@uga.edu
E-mail address: lywang@wayne.edu