Adjoint-Based Optimal Control of the Expected Exit Time for Stochastic Hybrid Systems

Robin L. Raffard[1], Jianghai Hu[2], and Claire J. Tomlin[1]

[1] Dept. of Aeronautics and Astronautics, Stanford University
{rraffard, tomlin}@stanford.edu
[2] School of Electrical and Computer Engineering, Purdue University
jianghai@ecn.purdue.edu

Abstract. In this paper, we study the problem of controlling the expected exit time from a region for a class of stochastic hybrid systems. That is, we find the least costly feedback control for a stochastic hybrid system that can keep its state inside a prescribed region for at least an expected amount of time. The stochastic hybrid systems considered are quite general: the continuous dynamics are governed by stochastic differential equations, and the discrete mode evolves according to a continuous time Markov chain. Instead of adopting the usual Hamilton-Jacobi viewpoint, we study the problem directly by formulating it as a PDE constrained optimization problem, and propose a solution using adjoint-based gradient descent methods. Numerical results of the proposed approach are presented for several representative examples, and, for the simple case, compared with analytical results.

1   Introduction

There has been considerable current research interest in stochastic hybrid systems (SHSs) [1, 2, 3, 4, 5] due to their ability to represent such systems as maneuvering aircraft [6], switching communication networks [7], etc. Most efforts have been devoted to the analysis of such systems: for control, the main approach to date relies on solving a dynamic programming problem using a Hamilton-Jacobi formulation [1]. In this paper, we propose an alternative method for optimal control of SHSs. The approach poses the optimal control problem as a partial differential equation (PDE) constrained optimization program, and uses an adjoint method to solve this optimization program. The adjoint method, introduced by Lions [8] and developed by Jameson [9] in the context of aerodynamic design, computes the gradient of an objective function whose variables are subject to PDE constraints. It is a powerful method, due mainly to the flexibility with which the optimal control problem can be formulated. Indeed, once the governing PDE, encoding the dynamics of the system, has been derived, many types of optimization problems can be posed. For instance, any constraints on the control input or on the state variable can be handled, contrary to Hamilton-Jacobi formulations.

In this paper, we focus on a particular practical optimal control problem. Given a domain of the state space of a SHS, we aim to maintain the expected sojourn time of the state within this domain above a certain threshold, while minimizing the cost of the control input. Different stochastic differential equations govern the continuous state in each mode, and a continuous time Markov chain dictates possible mode switches based on state and/or time. We first review the SHS model [1], and transform the optimal control problem into a PDE optimization problem. Then, we present and apply the adjoint method to this optimization problem. We present a set of interesting examples illustrating, through numerical solution, the resulting control policies, and we conclude with a note on further applications of the adjoint-based method in the context of SHSs.

2   Problem Formulation

2.1   Stochastic Hybrid Systems

General frameworks of stochastic hybrid systems have been proposed in [2, 3]. In this paper we focus on a special class called switched diffusions [1]. The state (X_t, m_t) of a switched diffusion system $\mathcal{H}$ consists of two parts: $X_t \in \mathbb{R}^n$ is the continuous state, and $m_t \in S = \{1, \ldots, M\}$ is the discrete state (or mode). The dynamics of the state (X_t, m_t) is characterized by the following:

– **Continuous Dynamics.** The continuous state X_t evolves according to a stochastic differential equation (SDE) whose drift and variance terms depend on the discrete mode m_t.

$$dX_t = u(X_t, m_t)\, dt + \sigma(X_t, m_t)\, dB_t. \qquad (1)$$

B_t is a d-dimensional Brownian motion in some probability space $(\Omega, \mathcal{F}, P)$, with Ω the sample space, $\mathcal{F}$ the σ-field, and P the probability measure, and $u : \mathbb{R}^n \times S \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \times S \to \mathbb{R}^{n \times d}$ are functions that are bounded and Lipschitz continuous in the first argument. Moreover, we assume that the possible values of σ are bounded away from zero.

– **Discrete Dynamics.** The discrete mode m_t evolves according to a continuous time Markov chain with a generator matrix $\Lambda(x) = [\lambda_{kl}(x)]_{1 \le k, l \le M}$ whose components depend on the continuous state $X_t = x$. Note that, $\forall k \ne l$, $\lambda_{kl}(x) \ge 0$ and $\lambda_{kk}(x) = -\sum_{l \ne k} \lambda_{kl}(x) \le 0$. Equivalently, for $\Delta t > 0$, we have

$$P(m_{t+\Delta t} = l \,|\, m_t = k, X_t = x) = \begin{cases} \lambda_{kl}(x)\Delta t + o(\Delta t), & \text{if } l \ne k, \\ 1 + \lambda_{kk}(x)\Delta t + o(\Delta t), & \text{if } l = k. \end{cases} \qquad (2)$$

Thus, given that $\mathcal{H}$ is in discrete mode k and continuous state x at time t, within a short time period Δt, $m_{t+\Delta t}$ jumps to a new mode $l \ne k$ with an approximate probability $\lambda_{kl}(x)\Delta t$.

– **Reset Condition.** For simplicity, we assume trivial reset condition. In other words, whenever a jump in m_t occurs, X_t remains unchanged. Note that the methodology developed in this paper is still applicable with general deterministic reset conditions.

We now outline the procedures to obtain stochastic solutions (executions) to the above stochastic hybrid system. Starting from an initial condition $X_0 = x$ and $m_0 = k$, the discrete state m_t remains in mode k for a random amount of time T_k while the continuous state evolves according to equation (1) with $m_t \equiv k$ until it reaches X_{T_k}. Then at time T_k the discrete state jumps to a new mode $l \neq k$ with probability $-\lambda_{kl}(X_{T_k})/\lambda_{kk}(X_{T_k})$ while the continuous state remains unchanged at X_{T_k}. This step is then repeated an infinite number of times. Note that if Λ is independent of x, the distribution of T_k is exponential with rate $-\lambda_{kk}$; however, in general, the random time T_k has a distribution dependent on the outcome of X_t.

In many practical applications, the variance term σ in the continuous dynamics (1) characterizing the environment noises and the λ_{kl} terms in (2) governing the transitions among operational modes are given and not adjustable, while the drift term u in (1) can be controlled by users to a certain degree. In this perspective, u can be treated as the control input of system $\mathcal{H}$.

In the following, we shall use $P^{(x,k)}$ and $E^{(x,k)}$ to denote the probability and expectation under the initial condition $X_0 = x$ and $m_0 = k$.

2.2   Optimal Exit Time Control

Given a switched diffusion system $\mathcal{H}$, we now formulate the problem studied in this paper. Let U be an open set of $\mathbb{R}^n$ with compact support. Let (X_t, m_t) be a stochastic solution to $\mathcal{H}$ starting from $X_0 = x$ and $m_0 = k$ at time 0. We consider the following stopping time:

$$\tau = \inf\{t > 0 : X_t \notin U\}, \tag{3}$$

which is called the exit time from U (or the sojourn time in U).

Remark 1. Note that the definition of τ in (3) does not involve m_t. Therefore, at exit time τ, the switched diffusion can be in any discrete mode $m_t \in S$.

Define $V(x,k)$ as the expected exit time from U, starting from (x,k):

$$V(x,k) = E^{(x,k)}[\tau]. \tag{4}$$

Treating u as the feedback control of the system $\mathcal{H}$, $V(x,k)$ is determined by the design of u. In practical situations, U is often referred to as the *safe set* in which one wants the system state to stay. Then a natural problem is to find the least expensive control $u : U \times S \to \mathbb{R}^n$ that can keep the system in U for at least an expected amount of time. Specifically, let $\rho, \xi : U \times S \to \mathbb{R}_+$ be two positive functions representing weights. The cost of the control, $J(u)$, is written as

$$J(u) \triangleq \sum_{k=1}^{M} \int_U \xi(x,k)\|u(x,k)\|^2 \, dx \,, \tag{5}$$

and the (weighted) cumulative expected exit time from U, $V_{cee}(u)$, is defined as

$$V_{cee}(u) = \sum_{k=1}^{M} \int_U \rho(x,k)V(x,k)\,dx. \tag{6}$$

In particular, if ρ is the probability density function of (X_0, m_0) over $U \times S$, then $V_{cee}(u)$ coincides with the expected exit time from U with uncertain (X_0, m_0).

Problem 1. The problem studied in this paper is

$$\text{Minimize } J(u) \text{ subject to } V_{cee}(u) \geq V_0, \tag{7}$$

for some constant $V_0 > 0$.

For simplicity, but without loss of generality, we shall assume that $\rho = \xi \equiv 1$, unless otherwise stated.

2.3   Reformulation as a PDE Constrained Optimization Problem

Problem (1) in its current form is not easy to analyze as the dependence of $V_{cee}(u)$ on the control u is implicit. In this section, we shall derive the PDE satisfied by $V(x,k)$ defined in (4), where u will appear as a coefficient to be controlled.

Definition 1 (Generator of the Switched Diffusion). *To the switched diffusion $(X_t, m_t)_{t \geq 0}$, we associate an operator L (referred to as the generator) that maps a function $f \in C_0^2(\mathbb{R}^n \times S)$ to a new function $Lf \in C_0^0(\mathbb{R}^n \times S)$ defined by*

$$Lf(x,m) \triangleq \sum_{i=1}^{n} u_i(x,m)\frac{\partial f(x,m)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n}(\sigma(x,m)\sigma(x,m)^T)_{ij}\frac{\partial^2 f(x,m)}{\partial x_i \partial x_j}$$
$$+ \sum_{k=1}^{M} \lambda_{mk}(x)f(x,k), \quad \forall x \in \mathbb{R}^n, \ \forall m = 1,\ldots,M. \tag{8}$$

Here $C_0^2(\mathbb{R}^n \times S)$ (resp. $C_0^0(\mathbb{R}^n \times S)$) denotes the set of functions on $\mathbb{R}^n \times S$ with compact support that are twice differentiable (resp. continuous) with respect to the first argument. $u_i(x,m)$ denotes the i-th component of the vector $u(x,m) \in \mathbb{R}^n$.

Lemma 1. *For all $(x,m) \in \mathbb{R}^n \times S$, and for any $f \in C_0^2(\mathbb{R}^n \times S)$, $M_t \triangleq f(X_t, m_t) - f(X_0, m_0) - \int_0^t Lf(X_s, m_s)\,ds$ is a Martingale on $(\Omega, \mathcal{F}, P^{(x,m)})$.*

A proof of the above lemma can be found in [3].

Theorem 1. *For all $(x,m) \in U \times S$, $V(x,m) = E^{x,m}[\tau]$ is finite, and is a solution of the following system of PDEs:*

$$\begin{aligned} LV(x,m) &= -1, \quad \forall x \in U, \ \forall m \in S, \\ V(x,m) &= 0, \quad \forall x \in \partial U, \ \forall m \in S. \end{aligned} \tag{9}$$

Proof. 1. By assumption, for each $m \in S$, $\sigma(\cdot, m) : x \in U \mapsto \sigma(x, m) \in \mathbb{R}^{n \times d}$ is a continuous function whose values are bounded away from zero. Therefore the PDE system (9) admits a unique solution V twice differentiable in its first argument [10]. Since $V_{\partial U \times S} \equiv 0$, we can construct $\tilde{V} \in C_0^2(\mathbb{R}^n \times S)$, such that $\tilde{V} = V$ on $\bar{U} \times S$. Applying Lemma 1 to $\tilde{V}$, we have that

$$M_t \triangleq \tilde{V}(X_t, m_t) - \tilde{V}(X_0, m_0) - \int_0^t L\tilde{V}(X_s, m_s) \, ds \qquad (10)$$

is a Martingale on $(\Omega, \mathcal{F}, P^{(x,m)})$.

Now, for each integer $N \in \mathbb{N}^+$, define $\tau_N \triangleq \tau \wedge N = \min(\tau, N)$. Then τ_N is a stopping time with finite expectation $E^{x,m}[\tau_N] \leq N < \infty$. Therefore, applying the optional sampling theorem to the Martingale M_t stopped at time τ_N, we deduce that $E^{x,m}[M_{\tau_N}] = 0$. In addition, $\forall s \in [0, \tau_N]$, $X_s \in \bar{U}$; therefore, $\int_0^{\tau_N} L\tilde{V}(X_s, m_s) \, ds = \int_0^{\tau_N} LV(X_s, m_s) \, ds = -\tau_N$, and thus,
$E^{x,m}[\tau_N] = E^{x,m}[\tilde{V}(X_0, m_0) - \tilde{V}(X_{\tau_N}, m_{\tau_N})] \leq 2 \sup_{(x,m) \in U \times S} \{|\tilde{V}(x,m)|\} < \infty.$
Applying the monotone convergence theorem, with $\tau_N \to \tau$ almost surely, we deduce that $E^{x,m}[\tau] = \sup_N \{E^{x,m}[\tau_N]\} \leq 2 \sup_{(x,m) \in U \times S} \{|\tilde{V}(x,m)|\} < \infty.$
2. Since $E^{x,m}[\tau] < \infty$, we can now apply the optional sampling theorem to M_t stopped at time τ. We obtain $E^{x,m}[\tilde{V}(X_\tau, m_\tau)] - E^{x,m}[\tilde{V}(x,m)] + E^{x,m}[\tau] = 0$. Given the boundary conditions satisfied by V, we have $\tilde{V}_{\partial U \times S} = 0$ and therefore, $\tilde{V}(X_\tau, m_\tau) \equiv 0$. It follows that $\forall (x,m) \in U \times S$, $V(x,m) = \tilde{V}(x,m) = E^{x,m}[\tau]$, which proves the theorem. $\qquad\qquad\square$

As a result of Theorem 1, Problem (1) can now be reformulated as follows:

$$
\begin{array}{ll}
\text{Minimize} & J(u) = \sum_{m=1}^M \int_U \|u(x,m)\|^2 \, dx \\[2mm]
\text{Subject to} & \sum_{m=1}^M \int_U V(x,m) \, dx \geq V_0; \\[2mm]
& \begin{cases} LV(x,m) = -1, & \forall x \in U, \ m \in S; \\ V(x,m) = 0, & \forall x \in \partial U, \ m \in S. \end{cases}
\end{array} \qquad (11)
$$

Note that the constraint on V in the above problem is written explicitly as a system of coupled PDEs with boundary condition. In the next section, we shall introduce tools to solve this kind of optimization problem.

3   PDE Constrained Optimization via Adjoint Method

The adjoint method is a gradient-based method which can numerically solve optimization problems subject to PDE constraints [9, 8]. In Section 3.1, we will first briefly review the adjoint method in its most general setting, and then apply it to the optimal exit time control problem for switched diffusions in Section 3.2.

3.1   Adjoint Method for Solving PDE Constrained Optimization

Consider the following general PDE constrained optimization program.

$$\begin{array}{ll} \text{Minimize} & J(u,v) \\ \text{Subject to} & \mathcal{N}(u,v) = 0 \\ & r(u,v) \le 0. \end{array} \tag{12}$$

Here, $\mathcal{N}(u,v) = 0$ denotes a PDE; v denotes the solution of the PDE; u is the control variable, which is in general an adjustable coefficient of the PDE or an adjustable term in the boundary conditions; J is the objective function of the optimization program and has to be real valued; r is a function of u and v which characterizes all the inequality constraints of the problem, such as bounds on the control variable u or on the solution v. J and r are assumed to be differentiable in u and v.

In order to apply the adjoint-method, we first need to reduce (12) into an unconstrained optimization problem. For this purpose, we use a (logarithm) barrier method:

$$\begin{array}{ll} \text{Minimize} & I(u,v) = J(u,v) - \epsilon \mathbf{1}^T \log(-r(u,v)) \\ \text{Subject to} & \mathcal{N}(u,v) = 0\,, \end{array} \tag{13}$$

where $\mathbf{1}^T \log(-r(u,v))$ represents the inner product between the identity $\mathbf{1}$ and $\log(-r(u,v))$. Problem (12) and Problem (13) are equivalent when $\epsilon \to 0$. Therefore solving (13) with $\epsilon \simeq 0$ will approximately solve (12). The adjoint method can then be used to derive the gradient of the cost function I with respect to the control input u, subject to the constraint $\mathcal{N}(u,v) = 0$, thus deriving a descent direction for u in Problem (13). First, take a first variation of I:

$$\delta I = \left(\frac{\partial I}{\partial u}\right)^T \delta u + \left(\frac{\partial I}{\partial v}\right)^T \delta v. \tag{14}$$

Similarly, a first order variation of the PDE gives the dependence of δv on δu:

$$\delta \mathcal{N} = \left(\frac{\partial \mathcal{N}}{\partial u}\right) \delta u + \left(\frac{\partial \mathcal{N}}{\partial v}\right) \delta v = 0\,, \tag{15}$$

in which $\frac{\partial \mathcal{N}}{\partial u}$ and $\frac{\partial \mathcal{N}}{\partial v}$ are linear operators. Taking the inner product of (15) with any differentiable function q (named *costate*) and subtracting it from (14), we have

$$\forall q, \quad \delta I = \left(\left(\frac{\partial I}{\partial u}\right)^T - q^T \frac{\partial \mathcal{N}}{\partial u}\right) \delta u + \left(\left(\frac{\partial I}{\partial v}\right)^T - q^T \frac{\partial \mathcal{N}}{\partial v}\right) \delta v. \tag{16}$$

Choosing q such that it satisfies the following adjoint PDE:

$$(\frac{\partial \mathcal{N}}{\partial v})^T q = \frac{\partial I}{\partial v}, \tag{17}$$

we derive δI as a function of δu only:

$$\delta I = \left(\left(\frac{\partial I}{\partial u}\right)^T - q^T \frac{\partial \mathcal{N}}{\partial u}\right) \delta u, \tag{18}$$

which precisely defines the gradient of I with respect to the control variable u:

$$\nabla I(u) = \frac{\partial I}{\partial u} - \left(\frac{\partial \mathcal{N}}{\partial u}\right)^T q. \tag{19}$$

With the gradient $\nabla I(u)$ in hand, one can then design an iterative algorithm using gradient descent to solve Problem (13) numerically. Note that in each step one has to solve the PDE and the adjoint PDE (17) in order to obtain the gradient.

3.2   Adjoint Method Applied to Optimal Sojourn Time Control

We now demonstrate how the adjoint method can be applied to solve the optimal sojourn time control Problem (11). We assume that $\xi = \rho \equiv 1$ on U.

First, we transform Problem (11) into the form of (13) as:

Minimize $I(u) = \sum_{m=1}^{M} \int_U \|u(x,m)\|^2 \, dx - \epsilon \log(\sum_{m=1}^{M} \int_U V(x,m) \, dx - V_0)$

Subject to $\begin{cases} LV(x,m) = -1, \ \forall x \in U, \ m \in S; \\ V(x,m) = 0, \ \forall x \in \partial U, \ m \in S. \end{cases}$
$$\tag{20}$$

Note that instead of writing $I(u, V)$, we omit V since it depends on u implicitly. A first variation of the cost function I gives:

$$\delta I(u) = 2 \sum_{m=1}^{M} \int_U u(x,m)^T \delta u(x,m) \, dx - \epsilon \frac{\sum_{m=1}^{M} \int_U \delta V(x,m) \, dx}{\sum_{m=1}^{M} \int_U V(x,m) \, dx - V_0}. \tag{21}$$

To compute δI as a function of δu only (and not as a function of δV, which is not directly controllable), one needs to eliminate δV. For this purpose, we take the first variation of the PDE constraint $LV(x,m) = -1$, which gives the dependence of δV on δu:

$$\sum_{i=1}^{n} \left[\delta u_i(x,m)\frac{\partial V(x,m)}{\partial x_i} + u_i(x,m)\frac{\partial \, \delta V(x,m)}{\partial x_i}\right] + $$
$$\sum_{i,j=1}^{n} \frac{1}{2}(\sigma(x,m)\sigma(x,m)^T)_{ij}\frac{\partial^2 \, \delta V(x,m)}{\partial x_i \partial x_j} + \sum_{k=1}^{M} \lambda_{mk}(x)\delta V(x,k) \; = \; 0. \tag{22}$$

For each m, multiplying (22) by a costate function $q(x,m)$ which is twice differentiable in x and which is identically zero on ∂U (required for the integration by parts in (24)), integrating over the domain U, and then summing over all m, we obtain:

$$\sum_{m=1}^{M} \int_U \left(q^m \sum_{i=1}^{n} u_i^m \frac{\partial \, \delta V^m}{\partial x_i} + \frac{q^m}{2} \sum_{i,j=1}^{n} (\sigma^m \sigma^{m\,T})_{ij} \frac{\partial^2 \, \delta V^m}{\partial x_i \partial x_j} \right.$$
$$\left. + q^m \sum_{k=1}^{M} \lambda_{mk}\delta V^k \right) dx = - \sum_{m=1}^{M} \int_U \sum_{i=1}^{n} q^m \frac{\partial V^m}{\partial x_i}\delta u_i^m \, dx, \tag{23}$$

where for simplicity we drop the explicit dependence on x, and use V^m for $V(\cdot, m)$, u_i^m for $u_i(\cdot, m)$, etc... Integrating by parts in (23) and using the boundary condition that, on ∂U, $V^m \equiv 0$ and $q^m \equiv 0$, we have

$$
\sum_{m=1}^{M} \int_U \left(-\sum_{i=1}^{n} \frac{\partial(q^m u_i^m)}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2[q^m(\sigma^m \sigma^{m\,T})_{ij}]}{\partial x_i \partial x_j} \right.
$$
$$
\left. + \sum_{k=1}^{M} \lambda_{km} q^k \right) \delta V^m \, dx = -\sum_{m=1}^{M} \int_U q^m \sum_{i=1}^{n} \frac{\partial V^m}{\partial x_i} \delta u_i^m \, dx. \tag{24}
$$

Suppose for each m we choose q^m so that the following adjoint PDE holds:

$$
\frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2[q^m(\sigma^m \sigma^{m\,T})_{ij}]}{\partial x_i \partial x_j} - \sum_{i=1}^{n} \frac{\partial(q^m u_i^m)}{\partial x_i} + \sum_{k=1}^{M} \lambda_{km} q^k
$$
$$
= \frac{\epsilon}{\sum_{m=1}^{M} \int_U V^m \, dx - V_0}. \tag{25}
$$

First substituting (25) into (24), and then the result into (21), we have

$$
\delta I(u) = \sum_{m=1}^{M} \int_U \left(2u^m + q^m \nabla V^m \right)^T \delta u^m \, dx.
$$

So the gradient of I with respect to the control u for the discrete mode m is

$$
\nabla I^m = 2u^m + q^m \nabla V^m, \quad m = 1, \dots, M. \tag{26}
$$

We emphasize here that q^m in equation (26) is the solution to the adjoint equation (25) with boundary condition $q^m \equiv 0$ on ∂U. Furthermore, the quantities q^m, u^m, etc. represent q, u, etc. in mode m; and not q or u to the power m.

Having obtained the gradient of I with respect to the control u, the gradient descent algorithm for finding the optimal u can be formulated as follows.

Algorithm 1 (Adjoint based algorithm). *Set $\epsilon = 1$ and guess an initial value for u.*
Repeat (loop a)
    Repeat (loop b)
        1. Solve equation (9) for V, using the current control u.
        2. Solve the adjoint equation (25) for q, using the current u and V.
        3. Determine the gradient ∇I according to equation (26).
        4. Line search: compute $\beta > 0$ so that $I(u - \beta \nabla I)$ is minimized.
        5. Update $u := u - \beta \nabla I$.
    Terminate loop b *when $||\nabla I||$ is smaller than the stopping criteria α_b.*
    Decrease ϵ by letting $\epsilon := \mu\epsilon$, where $\mu \in [0.1, 0.5]$.
Terminate loop a *when ϵ is smaller than the stopping criteria α_a.*
Return *$u_{opt} = u$.*

Remark 2. According to the analysis of Section 2, the admissible control u has to be bounded and Lipschitz continuous on U. Thus, to be rigorous, a constraint of the type $u_{\min} \leq u(x, m) \leq u_{\max}$ should be added to the optimization problem to ensure the boundedness of u. Regarding the Lipschitz continuity, the gradient ∇I should be projected on a functional subspace of Lipschitz continuous functions after step 3 of Algorithm 1.

3.3   Validation Against Analytical Solution in a Simple Case

It is possible to solve Problem (1) analytically in some simple cases. In this section, we shall present such a case studied in [11], and compare the analytical result with the numerical ones obtained using Algorithm 1.

Suppose that $M = 1$, $U = [-a, a] \subset \mathbb{R}$, and $\sigma = 1$ on U. Then the stochastic hybrid system degenerates into a simple diffusion on the interval $[-a, a]$:

$$dX_t = u(X_t)\, dt + dB_t.$$

The expected exit time $V(x)$ from U starting from $x \in U$ satisfies

$$\frac{1}{2}V''(x) + u(x)V'(x) + 1 = 0, \quad V(-a) = V(a) = 0. \tag{27}$$

Assume $\xi \equiv 1$, and $\rho(x)$ is a unit pulse centered at the origin. Then Problem (11), which is equivalent to Problem (1), can then be formulated as

$$\text{Minimize } \int_{-a}^{a} u^2(x)\, dx, \text{ subject to equation (27) and } V(0) \geq V_0.$$

Suppose in addition that the control u is odd in x, i.e., $u(-x) = -u(x)$, $\forall x \in U$. Then V as a solution to (27) is even in x. Because of the symmetry, it suffices to study the problem on the left half interval $[-a, 0]$ only:

$$\text{Minimize } \int_{-a}^{0} u^2(x)\, dx, \text{ subject to equation (27) and } V(0) \geq V_0. \tag{28}$$

Denote $y_1 = V$ and $y_2 = V'$. Then the above problem is equivalent to the following optimal control problem:

$$\text{Minimize } \int_{-a}^{0} u^2(x)\, dx, \text{ subject to } \begin{cases} y_1' = y_2, & y_1(-a) = 0, y_1(0) = V_0, \\ y_2' = -2uy_2 - 2, & y_2(0) = 0. \end{cases}$$

Using the Maximum Principle, and identifying two first integrals in the Hamiltonian equations, we can determine the optimal trajectory (y_1, y_2) as

$$y_2(x) = \Phi^{-1}(x), \quad y_1(x) = \int_{-a}^{x} y_2(x)\, dx, \tag{29}$$

where Φ is a function defined by $\Phi(y) = \int_0^y \frac{dx}{-2\sqrt{1 + x^2(c_1 x + c_2)}}$ for some suitably chosen constants c_1 and c_2. The optimal control u in this case can be determined from y_1 and y_2 accordingly. For more details, see [11].

In Figure 1, we plot the analytic and the numerical results for the above problem. One can see that Algorithm 1 generates results that fit the analytical one exceedingly well.
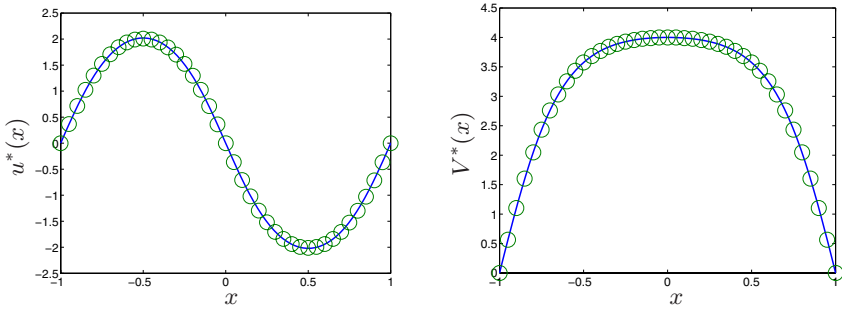
Fig. 1. *Validation against analytical results. Left*: Optimal control u^* returned by Algorithm 1 (solid) and by the analytical solution (circles). *Right*: Optimal expected sojourn time returned by Algorithm 1 (solid) and by analytical solution (circles)

4 Simulation Examples for Stochastic Hybrid Systems

We now apply our algorithm to investigate the optimal exit time control problem for switched diffusions. Two main categories of switched diffusions are considered: time switching and state switching.

4.1 Time Switching

The time switching case refers to the case in which the variance $\sigma(x, m)$ and the mode transition rate $\Lambda(x)$ are both independent of the continuous state x. Therefore, for each mode m, the variance $\sigma\sigma^T$ is a constant matrix, and the time the system spends in mode m before jumping to a new mode has an exponential distribution.

For simplicity, in this section we assume that there are only two modes: $S = \{1, 2\}$, and that, for each mode $m = 1, 2$, the variance $\sigma(\cdot, m) \equiv \sigma_m I_n$ for constants $\sigma_1^2 = 1$ and $\sigma_2^2 = 2$, where I_n is the n-by-n identity matrix. The problem is then to determine the control $u(\cdot, 1)$ and $u(\cdot, 2)$ for the two modes.

Example 1 (Switching Between Two Modes in a 2-D Disk). Suppose $U = B(0, a)$ is the disk of radius $a > 0$ around the origin in $\mathbb{R}^2$. In cylindrical coordinates $x = (r, \theta)$, the control $u = (u_r, u_\theta)$ is decomposed into the radial component u_r and ortho-radial component r_θ, and the PDE (9) governing the expected sojourn time $V(x, m) = V((r, \theta), m)$ in U becomes

$$
\frac{\sigma_m^2}{2}\left(\frac{\partial^2 V((r, \theta), m)}{\partial r^2} + \frac{1}{r}\frac{\partial V((r, \theta), m)}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V((r, \theta), m)}{\partial \theta^2}\right) +
$$

$$
u_r \frac{\partial V((r, \theta), m)}{\partial r} + \frac{u_\theta}{r}\frac{\partial V((r, \theta), m)}{\partial \theta} + \sum_{k=1}^{2}\lambda_{mk}V((r, \theta), k) + 1 = 0, \tag{30}
$$

$$
V((a, \theta), m) = 0, \,\forall \theta \in [0, 2\pi]; \quad V((r, 2\pi), m) = V((r, 0), m), \,\forall r \in [0, a].
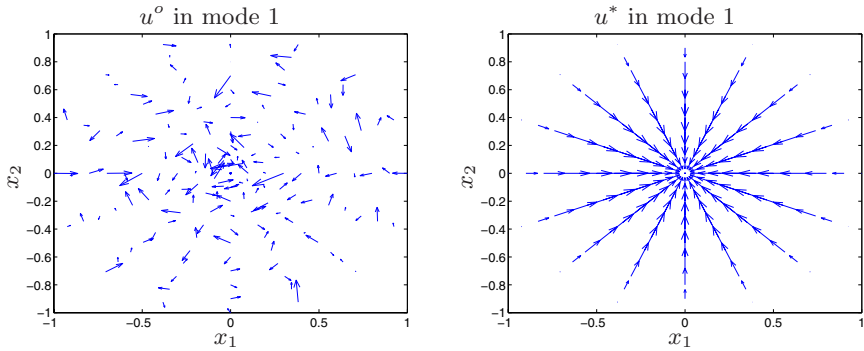$$

Fig. 2. *Left*: Initial guess for the control u in mode 1. Arrows represent the directions and the amplitudes of the randomly generated control u. *Right*: Optimal control u^* in mode 1 returned by Algorithm 1. Note that it is radially-invariant

So the equivalent PDE-constrained optimization Problem (11) becomes

$$\text{Minimize } \sum_{m=1}^{2} \int_0^{2\pi} \int_0^a \left(u_r^2((r,\theta),m) + u_\theta^2((r,\theta),m) \right) \, r \, dr \, d\theta$$
$$\text{Subject to } \sum_{m=1}^{2} \int_0^{2\pi} \int_0^a V((r,\theta),m) \, r \, dr \, d\theta \geq 2\,V_0; \tag{31}$$
$$\text{Equation (30)}.$$

Due to the rotational symmetry of the problem: the domain U, the cost function and the constraints are all invariant under the rotations around the origin, we conjecture that the optimal solutions are also radially symmetric:

Conjecture 1. *The optimal control u^* is of the form $u^* = u_r(r)e_r$, where e_r is the radius unit vector.*

We verify this conjecture numerically by applying the adjoint algorithm 1 using randomly generated initial guesses for the control u. We assume that the switching rates are given by $\Lambda = \begin{bmatrix} -10 & 10 \\ 10 & -10 \end{bmatrix}$. On the left of Fig. 2 we show a typical initial guess for the control u in mode 1 (guesses for control in mode 2 are similar). The optimal control u^* returned by the algorithm is shown on the right of the figure, which is radially symmetric. That the algorithm converges to the same solution from a wide selection of initial u demonstrates its robustness with respect to initial guesses.

Two more scenarios are also simulated, one with a higher switching rate $\Lambda = \begin{bmatrix} -20 & 20 \\ 20 & -20 \end{bmatrix}$, and the other with no switching at all: $\Lambda = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. In both cases, the algorithm produces radially symmetric solutions. Thus we only plot the radial component u_r^* of the solution u^* in Fig. 3, with controls in mode 1 and 2 on the left and right, respectively. The higher switching rate case is plotted in circles, and the no switching case in squares. For comparisons, we also plot the simulation results for the non-hybrid case (only one mode) for three different σ: $\sigma = \sigma_1$ (dash dot lines), $\sigma = \frac{\sigma_1 + \sigma_2}{2}$ (solid lines), and $\sigma = \sigma_2$ (dot lines).
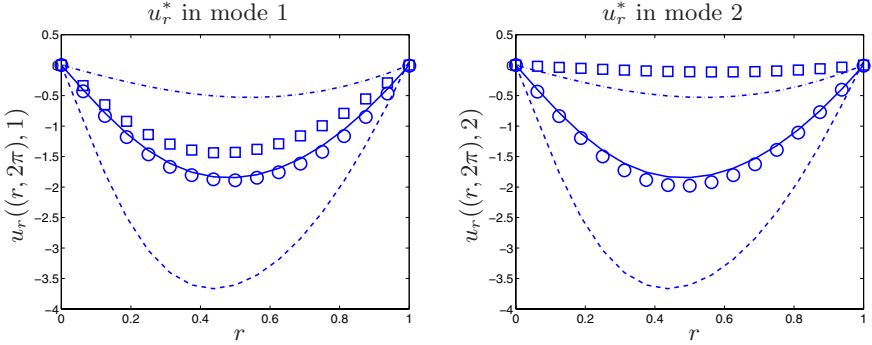
Fig. 3. Radial component u_r^* of the optimal control u^* in mode 1 (*left*) and mode 2 (*right*) for the high switching rate case (circles) and no switching case (squares). Also included are the u_r^* for the non-hybrid case (only one mode) with variance $\sigma = \sigma_1$ (dash dot), $\sigma = \frac{\sigma_1 + \sigma_2}{2}$ (solid), and $\sigma = \sigma_2$ (dot)

In the higher switching rate case, the optimal controls in the two modes are almost the same, and also close to the optimal control in the non-hybrid case with variance $\sigma = \frac{\sigma_1 + \sigma_2}{2}$. This is because, due to the very frequent switchings between the two modes, and the fact that quickly the distribution of m_t will converge to the stationary distribution of equal probability $\frac{1}{2}$ in each mode, the stochastic hybrid system switching between diffusions of variances σ_1 and σ_2, can be approximated by a single diffusion with variance $\sigma = \frac{\sigma_1 + \sigma_2}{2}$.

For the zero switching rate case, the optimal controls in the two modes are quite different from each other, and from the other cases as well. Barely any control is exerted in mode 2 compared with in mode 1. The reason is that it costs less control input to maintain the expected sojourn time above $2V_0$ in mode 1, than maintaining the expected sojourn time above V_0 in mode 2.

4.2 State Switching

Consider the following situation. The domain U is partitioned into a finite number of subdomains U_α. Consider a diffusion process X_t evolving on U accordingly to $dX_t = u(X_t)dt + \sigma(X_t)dB_t$, where the variance $\sigma(x)$ takes the constant value σ_α for $x \in U_\alpha$. Thus σ is piecewise constant on U. Assume that there is only one mode $m = 1$.

Rigorously speaking, for the above X_t to be well defined one needs $\sigma(x)$ to be Lipschitz continuous in x, which is not the case here. However, one can "smooth out" the transition of σ near the boundary of U_α to satisfy this requirement. Specifically, one can make σ to be constant σ_α in a compact subset of U_α approximately equal to U_α, and determine the value of σ for points near the boundary with other subdomains via interpolation. The σ thus obtained is Lipschitz continuous in x, and using it one gets an approximate of the original process. As an example, we partition the domain $U = [-a, a] \times [-b, b]$ into two subdomains: $U_a = [-a, 0] \times [-b, b]$, and $U_b = (0, a] \times [-b, b]$. We can choose the variance $\sigma(x)$
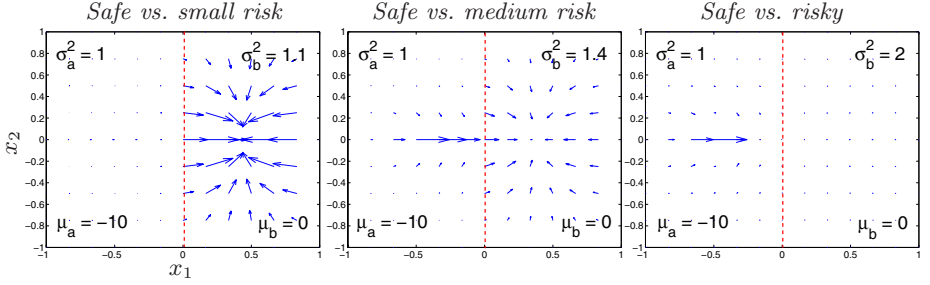
Fig. 4. Optimal control u^* for Example 2. A constant nominal drift exists in domain U_a. $\sigma_a^2 = 1$, while $\sigma_b^2 = 1.1$ (*left*); $\sigma_b^2 = 1.4$ (*center*); $\sigma_b^2 = 2$ (*right*). Note that in different figures the sizes of the arrows representing the magnitude of the control u^* are not in the same scale: *Left:* $\|u^*\|_{\max} = 4.3$. *Center:* $\|u^*\|_{\max} = 14.2$. *Right:* $\|u^*\|_{\max} = 16.1$

to be constant σ_a on $U_a^\epsilon = [-a, -\epsilon] \times [-b, b]$ and σ_b on $U_b^\epsilon = [\epsilon, a] \times [-b, b]$, with $\epsilon > 0$, and determine $\sigma(x)$ on $(-\epsilon, \epsilon) \times [-b, b]$ by interpolation.

Example 2 (Safe vs. Risky in a 2-D Box). Suppose that $U = [-a, a] \times [-b, b] = U_a \cup U_b$ as defined above. In this example, we would like to study whether it is more advantageous to try to stay in U_a with a low uncertainty (low variance) but a constant drift pushing toward the left boundary, or to try to stay in U_b with no uncontrollable drift but with high uncertainty (high variance). For this purpose, we set the following conditions. $\sigma(x)$ is constant equal to $\sigma_a I_2$ on U_a for $\sigma_a^2 = 1$, and constant equal to $\sigma_b I_2$ on U_b for $\sigma_b^2 = 2$. To smooth out the transition of $\sigma(x)$, we use the parameter $\epsilon = 0.02$. Furthermore, assume that on subdomain U_a there is an uncontrollable constant drift $(-10, 0)^T$ to the left, while on U_a there is no such drift. Thus X_t is governed by $dX_t = [\mu(X_t) + u(X_t)] dt + \sigma(X_t) dB_t$, with $\mu(x) = (-10, 0)^T$ if $x \in U_a$ and $\mu(x) = 0$ if $x \in U_b$.

First write the PDE governing the expected sojourn time $V(x)$ as

$$\frac{\sigma^2}{2}\left(\frac{\partial^2 V}{\partial x_1^2} + \frac{\partial^2 V}{\partial x_2^2}\right) + (\mu_1 + u_1)\frac{\partial V}{\partial x_1} + (\mu_2 + u_2)\frac{\partial V}{\partial x_2} = -1, \quad \forall x \in U, \quad (32)$$

with boundary condition $V(x) \equiv 0$ for $x \in \partial U$. Thus the problem becomes

$$\begin{array}{ll} \text{Minimize} & \int_{-a}^{a}\int_{-b}^{b}(u_1^2 + u_2^2)\,dx_1\,dx_2 \\ \text{Subject to} & \int_{-a}^{a}\int_{-b}^{b} V(x)\,dx_1\,dx_2 \geq V_0, \text{ and equation (32)}, \end{array} \quad (33)$$

Algorithm 1 can be extended easily to deal with the uncontrollable drift term. The results are shown in Fig. 4 and are quite interesting. For very small values of σ_b, $(\sigma_b^2 = 1.1)$, the optimal control concentrates all the energy in the domain U_b, trying to contain X_t near the center of U_b. However, as $\sigma_b^2 = 1.4$, the optimal control is distributed on both U_a and U_b. If σ_b^2 increases further to 2, the optimal control applies very little force on U_b and concentrates most of its energy near the center horizontal line in U_a.
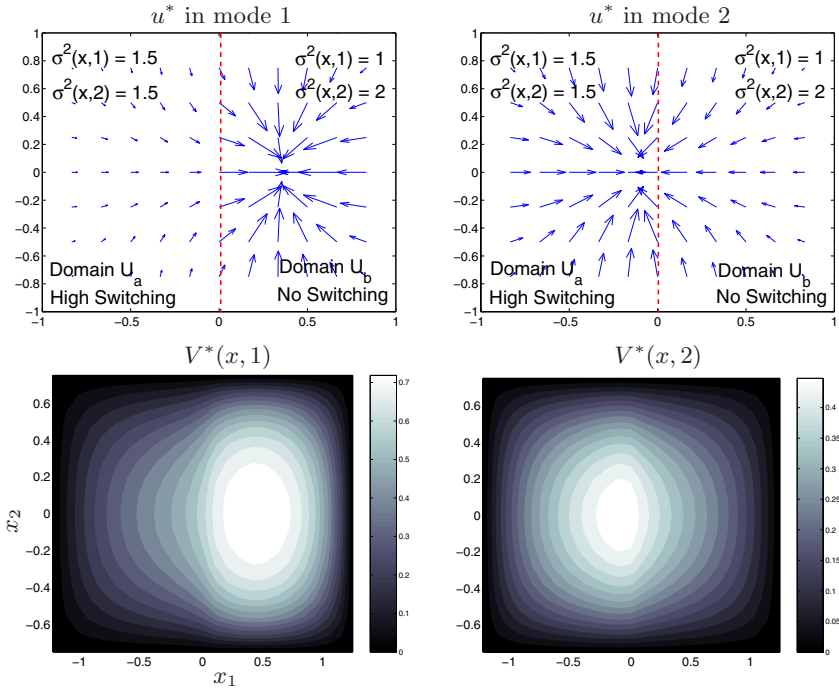
Fig. 5. *Top*: Optimal controls u^* in mode 1 (*left*) and mode 2 (*right*) for Example 3. *Bottom*: Optimal expected exit time $V^*(x, 1)$ (*left*) and $V^*(x, 2)$ (*right*)

4.3   General Switching

We now consider a general scenario that encompasses the previous two cases as special instances. In this case, sudden changes in variance can occur either due to the random discrete mode transitions or due to the state evolving into different subdomains of U.

Example 3 (General Switching in a 2-D box). Let $U = U_a \cup U_b$ be as before. Suppose that there are two discrete modes ($M = 2$). In subdomain U_a, the switching rate between the two modes due to the continuous time Markov chain m_t is high, so the diffusion switches rapidly between mode 1 and mode 2. On the other hand, the switching rate is zero on U_b so that once X_t enters U_b, m_t will remain in the same mode. In addition, assume that the variance on domain U_a is $\sigma_a I_2$ with $\sigma_a^2 = 1.5$, regardless of the discrete mode, while on U_b, the variance is $\sigma_b I_2$, with $\sigma_b^2 = 1$ if $m = 1$ and $\sigma_b^2 = 2$ if $m = 2$. To sum up, we have $\sigma(x, 1) = \sigma(x, 2) = \sqrt{1.5}I_2$, $\Lambda(x) = \begin{bmatrix} -20 & 20 \\ 20 & -20 \end{bmatrix}$, $\forall x \in U_a$; $\sigma(x, 1) = I_2$, $\sigma(x, 2) = \sqrt{2}I_2$, $\Lambda(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\forall x \in U_b$. The optimal control u^* for this problem as obtained by Algorithm 1 is shown in Fig. 5. Under mode 1, u^* tends

to drive X_t towards subdomain U_b where the variance is lower than in U_a. In mode 2, u^* tends to drive X_t towards U_a, in which the average variance is lower than the variance in U_b. However, the optimal control does not drive the diffusion near the center of U_a (as it drives the diffusion near the center of U_b, under mode 1). The interpretation is that the diffusion originally in mode 2, in U_b, will tend to be driven to U_a, switch mode and then go back to U_b in mode 1. This is a smart optimal control.

5 Conclusions and Extensions

In this paper we proposed to use the adjoint method to solve the optimal sojourn time control problem for a class of stochastic hybrid systems. We formulated the problem as a PDE-constrained optimization problem, and devised an algorithm to solve it using the gradients computed via the adjoint method.

The adjoint method is a powerful tool that can be applied to many other optimal control problems. Examples include optimal control of expected reward over an infinite time horizon: $V(x,m) = E^{x,m}[\int_0^\infty e^{-\alpha s} r(X_s, m)\, ds]$, which satisfies

$$LV(x,m) - \alpha V(x,m) = -r(x,m)\,, \quad m = 1, \ldots, M.$$

As another example, we can choose $V(x,t,m) = E^{x,m}[f(X_t, m)]$, which solves the backward Kolmogorov equation:

$$\frac{\partial V(x,t,m)}{\partial t} = LV(x,t,m). \quad V(x,0,m) = f(x,m).$$

Nevertheless, the adjoint method is not without its shortcomings. First, the dimension of the state of the stochastic hybrid system sets the dimension of the PDE to optimize. In implementation, PDEs can be numerically solved in dimension 3 or 4 – for higher dimensions, the memory requirement becomes problematic. Therefore, the adjoint method, which runs on a modern single processor computer in a few seconds for the application presented above, can be applied only for stochastic hybrid systems of continuous dimension say less than 4. Second, gradient descent methods for large scale optimization programs (n dimensions usually results in 100^n grid points) might be very slow. Therefore, second order optimization should be adopted. We are currently investigating these issues. The first might be overcome by applying directly the adjoint method in the Monte-Carlo space. Regarding the second, we are currently developing a Newton method implementation.

References

1. Ghosh, M., Arapostathis, A., Marcus, S.I.: Optimal control of switching diffusions with applications to flexible manufacturing systems. (SIAM J. Control Optim. 31 (1993), 1183–1204)

2. Hu, J., Lygeros, J., Sastry, S.: Towards a theory of stochastic hybrid systems. In Lynch, N.A., Krogh, B.H., eds.: Hybrid Systems: Computation and Control (HSCC). Lecture Notes in Computer Science 1790, (Springer-Verlag) 160–173.

3. Bujorianu, M.L., Lygeros, J.: General stochastic hybrid systems: Modelling and optimal control. (Proceedings of the IEEE Int. Conference on Decision and Control, Atlantis, Bahamas, 2004)

4. Bujorianu, M.L.: Extended stochastic hybrid systems and their reachability problem. In Alur, R., Pappas, G.J., eds.: Hybrid Systems: Computation and Control (HSCC). Lecture Notes in Computer Science (LNCS 2993), (Springer-Verlag) 234–249.

5. Yuan, C., Lygeros, J.: Asymptotic stability and boundeness of delay switching diffusions. In Alur, R., Pappas, G.J., eds.: Hybrid Systems: Computation and Control (HSCC). Lecture Notes in Computer Science (LNCS 2993), (Springer-Verlag) 646–659.

6. Hwang, I., Hwang, J., Tomlin, C.J.: Flight-mode-based aircraft conflict detection using a residual-mean interacting multiple model algorithm. In: Proceedings of the AIAA Guidance, Navigation, and Control Conference. AIAA-2003-5340 (2003)

7. Hespanha, J.P.: Stochastic hybrid systems: Application to communication networks. In Alur, R., Pappas, G.J., eds.: Hybrid Systems: Computation and Control (HSCC). Lecture Notes in Computer Science (LNCS 2993), (Springer-Verlag) 387–401.

8. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. (translated by S.K. Mitter, Springer Verlag, New York)

9. Jameson, A.: Aerodynamic design via control theory. (Princeton University Report MAE 1824, ICASE Report No. 88-64, November 1988, also, J. of Scientific Computing, Vol. 3, 1988, pp. 233-260)

10. Evans, L.: Partial Differential Equations. AMS Press (2002)

11. Hu, J., Sastry, S.: Optimal sojourn time control within an interval. (Proceedings of the American Control Conference, Denver, CO. 3478–3483. 2003)

12. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley (1986)

13. Gill, P.E., Murray, W., Wright, M.H.: Practical Optimization. Academic Press. Harcourt Brace and Company (1999)

14. Rogers, L., Williams, D.: Diffusions, Markov processes and Martingales. Vol.1, 2nd Ed., Cambridge (2000)

15. Durrett, R.: Stochastic Calculus: A Practical Introduction. CRC Press (1996)

16. Oksendal, B.: Stochastic Differential Equations. An Introduction with Applications. Springer-Verlag, Sixth Edition (2003)