

Optimal control of stochastic hybrid system with jumps: A numerical approximation



Büşra Zeynep Temoçin*, Gerhard-Wilhelm Weber

Middle East Technical University, Institute of Applied Mathematics, 06531 Ankara, Turkey

ARTICLE INFO

Article history:

Received 25 February 2013

Received in revised form 27 August 2013

Keywords:

Stochastic hybrid systems

Jump diffusions

Markov processes

Switching approximation

Optimal control

ABSTRACT

The generalized class of stochastic hybrid systems consists of models with regime changes including the occurrence of impulsive behavior. In this paper, the stochastic hybrid processes with jumps are approximated by locally consistent Markov decision processes that preserve local mean and covariance. We further apply a randomized switching policy for approximating the dynamics on the switching boundaries. To investigate the validity of the approximation, we study a stochastic optimal control problem. On the basis of the discrete approximation, we solve the optimal control problem and show that the solution obtained from the discretized problem converges to the solution of the original optimal control problem.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Stochastic hybrid systems (SHS) are dynamical systems with interacting continuous and discrete behaviors. This generalized class of systems arises in numerous applications of dynamical models with multiple modes. Engineering systems like air traffic management, biological networks and manufacturing systems are characterized by hybrid dynamics as they exhibit complex behaviors resulting from the interactions between heterogeneous components. The analysis and control of a stochastic hybrid system have attracted considerable interest from many researchers and constitute an important topic in hybrid system modeling. For continuous SHS, early contributions include the studies of [1–4], with [5,6] establishing theoretical foundations for the measurability of events for reachability problems. In [7], the authors focus on the reachability problem using the Markov chain approximation of [8] and apply the results to air traffic management problems. For various applications of stochastic hybrid systems we refer the reader to [9,10] and the references therein. Some other classes of stochastic hybrid systems which have been studied in the literature include diffusion processes with Markovian switching parameters [11,12], switching diffusions [1], and Markov decision drift processes [13].

In this paper, we concentrate on the approximation of SHS, where the continuous state is governed by a stochastic differential equation of Itô–Skorohod type with the drift coefficient, the diffusion matrix and a *jump* function depending on the discrete component. The main difference of our model from related SHS models is the presence of controlled jumps which represent the impulsive behavior in the continuous state. We apply the well-known approach presented in [8] to our extended SHS which incorporates jumps in the continuous state with autonomous switchings in the discrete state. The approach has been considered in many different classes of stochastic systems [14–16]; however, the extension to SHS with jumps (*SHSJ*) remains novel. The main contribution of this paper is the application of the weak approximation approach of [8] to controlled autonomous SHS with jumps, which is extended to autonomously switching controlled diffusions by [15] and to SHS by [14]. We extend the method employed in [14] to SHSJ where a randomized switching policy that guarantees the

* Corresponding author. Tel.: +90 5333086586.

E-mail address: btmocin@metu.edu.tr (B.Z. Temoçin).

continuity of switching times under certain conditions is introduced. Then, using an extension of the methodology presented in [8], we show that the approximating processes converge to the original stochastic hybrid systems.

The approximation of SHS is achieved by appropriately chosen Markov decision processes (MDPs) defined on a discrete state space. The discrete approximating processes are chosen in such a way that they preserve certain “local” properties that are similar to those of the original controlled process. Then, to investigate the validity of the approximation we consider the stochastic optimal control problem of minimizing a cost until a target set is reached. The advantage of the discretization method is that the cost of the approximating MDPs converges weakly to the original cost. Under rather broad conditions, we also prove that the solution of the problem for the sequence of the approximating processes converges to that for the underlying original process as the approximation parameter goes to zero.

The rest of the paper is organized as follows. Section 2 introduces our SHSJ model and Section 3 presents the stochastic optimal control problem as well as the approximation method. In Section 4, computational methods and the discretized optimal control problem are described. Convergence results and optimality discussions are presented in Section 5. A numerical example from finance is provided to illustrate the performance of the approximation method in Section 6. Finally, some additional remarks and ideas for future studies are provided in Section 7.

2. Stochastic hybrid systems with jumps

We define a stochastic hybrid system with jumps as follows:

$$(X, \Theta, U, \Omega, A, a, b, f, \delta, R, (x_0, \theta_0)), \quad (1)$$

where

- (i) $X \subseteq \mathbb{R}^d$ is the continuous state space,
- (ii) $\Theta = \{1, 2, \dots, M\}$ is a finite set of discrete states,
- (iii) $U = \{U_\theta\}_{\theta \in \Theta}$, with $U_\theta \subseteq \mathbb{R}^{m_\theta}$ ($\theta \in \Theta$), is a finite family of sets of continuous controls,
- (iv) the partition of X is denoted by $\Omega = \{\Omega_\theta\}_{\theta \in \Theta}$, with $\Omega_\theta \subseteq \mathbb{R}^d$ ($\theta \in \Theta$),
- (v) $A = \{A_\theta\}_{\theta \in \Theta}$, with $A_\theta \subset \partial\Omega_\theta$ ($\theta \in \Theta$), is the collection of switching points,
- (vi) $a : \mathbb{R}^d \times \Theta \times U_\theta \rightarrow \mathbb{R}^d$ is a controlled drift term,
- (vii) $b : \mathbb{R}^d \times \Theta \times U_\theta \rightarrow \mathbb{R}^d \times \mathbb{R}^d$ is a controlled diffusion term,
- (viii) $f : \mathbb{R}^d \times \Theta \times U_\theta \times \mathbb{R} \rightarrow \mathbb{R}$ is a controlled jump function,
- (ix) $\delta : \Theta \times A \rightarrow \Theta$ is an autonomous switching map,
- (x) $R : \Theta \times A \rightarrow P(X)$ is the reset map which assigns a reset probability kernel to each $x(\cdot) \in A_\theta$ and θ on X , and $P(X)$ is a family of probability measures on X ,
- (xi) (x_0, θ_0) is the initial state defined on $X \times \Theta$ at time $t_0 = 0$.

To describe the dynamics of an SHSJ, we need to consider an $\mathbb{R}^d$ -valued standard Wiener process $W(\cdot)$ and a Poisson random measure $\bar{N}(\cdot, \cdot)$. Let us suppose that at time t_i ($i \in \mathbb{N}_0$), the state of the SHSJ is $(x_i, \theta_i) = (x_{t_i}, \theta_{t_i})$ with $x_i \in \Omega_{\theta_i}^0$. Here, $\Omega_{\theta_i}^0$ denotes the interior of the set Ω . While the continuous state x_t stays in $\Omega_{\theta_i}^0$ and the discrete state θ_t remains constant, the evolution of x_t is given by the following stochastic differential equation (SDE):

$$dx_t = a(x_t, \theta_t, u) dt + b(x_t, \theta_t, u) dW_t + \int_{\mathbb{R}} f(x_t, \theta_t, u, y) \bar{N}(dt, dy), \quad (2)$$

where $\bar{N}(dt, dy)$ is defined as

$$\bar{N}(dt, dy) = \begin{cases} N(dt, dy) - \nu(dy)dt, & \text{if } |y| < R, \\ N(dt, dy), & \text{if } |y| \geq R, \end{cases}$$

for some $R \in [0, \infty]$. Here, $N(\cdot, \cdot)$ denotes a Poisson random measure with intensity $dt \times \nu(dy)$, where $\nu(\cdot)$ is the Lévy measure of x_t on a compact support $\Gamma \subset \mathbb{R}$.

Let us define the stopping time $t_{i+1} := \inf\{t > t_i : x_t \notin \Omega_{\theta_i}^0\}$. At time t_{i+1} , the discrete state switches autonomously, resulting in a reset in the continuous state. Following this switching, the new discrete state becomes $\theta_{i+1} = \delta(\theta_i, x(t_{i+1}^-))$ while the new continuous state $x_{t_{i+1}}$ is randomly determined according to the probability measure $\mathbb{P}(\theta_i, x_{t_{i+1}}^-)(\mathcal{E})$, where $\mathcal{E} \subseteq \Omega_{\theta_{i+1}} \subseteq \mathbb{R}^d$ is a measurable set. Here, $x(t^-)$ denotes the left limit of $x(t)$. The evolution of $x(t)$ is then described by the SDE (2) with $\theta(t) = \theta_{i+1}$ and initial condition $x(t_{i+1})$ until the next switching time. Thus, x_t switches from one jump diffusion path to another, as the discrete state θ_t moves from one state to another.

With these definition and this notation, a typical stochastic execution starts at time t_0 , from (x_0, θ_0) , and the continuous state x_t evolves according to the SDE (2) until time t_1 when x_t first hits $\partial\Omega_{\theta_0}$. Then, depending on the hitting position $x(t_1^-)$, the discrete state jumps to $\theta(t_1) = \theta_1$ and the continuous state is reset randomly to $x(t_1) = x_1$.

Considering the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we assume that the functions $a(x, \theta, u)$, $b(x, \theta, u)$ and $f(x, \theta, u, y)$ satisfy Lipschitz and growth conditions, for a given $\theta \in \Theta$. Moreover, we assume that a , b and f are continuous with respect to (x, θ, u) , and $f(x, \theta, u, \cdot)$ is bounded uniformly in $u \in U$, in a neighborhood of $y = 0$. These assumptions ensure that for

each admissible control $u \in U$, there exists a unique strong solution to (2) with an initial condition (see [10]). Therefore, the controlled SHS also has a unique strong solution between switching points (see [17]).

We make the assumption that every point in A_θ is a *regular* point, in which case a sample path of (2) that starts at $x \in A_\theta$ will not remain in A_θ for a nonempty time interval. We note that if $x \in A_\theta$ is regular for A_θ , then it is also *regular* for a neighborhood of A_θ around x and it can be concluded that

$$\lim_{\epsilon \rightarrow 0} \int_0^t \mathbb{P}(x_s \in N_\epsilon(A_\theta)) ds = 0,$$

where $N_\epsilon(A_\theta) = \{x \in \mathbb{R}^d : d(x, A_\theta) \leq \epsilon\}$ and $d(x, A_\theta) := \inf\{\|x - a\|_2 : a \in A_\theta\}$ is the Euclidean distance between x and A_θ . By assuming regularity, we guarantee instantaneous occurrences for autonomous switchings. For this assumption a sufficient condition is that the set A_θ is a surface of dimension $d - 1$ and the diffusion coefficient is non-degenerate. We also assume that there exists a $\delta > 0$ such that

$$\mathbb{P}\{\inf\{t > t_{i+1}, x_t \in A_\theta\} = 1$$

for $x(t_{i+1}) \in \mathcal{E} \subseteq \Omega_{\theta_{i+1}}$, $d(x(t_{i+1}), A_\theta) \geq \epsilon$. Therefore, the inequality $t_{i+1} - t_i > \delta$ ($i = 1, 2, \dots$) holds with probability 1. For the control policy, we consider an admissible continuous control $u(t)$ which is non-anticipative with respect to $W(\cdot)$ and $N(\cdot, \cdot)$ and takes values in a compact set U .

To define the jump term given in (2), we work with the jump times and sizes directly. Let v_n denote the time of the n th jump with $v_0 = 0$ and $y_n \in \Gamma$ denoting the jump size for $n \geq 1$. The jump term is represented by the function $f(\cdot)$ and the Poisson measure $N(\cdot)$ with intensity measure $\lambda dt \times F(dy)$. Here, $F(\cdot)$ is the common jump size distribution with a compact support Γ . The function $f(\cdot)$ is assumed to be bounded, measurable and continuous on X for each value of y and θ . Let $\{v_{n+1} - v_n, y_n : n \in \mathbb{N}\}$ be mutually independent random variables with $v_{n+1} - v_n$ being exponentially distributed with mean value $1/\lambda$. In addition, for each n , and let $\{v_{i+1} - v_i, y_i : i \geq n\}$ be independent of $\{x(s), v_{i+1} - v_i, y_i : s < v_n, i < n\}$. There exists a sequence $\{v_n, y_n : n \in \mathbb{N}\}$ such that we can represent the jump term of $x(\cdot)$ as

$$J(t) = \sum_{v_n \leq t} f(x(v_n^-), y_n). \quad (3)$$

Since $v_{n+1} - v_n$ is exponentially distributed, we can write

$$\mathbb{P}\{x(\cdot) \text{ jumps on } [t, t + \Delta) | x(s), W(s), N(s), s \leq t\} = \lambda \Delta + o(\Delta). \quad (4)$$

By the independence properties and the definition of y_n , we can write

$$\mathbb{P}\{x(t) - x(t^-) \in H | \text{jump at } t, x(t^-) = x, W(s), x(s), N(s), s < t\} = F\{y : f(x(t^-), y) \in H\},$$

for any subset $H \subset \Gamma$.

3. Stochastic optimal control

3.1. Foundations

The optimal control methodology includes a powerful recursive algorithm that selects the optimal policy for the current state on an optimized path of future policies. For detailed information on the classical theory, we refer the reader to [18,19,17].

In this section, we consider an optimal control problem of minimizing a cost until a target set is reached. Consider the SDE (2) that evolves in Ω_θ^0 and let $G \subseteq X \subseteq \mathbb{R}^d$ be a compact set which is the closure of its interior G^0 . For $\beta > 0$, we consider the discounted cost

$$H(x_0, \theta_0, u) := E \left[\int_0^\tau e^{-\beta s} k(x_s, \theta_s, u_s) ds + e^{-\beta \tau} g(x_\tau) \right], \quad (5)$$

where $k : \mathbb{R}^d \rightarrow \mathbb{R}$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}$ are continuous running and terminal cost functions, respectively. $\tau = \inf\{t \geq 0 : x(t) \notin G^0\}$ is the first time of escape of $x(\cdot)$ from G^0 . The optimal control problem is formulated as the minimization of the discounted cost (5) on the basis of dynamic programming arguments. To guarantee that the stopping time τ and the cost function (5) are well-defined and bounded, we assume that for every initial state (x_0, θ_0) , there exists an admissible control policy such that $x(t)$ will reach the target set G , in the case where $\beta = 0$. If $\beta > 0$, the cost will be bounded, even if the stopping time is not.

The value function is defined by

$$V(x_0, \theta_0) := \inf_{u \in U} H(x_0, \theta_0, u), \quad x_0 \in \Omega_{\theta_0}. \quad (6)$$

By standard dynamic programming, we can formulate the following verification theorem analogous to the classical *Hamilton–Jacobi–Bellman (HJB)* form (see [17, Theorem 4.1]). Let us generally denote the value function by $V(x, \theta)$, for any initial state (x, θ) .

We note that for a control $u = u(x)$, x_t is a jump diffusion with infinitesimal generator

$$A^u V(x, \theta) = \sum_{i=1}^k a_i(x, \theta, u) \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^k (bb^T)(x, \theta, u) \frac{\partial^2 V}{\partial x_i \partial x_j}(x, \theta) + \sum_{j=1}^l \int_{\mathbb{R}} \{V(x + f^{(j)}(x, \theta, u, y_j), \theta) - V(x, \theta) - \Delta V(x, \theta) f^{(j)}(x, \theta, u, y_j)\} \lambda F(dy_j). \quad (7)$$

Theorem 1 (The HJB Form for Optimal Control of Jump Diffusions).

(a) Suppose $V \in C^2(G^0) \cap C(G)$ satisfies the following:

- (i) $A^u V(x, \theta) + k(x, \theta, u) \leq 0$ for all $x \in \Omega_\theta^0$, $u \in U$, $\theta \in \Theta$.
- (ii) $x(\tau) \in \partial G$ a.s. on $\{\tau < \infty\}$ and $\lim_{t \rightarrow \tau^-} V(x, \theta) = g(x(\tau))$ a.s. for all admissible controls.
- (iii) $\mathbb{E}^x(|V(x(\tau), \theta)| + \int_0^\tau |AV(x_t, \theta_t)| dt) < \infty$ for all admissible controls.
- (iv) $V(x(\tau), \theta)$ is uniformly integrable for all admissible controls and $x \in G^0$. Then,

$$V(x', \theta') \leq V(x, \theta), \quad \forall \theta \in \Theta, x \in A_\theta, \theta = \delta(\theta, x), x' \sim R(\theta, x)(\Xi).$$

(b) Moreover, suppose that for each $x \in G$ there exists $\bar{u}(x) \in U$ such that

- (v) $A^{\bar{u}(x)} V(x, \theta) + k(x, \theta, \bar{u}(x)) = 0$ and
- (vi) $V(x^{\bar{u}}(\tau), \theta)$ is uniformly integrable,

where A is the infinitesimal generator given by (7). Then, $V(x, \theta)$ is the optimal cost and $\bar{u}(x)$ is the optimal control.

The computation of the optimal value function $V(x, \theta)$ can be very difficult for some cases and usually requires computational methods based on discretization of the state space. In this paper, we apply a discretization method for the approximation of the SHSJ by appropriately chosen Markov decision processes (MDPs) as described in [8].

3.2. Constructing the approximating MDPs

In this subsection, we discuss a method for approximating the continuous dynamics by discrete MDPs only in the interior regions of Ω_θ . The system is approximated at every location of θ by a controlled Markov process that evolves in a state space which is a discretization of the region Ω_θ . The criterion which must be satisfied by the approximating MDP is *local consistency*. Local consistency means that the conditional mean and covariance of the MDP are similar to the local mean and covariance of the original process. An approximation parameter h analogous to a finite difference interval parameterizes the approximating Markov process. As h goes to zero, the local properties of the MDP resemble the local properties of the original stochastic process. The approximation of the system at the switching points will be considered in the next subsection.

For each $h > 0$, let $\{\xi_n^h : n < \infty\}$ be a controlled discrete MDP on a discrete state space $S_h \subseteq \mathbb{R}$ with transition probabilities denoted by $p^h((x, \theta), (z, \theta)|u)$ for $u \in U_\theta$. Let us further assume that u_n^h is the actual control for the chain at time n . We consider an interpolation interval length of $\Delta t^h(X, \theta, u)$ and define $\Delta t_n^h := \Delta t^h(\xi_n^h, u_n^h)$. Let $\sup_{x,u} \Delta t^h(x, u) \rightarrow 0$ as $h \rightarrow 0$, but with $\inf_{x,u} \Delta t^h(x, u) > 0$ for each $h > 0$. Furthermore, let $f_h(\cdot)$ be a bounded measurable function such that

$$|f_h(x, \theta, u, y) - f((x, \theta, u, y))| \rightarrow 0$$

uniformly in x for each $y \in \Gamma$ and let it satisfy $x + f_h(x, \theta, u, y) \in S_h$ for $x \in S_h$. We will use the values of $f_h(x, \theta, u, y)$ to get approximations of the jump terms. In the case of a continuous diffusion process where the SDE is given as

$$dx_t = a(x_t, \theta_t, u) dt + b(x_t, \theta_t, u) dW_t, \quad (8)$$

a locally consistent MDP must satisfy the conditions

$$\mathbb{E}(\Delta \xi_n^h) = a(x, \theta, u) \Delta t^h(x, \theta, u) + o(t^h(x, \theta, u)), \quad (9)$$

$$\mathbb{E}((\Delta \xi_n^h - \mathbb{E}(\Delta \xi_n^h))(\Delta \xi_n^h - \mathbb{E}(\Delta \xi_n^h))^T) = b(x, \theta, u)b^T(x, \theta, u) \Delta t^h(x, \theta, u) + o(t^h(x, \theta, u)), \quad (10)$$

which represent both local mean and covariance, respectively, where $\Delta \xi_n^h = \xi_{n+1}^h - \xi_n^h$, $\xi_n^h = x$ (for details see [8]).

A controlled MDP is said to be *locally consistent* with (2) if there is an interpolation interval $\Delta t^h(X, \theta, u)$ which goes to zero as $h \rightarrow 0$, uniformly in X, θ and u , such that:

- (a) There is a transition probability $p_D^h((x, \theta), (z, \theta)|u)$ case which is locally consistent with (8) in the sense that the conditions (9) and (10) hold for an interpolation interval $\Delta t^h(x, \theta, u)$.
- (b) There is a function $\delta^h(x, \theta, u) = o(\Delta t^h(x, \theta, u))$ such that the one-step transition probability $p^h((x, \theta), (z, \theta)|u)$ for the chain can be represented in the factored form

$$p^h((x, \theta), (z, \theta)|u) = (1 - \lambda \Delta t^h(x, \theta, u) - \delta^h(x, \theta, u)) p_D^h((x, \theta), (z, \theta)|u) + (\lambda \Delta t^h(x, \theta, u) + \delta^h(x, \theta, u)) F(y|f_h(x, \theta, y, u) = z - x). \quad (11)$$

The rule (11) is equivalent to the experiment of flipping a coin with probability of heads being $(1 - \lambda \Delta t^h(x, \theta, u) - \delta^h(x, \theta, u))$. This probability depends only on ξ_n^h and the current control at time n , and not on the past. If heads occurs, then the next state ξ_{n+1}^h is determined using the diffusion transition probability $p_D^h((x, \theta), (z, \theta)|u)$. Hence, as the MDP evolves, there is a sequence of coin tosses, and heads and tails correspondingly.

3.3. Switching boundaries

The method discussed in Section 3.2 approximates the dynamics of the SDE (2) only in the interior of the regions Ω_θ . To approximate the whole SHSJ, we should define the MDP in the neighborhood of the switching boundaries in such a way that it preserves local consistency. To achieve this, we will follow and extend the approach of [14] for our SHSJ. The main idea consists of transforming the partition of the SHSJ to a cover and then defining appropriate random switching functions to approximate the switching behavior at the boundaries.

We consider the boundary set A_θ of Ω_θ as

$$A_\theta = \{x \in \mathbb{R}^d : a_\theta(x) = 0\},$$

where $a_\theta : \mathbb{R}^d \rightarrow \mathbb{R}$ is a smooth functional. To ensure that A_θ is a $(d - 1)$ -dimensional hypersurface, the functional a_θ must satisfy $\Delta a_\theta(x) \neq 0$, for all $x \in A_\theta$. The region Ω_θ is expanded to Ω'_θ , defined by

$$\Omega'_\theta := \{x \in \mathbb{R}^d : 0 \leq a_\theta(x) \leq \gamma(h)\},$$

where $\gamma(h) > h > 0$ for every $h > 0$ and $\gamma(h) \rightarrow 0$ as $h \rightarrow 0$. By expanding Ω_θ to Ω'_θ , we obtain a new family $\Omega' = \{\Omega'_\theta\}$ for all $\theta \in \Theta$. Herewith, Ω' is a covering of the continuous state space of the SHSJ as $\bigcup_{\theta \in \Theta} \Omega_\theta \subseteq \bigcup_{\theta \in \Theta} \Omega'_\theta$ for every θ . Let $\{(\xi_n^h, \theta_n^h) : n < \infty\}$ be an MDP on a discrete state space $S = \{(x, \theta) \in S^h \times \Omega : x \in \Omega'_\theta\}$ with transition probabilities denoted by $\tilde{p}^h((x, \theta), (z, \theta')|u)$. For all the states that are interior points of Ω_θ , that is $a_\theta(x) < 0$, the system does not switch, and the transition probabilities are computed as given in (11). Hence, for all $(x, \theta) \in S$ which satisfy $a_\theta(x) < 0$, we have

$$\begin{aligned} \tilde{p}^h((x, \theta), (z, \theta')|u) &= p^h((x, \theta), (z, \theta')|u), \\ \tilde{p}^h((x, \theta), (x, \theta')|u) &= p^h((x, \theta), (x, \theta')|u), \end{aligned} \quad (12)$$

and $\tilde{p}^h((x, \theta), (x, \theta')|u) = 0$ for all $\theta \neq \theta'$.

Extending the method presented in [14] to our SHSJ, the switching behavior of the system is approximated by introducing random switching times of the discrete state and discretizing the reset map of the continuous state. Considering the boundary sets $A'_\theta = \{x \in \mathbb{R}^d : a_\theta(x) = \gamma(h)\}$ for Ω'_θ , a switching rate function $\lambda_\theta(x)$ is defined for each boundary A'_θ , continuously on $L_\theta = \{x \in \mathbb{R}^d : 0 \leq a_\theta(x) \leq \gamma(h)\}$. Moreover, we assume that $\lambda_\theta(x) = 0$ if $a_\theta(x) = 0$, and $\lambda_\theta(x) \rightarrow \infty$ as $x \rightarrow A'_\theta$ in the Euclidean sense. The reset map $R(\theta, x)(\mathcal{E})$ is also approximated by a discrete transition probability kernel. If a discrete transition occurs, the next continuous state is determined randomly from the grid points that belong to $\mathcal{E}$ via a uniform distribution.

On the basis of the newly introduced random switching times and discretized reset maps, for any $(x, \theta) \in S$ such that $0 \leq a_\theta(x) \leq \gamma(h)$, the transition probabilities of the approximating MDP for states $(x, \theta), x \in L_\theta$, are defined as

$$\tilde{p}((x, \theta), (z, \theta')|u) := \begin{cases} (1 - e^{-\lambda_\theta(x) \Delta t^h(x, \theta, u)}) p((x, \theta), (z, \theta')|u), & \text{if } \theta \neq \theta' \text{ and } z \in \mathcal{E}, \\ e^{-\lambda_\theta(x) \Delta t^h(x, \theta, u)} p((x, \theta), (z, \theta')|u), & \text{if } \theta = \theta' \text{ and } z \notin \mathcal{E}. \end{cases}$$

4. Computational methods

Let N_h denote the stopping time when ξ_n^h reaches the target set G . Then, assuming constant β in the intervals $[t_n^h, t_{n+1}^h)$, the cost (5) can be approximated by

$$H^h(x_0, \theta_0, u) := E \left(\sum_{n=0}^{N_h} e^{-\beta t_n^h} c(\xi_n^h, \theta_n^h, u_n^h) + e^{-\beta t_{N_h}^h} g(\xi_{N_h}^h) \right),$$

where $c(\xi_n^h, \theta_n^h, u_n^h) := k(\xi_n^h, \theta_n^h, u_n^h) \Delta t_n^h$. Assuming that the above sum is well-defined and bounded, minimizing the cost becomes a discrete problem and it can be solved using standard dynamic programming algorithms based on value iteration methods. Defining the optimal value function of the approximated cost as

$$V^h(x_0, \theta_0) := \inf_u H^h(x_0, \theta_0, u),$$

and using standard programming arguments, we can derive the equation

$$V^h(x, \theta) = \min_u \left(\sum_{z, \theta'} \tilde{p}((x, \theta), (z, \theta')|u) V^h(z, \theta') + c(x, \theta, u) \right)$$

if $x \in X \setminus G$, and $V^h(x, \theta) = g(x)$ if $x \in G$. The optimal value function can be computed as follows from Jacobi iterations [8, Theorem 2.2]:

$$V_{n+1}^h(x, \theta) = \min_u \left(\sum_{z, \theta'} \tilde{p}((x, \theta), (z, \theta')|u) V_n^h(z, \theta') + c(x, \theta, u) \right).$$

The value function $V^h(x, \theta)$ given above is an approximation of the optimal cost (6) for the SHSJ. Given the discrete optimal value function, a feedback control scheme can be constructed to compute $u(x)$. For (x, θ) with $x \in \Omega_\theta$, the control law is given by

$$u(x) = \arg \min_{u \in U} \left(\sum_{z, \theta'} e^{-\beta \Delta^h(x, \theta, u)} \tilde{p}((x, \theta), (z, \theta')|u) H(x, \theta, u) + c(x, \theta, u) \right). \quad (13)$$

The cost of the discrete-time approximating processes converges weakly to the original cost as shown in the next section. This convergence supports the validity of the approximation approach employed.

5. Convergence

Consider the optimal control input $\{u_n^h : n < \infty\}$ and the locally consistent approximating process $\{(\xi_n^h, \theta_n^h) : n < \infty\}$. Let $\{t_i^h\}$ denote the sequence of switching times for $i \geq 0$. In [14], a continuous-time Markov interpolation $\{(\psi^h(t), \theta^h(t))\}$ is constructed such that $\{\psi^h(\cdot)\}$ is a Markov process for the construction of the continuous part, which is a Wiener process. We extend the approach and apply the method to our SHSJ. We define $\tau_0^h = 0$ and let $\{\tau_n^h : n \in \mathbb{N}_0\}$ denote the moments of change of $\{\psi^h(\cdot)\}$. Further, we put $v_n^h = \tau_{v_n^h+1}^h$ to distinguish the jump times of $\{\psi^h(\cdot)\}$ which are due to the approximation from the other jump times. By the definition of the approximating chain, we have

$$\psi^h(v_n^h) - \psi^h(v_n^{h-}) = f_h(\psi^h(v_n^{h-}), y_n).$$

Let $\xi_0^h = x$. Then we can write

$$\psi^h(t) = \psi^h(0) + \int_{t_0}^t a(\psi^h(s), \theta^h(s), u^h(s)) ds + M^h(t) + J^h(t) + \delta_1^h(t),$$

where $E[\sup_{s \leq t} \delta_1^h(s)] \rightarrow 0$ as $h \rightarrow 0$. Here, $M^h(\cdot)$ is a martingale whose discontinuities go to zero as $h \rightarrow 0$, with quadratic variation

$$\int_{t_0}^t (bb^T)(\psi^h(s), \theta^h(s), u^h(s)) ds + \delta_2^h(t),$$

where $E[\sup_{s \leq t} \delta_2^h(s)] \rightarrow 0$ as $h \rightarrow 0$. The computational methods will give an optimal control sequence of $\{u_n^h\}$. However, the optimal control may not exist as $h \rightarrow 0$. To show the convergence, a *relaxed control* representation is employed as introduced in [20,8]. We denote the space of the relaxed control by $\mathcal{A}$ and obtain a relaxed control representation by defining probability measures μ_t on $\mathcal{B}_{\mathcal{A}}$, and the Borel σ -algebra on $\mathcal{A}$. Letting μ_t^h denote the corresponding probability measure for the sequence $\psi^h(t)$, we can write $\psi^h(t)$ as

$$\psi^h(t) = \psi^h(0) + \int_{t_0}^t \int_{\mathcal{A}} a(\psi^h(s), \theta^h(s), u^h(s)) \mu_s^h(du) ds + M^h(t) + J^h(t) + \delta_1^h(t).$$

Next, we give a theorem which states that $\{\psi^h(t), \theta^h(t)\}$ converges weakly to the execution of the SHSJ based on the weak convergence results, as given in [21].

Theorem 2. Consider the locally consistent approximation process $\{(\xi_n^h, \theta_n^h) : n \in \mathbb{N}\}$ with an admissible control sequence $\{u_n^h : n \in \mathbb{N}\}$ and a sequence of random switching times $\{t_i^h : n \in \mathbb{N}\}$. Let $\{\psi^h(t), \theta^h(t)\}$ be the continuous-time Markov interpolation and $\mu^h(\cdot)$ a relaxed control representation of $\{u_n^h\}$ for $\{\psi^h(t)\}$. Then $\{\psi^h(t), \theta^h(t)\}$ converges weakly to the execution of the SHSJ.

The next theorem follows from Theorem 2 and the results of [8], and shows the convergence of the optimal cost. The cost for the continuous-time interpolation $\{(\psi^h(t), \theta^h(t))\}$ with control sequence $\mu^h(t)$ is

$$H^h(x_0, \theta_0, \mu^h) = E \left(\int_{t_0}^{\tau_h} e^{-\beta s} c(\psi^h(s), \theta^h(s), u) \mu_s^h(du) ds + e^{-\beta \tau_h} g(\psi_{\tau_h}^h) \right).$$

We define $V^h(x, \theta) := \inf_{\mu^h} H^h(x_0, \theta_0, \mu^h)$ and let $V^h(x, \theta)$ be the optimal value function for the process $\{(\xi_n^h, \theta_n^h) : n \in \mathbb{N}\}$ with an admissible control sequence $\{u_n^h : n \in \mathbb{N}\}$. Then, $|V^h(x, \theta) - V^h(x, \theta)| \rightarrow 0$ as $h \rightarrow 0$. We will use $V^h(x, \theta)$ to denote the optimal value function for both processes.

Theorem 3. Consider the locally consistent approximating process $\{(\xi_n^h, \theta_n^h) : n \in \mathbb{N}\}$ with an admissible control sequence $\{u_n^h : n \in \mathbb{N}\}$ and a sequence of random switching times $\{t_i^h\}$. Let $\{(\psi^h(t), \theta^h(t))\}$ be the continuous-time Markov interpolation and $\mu^h(\cdot)$ a relaxed control representation of $\{u_n^h : n \in \mathbb{N}\}$ for $\{\psi^h(t)\}$. If $\beta > 0$ or if $\beta = 0$ and the sequence τ_h is uniformly integrable, then $H^h(x_0, \theta_0, \mu^h) \rightarrow H(x_0, \theta_0, \mu)$ and $V^h(x, \theta) \rightarrow V(x, \theta)$.

5.1. Remarks on optimality

Here, provided that we are facing a minimization task on overall costs, an important role is played by the possible convexity of the value function and its admissible set. In the case of a maximization task on, e.g., expected overall utility, it is concavity. The two positions are often dual to each other. Whenever the conditions are attained from their strict counterpart, i.e., by strict convexity (or concavity), we aim to numerically find a *global minimum*. If a very “qualified” structure is given in terms of a global well-structuredness, as given in conic optimization and conic optimal control, we can study *duality theory* and follow up *optimality* or *sub-optimality* in the course of a running algorithm by following the length of the duality gap. Further cases of similar qualifications where primal–dual methods and interior point algorithms are applicable can be found in related literature and are often well-understood, e.g., with the help of so-called self-concordance. Discussions on optimality and sub-optimality related conditions in optimization and optimal control can be found in [23]. In general, an important concept for observing and, if needed, accelerating the convergence and, thereby, the sub-optimality of members of an iterative algorithm of optimization and optimal control is that of *merit functions*. Apart from the special cases of strict structure, local minimizers may exist which are not global minimizers, and then special attention has to be paid to the form of local a priori information. Furthermore, local approximations can be placed, established and left for the next ones, when incorporating elements of *trust-region algorithms* [24].

We will pay attention in the future, also, to convergence orders of algorithms, including varieties of our method. Our algorithms are founded in the discretization schemes related to the SDEs; after having addressed those, the convergence is influenced by the choice of the preferred optimality criterion.

The convergence level of approximation is controlled by a parameter ϵ and to investigate the sub-optimality level, we introduce an upper bound based on relaxed control which could be considered as an *a priori* measure of sub-optimality. Let $\mu^\epsilon(\cdot)$ denote the relaxed control of a piecewise constant control $u^\epsilon(\cdot)$. The optimality level of $\mu^\epsilon(\cdot)$ is given by ϵ and the solution $(x^\epsilon, \theta^\epsilon)$ based on $\mu^\epsilon(\cdot)$ satisfies $|H(x_0, \theta_0, \mu^\epsilon) - H(x_0, \theta_0, \mu)| \leq \epsilon$. It can be shown that the approximating process $(\psi^h(\cdot), \theta^h(\cdot))$ with control $\mu^h(\cdot)$ converges to the solution $(x^\epsilon, \theta^\epsilon)$ with control $\mu^\epsilon(\cdot)$ in the weak sense.

6. A numerical example

In this section, we illustrate the proposed approach using a simple surplus model of an insurance company presented in [22]. We have modified the model of [22] by considering a generator matrix for the discrete state that depends on the continuous state.

We consider the financial market with a risk-free asset $B(t)$ and a risky asset $S(t)$ with prices satisfying

$$\begin{aligned}\frac{dB(t)}{B(t)} &= l(\theta_t)dt, \\ \frac{dS(t)}{S(t)} &= b(\theta_t)dt + \sigma(\theta_t)dW_t,\end{aligned}$$

where for each $\theta_t \in \Theta$, $l(\theta_t)$ and $b(\theta_t)$ are the return rates of the risk-free and risky assets, respectively. We assume that $c(\theta_t)$ is the premium rate which satisfies $c(\theta_t) > 0$ for each $\theta_t \in \Theta$. The discrete state θ_t denotes the random environment and other random factors. W_t is a standard Brownian motion and $\sigma(\theta_t)$ is the corresponding volatility. The investment process of the insurer is modeled as a portfolio process $\varphi(t)$, where the proportional surplus $\varphi(t) \in [0, 1]$ was invested in the risky asset $S(t)$. The goal is to maximize the expected total discounted dividend payment until ruin. We consider a special case of SHS by assuming a Markov discrete state $\theta(t)$ that takes values in the set $\Theta = \{1, 2\}$ with the x -dependent generator matrix

$$Q(x) = \begin{pmatrix} -0.5 - \sin(x) \cos(x) & 0.5 + \sin(x) \cos(x) \\ 5 & -5 \end{pmatrix}.$$

$Z(t)$ denotes the dividend strategy for $t \geq 0$, corresponding to the accumulated amount of dividends paid up to time t such that it is a nonnegative and nondecreasing stochastic process that is right continuous with left limits. The investment strategy $\varphi(t)$ and dividend payment policy $Z(t)$ are regular and singular stochastic controls of the problem. The surplus process considering dividend payment and investment satisfies the stochastic differential equation

$$dx(t) = [l(\theta_t)(1 - \varphi(t)) + \varphi(t)b(\theta(t))x(t) + c(\theta(t))]dt + \varphi(t)\sigma(\theta(t))x(t)dW_t - dR(t) - dZ(t), \quad (14)$$

where $R(t) = \int_0^t \int_{\mathbb{R}_+} q(x(t^-), \theta(t), \rho)N(dt, d\rho)$ and $x(0) = x$ for all $t < \tau$. The function $q(x, \theta, \rho)$ is assumed to be the magnitude of the claim sizes, where ρ has the distribution $\Pi(\cdot)$. Then the Poisson measure $N(\cdot)$ has intensity $\lambda dt \times \Pi(d\rho)$. Assume

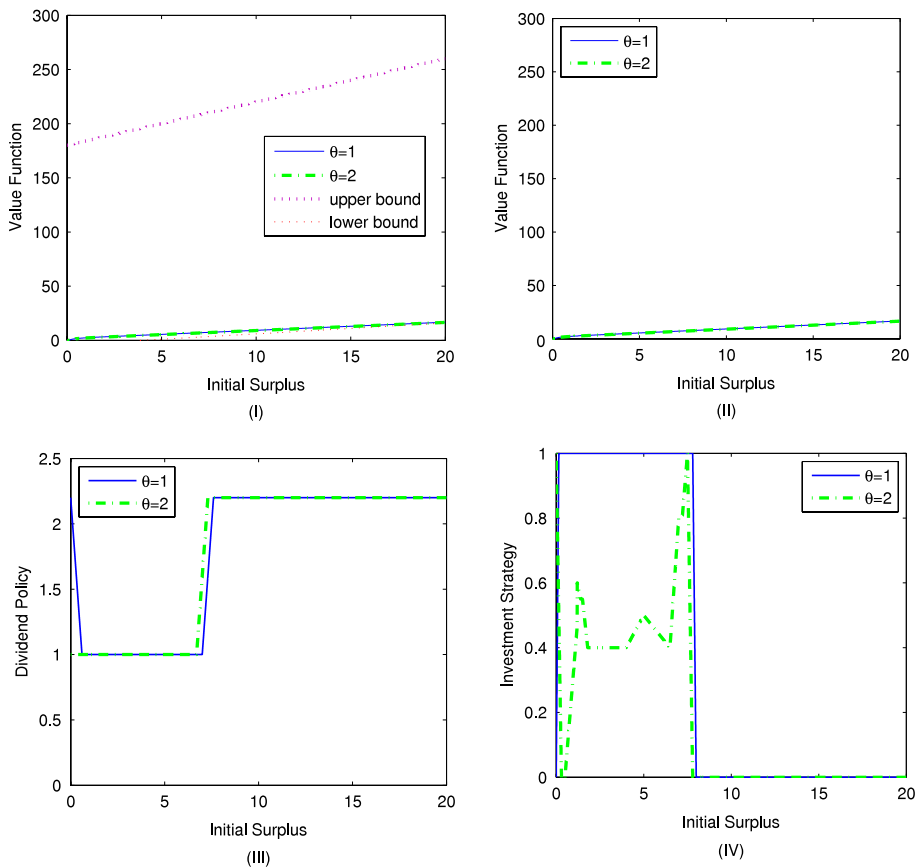


Fig. 1. Numerical results with uniform claim sizes.

that $q(\cdot, \theta_i, \rho)$ is continuous for each ρ and each $\theta_i \in \Theta$. We impose $x(t) = 0$ for all $t > \tau$, where $\tau = \inf\{t > 0 : x(t) \leq 0\}$ represents the time of ruin.

Denoting by $r > 0$ the discount factor, for an arbitrary admissible pair $\pi = (\varphi, Z)$, the expected discounted payoff is

$$J(x, \theta_i, \pi) := \mathbb{E}_{x, \theta_i} \left[\int_0^\tau e^{-rt} dZ(t) \right]. \quad (15)$$

Then the value function is

$$V(x, \theta_i) := \sup_{\pi \in \mathcal{A}} J(x, \theta_i, \pi), \quad (16)$$

where $\mathcal{A}$ denotes the collection of all admissible pairs.

The surplus process $x(t)$ under consideration is a jump diffusion process of Itô–Skorohod type under singular control π . We numerically solve the optimal control problem and obtain the optimal value function by using value iteration methods in value space. We assume a uniform distribution for claim sizes and consider the following parameters corresponding to the different discrete states:

$$\begin{aligned} c(1) &= 1 & c(2) &= 2 & l(1) &= 0.02 & l(2) &= 0.04 \\ b(1) &= 0.06 & b(2) &= 0.07 & \sigma(1) &= 0.2 & \sigma(2) &= 1.8. \end{aligned}$$

where $c(\cdot)$ denotes premium rate values, $l(\cdot)$ is the yield rate of the riskless asset, $b(\cdot)$ denotes the risky asset return rate and $\sigma(\cdot)$ is the volatility of the financial market. The portfolio rate $\varphi(t)$ is the control taking values in $[0, 1]$. Here, $R(t)$ is assumed to be a compound Poisson process with uniform claim sizes $\rho = 0.01$ and intensity $\lambda = 4$. Furthermore, the initial surplus x is supposed to have the maximum 20 and the minimum 0. We assume a discount rate of $r = 0.2$ and set the sub-optimality tolerance level to be 10^{-8} .

The computation results are given in Fig. 1. In Fig. 1(I) and (II), it is shown that the value function is a bounded concave function which increases with a linear trend after a certain threshold. Fig. 1(III) shows that the dividend policies vary for different values of discrete state while the optimal values of the discounted total dividend remain close. Finally, Fig. 1(IV) reveals that the proportion invested in the risky asset becomes zero after some barrier level and we observe

that the investment strategy changes due to Markov switching. The advantage of the discrete approximation is that the approximating process preserves the local mean and variance of the original system.

7. Conclusion and outlook

In this paper, we have developed a numerical approximation method for solving the optimal control problem for controlled autonomous stochastic hybrid systems with jumps. The approximation is done on the basis of Markov decision processes that preserve local consistency with the original solution. We recall the discussions provided in Section 5.1 which also include research suggestions and directions to be followed further. In future studies, we also aim to further introduce into our studies the tools of Lévy calculus, and interpretation of them in our generalized setting with jumps; we aim to develop our algorithm further and apply it to real-life problems of finance, economics and engineering.

References

- [1] M.K. Ghosh, A. Arapostathis, S.I. Marcus, Optimal control of switching diffusions with application to flexible manufacturing systems, *Stochastic Theory and Adaptive Control* 184 (1992) 210–224.
- [2] M. Davis, *Markov Models and Optimization*, Chapman & Hall, London, 1993.
- [3] J. Hu, J. Lygeros, S. Sastry, Towards a theory of stochastic hybrid systems, in: *Hybrid Systems: Computation and Control*, in: LNCS, vol. 1790, 2000, pp. 160–173.
- [4] M.L. Bujorianu, Extended stochastic hybrid systems and their reachability problem, in: *Hybrid Systems: Computation and Control*, in: LNCS, vol. 2993, 2004, pp. 234–249.
- [5] M.L. Bujorianu, J. Lygeros, Reachability questions in piecewise deterministic Markov processes, in: *Hybrid Systems: Computation and Control*, in: LNCS, vol. 2623, 2003, pp. 126–140.
- [6] A. Abate, M. Prandini, J. Lygeros, S. Sastry, Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems, *Automatica* 44 (2008) 2724–2734.
- [7] J. Hu, M. Prandini, S. Sastry, Probabilistic safety analysis in three dimensional aircraft flight, in: *42nd IEEE Conference on Decision and Control*, Vol. 5, 2003, pp. 5335–5340.
- [8] H.J. Kushner, P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, 2001.
- [9] M.K. Ghosh, A. Bagchi, Modeling stochastic hybrid systems, *System Modeling and Optimization* 166 (2005) 269–280.
- [10] H. Pham, Optimal stopping of controlled jump diffusion processes: a viscosity solution approach, *Journal of Mathematical Systems, Estimation and Control* 8 (1998) 1–27.
- [11] M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Dekker, 1990.
- [12] W.M. Wonham, Random differential equations in control theory, *Probability Analysis in Applied Mathematics* 2 (1970) 131–212.
- [13] F.A. van der Duyn Schouten, A. Hordijk, Average optimal policies in Markov decision drift processes with applications to a queueing and a replacement model, *Advances in Applied Probability* 15 (2) (1983) 274–303.
- [14] X. Koutsoukos, Optimal control of stochastic hybrid systems based on locally consistent Markov decision processes, *International Journal of Hybrid Systems* 4 (2004) 301–318.
- [15] M. Prandini, J. Hu, Application of reachability analysis for stochastic hybrid systems to aircraft conflict prediction, *IEEE Transactions on Automatic Control* 54 (4) (2009) 913–917.
- [16] C. Munk, Optimal consumption/investment policies with undiversifiable income risk and liquidity constraints, *Journal of Economic Dynamics & Control* 24 (2000) 1315–1343.
- [17] B. Øksendal, A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer, 2007.
- [18] R.C. Merton, *Continuous Time Finance*, Oxford, 1992.
- [19] R. Korn, *Optimal Portfolios: Stochastic Models for Optimal Investment and Risk Management in Continuous Time*, World Scientific, 1997.
- [20] L.C. Young, *Lectures on the Calculus of Variations and Optimal Control Theory*, AMS Chelsea, 1980.
- [21] S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, 1986.
- [22] Z. Jin, G. Yin, Numerical methods for optimal dividend payment and investment strategies for Markov-modulated jump diffusion models with regular and singular controls, *Journal of Optimization Theory and Applications* 54 (4) (2009) 913–917.
- [23] G.-W. Weber, *Generalized Semi-Infinite Optimization and Related Topics*, Heldermann Publishing House, 2003.
- [24] I. Griva, S.G. Nash, A. Sofer, *Linear and Nonlinear Optimization*, SIAM, 2009.