Automatica 46 (2010) 1951–1961

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem[☆]

Sean Summers*, John Lygeros

Automatic Control Laboratory, Department of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology (ETH) Zürich, Switzerland

ARTICLE INFO

Article history:
Received 16 August 2009
Received in revised form
27 March 2010
Accepted 15 July 2010
Available online 16 September 2010

Keywords:
Hybrid systems
Stochastic systems
Reachability
Safety
Target hitting
Optimal control
Dynamic programming

ABSTRACT

We present a dynamic programming based solution to a probabilistic reach-avoid problem for a controlled discrete time stochastic hybrid system. We address two distinct interpretations of the reach-avoid problem via stochastic optimal control. In the first case, a sum-multiplicative cost function is introduced along with a corresponding dynamic recursion which quantifies the probability of hitting a target set at some point during a finite time horizon, while avoiding an unsafe set during each time step preceding the target hitting time. In the second case, we introduce a multiplicative cost function and a dynamic recursion which quantifies the probability of hitting a target set at the terminal time, while avoiding an unsafe set during the preceding time steps. In each case, optimal reach while avoid control policies are derived as the solution to an optimal control problem via dynamic programming. Computational examples motivated by two practical problems in the management of fisheries and finance are provided.

1. Introduction

From air traffic control and chemical processes to biological networks and financial markets, many systems exhibit a behavior that cannot easily be captured by simple deterministic models. In light of this, Stochastic Hybrid System (SHS) models have become a common mechanism for the analysis and design of complex systems. Their ability to integrate continuous and discrete dynamics with intrinsic uncertainty enables SHS to naturally capture the variable temporal and spatial behavior often found in realistic systems.

Methods and numerical tools for reachability of continuous time deterministic systems have been well researched over the years (see Aubin (1991), Lygeros (2004) and Mitchell (2008)). In Gao, Lygeros, and Quincampoix (2007), the authors considered a reachability problem for uncertain hybrid systems in which the problem is formulated as a dynamic game and solved using methods from nonsmooth analysis. In the deterministic setting, reach-avoid problems have also been explored in Mitchell and Tomlin (2000) and Tomlin, Lygeros, and Sastry (2000), where the solution to these pursuit-evasion games is characterized through the level sets of an appropriate value function using dynamic programming (Mitchell & Tomlin, 2000).

For continuous stochastic hybrid systems, early contributions include the works of Davis (1993), Ghosh, Arapostathis, and Marcus (1997) and Hu, Lygeros, and Sastry (2000), with Bujorianu and Lygeros (2003) establishing a theoretical foundation for the measurability of events for reachability problems. In Hu, Prandini, and Sastry (2003, 2005), the authors address the reachability problem using a Markov chain approximation Kushner and Dupuis (1992), and apply the results to air traffic control studies. Recently, Koutsoukos and Riley (2006, 2008) have investigated a probabilistic reach while avoid problem similar to the problem we consider here, albeit for the verification of continuous time SHS without explicit mathematical consideration of the existence of optimal control policies.

Consideration of discrete time stochastic hybrid systems has recently attracted attention, partly because technical issues such as measurability are easier to resolve in this case. Here, we are interested in controlled discrete time stochastic hybrid systems (DTSHS) as defined in Abate, Prandini, Lygeros, and Sastry (2008). Probabilistic reachability for the safety of DTSHS was addressed in Abate et al. (2008) based on a theoretical foundation for the solution of stochastic optimal control problems of general discrete time systems of Bertsekas and Shreve (2007).

☆ This research was partially supported by the Swiss National Science Foundation, grant 200021-122072, and the ETH internal grant 0-20587-09. The material in this paper was partially presented at the IFAC Conference on Analysis and Design of Hybrid Systems, September 16–18, 2009, Zaragoza, Spain. This paper was recommended for publication in revised form by Associate Editor George Yin under the direction of Editor Ian R. Petersen.

* Corresponding author. Tel.: +41 076 232 7728; fax: +41 44 632 12 11.
E-mail addresses: summers@control.ee.ethz.ch, seansummers@hotmail.com (S. Summers), lygeros@control.ee.ethz.ch (J. Lygeros).

0005-1098/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2010.08.006

In this paper we extend recent results in the area of probabilistic safety (Abate et al., 2008) of DTSHS to a more general class of reachability problems. We consider a probabilistic reach while avoid problem where the objective is to maximize or minimize the probability that a system starting at a specific initial condition will hit a target set while avoiding an unsafe set over a finite time horizon. We consider two interpretations of the finite time horizon reach-avoid problem. In the first interpretation, which we will refer to as the first hitting time reach-avoid problem, we consider the case where the system hits the target set at some time during a finite time horizon, while avoiding the unsafe set up to the hitting time. Following the methods of Abate et al. (2008), we formulate the problem as a finite horizon stochastic optimal control problem with a max-multiplicative (or equivalently, sum-multiplicative) cost-to-go function. Dynamic programming can then be used to compute the Markov control policy that maximizes or minimizes the cost. We show that the dynamic program developed here can easily be adapted to solve the probabilistic safety problem of Abate et al. (2008) as well as the stochastic target hitting time problem of Boda, Filar, Lin, and Spanjers (2004), extending the breadth of problems which can be solved by these methods. Finally, we study the reach-avoid problem in the infinite time horizon setting, and argue that the problem remains relevant as the horizon is taken to infinity, as opposed to the safety and target hitting problems.

In the second interpretation, which we will refer to as the terminal hitting time reach-avoid problem, we consider the case where the system is inside the target set at the terminal time, while avoiding the unsafe set up to the terminal time. We again formulate the reach-avoid problem as a finite time horizon stochastic optimal control problem, although in this case we consider a multiplicative cost-to-go function (Abate et al., 2008). Dynamic programming can again be used to compute the Markov control policy that maximizes the cost. Unlike the first interpretation, the infinite horizon problem is not relevant in this case.

To exhibit the value of the proposed methodology, we demonstrate its effectiveness on two applied problems. The first is a fishery management problem. A bioeconomic model for a single species fishery is proposed and optimal management policies pertaining to a two-patch fishing area are explored. The second problem is an insurance company ruin problem. A discrete time Markov model for the company profit is coupled with a stochastic model of the overall market share, and policies regarding capital investment and advertising are explored. The problems are solved numerically using the computational approach introduced in Abate, Amin, Prandini, Lygeros, and Sastry (2007). While the examples refer to practical problems, we emphasize that the two numerical examples are meant to demonstrate the methods and not to reflect all subclasses of realistic applications.

The rest of the paper is arranged as follows. In Section 2 we briefly recall the DTSHS model of Abate et al. (2008). In Sections 3 and 4 we introduce the two notions of probabilistic reach while avoid over a finite time horizon. We also develop mechanisms to determine optimal Markov control policies based on dynamic programming for each case. In Section 5 we offer a comparison of the current reach-avoid formulations with the safety problem and the stochastic target hitting problem. In Section 6 we provide two numerical examples.

2. Stochastic hybrid system model

Here we recall the DTSHS model and associated semantics introduced in Abate et al. (2008). Throughout $\mathcal{B}(\cdot)$ denotes the Borel σ-algebra.

Definition 1 (*DTSHS*)**.** A discrete time stochastic hybrid system, $\mathcal{H} = (\mathcal{Q}, n, \mathcal{A}, T_v, T_q, R)$, comprises

- A discrete state space $\mathcal{Q} := \{q_1, q_2, \dots, q_m\}$, for some $m \in \mathbb{N}$;
- A map $n : \mathcal{Q} \to \mathbb{N}$ which assigns to each discrete state value $q \in \mathcal{Q}$ the dimension of the continuous state space $\mathbb{R}^{n(q)}$. The hybrid state space is then given by $X := \bigcup_{q\in\mathcal{Q}} \{q\} \times \mathbb{R}^{n(q)}$;
- A compact Borel space $\mathcal{A}$ representing the control space;
- A Borel-measurable stochastic kernel on $\mathbb{R}^{n(\cdot)}$ given $X \times \mathcal{A}$, $T_v : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times X \times \mathcal{A} \to [0, 1]$, which assigns to each $x = (q, v) \in X$ and $a \in \mathcal{A}$ a probability measure $T_v(\cdot|x, a)$ on the Borel space $(\mathbb{R}^{n(q)}, \mathcal{B}(\mathbb{R}^{n(q)}))$;
- A discrete stochastic kernel on $\mathcal{Q}$ given $X \times \mathcal{A}$, $T_q : \mathcal{Q} \times X \times \mathcal{A} \to [0, 1]$, which assigns to each $x \in X$ and $a \in \mathcal{A}$ a probability distribution $T_q(\cdot|x, a)$ over $\mathcal{Q}$;
- A Borel-measurable stochastic kernel on $\mathbb{R}^{n(\cdot)}$ given $X \times \mathcal{A} \times \mathcal{Q}$, $R : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times X \times \mathcal{A} \times \mathcal{Q} \to [0, 1]$, which assigns to each $x \in X$, $a \in \mathcal{A}$, and $q' \in \mathcal{Q}$ a probability measure $R(\cdot|x, a, q')$ on the Borel space $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$.

Consider the DTSHS evolving over the finite time horizon $k = 0, 1, \dots, N$ with $N \in \mathbb{N}$. We specify the initial state as $x_0 \in X$ at time $k = 0$, and define the notion of a Markov policy.

Definition 2 (*Markov Policy*)**.** A Markov Policy for a DTSHS, $\mathcal{H}$, is a sequence $\mu = (\mu_0, \mu_1, \dots, \mu_{N-1})$ of universally measurable maps $\mu_k : X \to \mathcal{A}$, $k = 0, 1, \dots, N-1$. The set of all admissible Markov policies is denoted by $\mathcal{M}_m$.

Let $\tau_v : \mathcal{B}(\mathbb{R}^{n(\cdot)}) \times X \times \mathcal{A} \times \mathcal{Q} \to [0, 1]$ be a stochastic kernel on $\mathbb{R}^{n(\cdot)}$ given $X \times \mathcal{A} \times \mathcal{Q}$, which assigns to each $x \in X$, $a \in \mathcal{A}$, and $q' \in \mathcal{Q}$, a probability measure on the Borel space $(\mathbb{R}^{n(q')}, \mathcal{B}(\mathbb{R}^{n(q')}))$ given by

$$\tau_v(\mathrm{d}v'|(q, v), a, q') = \begin{cases} T_v(\mathrm{d}v'|(q, v), a), & \text{if } q' = q \\ R(\mathrm{d}v'|(q, v), a, q'), & \text{if } q' \neq q. \end{cases}$$

Based on τ_v we introduce the kernel $Q : \mathcal{B}(X) \times X \times \mathcal{A} \to [0, 1]$:

$$Q(\mathrm{d}x'|x, a) = \tau_v(\mathrm{d}v'|x, a, q')T_q(q'|x, a).$$

Definition 3 (*DTSHS Execution*)**.** Consider the DTSHS, $\mathcal{H}$, and time horizon $N \in \mathbb{N}$. A stochastic process $\{x_k, k = 0, \dots, N\}$ with values in X is an execution of $\mathcal{H}$ associated with a Markov policy $\mu \in \mathcal{M}_m$ and an initial condition $x_0 \in X$ if and only if its sample paths are obtained according to the DTSHS Algorithm.

Algorithm 1 DTSHS Algorithm

Require: Sample Path $\{x_k, k = 0, \dots, N\}$
Ensure: Initial hybrid state $x_0 \in X$ at time $k = 0$, and Markov control policy $\mu = (\mu_0, \mu_1, \dots, \mu_{N-1}) \in \mathcal{M}_m$

```
1: while k < N do
2:   Set a_k = μ_k(x_k)
3:   Extract from X a value x_{k+1} according to Q(·|x_k, a_k)
4:   Increment k
5: end while
```

Equivalently, the DTSHS $\mathcal{H}$ can be described as a Markov control process with state space X, control space $\mathcal{A}$, and controlled transition probability function Q. Further, given a specific control policy $\mu \in \mathcal{M}_m$ and initial state $x_0 \in X$, the execution $\{x(k), k = 0, \dots, N\}$ is a time inhomogeneous stochastic process defined on the canonical sample space $\Omega = X^{N+1}$, endowed with its product σ-algebra $\mathcal{B}(\Omega)$. The probability measure $P^{\mu}_{x_0}$ is uniquely defined by the transition kernel Q, the Markov policy $\mu \in \mathcal{M}_m$, and the initial condition $x_0 \in X$ (see Bertsekas and Shreve (2007)). From now on, we will use interchangeably the notation $Q(\cdot|x, \mu_k(x))$ and $Q^{\mu_k}_x(\cdot|x)$ to represent the one-step transition kernel.

3. Finite horizon reach while avoiding — The first hitting time case

Let $K, K' \in \mathcal{B}(X)$, with $K \subseteq K'$. We define the stopping time associated with hitting K as $\tau := \inf\{j \geq 0 | x_j \in K\}$, and the stopping time associated with hitting $X \setminus K'$ as $\tau' := \inf\{j \geq 0 | x_j \in X \setminus K'\}$; if a set is empty we set its infimum equal to $+\infty$. Our goal is to evaluate the probability that the execution of the Markov control process associated with the Markov policy $\mu \in \mathcal{M}_m$ and the initial condition x_0 will hit K before hitting $X \setminus K'$ during the time horizon N. The probability that the system initialized at $x_0 \in X$ with control policy $\mu \in \mathcal{M}_m$ reaches K while avoiding $X \setminus K'$ is given by

$$\begin{aligned} r_{x_0}^{\mu}(K, K') &:= P_{x_0}^{\mu}\{\exists j \in [0, N] : x_j \in K \wedge \\ &\qquad \forall i \in [0, j-1] x_i \in K' \setminus K\}, \\ &= P_{x_0}^{\mu}\{\{\tau < \tau'\} \wedge \{\tau \leq N\}\}, \end{aligned} \tag{1}$$

where $\wedge$ denotes the logical AND, and we operate under the assumption that the requirement on i is automatically satisfied when $x_0 \in K$; subsequently we will use the same convention for products, i.e. $\prod_{i=k}^{j}(\cdot) = 1$ if $k > j$.

3.1. Cost and dynamic recursion

We can formulate $r_{x_0}^{\mu}(K, K')$ according to a max-multiplicative (and equivalently sum-multiplicative) cost function by using indicator functions. In this respect, we exploit the fact that the indicator function serves as a natural mechanism for logical arguments (i.e. either an argument is true or it is false). For some set $D \subseteq X$, let $\mathbf{1}_D(\cdot) : X \to \{0, 1\}$ denote the indicator function. Consider now

$$\max_{j \in [0,N]} \left(\prod_{i=0}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j) = \begin{cases} 1, & \text{if } \exists j \in [0, N] : x_j \in K \wedge \forall i \in [0, j-1] x_i \in K' \setminus K \\ 0, & \text{otherwise} \end{cases}$$

where $x_j \in X, j \in [0, N]$. It is easy to show that we may equivalently express the max-multiplicative cost as a sum-multiplicative cost

$$\max_{j \in [0,N]} \left(\prod_{i=0}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j) = \sum_{j=0}^{N} \left(\prod_{i=0}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j).$$

Hence $r_{x_0}^{\mu}(K, K')$ can be expressed as the expectation

$$r_{x_0}^{\mu}(K, K') = E_{x_0}^{\mu}\left[\sum_{j=0}^{N} \left(\prod_{i=0}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j) \right].$$

This is a significant difference between the formulation here and that of Abate et al. (2008). In Abate et al. (2008) a max-cost is introduced to quantify a logical OR formulation of possibly dependent events. In our formulation the logical OR involves disjoint events. Therefore the max-cost formulation can be equivalently represented through a sum-cost, which is in turn equivalent to the cost associated with the hitting times in (1). Thus, we can solve a hitting time problem without explicitly defining the hitting times in the cost function or dynamic program.

For a fixed Markov policy $\mu \in \mathcal{M}_m$, let us define the functions $V_k^{\mu} : X \to [0, 1], k = 0, \ldots, N$ as

$$V_N^{\mu}(x) = \mathbf{1}_K(x), \tag{2}$$

$$\begin{aligned} V_k^{\mu}(x) \;=\; &\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) \int_{X^{N-k}} \sum_{j=k+1}^{N} \left(\prod_{i=k+1}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \\ &\times \mathbf{1}_K(x_j) \prod_{j=k+1}^{N-1} Q^{\mu_j}(\mathrm{d}x_{j+1}|x_j) Q^{\mu_k}(\mathrm{d}x_{k+1}|x). \end{aligned} \tag{3}$$

Note that

$$\begin{aligned} V_0^{\mu}(x_0) &= E_{x_0}^{\mu}\left[\sum_{j=0}^{N} \left(\prod_{i=0}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j) \right] \\ &= r_{x_0}^{\mu}(K, K'). \end{aligned}$$

Let $\mathcal{F}$ denote the set of functions from X to $\mathbb{R}$ and define the operator $H : X \times \mathcal{A} \times \mathcal{F} \to \mathbb{R}$ as

$$H(x, a, Z) := \int_X Z(y) Q(\mathrm{d}y|x, a). \tag{4}$$

The following lemma shows that $r_{x_0}^{\mu}(K, K')$ can be computed via a backwards recursion.

Lemma 4. *Fix a Markov policy $\mu = (\mu_0, \mu_1, \ldots) \in \mathcal{M}_m$. The functions $V_k^{\mu} : X \to [0, 1], k = 0, 1, \ldots, N-1$ can be computed by the backward recursion:*

$$V_k^{\mu}(x) = \mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, \mu_k(x), V_{k+1}^{\mu}), \tag{5}$$

initialized with $V_N^{\mu}(x) = \mathbf{1}_K(x), x \in X$.

Proof. By induction. First, due to the definition of (2) and (3), we have that

$$V_{N-1}^{\mu}(x) = \mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) \int_X V_N^{\mu}(x_N) Q^{\mu_{N-1}}(\mathrm{d}x_N|x),$$

so that (5) is proven for $k = N - 1$. For $k < N - 1$ we can separate the terms associated with x_{k+1} as follows

$$\begin{aligned} V_k^{\mu}(x) \;=\; &\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) \int_X \Bigg(\mathbf{1}_K(x_{k+1}) + \mathbf{1}_{K' \setminus K}(x_{k+1}) \\ &\times \int_{X^{N-k-1}} \sum_{j=k+2}^{N} \left(\prod_{i=k+2}^{j-1} \mathbf{1}_{K' \setminus K}(x_i) \right) \mathbf{1}_K(x_j) \\ &\times \prod_{j=k+2}^{N-1} Q^{\mu_j}(\mathrm{d}x_{j+1}|x_j) Q^{\mu_{k+1}}(\mathrm{d}x_{k+2}|x_{k+1}) \Bigg) \\ &\times Q^{\mu_k}(\mathrm{d}x_{k+1}|x) \\ =\; &\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) \int_X V_{k+1}^{\mu}(x_{k+1}) Q^{\mu_k}(\mathrm{d}x_{k+1}|x) \end{aligned}$$

which concludes the proof. □

3.2. The dynamic programming algorithm

Definition 5 (*Maximal in the First Sense*)**.** Let $\mathcal{H}$ be a Markov control process, $K \in \mathcal{B}(X)$, $K' \in \mathcal{B}(X)$, with $K \subseteq K'$. A Markov policy μ^* is a maximal reach-avoid policy in the first sense if and only if $r_{x_0}^{\mu^*}(K, K') = \sup_{\mu \in \mathcal{M}_m} r_{x_0}^{\mu}(K, K')$, for all $x_0 \in X$.

Theorem 6. *Define $V_k^* : X \to [0, 1], k = 0, 1, \ldots, N$, by the backward recursion:*

$$V_k^*(x) = \sup_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V_{k+1}^*)\} \tag{6}$$

$x \in X$, initialized with $V_N^*(x) = \mathbf{1}_K(x), x \in X$. Then, $V_0^*(x_0) = \sup_{\mu \in \mathcal{M}_m} r_{x_0}^{\mu}(K, K'), x_0 \in X$. If $\mu_k^* : X \to \mathcal{A}, k \in [0, N-1]$, is such that for all $x \in X$

$$\mu_k^*(x) = \arg\sup_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V_{k+1}^*)\} \tag{7}$$

then $\mu^* = (\mu_0^*, \mu_1^*, \ldots, \mu_{N-1}^*)$ is a maximal reach-avoid Markov policy in the first sense. A sufficient condition for the existence of μ^* is that $U_k(x, \lambda) = \{a \in \mathcal{A} | H(x, a, V_{k+1}^*) \geq \lambda\}$ is compact for all $x \in X, \lambda \in \mathbb{R}, k \in [0, N-1]$.

Note that Theorem 6 gives a sufficient condition for the existence of an optimal nonrandomized Markov policy. While the consideration of randomized Markov policies is indeed interesting in the event that an optimal nonrandomized Markov policy does not exist, in most cases the "best" policy can be taken to be nonrandomized (Bertsekas & Shreve, 2007). In light of this fact and in the interest of space, we do not consider randomized Markov policies in the present work and urge the interested reader to consider Chapter 8 in Bertsekas and Shreve (2007) for additional details. A proof of Theorem 6 is provided in Appendix. Taking the infimum over the reach-avoid probability (2) may also be useful in some cases. For completeness we state without proof the following.

Definition 7 (*Minimal in the First Sense*)**.** Let $\mathcal{H}$ be a Markov control process, $K \in \mathcal{B}(X)$, $K' \in \mathcal{B}(X)$, with $K \subseteq K'$. A Markov policy μ^- is a minimal reach-avoid policy in the first sense if $r_{x_0}^{\mu^-}(K, K') = \inf_{\mu \in \mathcal{M}_m} r_{x_0}^{\mu}(K, K')$, for all $x_0 \in K' \setminus K$.

Theorem 8. *Define $V_k^- : X \to [0, 1], k = 0, 1, \ldots, N$, by the backward recursion:*

$$V_k^-(x) = \inf_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V_{k+1}^-)\} \tag{8}$$

$x \in X$, initialized with $V_N^-(x) = \mathbf{1}_K(x), x \in X$. Then, $V_0^-(x_0) = \inf_{\mu \in \mathcal{M}_m} r_{x_0}^{\mu}(K, K'), x_0 \in X$. If $\mu_k^- : X \to \mathcal{A}, k \in [0, N-1]$, is such that $\forall x \in X$

$$\mu_k^-(x) = \arg \inf_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V_{k+1}^-)\} \tag{9}$$

then $\mu^- = (\mu_0^-, \mu_1^-, \ldots, \mu_{N-1}^-)$ is a minimal reach-avoid Markov policy in the first sense. A sufficient condition for the existence of μ^- is that $U_k(x, \lambda) = \{a \in \mathcal{A} | H(x, a, V_{k+1}^-) \leq \lambda\}$ is compact for all $x \in X, \lambda \in \mathbb{R}, k \in [0, N-1]$.

3.3. Extension to the infinite time horizon

By Theorems 6 and 8, and the fact that the value function is monotonically increasing and bounded above by one, we have that

$$0 \leq V_N^*(x) \leq V_{N-1}^*(x) \leq \cdots \leq V_0^*(x) \leq 1,$$

for any time horizon N and for every $x \in X$. Also, by the results in Appendix and Propositions 7.45 and 7.46 in Bertsekas and Shreve (2007), we have that

$$V_0^*(x) \to V^*(x) \quad \text{as } N \to \infty,$$

for every $x \in X$, where $V^*(x)$ is the infinite horizon optimal cost at x (for details see Chapter 9 in Bertsekas and Shreve (2007)). Thus, by the Monotone Convergence Theorem (Rudin, 1987), V^* is universally measurable and satisfies the Bellman equation

$$V^*(x) = \sup_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V^*)\}$$

for the Maximal Reach-Avoid Infinite Horizon problem. A similar discussion for the Minimal Reach-Avoid problem shows that the infinite horizon value function is universally measurable and satisfies

$$V^-(x) = \inf_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V^-)\}.$$

The resulting Markov control policies (if they exist) are stationary and deterministic and satisfy the optimality conditions

$$\mu^*(x) = \arg \sup_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V^*)\}$$

$$\mu^-(x) = \arg \inf_{a \in \mathcal{A}} \{\mathbf{1}_K(x) + \mathbf{1}_{K' \setminus K}(x) H(x, a, V^-)\}.$$

4. Finite Horizon reach while avoiding — The terminal hitting time case

As before let $K, K' \in \mathcal{B}(X)$ and $\tau' = \inf\{j \geq 0 | x_j \in X \setminus K'\}$. Our goal is to evaluate the probability that the execution of the Markov control process associated with the Markov policy $\mu \in \mathcal{M}_m$ and the initial condition x_0 will be inside K at the terminal time N, while avoiding the unsafe set $X \setminus K'$ for the duration of the time horizon $[0, N]$. We note that the restriction that $K \subseteq K'$ is not necessary in this case. The probability that the system initialized at $x_0 \in X$ with control policy $\mu \in \mathcal{M}_m$ reaches K at terminal time N while avoiding $X \setminus K'$ is given by

$$\begin{aligned} \hat{r}_{x_0}^{\mu}(K, K') &= P_{x_0}^{\mu}\{x_N \in K \wedge \forall i \in [0, N-1] x_i \in K'\}, \\ &= P_{x_0}^{\mu}\{\{N < \tau'\} \wedge \{x_N \in K\}\}. \end{aligned}$$

We can formulate $\hat{r}_{x_0}^{\mu}(K, K')$ according to a multiplicative cost function by using indicator functions. Consider

$$\left(\prod_{i=0}^{N-1} \mathbf{1}_{K'}(x_i)\right) \mathbf{1}_K(x_N) = \begin{cases} 1, & \text{if } x_N \in K \wedge \forall i \in [0, N-1] x_i \in K' \\ 0, & \text{otherwise} \end{cases}$$

where $x_j \in X, j \in [0, N]$. $\hat{r}_{x_0}^{\mu}(K, K')$ can now be expressed as the expectation

$$\hat{r}_{x_0}^{\mu}(K, K') = E_{x_0}^{\mu}\left[\left(\prod_{i=0}^{N-1} \mathbf{1}_{K'}(x_i)\right) \mathbf{1}_K(x_N)\right]. \tag{10}$$

Based on the formulation of the multiplicative cost in (10) and a fixed Markov policy $\mu \in \mathcal{M}_m$, let us define the functions $\hat{V}_k^{\mu} : X \to [0, 1], k = 0, 1, \ldots, N$ as

$$\hat{V}_N^{\mu}(x) = \mathbf{1}_K(x),$$

$$\begin{aligned} \hat{V}_k^{\mu}(x) \;=\; &\mathbf{1}_{K'}(x) \int_{X^{N-k}} \left(\prod_{i=k+1}^{N-1} \mathbf{1}_{K'}(x_i)\right) \\ &\times \mathbf{1}_K(x_N) \prod_{j=k+1}^{N-1} Q^{\mu_j}(\mathrm{d}x_{j+1}|x_j) Q^{\mu_k}(\mathrm{d}x_{k+1}|x). \end{aligned}$$

It can be seen that

$$\hat{V}_0^{\mu}(x_0) = E_{x_0}^{\mu}\left[\left(\prod_{i=0}^{N-1} \mathbf{1}_{K'}(x_i)\right) \mathbf{1}_K(x_N)\right] = \hat{r}_{x_0}^{\mu}(K, K').$$

Let $\mathcal{F}$ denote the set of functions from X to $\mathbb{R}$ and define the operator $H : X \times \mathcal{A} \times \mathcal{F} \to \mathbb{R}$ as in (4). The following lemma shows that $\hat{r}_{x_0}^{\mu}(K, K')$ can be computed via a backwards recursion; the proof is similar to that of Lemma 4.

Lemma 9. *Fix a Markov policy $\mu = (\mu_0, \mu_1, \ldots) \in \mathcal{M}_m$. The functions $\hat{V}_k^{\mu} : X \to [0, 1], k = 0, 1, \ldots, N-1$ can be computed by the backward recursion:*

$$\hat{V}_k^{\mu}(x) = \mathbf{1}_{K'}(x) H(x, \mu_k(x), \hat{V}_{k+1}^{\mu}),$$

initialized with $\hat{V}_N^{\mu}(x) = \mathbf{1}_K(x), x \in X$.

Definition 10 (*Maximal in the Terminal Sense*)**.** Let $\mathcal{H}$ be a Markov control process, $K \in \mathcal{B}(X)$, and $K' \in \mathcal{B}(X)$. A Markov policy μ^* is a maximal reach-avoid policy in the terminal sense if and only if $\hat{r}_{x_0}^{\mu^*}(K, K') = \sup_{\mu \in \mathcal{M}_m} \hat{r}_{x_0}^{\mu}(K, K')$, for all $x_0 \in X$.

As it turns out, apart from the terminal value function, the dynamic recursion for the current reach-avoid problem is equivalent to the recursion introduced for the safety problem in Abate et al. (2008). This should come as no surprise given that both problems rely on a multiplicative cost. By Theorem 1 in Abate et al. (2008) we therefore have the following.

Theorem 11. *Define $\hat{V}_k^* : X \to [0, 1], k = 0, 1, \ldots, N$, by the backward recursion:*

$$\hat{V}_k^*(x) = \sup_{a \in \mathcal{A}} \{\mathbf{1}_{K'}(x) H(x, a, \hat{V}_{k+1}^*)\} \tag{11}$$

$x \in X$, initialized with $\hat{V}_N^*(x) = \mathbf{1}_K(x), x \in X$. Then, $\hat{V}_0^*(x_0) = \sup_{\mu\in\mathcal{M}_m} \hat{r}_{x_0}^{\mu}(K, K'), x_0 \in X$. If $\hat{\mu}_k^* : X \to \mathcal{A}, k \in [0, N-1]$, is such that $\forall x \in X$

$$\hat{\mu}_k^*(x) = \arg\sup_{a\in\mathcal{A}}\{\mathbf{1}_{K'}(x)H(x, a, \hat{V}_{k+1}^*)\} \tag{12}$$

then $\hat{\mu}^* = (\hat{\mu}_0^*, \hat{\mu}_1^*, \ldots, \hat{\mu}_{N-1}^*)$ is a maximal reach-avoid Markov policy in the terminal sense. A sufficient condition for the existence of $\hat{\mu}^*$ is that $U_k(x, \lambda) = \{a \in \mathcal{A} | H(x, a, \hat{V}_{k+1}^*) \geq \lambda\}$ is compact for all $x \in X, \lambda \in \mathbb{R}, k \in [0, N-1]$.

Minimal reach avoid policies can also be defined in this context. The discussion is omitted in the interest of space.

5. Discussion

5.1. The safety problem

Consider a safe set $A \in \mathcal{B}(X)$ and an unsafe set $X \setminus A$. In Abate et al. (2008), the authors consider the probability that a DTSHS with Markov policy $\mu \in \mathcal{M}_m$ remains inside A for a finite horizon $[0, N]$, and denote this probability $p_{x_0}^{\mu}(A)$. Specifically, they consider two dynamic programs, one based on a multiplicative cost-to-go function and the other based on a max cost-to-go function, which are shown to solve the same problem. Considering the max-cost approach, for a fixed Markov policy $\mu \in \mathcal{M}_m$ the associated dynamic recursion is given as

$$W_k^{\mu}(x) = \mathbf{1}_{X\setminus A}(x) + \mathbf{1}_A(x)H(x, \mu_k(x), W_{k+1}^{\mu}), \tag{13}$$

where $x \in X, k = 0, 1, \ldots, N$ and $W_N(x) = \mathbf{1}_A(x)$. It is shown that the initial value function satisfies

$$W_0^{\mu}(x_0) = E_{x_0}^{\mu}\left[\max_{k\in[0,N]} \mathbf{1}_{X\setminus A}(x_k)\right] = 1 - p_{x_0}^{\mu}(A),$$

such that minimizing $W_0^{\mu}(x)$ is equivalent to maximizing $p_{x_0}^{\mu}(A)$ and therefore solves the safety problem. In the context of the reach-avoid problem, it is clear that if we set $K = X \setminus A$ and $K' = X$, then (5) and (13) are equivalent. Thus, defining the safe set and unsafe set as above, Theorem 8 solves the safety problem of Abate et al. (2008). Likewise, we have the trivial result that taking $K' = K = A$ and applying Theorem 11 also solves the safety problem of Abate et al. (2008).

5.2. The stochastic hitting time problem

Consider a target set $B \in \mathcal{B}(X)$ and its complement $X \setminus B$. In Boda et al. (2004), the authors consider the problem of early retirement, where the system process is modeled as a discrete-time Markov decision process (MDP) with a finite state space. Note that this falls under the umbrella of DTSHS. They pose the problem as a target hitting time problem, and consider Markov policies which maximize the probability $P_{x_0}^{\mu}\{\tau \leq N\}$, i.e. the probability that the MDP with Markov policy $\mu \in \mathcal{M}_m$ enters the target set at some point before or including the terminal time N, where τ is defined as the first target hitting time. In order to apply standard techniques, they consider an equivalent problem which minimizes the probability of staying outside the target set, i.e. they minimize the safety of the system outside the target set. For a fixed Markov policy $\mu \in \mathcal{M}_m$, they consider a multiplicative cost as in Abate et al. (2008) with dynamic recursion (reconfigured to match the uncountable state space considered in the current paper) given by

$$G_k^{\mu}(x) = \mathbf{1}_{X\setminus B}(x)H(x, \mu_k(x), G_{k+1}^{\mu}(x)),$$

where $x \in X, k = 0, 1, \ldots, N$, and $G_N^{\mu}(x) = \mathbf{1}_{X\setminus B}(x)$. Note that this is the exact dynamic recursion that was introduced for the multiplicative cost in Abate et al. (2008). Thus, the stochastic hitting problem Boda et al. (2004) can be solved by setting $K' = X, K = B$, and applying Theorem 6.

5.3. On the infinite horizon

Under weak assumptions on the stochastic kernel, the probability of staying within the safe set goes to 0 in both the safety problem and the terminal hitting time reach-avoid problem. Likewise, the probability of hitting the target set goes to 1 in the target hitting problem. For example, when considering general (unbounded) probability distributions, there is always a nonzero probability of jumping to the unsafe (or target) set regardless of location in the state space. Thus, it is only a matter of time before the system enters the safe set (or target set) and the process is stopped. This is not necessarily the case in the first hitting time reach-avoid formulation, and therefore an infinite time representation makes sense.

Considering the first hitting time reach-avoid formulation for the infinite horizon case, it is clear that the goal of the resulting DP is to choose a control policy which minimizes or maximizes the probability of hitting K before hitting $X \setminus K'$, i.e.

$$r_{x_0}^{\mu}(K, K') = P_{x_0}^{\mu}\{\tau < \tau'\},$$

regardless of when the hitting time occurs. Under weak assumptions on the stochastic kernel, the probability of leaving $K' \setminus K$ goes to 1, which is equivalent to $P_{x_0}^{\mu}\{\tau < \tau'\} + P_{x_0}^{\mu}\{\tau' < \tau\} = 1$, and therefore $r_{x_0}^{\mu}(K, K')$ may take a value between 0 and 1 which has meaning in the infinite horizon. We note that it is because of this fact that the infinite horizon case makes sense here, as opposed to the safety problem Abate et al. (2008), the target hitting problem Boda et al. (2004), or even the terminal hitting time reach-avoid formulation introduced in Section 4.

6. Computational examples

6.1. Fishery management problem

Over-exploitation of fisheries may lead to a decrease in the fish stock to a level below that which would support maximum sustainable yield (MSY) and/or to a level where net revenue has been driven to zero (Clark, 1990). In order to avoid over-exploitation, fishery management policy must take into consideration both the survival of the fish species and the economic well-being of the fishery. To this end, bioeconomic models have become an important tool in fishery management Chae and Pascoe (2005) and Sethi, Costello, Fisher, Hanemann, and Karp (2005).

Historically, catch quotas based on the analysis of deterministic models and the MSY have been used to influence the interaction between fish species and the fishery. However, with the recent collapse of several important fisheries, there has been an increased level of research into alternate management schemes at both the fishery and government level (Sainsbury, Punt, & Smith, 2000). Examples include advanced quota control methods, such as the harvest control rule (HCR), as well as policies involving marine reserves, or marine protected areas (MPA) (Conrad, 1999; Costello & Polasky, 2008; Lauck, Clark, Mangel, & Munro, 1998; White, Kendall, Gaines, Siegel, & Costello, 2008). Additionally, researchers are exploring bioeconomic models with stochastic effects (Brasao, 2000; Nostbakken, 2008) in an effort to capture fishery dynamics more accurately.

Here, we consider a discrete time stochastic hybrid bioeconomic model of a single species fishery where the goal is to obtain a management policy which maximizes the probability of surpassing a target average yield (i.e. fish biomass caught each time unit) while maintaining a safe overall fish biomass at the end of the finite time horizon. We consider a total fishable area comprising two smaller, equally sized patches. For a time horizon $k = 0, 1, \ldots, N$, the evolution of the fish biomass and the fishery surplus is given by the stochastic difference equations (Pitchford, Codling, & Psarra, 2007)

$$\begin{aligned} x_1(k+1) &= (1-\nu_1(k))x_1(k) + \gamma_1(k)R(x_1(k)) \\ &\quad - C_1(x_1(k)+x_2(k), d_1(k)) \\ &\quad - \lambda(k)M(x_1(k), x_2(k), d_1(k), d_2(k)), \end{aligned} \tag{14}$$

$$\begin{aligned} x_2(k+1) &= (1-\nu_2(k))x_2(k) + \gamma_2(k)R(x_2(k)) \\ &\quad - C_2(x_1(k)+x_2(k), d_2(k)) \\ &\quad + \lambda(k)M(x_1(k), x_2(k), d_1(k), d_2(k)), \end{aligned} \tag{15}$$

$$\begin{aligned} s(k+1) &= s(k) + C_1(x_1(k)+x_2(k), d_1(k)) \\ &\quad + C_2(x_1(k)+x_2(k), d_2(k)), \end{aligned} \tag{16}$$

where x_1 and x_2 represent the fish biomass in each patch and s is the fishery surplus. $R(\cdot)$ and $M(\cdot)$ are functions representing recruitment and migration of fish, $C_1(\cdot)$ and $C_2(\cdot)$ are the catch functions for each patch, $\nu_i(k)$ is a random variable that represents natural fish mortality, $d_i(k) \in \left\{0, \frac{1}{2}, 1\right\}$ is an input variable representing the catch effort in each patch (if $d_i(k) = 0$, then patch i is a MPA), $\lambda(k)$ is a random variable representing the migration between the two patches, and $\gamma_i(k)$ is a random variable representing the variability in the recruitment of the fish population. The species escapement (i.e. the resulting adult population at the end of each time period) function for patch $i \in \{1, 2\}$ is

$$\begin{aligned} &E_i(x_1(k), x_2(k), d_i(k)) \\ &\quad = (1-\nu_i(k))x_i(k) - C_i(x_1(k)+x_2(k), d_i(k)), \end{aligned}$$

the species migration function is

$$\begin{aligned} &M(x_1(k), x_2(k), d_1(k), d_2(k)) \\ &\quad = E_1(x_1(k), x_2(k), d_1(k)) - E_2(x_1(k), x_2(k), d_2(k)), \end{aligned}$$

and the species recruitment function (i.e. single species logistic equation) is

$$R(x(k)) = \max\left\{ r(k)x(k)\left(1 - \frac{x(k)}{L}\right), 0 \right\},$$

where $r(k) \in [0, 1]$ is the per-capita recruitment at time step k and L is the biomass limit (i.e., an upper bound on the allowable fish population) for each patch. The catch function for each patch $i \in \{1, 2\}$ is defined according to a government imposed catch quota (for the entire fishable area) according to the HCR and is given by the function

$$C_i(x(k), d_i(k)) = \begin{cases} \delta_i(k)d_i(k)C\dfrac{x(k)}{L} & \text{if } x(k) < L \\ \delta_i(k)d_i(k)C & \text{otherwise,} \end{cases}$$

where $\delta_i(k)$ is a random variable representing the variability in the catch and C is the maximum target catch.

The resulting system ((14)–(16)) is a DTSHS consisting of three continuous states $\{x_1, x_2, s\} \in \mathbb{R}^3$, two modes of operation defined according to the HCR, and significant uncertainty. For each time step k, the fishery manager has the ability to assign the input variables $d_1(k)$ and $d_2(k)$, effectively controlling where to fish and how much to fish (in accordance with the HCR, $d_1(k) + d_2(k) \leq 1$). For the numerical implementation all parameters for the DTSHS were chosen according to Pitchford et al. (2007) and the units for time and biomass taken as arbitrary. For the constant variables, we assign the values $N = 10, L = 200, r = 1$, and $C = 70$. All random variables are assumed to be i.i.d. and occur according to the following distributions $\nu_i \sim \mathcal{N}(0.2, 0.1^2)$, $\lambda \sim \mathcal{N}(0.25, 0.1^2)$, $\gamma_i \sim \mathcal{N}(1, 0.6^2)$, and $\delta_i \sim \mathcal{N}(1.1, 0.2^2)$ for $i \in \{1, 2\}$. Also, as in Pitchford et al. (2007), given the generality of the system of equations, we take certain actions to guarantee that the population does not go above the biomass limit in each patch, recruitment is positive and bounded ($\nu_i(k) \in [0, 1]$), migration is positive and bounded ($\lambda(k) \in [0, 1]$), and catch is positive ($\delta_i(k) > 0$) and has an upward bias.

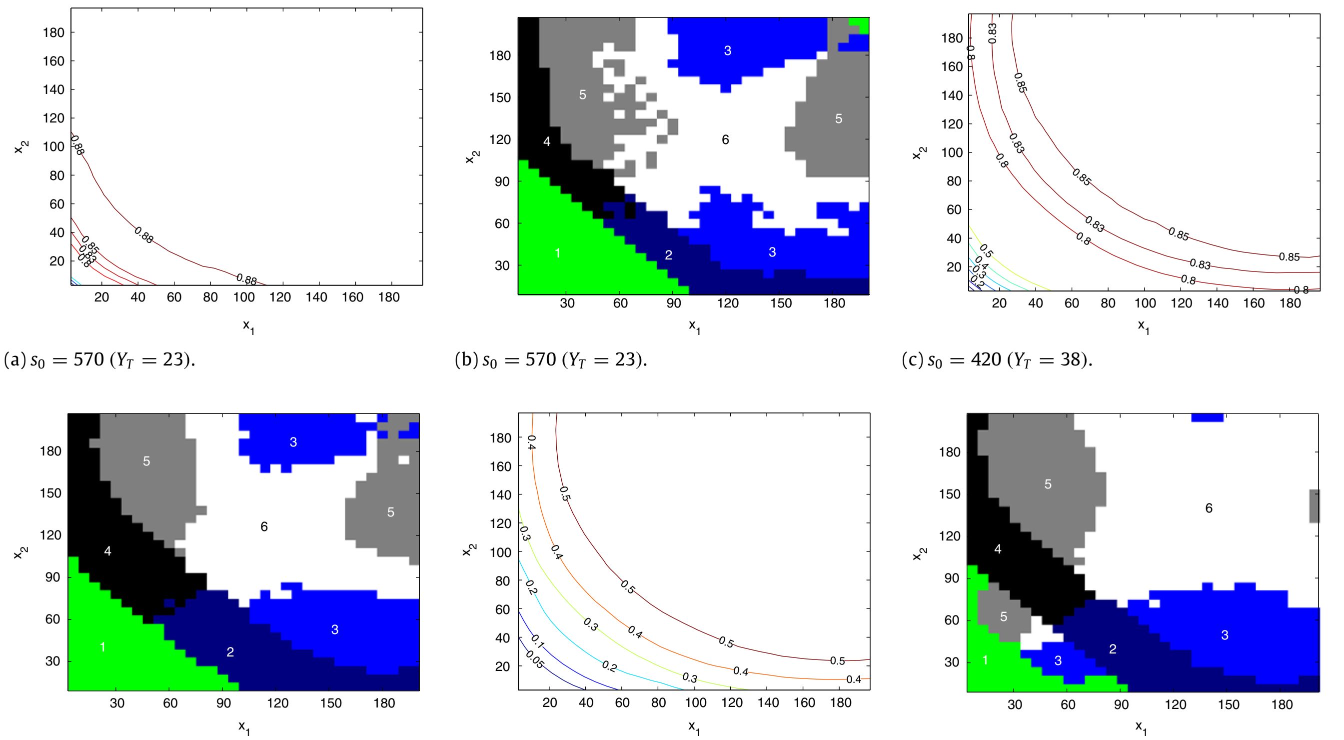
Consider a target operating region $K = [100, 200] \times [100, 200] \times [800, 1200]$ and a safe operating region $K' = (0, 200] \times (0, 200] \times (0, 1200]$. The fishery manager would like to maximize the probability that the state of the fishery is within the target operating region at the end of the finite time interval, while avoiding population ruin (escape from the safe region). To this end, we solve the terminal hitting time reach-avoid dynamic program. Given an initial surplus $s(0) = s_0$ at $k = 0$, the average yield in terms of fish biomass over a finite horizon is simply $Y_{\text{avg}} = (s(N) - s(0))/N$. Therefore, the probability of hitting a value greater than or equal to a target threshold s_T (in this case $s_T = 800$) at time N, i.e. $s(N) \geq s_T$, is equivalent to the probability of attaining an average yield greater than or equal to a target average yield according to the relationship $Y_{\text{avg}} \geq Y_T = (s_T - s_0)/N$. As a result of the relationship between s_0, s_T, and Y_T, the solution enables the fishery manager to evaluate the probability of success and the optimal management policy for any target yield between $Y_T = 0$ and $Y_T = 80$, and any initial fish population.

Computational results for the fishery management problem using the numerical methods proposed in Abate et al. (2007) are given in Fig. 1. All computations were performed on a $33 \times 33 \times 33$ grid in the continuous state space and the transition probabilities associated with each grid point were obtained via 30 000 Monte Carlo samples. We additionally discretized in the mode space and the control space, although this was trivial given that both are finite in this example. Note that, since numerical methods based on gridding suffer from the curse of dimensionality (i.e. the number of grid points grows exponentially with dimension), this problem is at the upper limit (in terms of dimensionality) of the problems we can currently solve. Level sets of the optimal value function and the corresponding optimal policy (regarding the MPA) at the initial time $k = 0$ are provided for select values of s_0 (equivalently Y_T). In this way, we are able to quantify the maximum probability of achieving an average biomass yield of $Y_{\text{avg}} \geq (800 - s_0)/10$ at time $N = 10$, with the additional constraint that the fish population $(x_1(N), x_2(N)) \in [100\times200]\times[100\times200]$. Specifically, in Fig. 1 we see results for an average biomass yield of $Y_T = 23$, $Y_T = 38$, and $Y_T = 52$ units of biomass respectively. As can be expected, an increase in the target average yield results in a decrease of the probability of success. However, it is interesting to note the significant difference between the drop in probability between $Y_T = 23$ and $Y_T = 38$, and the drop between $Y_T = 38$ and $Y_T = 52$, which emphasizes a nonlinear dependence of the success probability on the target average yield. Also of interest is the shape of the value function level sets, which suggests that a higher success probability may be achieved when the overall population is balanced between the two patches. A byproduct of this can be seen in the control policies for $Y_T = 23$ and $Y_T = 38$, where at high values of x_1 and $x_2 \in [100, 200]$ (approximately), the optimal policy is to fish x_2 at full effort. While this may seem counterintuitive, a deterministic analysis of the system of equations shows that this leads to an increased balance in the two patches, and therefore can lead to a higher probability of success over the finite horizon.

6.2. *Economic ruin model with investment income*

In mathematical finance, and specifically actuarial science and risk theory, the problem of ruin has a long history going back to Lundberg's classic model (Lundberg, 1903). In Lundberg's model the company did not earn any investment on its capital, however, mathematical advances over the years have led to the consideration of more difficult, and hopefully more realistic, models including those with investment return. Comprehensive surveys on ruin models with investment income can be found in Paulsen (1998) and Paulsen (2008).

Recently, increased attention has been given to stochastic risk models with investment income in the discrete time setting



(a) $s_0 = 570$ ($Y_T = 23$). (b) $s_0 = 570$ ($Y_T = 23$). (c) $s_0 = 420$ ($Y_T = 38$).

(d) $s_0 = 420$ ($Y_T = 38$). (e) $s_0 = 280$ ($Y_T = 52$). (f) $s_0 = 280$ ($Y_T = 52$).

Fig. 1. Results for the fishery management problem at time $k = 0$. Probability level sets of the optimal value function are given in (a), (c), and (e). Optimal policies for the catch and MPA are given in (b), (d), and (f), where 1-green, 2-dark blue, 3-light blue, 4-black, 5-gray, and 6-white represent $d = \{0, 0\}$, $d = \left\{\frac{1}{2}, 0\right\}$, $d = \{1, 0\}$, $d = \left\{0, \frac{1}{2}\right\}$, $d = \{0, 1\}$, and $d = \left\{\frac{1}{2}, \frac{1}{2}\right\}$, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Cai, 2002; Cai & Dickson, 2004; de Kok, 2003; Nyrhinen, 1999; Tang & Tsitsiashvili, 2003; Wei & Hu, 2008; Yang & Zhang, 2006). In most cases the probability of ruin over a finite or infinite time horizon is the main area of interest, with the infinite horizon case being mathematically easier and thus more popular in the literature (Paulsen, 2008). In other cases, authors have addressed the possibility of influencing the ruin probability e.g. by control of risky investments (Schael, 2005).

Motivated by de Kok (2003), we consider a discrete time stochastic hybrid insurance model of market and investment growth for a new insurance product over a finite time period. We assume that an insurance company would like to introduce a new product to the market, and has determined a target market share $Y_T \subseteq [0, 1]$ and a target reserve requirement of $X_T \subseteq \mathbb{R}$ such that profits are sustainable. Conservative estimates show that the reserve requirement is linearly dependent on the market share, and that the initial investment required is impossible due to immediate cash-flow problems. In order to obtain the target market share and reserve requirement, the company must raise capital through investment strategies and increase its market share through advertisement strategies. Based on a common extension of the Lundberg model (Paulsen, 2008), the evolution of the capital X_k and market share Y_k over a finite horizon $k = 0, 1, \ldots, N$ is given according to the stochastic difference equation

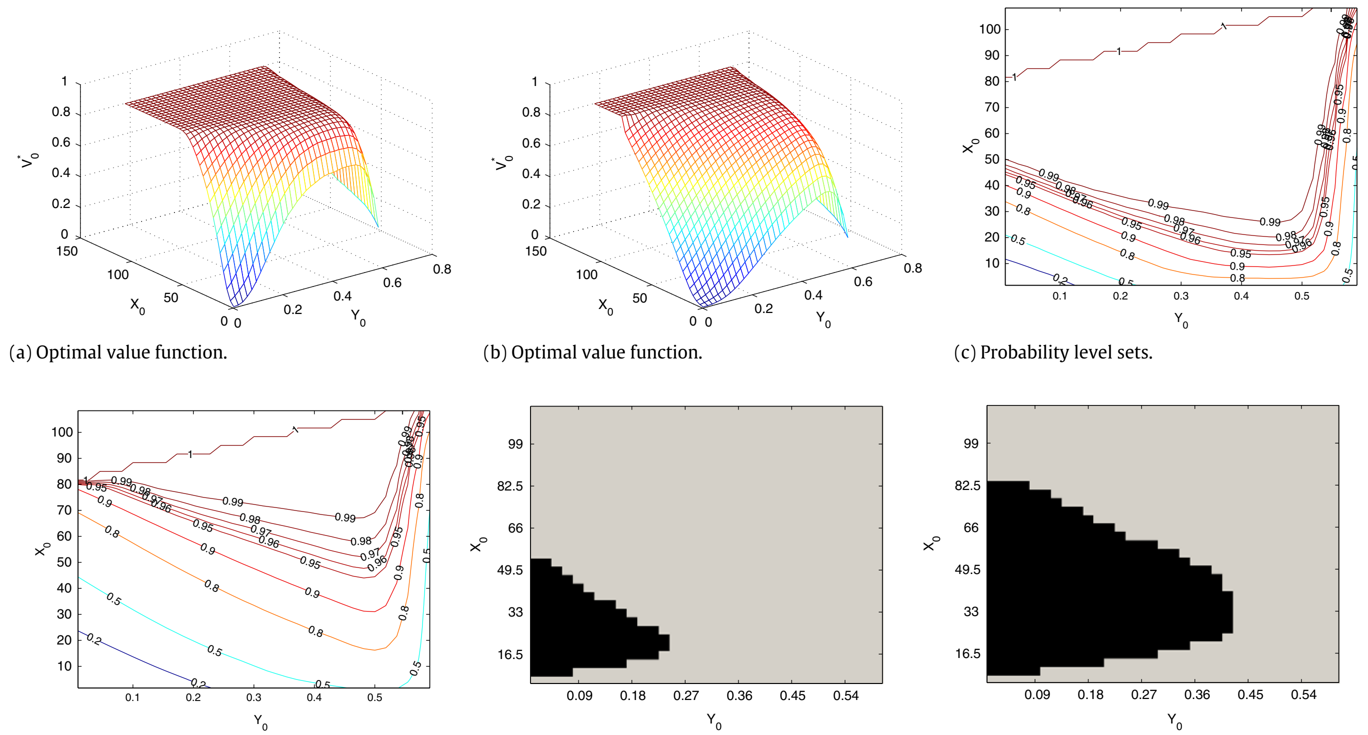
$$\begin{aligned} X_{k+1} &= a_k(X_k + P(Y_k) - c_k A)(1 + S_k) \\ &\quad + (1 - a_k)(X_k + P(Y_k) - c_k A)(1 + R_k) - C(Y_k) \\ Y_{k+1} &= Y_k + c_k G_k + (1 - c_k) F_k, \end{aligned}$$

where S_k and R_k are i.i.d. random variables representing the average rates of return for a safe investment and a risky investment option over one time interval, $P(\cdot)$ and $C(\cdot)$ give the policy premiums (i.e. customer payments) and policy claims (i.e. company payouts) as a function of market share, A is a constant variable representing the investment in advertising, and G_k and F_k are i.i.d. random variables representing market share growth in the presence and in the absence of advertising. Further, a_k is the percent of capital invested in the safe asset, and c_k is a Boolean variable representing whether advertising is used ($c_k = 1$) or not used ($c_k = 0$) during the current time horizon. We assume that the policy premium function and the policy claims function are given by $P(Y_k) = \alpha_k Y_k$ and $C(Y_k) = \beta_k Y_k$, where $k \in \{0, N\}$, and α_k and β_k are i.i.d. random variables. Notice that we assume that premiums are paid at the beginning of each period and therefore may be invested along with the initial surplus. We also assume that advertising costs are deducted at the start of the period and claims are deducted at the end of the period.

We introduce parameter values as follows. Let the advertising cost over one period be $A = 5$. All random variables are assumed to be i.i.d. and occur according to the following Gaussian distributions $S_k \sim \mathcal{N}(0.01, 0.005^2)$, $R_k \sim \mathcal{N}(0.03, 0.02^2)$, $G_k \sim \mathcal{N}(0.06, 0.01^2)$, $F_k \sim \mathcal{N}(0, 0.01^2)$, $\alpha_k \sim \mathcal{N}(40, 5^2)$, and $\beta_k \sim \mathcal{N}(30, 15^2)$ for all $k = 0, 1, \ldots, N$. Also, given the generality of the random variables, we take action to ensure that $Y_k \in [0, 1]$.

Consider the target set $K = \{(x, y) \in \mathbb{R}^2 | x \geq 80 + 40y, 0 \leq y \leq 0.6\}$ and the safe set $K' = (0, +\infty) \times [0, 0.6]$, where 60% represents the largest market share that can be handled by the company. For some initial investment $X_0 \in \mathbb{R}$ and initial market share $Y_0 \in [0, 1]$, we would like to maximize the probability of hitting the safe operating region defined by K, while avoiding financial ruin, i.e. without losing the initial investment over the finite time horizon with $N = 24$. To this end, we use the dynamic program associated with the first hitting time reach-avoid formulation.

Computational results with and without (i.e. $S_k = 0$ and $R_k = 0$) investment income are given in Fig. 2. All computations were performed on a 33×33 grid using the methods introduced in Abate et al. (2007). The transition probabilities associated with each grid point were obtained via 20 000 Monte Carlo samples. For each case, the value function at time zero (V_0^*), a contour plot representing the level sets of the value function at time zero, and the optimal policy regarding the advertising decision at time zero are provided. It is



(a) Optimal value function. (b) Optimal value function. (c) Probability level sets.

(d) Probability level sets. (e) Optimal advertising policy. (f) Optimal advertising policy.

Fig. 2. Results for the Economic Ruin problem at time $k = 0$. Optimal value functions V_0^* with and without investment return are given in (a) and (b), respectively. Probability level sets of the optimal value functions with and without investment return are given by the contour plots (c) and (d), respectively. The optimal advertising policies with and without investment return are given in (e) and (f), respectively, where black calls for advertising, and gray calls for no advertising.

interesting to point out that, given an initial market share of zero and an available surplus that is below the target investment, the company essentially has two routes that it can take in order to raise the necessary capital. It can either invest in the available assets, and hope to gain the capital without officially introducing the product, or it can both invest and advertise, in which case customer premiums will also be used as a mechanism to raise capital. For the finite time horizon case under the given assumptions, the threshold regarding the advertisement decision is clear in the results.

7. Conclusion

We have introduced two interpretations of the probabilistic reach-avoid problem. In the first case, we considered a problem where the objective was to maximize or minimize the probability that a system starting at a specific initial condition will hit a target set while avoiding an unsafe set over a finite time horizon. In the second case, we considered the case where the system will be inside the target set at the terminal time, while avoiding the unsafe set up to the terminal time. Optimal reach-avoid policies for each case were derived as the solution to an optimal control problem using dynamic programming. Further, we considered the first hitting time reach-avoid problem in the infinite horizon and discussed the relevance of this problem in the infinite horizon as opposed to the safety, target hitting, and the terminal hitting time problems.

Given the natural progression from a cost function defining a path over the state space to a dynamic program, it is only natural that a model checking language for SHS be developed in the future and provide researchers with a powerful tool for the verification of SHS. In light of this, future work includes model checking PCTL for DTSHS utilizing the dynamic programs developed here, as well as the consideration of extensions to the language such that more complicated (interesting) trajectories may be captured. Additionally, research is ongoing regarding high-dimensional integration and simplifying abstractions of DTSHS in an effort to combat the computational complexity of the numerical methods based on gridding.

Acknowledgements

The authors would like to thank Maria Prandini, Eugenio Cinquemani, and Debasish Chatterjee for helpful discussions and Alessandro Abate for providing MatLab code for the probabilistic safety problem.

Appendix. Proofs

To facilitate the application of the theoretical results in Bertsekas and Shreve (2007), we first define an equivalent minimization (rather than maximization) control problem. Consider the value function $J_k^\mu := -V_{N-k}^\mu$, $k = 0, 1, \ldots, N$ for a control policy $\mu \in \mathcal{M}_m$. It can be shown that the resulting forward dynamic recursion (with respect to Lemma 4) is given by

$$J_k^\mu(x) = -\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x) H(x, \mu_{N-k}(x), J_{k-1}^\mu),$$

initialized with $J_0^\mu(x) = -\mathbf{1}_K(x)$, $x \in X$. Consequently,

$$\begin{aligned} J_N^\mu(x_0) &= E_{x_0}^\mu\left[-\sum_{j=0}^{N}\left(\prod_{i=0}^{j-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_j)\right] \\ &= -r_{x_0}^\mu(K, K'). \end{aligned} \tag{A.1}$$

We define the optimal one-step cost and input as

$$J_k^*(x) = \inf_{a \in \mathcal{A}}\{-\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x) H(x, a, J_{k-1}^*)\},$$

$$\mu_k^*(x) = \arg\inf_{a \in \mathcal{A}}\{-\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x) H(x, a, J_{N-k-1}^*)\},$$

whenever the infimum is attained, with $J_0^*(x) = -\mathbf{1}_K(x)$, $x \in X$.

Next, consider a (universally measurable) function $\mu : X \to \mathcal{A}$ and define the map $T_\mu[J](x) = G(x, \mu(x), J)$, $x \in X$, where $G(x, a, J) = -\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x)H(x, a, J)$, $x \in X$, $a \in \mathcal{A}$, and $J \in \mathcal{F}$. Let $\overline{\mathcal{F}} \subset \mathcal{F}$ denote the set of universally measurable real functions $J : X \to \mathbb{R}$. The following follows from Proposition 7.46 in Bertsekas and Shreve (2007).

Lemma 12. *The map T_μ preserves the universal measurability property: i.e., if $J \in \overline{\mathcal{F}}$, then, $T_\mu[J] \in \overline{\mathcal{F}}$.*

Since (A.1) can be rewritten $J_k^\mu = T_{\mu_{N-k}}[J_{k-1}^\mu]$ and $J_0^\mu \in \overline{\mathcal{F}}$, by Lemma 12 we have $J_k^\mu \in \overline{\mathcal{F}}$, $k = 1, 2, \ldots, N$.

Lemma 13. *The map T_μ satisfies the following:*

(1) *For all $J, J' \in \overline{\mathcal{F}}$ such that $J(x) \le J'(x)$, $\forall x \in X$, then $T_\mu[J](x) \le T_\mu[J'](x)$, $\forall x \in X$,*
(2) *For any $J \in \overline{\mathcal{F}}$, $x \in X$, and any real number $r > 0$,*

$$T_\mu[J](x) \le T_\mu[J + r](x) \le T_\mu[J](x) + r. \tag{A.2}$$

Proof. Part 1 immediately follows from the definition of T_μ. For part 2, the following holds

$$\begin{aligned} G(x, a, J + r) &= -\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x)H(x, a, J + r), \\ &= -\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x)H(x, a, J) \\ &\quad + r\mathbf{1}_{K'\setminus K}(x)\int_X Q(dy|x, a), \end{aligned}$$

for $x \in X$, $a \in \mathcal{A}$. Since $\mathbf{1}_{K'\setminus K}(x)\int_X Q(dy|x, a) = \mathbf{1}_{K'\setminus K}(x)$ is upper bounded by 1 and lower bounded by 0, we can conclude that

$$G(x, a, J) \le G(x, a, J + r) \le G(x, a, J) + r,$$

for $x \in X$, $a \in \mathcal{A}$. □

We now define the map $T : \mathcal{F} \to \mathcal{F}$ as $T[J](x) = \inf_{a\in\mathcal{A}} G(x, a, J)$, $x \in X$. The recursion defined in (A.2) can be re-expressed as $J_k^* = T[J_{k-1}^*]$. It follows that $J_k^* = T^k[J_0^*]$, $k = 0, 1, \ldots, N$, where $T^0[J] = J$ and $T^k[J] = T[T^{k-1}[J]]$. Let $\mathcal{F}^* \subset \overline{\mathcal{F}}$ denote the set of lower-semianalytic functions. The following fact follows from Propositions 7.47, 7.48, and Lemma 7.30 in Bertsekas and Shreve (2007).

Lemma 14. *The map T preserves the lower-semianalytic property, i.e. if $J \in \mathcal{F}^*$, then $T[J] \in \mathcal{F}^*$.*

Since $J_k^* = T[J_{k-1}^*]$ and $J_0^* \in \mathcal{F}^*$, by Lemma 14 we have $J_k^* \in \mathcal{F}^*$, $k = 1, 2, \ldots, N$.

Next, we show that the optimal cost at time N can be defined in terms of the map T^N, and that the optimal cost achieved is $\inf_\mu\{-r_{x_0}^\mu(K, K')\}$. Note that in the sequel we denote $x_0 = x$.

Proposition 15. *For any time horizon $N \in \mathbb{N}$ and control policy $\mu \in \mathcal{M}_m$, the following holds*

$$\inf_\mu E_x^\mu\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] = T^N[J_0^*](x),$$

with $J_0^*(x) = -\mathbf{1}_K(x)$, $x \in X$.

Proof. By induction on N. Clearly, the statement holds for $N = 0$. Suppose that the statement holds for $N = h$. By the induction hypothesis and (A.1), for all $\epsilon > 0$, there exists $\mu_\epsilon = (\mu_{\epsilon,0}, \mu_{\epsilon,1}, \ldots) \in \mathcal{M}_m$ such that

$$J_h^{\mu_\epsilon}(x) \le \inf_\mu E_x^\mu\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] + \epsilon.$$

For any universally measurable function $\pi : X \to \mathcal{A}$, we then have that, by the monotonicity of T_π and (A.2),

$$\begin{aligned} &T_\pi[J_h^{\mu_\epsilon}](x) \\ &\le T_\pi\left[\inf_\mu E_x^\mu\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] + \epsilon. \end{aligned}$$

If we consider policy $\overline{\mu}_\epsilon = (\pi, \mu_{\epsilon,0}, \mu_{\epsilon,1}, \ldots)$,

$$\begin{aligned} \inf_\mu E_x^\mu\left[-\sum_{k=0}^{h+1}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] &\le J_{h+1}^{\overline{\mu}_\epsilon}(x) \\ &= T_\pi[J_h^{\mu_\epsilon}](x). \end{aligned}$$

Combined with the argument above, this leads to

$$\begin{aligned} &\inf_\mu E_x^\mu\left[-\sum_{k=0}^{h+1}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] \\ &\le T_\pi\left[\inf_\mu E_x^\mu\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] + \epsilon, \end{aligned}$$

for any universally measurable function $\pi : X \to \mathcal{A}$. It follows that

$$\begin{aligned} &\inf_\mu E_x^\mu\left[-\sum_{k=0}^{h+1}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] \\ &\le T\left[\inf_\mu E_x^\mu\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] \\ &= T^{h+1}[J_0^*](x), \end{aligned}$$

by the induction hypothesis. On the other hand, for any $\mu \in \mathcal{M}_m$ and $N = h + 1$, we have

$$\begin{aligned} J_{h+1}^\mu &= (T_{\mu_0}\cdots T_{\mu_h})[J_0^\mu] \\ &\ge (T_{\mu_0}\cdots T_{\mu_{h-1}}T)[J_0^\mu] \ge T^{h+1}[J_0^\mu]. \end{aligned}$$

Taking the infimum over $\mu \in \mathcal{M}_m$, we obtain

$$T^{h+1}[J_0^*](x) \le \inf_\mu E_x^\mu\left[-\sum_{k=0}^{h+1}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right].$$

Which concludes the proof. □

We continue with a proof pertaining to the existence of ϵ-optimal control policies.

Proposition 16. *For any time horizon $N \in \mathbb{N}$ and for all $\epsilon > 0$, there exists $\mu_\epsilon = (\mu_{\epsilon,0}, \mu_{\epsilon,1}, \ldots) \in \mathcal{M}_m$ such that for all $x \in X$, the following holds*

$$\begin{aligned} &\inf_\mu E_x^\mu\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] \\ &\le J_N^{\mu_\epsilon}(x) \le \inf_\mu E_x^\mu\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] + \epsilon. \end{aligned}$$

Proof. For any $\mu_\epsilon \in \mathcal{M}_m$,

$$\begin{aligned} J_N^{\mu_\epsilon}(s) &= E_x^{\mu_\epsilon}\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] \\ &\ge \inf_\mu E_x^\mu\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right], \end{aligned}$$

which satisfies the left inequality. The proof of the right inequality is by induction on N. Clearly, the inequality holds for $N = 0$.

Assume that the inequality holds for $N = h$. Then, for any $\overline{\epsilon} > 0$, there exists $\overline{\mu} = (\overline{\mu}_0, \overline{\mu}_1, \ldots) \in \mathcal{M}_m$ such that

$$J_h^{\overline{\mu}}(x) \leq \inf_{\mu} E_x^{\mu}\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] + \frac{\overline{\epsilon}}{2}.$$

By Proposition 7.50 in Bertsekas and Shreve (2007), there exists a universally measurable function $\overline{\pi} : X \to \mathcal{A}$ such that

$$\begin{aligned} T_{\overline{\pi}}&\left[E_x^{\mu}\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] \\ &\leq T\left[E_x^{\mu}\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] + \frac{\overline{\epsilon}}{2}. \end{aligned}$$

If we consider $\mu_{\overline{\epsilon}} = (\overline{\pi}, \overline{\mu}_0, \overline{\mu}_1, \ldots)$, by the monotonicity of $T_{\overline{\pi}}$ and (A.2), we obtain

$$\begin{aligned} J_{h+1}^{\mu_{\overline{\epsilon}}}(x) &= T_{\overline{\pi}}[J_h^{\overline{\mu}}](x) \\ &\leq T_{\overline{\pi}}\left[\inf_{\mu} E_x^{\mu}\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] + \frac{\overline{\epsilon}}{2} \\ &\leq T\left[\inf_{\mu} E_x^{\mu}\left[-\sum_{k=0}^{h}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right]\right] + \overline{\epsilon}. \end{aligned}$$

By the induction hypothesis and Proposition 15, we come to the conclusion

$$\begin{aligned} J_{h+1}^{\mu_{\overline{\epsilon}}}(x) &\leq T^{h+1}[J_0^*] + \overline{\epsilon} \\ &= \inf_{\mu} E_x^{\mu}\left[-\sum_{k=0}^{h+1}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] + \overline{\epsilon}, \end{aligned}$$

which concludes the proof. □

Next, assuming that the infima are all attained, we show that the control policy defined by $\mu^* = (\mu_0^*, \mu_1^*, \ldots)$ satisfying argument (A.2) is a Markov policy and that it is N-stage optimal (Bertsekas & Shreve, 2007).

Lemma 17. *If $\mu_k^* : X \to \mathcal{A}, k \in [0, N-1]$, is such that $\forall x \in X$, $\mu_k^*(x)$ satisfies (A.2), then $\mu^* = (\mu_0^*, \mu_1^*, \ldots, \mu_{N-1}^*)$ is an N-stage optimal Markov policy. A sufficient condition for the existence of such a μ^* is that $U_k(x, \lambda) = \{a \in \mathcal{A} | G(x, a, J_{N-k-1}^*) \leq \lambda\}$ is compact for all $x \in X$, $\lambda \in \mathbb{R}$, $k \in [0, N-1]$.*

Proof. Note that μ_k^* satisfying (A.2) can be characterized through

$$T_{\mu_k^*}[J_{N-k-1}^*](x) = \inf_{a\in\mathcal{A}} G(x, a, J_{N-k-1}^*) = J_{N-k}^*(x),$$

$x \in X$. Since $G(x, a, J_{N-k}^*)$ is lower-semianalytic by Lemma 14, if its infimum with respect to $a \in \mathcal{A}$ is attained for any $x \in X$, the resulting function $\mu_k^* : X \to \mathcal{A}$ is universally measurable (Bertsekas & Shreve, 2007). Observe that

$$\begin{aligned} \inf_{\mu} & E_x^{\mu}\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right] \\ &= J_N^*(x) = T_{\mu_0^*}[J_{N-1}^*](x) \\ &= T_{\mu_0^*}[T_{\mu_1^*}[J_{N-2}^*]](x) \\ &= \cdots = T_{\mu_0^*}[T_{\mu_1^*}[\cdots T_{\mu_{N-1}}^*[J_0^*]]](x) \\ &= J_N^{\mu^*}(x) \\ &= E_x^{\mu^*}\left[-\sum_{k=0}^{N}\left(\prod_{i=0}^{k-1}\mathbf{1}_{K'\setminus K}(x_i)\right)\mathbf{1}_K(x_k)\right], \end{aligned}$$

$x \in X$, thus, μ^* is an N-stage optimal Markov policy. The second part of the lemma is proved by applying Lemma 3.1 of Bertsekas and Shreve (2007). □

Finally, we present the proof of Theorem 6.

Proof. It directly follows from the definition of the cost function J, (A.2), (6), and Proposition 15 that the dynamic programming algorithm holds, i.e.

$$V_0^*(x_0) = -J_N^*(x_0) = -T^N[J_0^*](x_0) = \sup_{\mu\in\mathcal{M}_m} r_{x_0}^{\mu}(K, K'),$$

for any $x_0 \in X$, so the first part of the theorem is proven. For the second part, we note that since $J_{N-k-1}^* = -V_{k+1}^*$, then

$$\begin{aligned} G(x, a, J_{N-k-1}^*) &= -\mathbf{1}_K(x) + \mathbf{1}_{K'\setminus K}(x)H(x, a, J_{N-k-1}^*) \\ &= -\mathbf{1}_K(x) - \mathbf{1}_{K'\setminus K}(x)H(x, a, V_{k+1}^*). \end{aligned}$$

Thus, the expressions (A.2) and (7) are equivalent, and by Lemma 17 $\mu^* = (\mu_0^*, \mu_1^*, \ldots, \mu_{N-1}^*)$ is an optimal reach-avoid Markov policy. The condition on $U_k(x, \lambda)$ follows by Lemma 17 and Lemma 3.1 of Bertsekas and Shreve (2007). □

References

Abate, A., Amin, S., Prandini, M., Lygeros, J., & Sastry, S. (2007). Computational approaches to reachability analysis of stochastic hybrid systems. In A. Bemporad, A. Bicchi, & G. C. Buttazzo (Eds.), *Lecture notes in computer science*: *Vol. 4416*. *Hybrid systems: computation and control* (pp. 4–17). Springer.

Abate, A., Prandini, M., Lygeros, J., & Sastry, S. (2008). Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, *44*(11), 2724–2734.

Aubin, J.-P. (1991). *Viability theory*. Cambridge, MA, USA: Birkhauser Boston Inc.

Bertsekas, D. P., & Shreve, S. E. (2007). *Stochastic optimal control: the discrete-time case*. Athena Scientific.

Boda, K., Filar, J. A., Lin, Y., & Spanjers, L. (2004). Stochastic target hitting time and the problem of early retirement. *IEEE Transactions on Automatic Control*, *49*(3), 409–419.

Brasao, A. (2000). Stochastic bio-economic model of northern atlantic and mediterranean bluefin tuna. *Technical report*.

Bujorianu, M. L., & Lygeros, J. (2003). Reachability questions in piecewise deterministic Markov processes. In O. Maler, & A. Pnueli (Eds.), *Lecture notes in computer science*: *Vol. 2623*. *Hybrid systems: computation and control* (pp. 126–140). Springer.

Cai, J. (2002). Discrete time risk models under rates of interest. *Probability in the Engineering and Informational Sciences*, *16*(3), 309–324.

Cai, J., & Dickson, D. C. (2004). Ruin probabilities with a Markov chain interest model. *Insurance: Mathematics and Economics*, *35*(3), 513–525.

Chae, D.-R., & Pascoe, S. (2005). Use of simple bioeconomic models to estimate optimal effort levels in the Korean coastal flounder fisheries. *Aquatic Living Resources*, *18*(2), 93–101.

Clark, C. W. (1990). *Mathematical bioeconomics: the optimal management of renewable resources* (2nd ed.). Wiley-Interscience.

Conrad, J. M. (1999). The bioeconomics of marine sanctuaries. *Journal of Bioeconomics*, *1*(2), 205–217.

Costello, C., & Polasky, S. (2008). Optimal harvesting of stochastic spatial resources. *Journal of Environmental Economics and Management*, *56*(1), 1–18.

Davis, M. (1993). *Markov models and optimization*. London: Chapman & Hall.

de Kok, T. G. (2003). Ruin probabilities with compounding assets for discrete time finite horizon problems, independent period claim sizes and general premium structure. *Insurance: Mathematics and Economics*, *33*(3), 645–658.

Gao, Y., Lygeros, J., & Quincampoix, M. (2007). On the reachability problem for uncertain hybrid systems. *IEEE Transactions on Automatic Control*, *52*(9), 1572–1586.

Ghosh, M. K., Arapostathis, A., & Marcus, S. I. (1997). Ergodic control of switching diffusions. *SIAM Journal on Control and Optimization*, *35*(6), 1952–1988.

Hu, J., Lygeros, J., & Sastry, S. (2000). Towards a theory of stochastic hybrid systems. In N. A. Lynch, & B. H. Krogh (Eds.), *Lecture notes in computer science*: *Vol. 1790*. *Hybrid systems: computation and control* (pp. 160–173). Springer.

Hu, J., Prandini, M., & Sastry, S. (2003). Probabilistic safety analysis in three dimensional aircraft flight. In *42nd IEEE conference on decision and control: Vol. 5. December* (pp. 5335–5340).

Hu, J., Prandini, M., & Sastry, S. (2005). Aircraft conflict prediction in the presence of a spatially correlated wind field. *IEEE Transactions on Intelligent Transportation Systems*, *6*(3), 326–340.

Koutsoukos, X. D., & Riley, D. (2006). Computational methods for reachability analysis of stochastic hybrid systems. In J. P. Hespanha, & A. Tiwari (Eds.), *Lecture notes in computer science*: *Vol. 3927*. *Hybrid systems: computation and control* (pp. 377–391). Springer.

Koutsoukos, X. D., & Riley, D. (2008). Computational methods for verification of stochastic hybrid systems. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, *38*(2), 385–396.
Kushner, H. J., & Dupuis, P. G. (1992). *Numerical methods for stochastic control problems in continuous time*. London, UK: Springer-Verlag.
Lauck, T., Clark, C. W., Mangel, M., & Munro, G. R. (1998). Implementing the precautionary principle in fisheries management through marine reserves. *Ecological Applications*, *8*(sp1), S72–S78.
Lundberg, F. (1903). *Approximerad framstilling av sannolikhetsfunktionen II. Aterforsaekring av kollektivrisker*. Uppsala: Almquist & Wiksell.
Lygeros, J. (2004). On reachability and minimum cost optimal control. *Automatica*, *40*(6), 917–927.
Mitchell, I. (2008). The flexible, extensible and efficient toolbox of level set methods. *Journal of Scientific Computing*, *35*(2–3), 300–329.
Mitchell, I., & Tomlin, C. (2000). Level set methods for computation in hybrid systems. In *Hybrid systems: computation and control* (pp. 310–323). London, UK: Springer-Verlag.
Nostbakken, L. (2008). Stochastic modelling of the north sea herring fishery under alternative management regimes. *Marine Resource Economics*, *23*(1).
Nyrhinen, H. (1999). On the ruin probabilities in a general economic environment. *Stochastic Processes and their Applications*, *83*(2), 319–330.
Paulsen, J. (1998). Ruin theory with compounding assets — a survey. *Insurance: Mathematics and Economics*, *22*(1), 3–16 [special issue on the interplay between insurance, finance and control].
Paulsen, J. (2008). Ruin models with investment income. *Probability Surveys*, *5*, 416.
Pitchford, J. W., Codling, E. A., & Psarra, D. (2007). Uncertainty and sustainability in fisheries and the benefit of marine protected areas. *Ecological Modelling*, *207*(2–4), 286–292.
Rudin, W. (1987). *Real and complex analysis* (3rd ed.). New York, NY, USA: McGraw-Hill, Inc.
Sainsbury, K. J., Punt, A. E., & Smith, A. D. M. (2000). Design of operational management strategies for achieving fishery ecosystem objectives. *ICES Journal of Marine Science*, *57*(3), 731–741.
Schael, M. (2005). Control of ruin probabilities by discrete-time investments. *Mathematical Methods of Operations Research*, *62*(1), 141–158.
Sethi, G., Costello, C., Fisher, A., Hanemann, M., & Karp, L. (2005). Fishery management under multiple uncertainty. *Journal of Environmental Economics and Management*, *50*(2), 300–318.
Tang, Q., & Tsitsiashvili, G. (2003). Precise estimates for the ruin probability in finite horizon in a discrete-time model with heavy-tailed insurance and financial risks. *Stochastic Processes and their Applications*, *108*(2), 299–325.
Tomlin, C., Lygeros, J., & Sastry, S. (2000). A game theoretic approach to controller design for hybrid systems. *Proceedings of the IEEE*, *88*, 949–969.
Wei, X., & Hu, Y. (2008). Ruin probabilities for discrete time risk models with stochastic rates of interest. *Statistics & Probability Letters*, *78*(6), 707–715.
White, C., Kendall, B., Gaines, S., Siegel, D., & Costello, C. (2008). Marine reserve effects on fishery profit. *Ecology Letters*, *11*(4), 370–379.
Yang, H., & Zhang, L. (2006). Ruin problems for a discrete time risk model with random interest rate. *Mathematical Methods of Operations Research*, *63*(2), 287–299.

Sean Summers received the B.S. degree in Aerospace Engineering in 2004 and the M.S. degree in Mechanical Engineering in 2007, both from the Department of Mechanical and Aerospace Engineering at the University of California, San Diego. He is currently a Ph.D. candidate in the Department of Information Technology and Electrical Engineering, Swiss Federal Institute of Technology (ETH), Zurich. His research interests include modeling and control of stochastic hybrid systems, reachability analysis, model predictive control, and applications of control theory in systems biology.

John Lygeros received a B.Eng. degree in Electrical Engineering and an M.Sc. degree in Automatic Control from Imperial College, London, UK, in 1990 and 1991, respectively. He then received a Ph.D. degree from the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California, USA in 1996. He held postdoctoral research appointments with the National Automated Highway Systems Consortium, the Massachusetts Institute of Technology, and the University of California, Berkeley. In parallel, he was also a part-time Research Engineer with SRI International, Menlo Park, California, USA and a Visiting Professor with the Department of Mathematics, Universite de Bretagne Occidentale, Brest, France. Between 2000 and 2003, he was a university lecturer with the Department of Engineering, University of Cambridge, Cambridge, UK and a fellow of Churchill College. Between 2003 and 2006, he was an assistant professor with the Department of Electrical and Computer Engineering, University of Patras, Patras, Greece. In July 2006, he joined the Automatic Control Laboratory, ETH Zurich, Switzerland, where he is currently serving as the Head of the laboratory. His research interests include modeling, analysis, and control of hierarchical, hybrid and stochastic systems with applications to biochemical networks and large-scale engineering systems such as power networks and air traffic management.