A Risk-Sensitive Finite-Time Reachability Approach for Safety of Stochastic Dynamic Systems

Margaret P. Chapman[1,2], Jonathan Lacotte[3], Aviv Tamar[1], Donggun Lee[4], Susmit Jha[2], Kevin Smith[5], Victoria Cheng[6], Jaime Fernandez-Fisac[1], Marco Pavone[7], Claire J. Tomlin[1]

***Abstract*— A classic reachability problem for safety of dynamic systems is to compute the set of initial states from which the state trajectory is guaranteed to stay inside a given constraint set over some time horizon. In this paper, we leverage existing theory of reachability analysis and risk measures to devise a *risk-sensitive* reachability approach for safety of *stochastic* dynamic systems under non-adversarial disturbances over a finite time horizon. Specifically, we first introduce the notion of a *risk-sensitive safe set* as the set of initial states from which the risk of extreme constraint violations can be made small via an appropriate control policy, where risk is quantified using the *Conditional Value-at-Risk* (CVaR) measure. Second, we show how the computation of a risk-sensitive safe set can be reduced to the solution to a Markov Decision Process (MDP), where cost is assessed according to CVaR. Third, leveraging this reduction, we devise a tractable algorithm to approximate a risk-sensitive safe set, and provide theoretical arguments about its correctness. Finally, we present numerical experiments that demonstrate the utility of risk-sensitive reachability analysis. In particular, our approach allows a practitioner to tune the level of risk sensitivity from worst-case (which is typical for Hamilton-Jacobi reachability analysis) to risk-neutral (which is the case for stochastic reachability analysis).**

I. Introduction

Reachability analysis is a formal verification method based on optimal control theory that is used to prove safety or performance properties of dynamic systems [1]. A classic reachability problem for safety is to compute the set of initial states from which the state trajectory is guaranteed to stay inside a given constraint set over some time horizon. This problem was first considered for discrete-time dynamic systems by Bertsekas and Rhodes under the assumption that disturbances are uncertain but belong to known sets [2], [3], [4]. In this context, the problem is solved using a minimax formulation, in which disturbances behave adversarially and safety is described as a binary notion based on set membership [2], [3], [4, Sec. 3.6.2].

In practice, minimax formulations can yield overly conservative solutions, particularly because disturbances are not often adversarial. Most storms do not cause major floods, and most vehicles are not involved in pursuit-evader games. If there are enough observations of the system, one can estimate a probability distribution for the disturbance (e.g., see [5]), and then assess safety properties of the system in a more realistic context. For stochastic discrete-time dynamic systems, Abate et al. [6] developed an algorithm that computes the set of initial states from which the probability of safety of the state trajectory can be made large by an appropriate control policy.[1] Summers and Lygeros [7] extended the algorithm of Abate et al. to quantify the probability of safety and performance of the state trajectory, by specifying that the state trajectory should also reach a target set.

Both the stochastic reachability methods [6], [7] and the minimax reachability methods [2], [3], [4] for discrete-time dynamic systems describe safety as a binary notion based on set membership. In Abate et al., for example, the probability of safety to be optimized is formulated as the expectation of the product (or maximum) of indicator functions, where each indicator encodes the event that the state at a particular time point is inside a given set [6]. Thus, the stochastic reachability methods [6], [7] do not generalize to quantify the random distance between the state trajectory and the boundary of the constraint set, since they use indicator functions to convert probabilities to expectations to be optimized.

In contrast, Hamilton-Jacobi (HJ) reachability methods quantify the deterministic analogue of this distance for continuous-time systems subject to adversarial disturbances (e.g., see [1], [8], [9], [10]). Quantifying the distance between the state trajectory and the boundary of the constraint set in a non-binary fashion may be important in applications where the boundary is not known exactly, or where mild constraint violations are inevitable, but extreme constraint violations must be avoided.

It is imperative that reachability methods for safety take into account the possibility that rare events can occur with potentially damaging consequences. Reachability methods that assume adversarial disturbances (e.g., [1], [3]) suppose that harmful events can always occur, which may yield solutions with limited practical utility, especially in applications with large uncertainty sets. Stochastic reachability methods [6], [7] do not explicitly account for rare high-

[1]M.C., J.F., A.T., and C.T. are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA. chapmanm@berkeley.edu

[2]S.J. is with SRI International, Menlo Park, California, USA. M.C. was a Student Associate at SRI International.

[3]J.L. is with the Department of Electrical Engineering, Stanford University, USA.

[4]D.L. is with the Department of Mechanical Engineering, University of California, Berkeley, USA.

[5]K.S. is with OptiRTC, Inc. and the Department of Civil and Environmental Engineering, Tufts University, USA.

[6]V.C. is with the Department of Civil and Environmental Engineering, University of California, Berkeley, USA.

[7]M.P. is with the Department of Aeronautics and Astronautics, Stanford University, USA.

[1]Safety of the state trajectory is the event that the state trajectory stays in the constraint set over a finite time horizon.

consequence events, as costs are evaluated according to a risk-neutral expectation.

In contrast, in this paper, we harness the tool of *risk measure theory* to formulate a reachability analysis approach that explicitly accounts for the possibility of rare events with negative consequences: harmful events are likely to occur at some point, but they are unlikely to occur always. Specifically, a *risk measure* is a function that maps a random variable, Z, representing a loss, or a cost, into the real line, according to the risk associated with Z [11, Sec. 6.3], [12, Sec. 2.2]. Risk-sensitive optimization has been studied in applied mathematics [13], reinforcement learning [14], [15], [16], and optimal control [17]. Risk-sensitive formulations have the potential to inform practical decision-making that also protects against damaging outcomes, where the level of conservatism can be modified as needed.

In this paper, we use a particular risk measure, called *Conditional Value-at-Risk* (CVaR). If Z is a random cost with finite expectation, then the Conditional Value-at-Risk of Z at the confidence level, $\alpha \in (0,1]$, is defined as [11, Equation 6.22],[2]

$$\text{CVaR}_\alpha[Z] := \min_{t\in\mathbb{R}} \Big\{ t + \frac{1}{\alpha}\mathbb{E}\big[\max\{Z-t,0\}\big]\Big\}. \quad (1)$$

Note that $\text{CVaR}_\alpha[Z]$ increases from $\mathbb{E}[Z]$ to $\operatorname{ess\,sup} Z$, as α decreases from 1 to 0. Thus, CVaR captures a full spectrum of risk assessments, from risk-neutral to worst-case, according to the value of α. Further, there is a well-established relationship between CVaR and chance constraints that we will use to obtain probabilistic safety guarantees.

Statement of Contributions. This paper introduces a *risk-sensitive* reachability approach for safety of stochastic dynamic systems under non-adversarial disturbances over a finite-time horizon. Specifically, the contributions are four-fold. First, we introduce the notion of a *risk-sensitive safe set* as the set of initial states from which the risk of extreme constraint violations can be made small via an appropriate control policy, where risk is quantified using the *Conditional Value-at-Risk* (CVaR) measure. Our formulation explicitly assesses the distance between the boundary of the constraint set and the state trajectory of a stochastic dynamic system. This is an extension of stochastic reachability methods (e.g., [6], [7]), which replace this distance with a binary random variable. Further, in contrast to stochastic reachability methods, our formulation explicitly accounts for rare high-consequence events, by posing the optimal control problem in terms of CVaR, instead of a risk-neutral expectation. To the best of our knowledge, this paper is the first to use the Conditional Value-at-Risk measure in the reachability literature. Second, we show how the computation of a risk-sensitive safe set can be reduced to the solution to a Markov Decision Process (MDP), where cost is assessed according to CVaR. Third, leveraging this reduction, we devise a tractable algorithm to approximate a risk-sensitive safe set, and provide theoretical arguments about its correctness. Finally, we present numerical experiments that demonstrate the utility of risk-sensitive reachability analysis.

[2]Conditional Value-at-Risk is also called *Average Value-at-Risk*, which is abbreviated as AV@R in [11].

Organization. The rest of this paper is organized as follows. We present the problem formulation and define risk-sensitive safe sets in Sec. II. In Sec. III, we show how the computation of a risk-sensitive safe set can be reduced to the solution to a CVaR-MDP problem, i.e., an MDP where cost is assessed according to CVaR. In Sec. IV, we present a value-iteration algorithm to approximate risk-sensitive safe sets, along with theoretical arguments that support its correctness. In Sec. V, we provide a numerical example in the domain of stormwater infrastructure design. Finally, in Sec. VI, we draw conclusions and discuss directions for future work.

II. Problem Formulation

A. System Model

We consider a fully observable stochastic discrete-time dynamic system over a finite-time horizon [4, Sec. 1.2],

$$x_{k+1} = f(x_k, u_k, w_k), \quad k = 0, 1, \dots, N-1, \quad (2)$$

such that $x_k \in \mathcal{X} \subseteq \mathbb{R}^n$ is the state of the system at time k, $u_k \in U$ is the control at time k, and $w_k \in D$ is the random disturbance at time k. The control space, U, and disturbance space, D, are finite sets of real-valued vectors. The function, $f : \mathcal{X} \times U \times D \to \mathcal{X}$, is bounded and Lipschitz continuous. The probability that the disturbance equals $d_j \in D$ at time k is, $\mathbb{P}[w_k = d_j] = p_j$, where $0 \leq p_j \leq 1$ and $\sum_{j=1}^W p_j = 1$. We assume that w_k is independent of x_k, u_k, and disturbances at any other times. The only source of randomness in the system is the disturbance. In particular, the initial state, x_0, is not random. The set of *admissible, deterministic, history-dependent control policies* is,

$$\Pi := \big\{(\mu_0, \mu_1, \dots, \mu_{N-1}) \mid \mu_k : H_k \to U\big\}, \quad (3)$$

where $H_k := \underbrace{\mathcal{X} \times \ldots \times \mathcal{X}}_{\text{(k+1) times}}$ is the set of state histories up to time k. We are given a constraint set, $\mathcal{K} \subseteq \mathcal{X}$, and the *safety criterion* that the state of the system should stay inside $\mathcal{K}$ over time. For example, if the system is a pond, then x_k may be the water level of the pond at time k, and $\mathcal{K} := [0, 5\text{ft})$ indicates that the pond overflows if the water level exceeds 5ft. We quantify the extent of constraint violation/satisfaction using a surface function that characterizes the constraint set. Specifically, similar to [9, Eq. 2.3], let $g : \mathcal{X} \to \mathbb{R}$ satisfy,

$$x \in \mathcal{K} \iff g(x) < 0. \quad (4)$$

For example, we may choose $g(x) := x - 5$ to characterize $\mathcal{K} := [0, 5\text{ft})$ on the state space, $\mathcal{X} = [0, \infty)$.

B. Risk-Sensitive Safe Sets

A *risk-sensitive safe set* is a set of initial states from which the risk of extreme constraint violations over time can be made small via an admissible control policy, where risk is quantified using the CVaR measure. We use the term, *risk level*, to mean the allowable level of risk of constraint violations. Formally, the risk-sensitive safe set at the confidence level, $\alpha \in (0,1]$, and the risk level, $r \in \mathbb{R}$, is defined as,

$$\mathcal{S}_\alpha^r := \{x \in \mathcal{X} \mid W_0^*(x,\alpha) \leq r\}, \tag{5a}$$

where

$$\begin{aligned} W_0^*(x,\alpha) &:= \min_{\pi \in \Pi} \text{CVaR}_\alpha\big[Z_x^\pi\big], \\ Z_x^\pi &:= \max\big\{g(x_k) \mid k = 0,\dots,N\big\}, \end{aligned} \tag{5b}$$

such that the state trajectory, $(x_0, x_1, ..., x_N)$, evolves according to the dynamics model (2) with the initial state, $x_0 := x$, under the policy, $\pi \in \Pi$. The surface function, g, characterizes distance to the constraint set, $\mathcal{K}$, according to (4). Note that the minimum in the definition of $W_0^*(x,\alpha)$ is attained, as the next lemma states.

Lemma 1 (Existence of a minimizer): For any initial state, $x_0 \in \mathcal{X}$, and any confidence level, $\alpha \in (0,1]$, there exists a policy, $\pi^* \in \Pi$, such that

$$\text{CVaR}_\alpha\big[Z_x^{\pi^*}\big] = \inf_{\pi\in\Pi} \text{CVaR}_\alpha\big[Z_x^\pi\big] = \min_{\pi\in\Pi} \text{CVaR}_\alpha\big[Z_x^\pi\big].$$

Proof: Fix the initial state, x_0. Since the control and disturbance spaces are finite, the set of states that could be visited (starting from x_0) is finite. Therefore, the corresponding set of policies, Π, is finite. Hence, the infimum must be attained by some policy, π^*. ■

In the next sections, we will present a tractable algorithm to approximately compute risk-sensitive safe sets at different levels of confidence and risk.

C. Discussion

Computing risk-sensitive safe sets, as defined by (5), is well-motivated for several reasons. Our formulation incorporates different confidence levels and non-binary distance to the constraint set. In contrast, the stochastic reachability problem addressed by Abate et al. [6] uses a single confidence level and an indicator function to measure distance to the constraint set, in order to quantify the probability of constraint violation. Specifically, let $\epsilon \in [0,1]$ be the maximum tolerable probability of constraint violation (called *safety level* in [6]), and choose $\alpha := 1$, $r := \epsilon - \frac{1}{2}$, and $g(x) := \mathbf{1}_{\bar{\mathcal{K}}}(x) - \frac{1}{2}$, where

$$\mathbf{1}_{\bar{\mathcal{K}}}(x) := \begin{cases} 1 \text{ if } x \notin \mathcal{K} \\ 0 \text{ if } x \in \mathcal{K} \end{cases}. \tag{6a}$$

Then, the risk-sensitive safe set (5) is equal to,

$$\Big\{x \in \mathcal{X} \;\Big|\; \min_{\pi\in\Pi} \mathbb{E}\big[\max_{k=0,\dots,N} \mathbf{1}_{\bar{\mathcal{K}}}(x_k)\big] \leq \epsilon\Big\}, \tag{6b}$$

which is the *maximal probabilistic safe set* at the ϵ-safety level [6, Eqs. 11 and 13], if we consider non-hybrid dynamic systems that evolve under history-dependent policies.[3]

Risk-sensitive safe sets have two desirable mathematical properties. The first property is that $\mathcal{S}_\alpha^r$ shrinks as the risk level, r, or the confidence level, α, decrease. Since $\mathcal{S}_\alpha^r$ is an r-sublevel set, and CVaR_α increases as α decreases, one can show that

$$\begin{aligned} &\mathcal{S}_{\alpha_2}^{r_2} \subseteq \mathcal{S}_{\alpha_1}^{r_2} \subseteq \mathcal{S}_{\alpha_1}^{r_1}, \text{ and} \\ &\mathcal{S}_{\alpha_2}^{r_2} \subseteq \mathcal{S}_{\alpha_2}^{r_1} \subseteq \mathcal{S}_{\alpha_1}^{r_1} \end{aligned} \tag{7}$$

hold for any $r_1 \geq r_2$ and $1 \geq \alpha_1 \geq \alpha_2 > 0$. In other words, as the allowable level of risk of constraint violation, r, decreases, or as the fraction of damaging outcomes that are not fully addressed, α, decreases, $\mathcal{S}_\alpha^r$ encodes a higher degree of safety.

The second property is that risk-sensitive safe sets at the risk level, $r := 0$, enjoy probabilistic safety guarantees.

Lemma 2 (Link to a probabilistic safety guarantee): If $x \in \mathcal{S}_\alpha^0$, then the probability that the state trajectory initialized at x exits the constraint set can be made less than or equal to α, by an admissible control policy.

Proof: The proof follows from the fact that $\text{CVaR}_\alpha[Z_x^\pi] \leq 0 \implies \mathbb{P}[Z_x^\pi \geq 0] \leq \alpha$ [11, Sec. 6.2.4, pp. 257-258]. The event, $Z_x^\pi \geq 0$, is equivalent to the event that there is a state, x_k, of the associated trajectory that exits the constraint set, since $g(x_k) \geq 0 \iff x_k \notin \mathcal{K}$. ■

Lemma 2 indicates that $\mathcal{S}_\alpha^0$ is a subset of the *maximal probabilistic safe set* at the safety level, $\alpha \in (0,1]$, if we consider non-hybrid dynamic systems that evolve under history-dependent policies [6, Eqs. 9 and 11].

III. Reduction of Risk-Sensitive-Safe-Set Computation to CVaR-MDP

Computing risk-sensitive safe sets is challenging since the computation involves a maximum of costs (as opposed to a summation of costs) and the Conditional Value-at-Risk measure (as opposed to an expectation). In this section, we show how computing an under-approximation of a risk-sensitive safe set can be reduced to solving a CVaR-MDP, which has been studied, for example, by [15] and [18]. Such a reduction will be leveraged in Section IV to devise a value-iteration algorithm to compute tractable approximations of risk-sensitive safe sets.

A. Preliminaries

The reduction procedure is inspired by Chow et al. [15]. Specifically, we consider an augmented state space, $\mathcal{X} \times \mathcal{Y}$, that consists of the original state space, $\mathcal{X}$, and the space of confidence levels, $\mathcal{Y} := (0,1]$. The under-approximations of risk-sensitive safe sets will be defined in terms of the dynamics of the augmented state, $(x,y) \in \mathcal{X} \times \mathcal{Y}$, as explained next.

Let $(x_0, y_0) := (x, \alpha)$ be a given initial condition. The augmented state at time $k+1$, (x_{k+1}, y_{k+1}), depends on the augmented state at time k, (x_k, y_k), as follows. Given a control, $u_k \in U$, and a sampled disturbance, $w_k \in D$, the next state, $x_{k+1} \in \mathcal{X}$, satisfies the dynamics model (2). The next confidence level, $y_{k+1} \in \mathcal{Y}$, is given by,

$$y_{k+1} = \bar{R}_{x_k,y_k}(w_k) \cdot y_k, \tag{8}$$

where $\bar{R}_{x_k,y_k} : D \to (0, \frac{1}{y_k}]$ is a known deterministic function, which we will specify in Lemma 3. The augmented state space, $\mathcal{X} \times \mathcal{Y}$, is fully observable. Indeed, the history of

[3]Abate et al. [6] considers hybrid dynamic systems that evolve under Markov policies.

states and actions, $(x_0, u_0, \ldots, x_{k-1}, u_{k-1}, x_k)$, is available at time k by (2). Also, the history of confidence levels, $(y_0, \ldots, y_k)$, is available at time k, since the functions, $\bar{R}_{x_k,y_k}$, and the initial confidence level, $y_0 = \alpha$, are known.

We define the set of *deterministic, Markov* control policies in terms of the augmented state space as follows,

$$\begin{aligned} &\bar{\Pi}_t := \{(\bar{\mu}_t, \bar{\mu}_{t+1}, \ldots, \bar{\mu}_{N-1}) \mid \bar{\mu}_k : \mathcal{X} \times \mathcal{Y} \to U\}, \\ &t = 0, \ldots, N-1. \end{aligned} \tag{9}$$

There is an important distinction between the set of policies, $\bar{\Pi}_0$, as defined above, and the set of policies, Π, as defined in (3). Given $\bar{\pi}_0 \in \bar{\Pi}_0$, the control law at time k, $\bar{\mu}_k \in \bar{\pi}_0$, only depends on the current state, $x_k \in \mathcal{X}$, and the current confidence level, $y_k \in \mathcal{Y}$. However, given $\pi \in \Pi$, the control law at time k, $\mu_k \in \pi$, depends on the state history up to time k, $(x_0, \ldots, x_k) \in H_k$. In particular, the set of policies, $\bar{\Pi}_0$, is included in the set of policies, Π. This is because the augmented state at time k is uniquely determined by the initial confidence level and the state history up to time k.[4]

The benefits of considering $\bar{\Pi}_0$ instead of Π are two-fold. First, the computational requirements are reduced when the augmented state at time k, (x_k, y_k), is processed instead of the initial confidence level and the state history up to time k, $(y_0, x_0, x_1, \ldots, x_k)$. Second, we are able to define an under-approximation of the risk-sensitive safe set given by (5), using $\bar{\Pi}_0$, which we explain below.

B. Under-Approximation of Risk-Sensitive Safe Set

Define the set, $\mathcal{U}_\alpha^r \subseteq \mathcal{X}$, at the confidence level, $\alpha \in (0, 1]$, and the risk level, $r \in \mathbb{R}$,

$$\mathcal{U}_\alpha^r := \{x \in \mathcal{X} \mid J_0^*(x, \alpha) \leq \beta e^{m \cdot r}\}, \tag{10}$$

where

$$\begin{aligned} J_0^*(x, \alpha) &:= \min_{\pi \in \bar{\Pi}_0} \text{CVaR}_\alpha\left[Y_x^\pi\right], \\ Y_x^\pi &:= \textstyle\sum_{k=0}^N c(x_k), \end{aligned} \tag{11}$$

such that $c : \mathcal{X} \to \mathbb{R}$ is a stage cost, and the augmented state trajectory, $(x_0, y_0, \ldots, x_{N-1}, y_{N-1}, x_N)$, satisfies (2) and (8) with the initial condition, $(x_0, y_0) := (x, \alpha)$, under the policy, $\pi \in \bar{\Pi}_0$. The next theorem, whose proof is provided in the Appendix, states that if the stage cost takes a particular form, then $\mathcal{U}_\alpha^r$ is an under-approximation of the risk-sensitive safe set, $\mathcal{S}_\alpha^r$.

Theorem 1 (Reduction to CVaR-MDP): Choose the stage cost, $c(x) := \beta e^{m \cdot g(x)}$, where $\beta > 0$ and $m > 0$ are constants, and g satisfies (4). Then, $\mathcal{U}_\alpha^r$, as defined in (10), is a subset of $\mathcal{S}_\alpha^r$, as defined in (5). Further, the gap between $\mathcal{U}_\alpha^r$ and $\mathcal{S}_\alpha^r$ can be reduced by increasing m.

In the definition of the state costs, the parameter, β, is included to counter numerical issues that may arise, if m is set to a very large number.

[4] More formally, there exists an injective function, $h_k : \mathcal{Y} \times H_k \to \mathcal{X} \times \mathcal{Y}$, such that $h_k(y_0, x_0, x_1, \ldots, x_k) = (x_k, y_k)$; see (2) and (8). Given $\bar{\pi}_0 \in \bar{\Pi}_0$, the control at time k is $\bar{\mu}_k(x_k, y_k)$, which equals $\bar{\mu}_k(h_k(y_0, x_0, x_1, \ldots, x_k))$. Define $\mu_k(x_0, x_1, \ldots, x_k) := \bar{\mu}_k(h_k(y_0, x_0, x_1, \ldots, x_k))$ for all $y_0 \in \mathcal{Y}$. Note that μ_k is the control law at time k for a particular $\pi \in \Pi$. Thus, there is an injective function that maps $\bar{\Pi}_0$ to Π.

IV. A Value-Iteration Algorithm to Approximate Risk-Sensitive Safe Sets

By leveraging Theorem 1, one can use existing CVaR-MDP algorithms to compute under-approximations of risk-sensitive safe sets. In this paper, we adapt a value-iteration algorithm from Chow et al. [15] to compute tractable approximations of the risk-sensitive-safe-set under-approximations, $\{\mathcal{U}_\alpha^r\}$. We start by stating an existing temporal decomposition result for CVaR that will be instrumental to devising the value-iteration algorithm.

A. Temporal Decomposition of Conditional Value-at-Risk

In this section, we present an existing result (namely, Lemma 22 in [19]) that specifies how the Conditional Value-at-Risk of a sum of costs can be partitioned over time, and how the confidence level evolves over time, which motivates the choice of the update function (8).

Lemma 3 (Temporal decomposition of CVaR): At time k, suppose that the system (2) is at the state, $x_k \in \mathcal{X}$, with the confidence level, $y_k \in \mathcal{Y}$, and is subject to a policy, $\pi_k := (\mu_k, \pi_{k+1}) \in \bar{\Pi}_k$. Then,

$$\begin{aligned} &\text{CVaR}_{y_k}[Z|x_k, \pi_k] = \max_{R \in \mathcal{R}(y_k, \mathbb{P})} C(R, Z; x_k, y_k, \pi_k), \\ &C(R, Z; x_k, y_k, \pi_k) := \\ &\mathbb{E}_{w_k \sim \mathbb{P}}\big[R(w_k) \cdot \text{CVaR}_{y_k R(w_k)}[Z|x_{k+1}, \pi_{k+1}]\big|x_k, \mu_k\big], \end{aligned} \tag{12a}$$

where

$$\begin{aligned} &\mathcal{R}(y_k, \mathbb{P}) := \big\{R : D \to \big(0, \tfrac{1}{y_k}\big] \mid \mathbb{E}_{w_k \sim \mathbb{P}}\big[R(w_k)\big] = 1\big\}, \\ &Z := \textstyle\sum_{i=k+1}^N c(x_i), \end{aligned} \tag{12b}$$

such that $c : \mathcal{X} \to \mathbb{R}$ is a stage cost. Further, given the current state, (x_k, y_k), the current control, $u_k := \mu_k(x_k, y_k)$, and the next state, x_{k+1}, the function that was introduced in (8), $\bar{R}_{x_k,y_k} : D \to (0, \frac{1}{y_k}]$, is defined as,

$$\bar{R}_{x_k,y_k}(w_k) = \arg\max_{R \in \mathcal{R}(y_k, \mathbb{P})} C(R, Z; x_k, y_k, \pi_k). \tag{13}$$

Remark 1: The proof of Lemma 3 is a consequence of Lemma 22 in [19], and its proof is omitted for brevity.

Remark 2: If we do not have access to w_k, but only to (x_k, y_k, u_k, x_{k+1}), then the next confidence level is defined as, $y_{k+1} := \bar{R}_{x_k,y_k}(w)$, where $w \in D$ is a disturbance that satisfies $x_{k+1} = f(x_k, u_k, w)$.

Remark 3: $\text{CVaR}_{y_k}[Z|x_k, \pi_k]$ is the risk of the cumulative cost of the trajectory, $(x_{k+1}, \ldots, x_N)$, that is initialized at the state, x_k, with the confidence level, y_k, and is subject to the policy, $\pi_k \in \bar{\Pi}_k$.

B. Value-Iteration Algorithm

Using Lemma 3, we will devise a dynamic programming value-iteration algorithm to compute an approximation, J_0, of J_0^*, and thus, an approximation of $\mathcal{U}_\alpha^r$ at different levels of confidence, α, and risk, r.

Specifically, compute the functions, $J_{N-1}, \ldots, J_0$, recursively as follows: for all $z_k := (x_k, y_k) \in \mathcal{X} \times \mathcal{Y}$,

$$\begin{aligned} &J_k(z_k) \\ &:= \min_{u \in U} \Big\{ c(x_k) + \max_{R \in \mathcal{R}(y_k, \mathbb{P})} \mathbb{E}_{w_k \sim \mathbb{P}} \big[R J_{k+1}(x', y_k R) \big| z_k, u \big] \Big\}, \\ &\text{for } k = N-1, \ldots, 0, \end{aligned} \tag{14}$$

where $J_N(x_N, y_N) := c(x_N)$, $c(x) := \beta e^{m \cdot g(x)}$, $x' := x_{k+1}$ satisfies (2), and $\mathcal{R}(y_k, \mathbb{P})$ is defined in (12).

Then, we approximate the set, $\mathcal{U}_\alpha^r$, as $\widehat{\mathcal{U}}_\alpha^r := \{x \in \mathcal{X} \mid J_0(x, \alpha) \leq \beta e^{m \cdot r}\}$, where we have replaced J_0^* in (10) with J_0. The function, J_0, is obtained from the last step of the value iteration (14).

We present theoretical arguments, inspired by [15] and [4, Sec. 1.5], in the Appendix that justify such an approximation. In particular, we provide theoretical evidence for the following conjecture.

Conjecture (C): Assume that the functions, $J_{N-1}, \ldots, J_0$, are computed recursively as per the value-iteration algorithm (14). Then, for any $(x, \alpha) \in \mathcal{X} \times \mathcal{Y}$,

$$J_0(x, \alpha) = J_0^*(x, \alpha), \tag{15}$$

where J_0^* is given by (11).

This conjecture is further supported by the numerical experiments presented next.

V. Numerical Experiments

In this section, we provide empirical results that demonstrate the following: (1) a value-iteration estimate of J_0 is close to a Monte Carlo estimate of J_0^*, (2) a value-iteration estimate of $\hat{\mathcal{U}}_y^r$ is an under-approximation of a Monte Carlo estimate of $\mathcal{S}_y^r$, and (3) estimating J_0 (and $\hat{\mathcal{U}}_y^r$) via the value-iteration algorithm is tractable on a moderately large example. Item (1) provides empirical support for the conjecture. Items (2) and (3) provide empirical support for reducing the risk-sensitive-safe-set computation to a CVaR-MDP.

We demonstrate the utility of computing approximate risk-sensitive safe sets in a practical setting: to evaluate the design of a stormwater retention pond. We consider a retention pond from our prior work [20] as a stochastic discrete-time dynamic system,[5]

$$\begin{aligned} x_{k+1} &= x_k + \frac{\triangle t}{A}(w_k - q_p(x_k, u_k)), \;\; k = 0, \ldots, N-1, \\ q_p(x_k, u_k) &:= \begin{cases} C_d \pi r^2 u_k \sqrt{2\eta(x - E)} & \text{if } x_k \geq E \\ 0 & \text{if } x_k < E, \end{cases} \end{aligned} \tag{16}$$

where $x_k \geq 0$ is the water level of the pond in feet at time k, $u_k \in U := \{0, 1\}$ is the valve setting at time k, and $w_k \in D := \{d_1, \ldots, d_{10}\}$ is the random surface runoff in feet-cubed-per-second at time k. We estimated a finite probability distribution for w_k using the surface runoff samples that we previously generated from a time-varying design storm (a synthetic storm based on historical rainfall) [20]. We averaged each sample over time and solved for a distribution that satisfied the empirical statistics of the time-averaged samples (Table I). We set $\triangle t := 300$ seconds, and $N := 48$ to yield a 4-hour horizon. We set the constraint set, $\mathcal{K} := [0, 5\text{ft})$, and $g(x) := x - 5$. We computed over a grid of states and confidence levels, $G := G_s \times G_c$, where $G_s := \{0, 0.1\text{ft}, \ldots, 6.4\text{ft}, 6.5\text{ft}\}$, and $G_c := \{0.999, 0.95, 0.80, \ldots, 0.20, 0.05, 0.001\}$. Since the initial state, x_0, is non-negative and the smallest realization of w_k is about $8.5\frac{\text{ft}^3}{\text{s}}$, $x_{k+1} \geq x_k$ for all k. If $x_{k+1} > 6.5$ft, we set $x_{k+1} := 6.5$ft to stay within the grid.

[5] $\eta = 32.2\frac{\text{ft}}{\text{s}^2}$ is the acceleration due to gravity, $\pi \approx 3.14$, $r = \frac{1}{3}$ft is the outlet radius, $A = 28,292\text{ft}^2$ is the pond surface area, $C_d = 0.61$ is the discharge coefficient, and $E = 1$ft is the elevation of the outlet.

TABLE I

Sample moment	Value
Mean	12.16 ft^3/s
Variance	3.22 ft^6/s^2
Skewness	1.68 ft^9/s^3
Disturbance sample, d_j ft^3/s	**Probability**, $\mathbb{P}[w_k = d_j]$
8.57	0.0236
9.47	10^{-4}
10.37	10^{-4}
11.26	0.5249
12.16	0.3272
13.06	10^{-4}
13.95	10^{-4}
14.85	10^{-4}
15.75	10^{-4}
16.65	0.1237

We were able to empirically assess the accuracy of our proposed approach because an optimal control policy is known *a priori* for the one-pond system. Since $x_{k+1} \geq x_k$ for all k, and the only way to exit the constraint set is if $x_k \geq 5$ft, an optimal policy is to keep the valve open over time, regardless of the current state, the current confidence level, or the state history up to the current time.

A value-iteration estimate of the function, J_0, is shown in Fig. 1, and a Monte Carlo estimate of the function, J_0^*, is shown in Fig. 2. The estimates of J_0 and J_0^* are similar in most regions of the grid (except near the smaller confidence levels), which provides empirical support for the conjecture.

A value-iteration estimate of the set, $\widehat{\mathcal{U}}_\alpha^r$, and a Monte Carlo estimate of the set, $\mathcal{S}_y^r$, are shown in Fig. 3 at different levels of confidence, y, and risk, r. The empirical results indicate that $\widehat{\mathcal{U}}_\alpha^r$ is an under-approximation of $\mathcal{S}_y^r$. We estimated the sets, $\{\mathcal{S}_y^r\}$, using a Monte Carlo estimate of the function, W_0^*, which is shown in Fig. 4.

The computation time for our value-iteration estimate of J_0 was roughly 3h 6min. We deem this performance to be acceptable because 1) computations to evaluate design choices are performed off-line, 2) the problem instance entailed a moderately large state space ($|G_s| \cdot |G_c| = 594$ grid points) and time horizon ($N = 48$ time points), and 3) our implementation is not yet optimized. We used MATLAB R2016b (The MathWorks, Inc., Natick, MA) and MOSEK (Copenhagen, Denmark) with CVX [21] on a standard laptop (64-bit OS, 16.0 GB RAM, Intel®

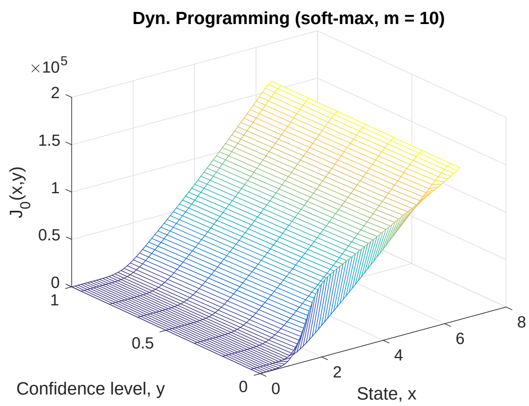

Fig. 1. A value-iteration estimate of $J_0(x, \alpha)$ versus $(x, \alpha) \in G$ for the pond system, see (14). $c(x) := \beta e^{m \cdot g(x)}$, $\beta := 10^{-3}$, $m := 10$, and $g(x) := x - 5$.

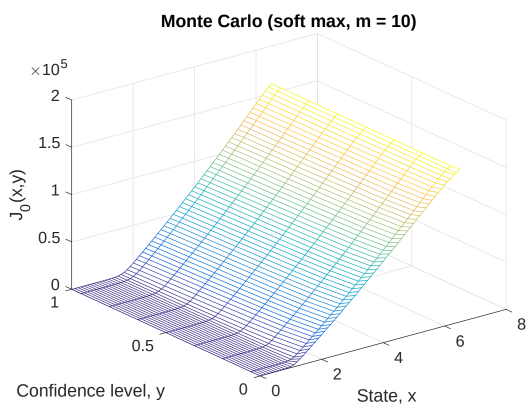

Fig. 2. A Monte Carlo estimate of $J_0^*(x, \alpha)$ versus $(x, \alpha) \in G$ for the pond system. $c(x) := \beta e^{m \cdot g(x)}$, $\beta := 10^{-3}$, $m := 10$, and $g(x) := x - 5$. 100,000 samples were generated per grid point. See also Fig. 1.

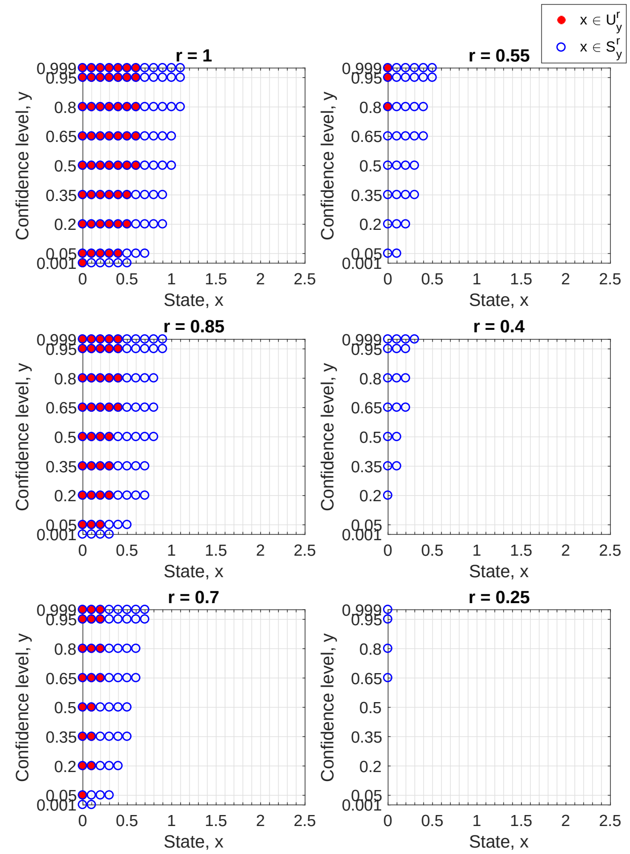

Fig. 3. Approximations of $\{\widehat{\mathcal{U}}_y^r\}$ and $\{\mathcal{S}_y^r\}$ are shown for the pond system at various levels of confidence, y, and risk, r. Approximations of $\{\widehat{\mathcal{U}}_y^r\}$ were obtained from a value-iteration estimate of J_0 (see Fig. 1). Approximations of $\{\mathcal{S}_y^r\}$ were obtained from a Monte Carlo estimate of W_0^* (see Fig. 4).

Core™ i7-4700MQ CPU @ 2.40GHz). Our code is available at `https://github.com/chapmanmp/Risk_Sensitive_Reachability_Project/tree/stormwater_example/MATLAB_Code`.

Next, we provide implementation details.

Value-iteration implementation. To implement the value-iteration algorithm, we used the interpolation method over the confidence levels proposed by Chow et al. [15] to approximate the expectation in (14) as a piecewise linear concave function, which we maximized by solving a linear program. Further, at each $\alpha \in G_c$, we used multi-linear interpolation to approximate the value of $J_{k+1}(x_{k+1}, \alpha)$. We set $J_{k+1}(x_{k+1}, \alpha) := \frac{(x_{k+1} - x_i) \cdot J_{k+1}(x^{i+1}, \alpha) + (x^{i+1} - x_{k+1}) \cdot J_{k+1}(x^i, \alpha)}{x_{i+1} - x_i}$, where $x^i \in G_s$ and $x^{i+1} \in G_s$ are the two nearest grid points to x_{k+1} that satisfy $x^i \leq x_{k+1} \leq x^{i+1}$. Fig. 1 shows the approximation of J_0, generated by the value-iteration algorithm (14), over the grid, G, using the interpolations just described.

Monte Carlo implementation. For each $(x, \alpha) \in G$, we sampled 100,000 trajectories starting from $x_0 := x$, subject to keeping the valve open over time, as this is an optimal

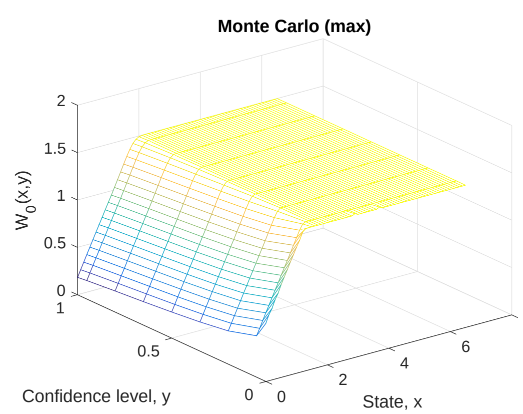

Fig. 4. A Monte Carlo estimate of $W_0^*(x, \alpha)$, as defined in (5), versus $(x, \alpha) \in G$ for the pond system. 100,000 samples were generated per grid point, and $g(x) := x - 5$. The maximum is 1.5ft because the system state was prevented from exceeding 6.5ft.

policy. For each trajectory sample i, we computed the cost sample, $z_i := \max\{g(x_k^i)\}$, and estimated the Conditional Value-at-Risk of the 100,000 cost samples at the confidence level, α. We used the CVaR estimator, $\widehat{\text{CVaR}}_\alpha[Z] := \frac{1}{\alpha M}\sum_{i=1}^{M} z_i \mathbf{1}_{\{z_i \geq \hat{Q}_\alpha\}}$, where $\hat{Q}_\alpha$ is the $(1-\alpha)$-quantile of the empirical distribution of the samples, $\{z_i\}_{i=1}^{M}$, and $M := 100,000$ is the number of samples [11, Sec. 6.5.1]. Since this estimator is designed for continuous distributions, we added zero-mean Gaussian noise with a small standard deviation, $\sigma := 10^{-12}$, to each cost sample prior to computing the CVaR. Fig. 4 provides a Monte Carlo estimate of W_0^*. To obtain a Monte Carlo estimate of J_0^*, we used an analogous procedure with the cost sample, $z_i := \xi_i + \sum_{k=0}^{N} \beta e^{m\cdot g(x_k^i)}$, $\xi_i \sim \mathcal{N}(0, \sigma := 10^{-7})$, $m := 10$, and $\beta = 10^{-3}$; see Fig. 2.

VI. Conclusion

In this paper, we propose the novel idea of a risk-sensitive safe set to encode safety of a stochastic dynamic system in terms of an allowable level of risk of constraint violations, r, in the α-fraction of the most damaging outcomes. We show how the computation of a risk-sensitive safe set can be reduced to the solution to a Markov Decision Process, where cost is assessed according to the Conditional Value-at-Risk measure. Further, we devise a tractable algorithm to approximate a risk-sensitive safe set, and provide theoretical and empirical arguments about its correctness.

Risk-sensitive safe sets have the potential to inform the design of safety-critical infrastructure systems, by revealing trade-offs between the risk of damaging outcomes and design choices, at different levels of confidence. We illustrate our risk-sensitive reachability approach on a stormwater retention pond that must be designed to operate safely in the presence of uncertain rainfall. Our results reveal that the current design of the pond is likely undersized: even if the pond starts empty, there is a risk of at least 0.25ft of overflow, at most levels of confidence, under the random surface runoff of the design storm (see Fig. 3, $r = 0.25$ plot at $x = 0$).

On the methodological side, future steps include: 1) formally prove the correctness of the value-iteration algorithm, 2) devise approximate value-iteration algorithms to improve scalability, and 3) consider a broader class of risk measures. On the applications side, future steps include: 1) adjust the parameters of the dynamics model (e.g., outlet radius) to reduce the risk of extreme overflows, 2) apply our method to a more realistic stormwater system that consists of two ponds in series on a larger grid, and 3) develop an optimized toolbox for the computation of risk-sensitive safe sets. We are hopeful that with further development the concept of risk-sensitive reachability will become a valuable tool for the design of safety-critical systems.

Acknowledgments

We thank Dr. Sumeet Singh, Dr. Mo Chen, Dr. Murat Arcak, and Dr. Alessandro Abate for discussions. M.C. is supported by an NSF Graduate Research Fellowship and was supported by a Berkeley Fellowship for Graduate Studies. This work is supported by NSF CPS 1740079 and NSF PIRE.

Appendix

Proof: [Proof of Theorem 1] The proof relies on two facts. The first fact is,

$$\begin{aligned}\max\{y_1,\dots,y_p\} &\leq \frac{1}{m}\log(e^{my_1}+\dots+e^{my_p}) \\ &\leq \max\{y_1,\dots,y_p\} + \frac{\log p}{m},\end{aligned} \tag{17a}$$

for any $y \in \mathbb{R}^p$, $m > 0$.[6] So, as $m \to \infty$,

$$\frac{1}{m}\log(e^{my_1}+\dots+e^{my_p}) \to \max\{y_1,\dots,y_p\}. \tag{17b}$$

The second fact is that Conditional Value-at-Risk is a *coherent risk measure*, so it satisfies useful properties. In particular, CVaR is *positively homogeneous*, $\text{CVaR}_\alpha[\lambda Z] = \lambda\text{CVaR}_\alpha[Z]$, for any $\lambda \geq 0$, and *monotonic*, $\text{CVaR}_\alpha[Y] \leq \text{CVaR}_\alpha[Z]$, for any random variables, $Y \leq Z$ [12, Sec. 2.2]. Also, CVaR can be expressed as the supremum expectation over a particular set of probability density functions [11, Eqs. 6.40 and 6.70]. Using this property and the fact, $\mathbb{E}[\log(Z)] \leq \log(\mathbb{E}[Z])$, one can show,

$$\text{CVaR}_\alpha[\log(Z)] \leq \log\left(\text{CVaR}_\alpha[Z]\right), \tag{18}$$

for any random variable, Z, with finite expectation.

By monotonicity, positive homogeneity, (17), and (18),

$$\text{CVaR}_\alpha\big[Z_x^\pi\big] \leq \tfrac{1}{m}\text{CVaR}_\alpha\big[\log\big(\bar{Y}_x^\pi\big)\big] \leq \tfrac{1}{m}\log\big(\text{CVaR}_\alpha\big[\bar{Y}_x^\pi\big]\big), \tag{19}$$

where $\bar{Y}_x^\pi := Y_x^\pi/\beta$. Now, if $x \in \mathcal{U}_\alpha^r$, then

$$e^{m\cdot r} \geq \min_{\pi\in\bar{\Pi}_0} \text{CVaR}_\alpha\big[Y_x^\pi/\beta\big] \geq \min_{\pi\in\Pi} \text{CVaR}_\alpha\big[Y_x^\pi/\beta\big],$$

since $\bar{\Pi}_0$ is included in Π. By Lemma 1, there exists $\pi \in \Pi$ such that,

$$r \geq \tfrac{1}{m}\log\big(\text{CVaR}_\alpha\big[Y_x^\pi/\beta\big]\big) \geq \text{CVaR}_\alpha\big[Z_x^\pi\big],$$

where the second inequality holds by (19). So, $x \in \mathcal{S}_\alpha^r$. ■

Proof: [Theoretical Justification of Conjecture (C)] Let $\epsilon > 0$. For all $k = 0,\dots,N-1$ and $z_k := (x_k, y_k) \in \mathcal{X}\times\mathcal{Y}$, let $\mu_k^\epsilon : \mathcal{X}\times\mathcal{Y} \to U$ satisfy,

$$c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}[RJ_{k+1}(x_{k+1}, y_k R)|z_k, \mu_k^\epsilon] \leq J_k(z_k) + \epsilon. \tag{20}$$

Let J_k^ϵ be a sub-optimal cost-to-go starting at time k,

$$J_k^\epsilon(z_k) := \text{CVaR}_{y_k}\big[\textstyle\sum_{i=k}^{N} c(x_i)\big|z_k, \pi_k^\epsilon\big], \tag{21}$$

where $\pi_k^\epsilon := (\mu_k^\epsilon,\dots,\mu_{N-1}^\epsilon) = (\mu_k^\epsilon, \pi_{k+1}^\epsilon)$. Recall J_k, as defined in (14). Define $J_k^*(z_k) := \min_{\pi\in\bar{\Pi}_k} \text{CVaR}_{y_k}\big[\sum_{i=k}^{N} c(x_i)\big|z_k, \pi\big]$. To prove the conjecture, we would like to show by induction that for all $z_k := (x_k, y_k) \in \mathcal{X}\times\mathcal{Y}$ and $k = N-1,\dots,0$,

$$J_k(z_k) \leq J_k^\epsilon(z_k) \leq J_k(z_k) + (N-k)\epsilon, \tag{22a}$$

$$J_k^*(z_k) \leq J_k^\epsilon(z_k) \leq J_k^*(z_k) + (N-k)\epsilon, \tag{22b}$$

$$J_k(z_k) = J_k^*(z_k), \tag{22c}$$

[6] Use the log-sum-exp relation stated in [22, Sec. 3.1.5]

which is the proof technique in [4, Sec. 1.5]. One can show (22) for the base case, $k := N-1$, since J_N is known. Assuming that (22) holds for index $k+1$ (induction hypothesis), we want to show that (22) holds for index k (induction step). The key idea is to use the following recursion,

$$J_k^\epsilon(z_k) = c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}[R\cdot J_{k+1}^\epsilon(x_{k+1}, y_k R)|z_k, \mu_k^\epsilon], \tag{23}$$

which we justify next. Let $Z := \sum_{i=k+1}^N c(x_i)$.

$$\begin{aligned} J_k^\epsilon(z_k) - c(x_k) &= \text{CVaR}_{y_k}\big[Z\big|z_k, \pi_k^\epsilon\big] \\ &= \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R\cdot \text{CVaR}_{y_k R}[Z|x_{k+1}, \pi_{k+1}^\epsilon]\big|z_k, \mu_k^\epsilon\big] \\ &\overset{(a)}{=} \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R\cdot J_{k+1}^\epsilon(x_{k+1}, y_k R)\big|z_k, \mu_k^\epsilon\big], \end{aligned}$$

where we use (12), (21), and translation invariance[7]. The last equality (a) is *the crux of our conjecture*, as one needs to justify why the worst-case density, R, is equal to the *a priori* chosen density, $\bar{R}$, that defines the dynamics of the confidence level. Based on [15], we believe this equality to be correct, but we leave its formal proof for future research. Assuming the aforementioned equality is correct, then we show (22a) for index k using (23) and the induction hypothesis. Let $\bar{\epsilon}_k := (N-k-1)\epsilon$, and $x' := x_{k+1}$.

$$\begin{aligned} J_k^\epsilon(z_k) &\leq c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R\big(J_{k+1}(x', y_k R) + \bar{\epsilon}_k\big)\big|z_k, \mu_k^\epsilon\big] \\ &= c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R J_{k+1}(x_{k+1}, y_k R)\big|z_k, \mu_k^\epsilon\big] + \bar{\epsilon}_k \\ &\leq J_k(z_k) + (N-k)\epsilon, \end{aligned}$$

since $\mathbb{E}[R] = 1$, and by (20). By (14), sub-optimality of $\mu_k^\epsilon(z_k) \in U$, $J_{k+1} \leq J_{k+1}^\epsilon$, and (23),

$$\begin{aligned} J_k(z_k) &\leq c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R J_{k+1}^\epsilon(x_{k+1}, y_k R)\big|z_k, \mu_k^\epsilon\big] \\ &\leq J_k^\epsilon(z_k). \end{aligned}$$

The induction step for (22a) would be complete. Next, we show (22b) for index k. By definition, J_k^* is the optimal risk-sensitive cost-to-go from stage k, thus, $J_k^* \leq J_k^\epsilon$. Let $\hat{\epsilon}_k := (N-k)\epsilon$, $x' := x_{k+1}$, $y' := y_k R$, and $Z := \sum_{i=k+1}^N c(x_i)$. For any $\pi_k := (\mu_k, \pi') \in \bar{\Pi}_k$,

$$\begin{aligned} J_k^\epsilon(z_k) &\leq J_k(z_k) + \hat{\epsilon}_k \\ &\leq c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R J_{k+1}^*(x_{k+1}, y_k R)\big|z_k, \mu_k\big] + \hat{\epsilon}_k \\ &\leq c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}\big[R\text{CVaR}_{y'}[Z|x', \pi']\big|z_k, \mu_k\big] + \hat{\epsilon}_k \\ &= c(x_k) + \text{CVaR}_{y_k}[Z|z_k, \pi_k] + \hat{\epsilon}_k \\ &= \text{CVaR}_{y_k}\big[\textstyle\sum_{i=k}^N c(x_k)|z_k, \pi_k\big] + (N-k)\epsilon. \end{aligned}$$

The above statement implies

$$\begin{aligned} J_k^\epsilon(z_k) &\leq \min_{\pi\in\bar{\Pi}_k} \text{CVaR}_{y_k}\big[\textstyle\sum_{i=k}^N c(x_k)|z_k, \pi_k\big] + (N-k)\epsilon \\ &= J_k^*(z_k) + (N-k)\epsilon, \end{aligned}$$

which would complete the induction step for (22b).

[7]For $a \in \mathbb{R}$, $\text{CVaR}_\alpha[a+Z] = a + \text{CVaR}_\alpha[Z]$; see [12, Sec. 2.2].

We have shown that (22a) and (22b) hold for index k, for any $\epsilon > 0$. So, (22c) holds for index k. Assuming the conjectured equality (a) is correct, this would complete the proof of the conjecture. ■

REFERENCES

[1] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi Reachability: A Brief Overview and Recent Advances," *arXiv preprint arXiv:1709.07523*, 2017.

[2] D. P. Bertsekas, "Control of Uncertain Systems with a Set-Membership Description of the Uncertainty," Ph.D. dissertation, Massachusetts Institute of Technology, 1971.

[3] D. P. Bertsekas and I. B. Rhodes, "On the Minimax Reachability of Target Sets and Target Tubes," *Automatica*, vol. 7, no. 2, pp. 233–247, 1971.

[4] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 4th ed. Athena Scientific, 2017, vol. 1.

[5] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, 1998.

[6] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724–2734, 2008.

[7] S. Summers and J. Lygeros, "Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem," *Automatica*, vol. 46, no. 12, pp. 1951–1961, 2010.

[8] S. L. Herbert, M. Chen, S. Han, S. Bansal, J. F. Fisac, and C. J. Tomlin, "FaSTrack: A Modular Framework for Fast and Guaranteed Safe Motion Planning," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*. IEEE, 2017, pp. 1517–1522.

[9] A. Akametalu, "A learning-based approach to safety for uncertain robotic systems," Ph.D. dissertation, EECS Department, University of California, Berkeley, May 2018. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-41.html

[10] I. M. Mitchell and J. A. Templeton, "A Toolbox of Hamilton-Jacobi Solvers for Analysis of Nondeterministic Continuous and Hybrid Systems," in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2005, pp. 480–494.

[11] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial and Applied Mathematics, Mathematical Programming Society, 2009.

[12] J. Kisiala, "Conditional Value-at-Risk: Theory and Applications," Master's thesis, The School of Mathematics, The University of Edinburgh, August 2015.

[13] A. Ruszczyński, "Risk-averse dynamic programming for Markov decision processes," *Mathematical Programming*, vol. 125, no. 2, pp. 235–261, 2010.

[14] T. Osogami, "Robustness and Risk-Sensitivity in Markov Decision Processes," in *Advances in Neural Information Processing Systems*, 2012, pp. 233–241.

[15] Y. Chow, A. Tamar, S. Mannor, and M. Pavone, "Risk-Sensitive and Robust Decision-Making: a CVaR Optimization Approach," in *Advances in Neural Information Processing Systems*, 2015, pp. 1522–1530.

[16] L. J. Ratliff and E. Mazumdar, "Risk-sensitive inverse reinforcement learning via gradient methods," *arXiv preprint arXiv:1703.09842*, 2017.

[17] Y.-L. Chow and M. Pavone, "A Framework for Time-consistent, Risk-Averse Model Predictive Control: Theory and Algorithms," in *American Control Conference*. IEEE, 2014, pp. 4204–4211.

[18] W. B. Haskell and R. Jain, "A convex analytic approach to risk-aware markov decision processes," *SIAM Journal on Control and Optimization*, vol. 53, no. 3, pp. 1569–1598, 2015.

[19] G. C. Pflug and A. Pichler, "Time-consistent decisions and temporal decomposition of coherent risk functionals," *Mathematics of Operations Research*, vol. 41, no. 2, pp. 682–699, 2016.

[20] M. P. Chapman, K. M. Smith, V. Cheng, D. Freyberg, and C. J. Tomlin, "Reachability Analysis as a Design Tool for Stormwater Systems," in *6th IEEE Conference on Technologies for Sustainability*, November 2018.

[21] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab Software for Disciplined Convex Programming," 2008.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.