A Risk-Sensitive Finite-Time Reachability Problem 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 formulate a *risk-sensitive* reachability problem for safety of stochastic dynamic systems under non-adversarial disturbances over a finite time horizon. We provide two key contributions to the reachability literature. First, our formulation approximates the distance between the boundary of the constraint set and the state trajectory for a stochastic dynamic system. In the literature, Hamilton-Jacobi (HJ) reachability methods quantify this distance for non-deterministic systems subject to adversarial disturbances, while stochastic reachability methods replace the distance with a binary random variable in order to quantify the probability of safety. Second, our formulation accounts for rare high-consequence events by posing the optimal control problem in terms of a risk measure, called *Conditional Value-at-Risk* (CVaR). HJ reachability assumes that high-consequence events occur always, which may yield overly conservative solutions in practice, whereas stochastic reachability does not explicitly account for rare high-consequence events, since the optimal control problem is posed in terms of the expectation operator. We define 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 CVaR. We show that certain risk-sensitive safe sets enjoy probabilistic safety guarantees. We provide a dynamic programming algorithm with theoretical justification to compute tractable approximations of risk-sensitive safe sets. Finally, we demonstrate the utility of risk-sensitive reachability analysis on a numerical example.**

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].[1]

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, and then assess safety properties of the system in a more realistic context.[2] For stochastic discrete-time dynamic systems, Abate et al. 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 [6].[3] Summers and Lygeros 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 [7].

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 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]. 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.

[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]in ref. [4], see Sec. 3.6.2, "Control within a Target Tube"

[2]Ref. [5] presents methods for estimating probability distributions.

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

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-consequence events because the optimal control problem is expressed as an expectation.

In contrast, we leverage existing results on *risk measures* to formulate an optimal control problem that explicitly encodes a realistic viewpoint on the possibility of rare high-consequence events: harmful events are likely to occur at some point, but they are unlikely to occur always. A *risk measure* is a function that maps a random variable, Z, representing loss, into the real line, according to the risk associated with Z (see [11], Sec. 6.3; see [12], Sec. 2.2). Risk-sensitive optimization is being 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.

We use a particular risk measure, called *Conditional Value-at-Risk* (CVaR), in this paper. If Z is a random cost with finite expectation, then the Conditional Value-of-Risk of Z at confidence level $\alpha \in (0, 1]$ is,

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

see [11], Equation 6.22.[4] Note that $\text{CVaR}_\alpha[Z]$ increases from $\mathbb{E}[Z]$ to $\operatorname{ess\,sup} Z$, as α decreases from 1 to 0.[5] Further, there is a well-established relationship between CVaR and chance constraints that we use to obtain probabilistic safety guarantees. Chow et al. provides tractable methods to compute the CVaR of a cumulative cost incurred by a Markov Decision Process [15] that we also leverage. CVaR has additional desirable properties that are of particular interest to researchers in financial risk management and are summarized in ref. [18].

The key contributions of this paper follow. We formulate a risk-sensitive reachability problem for safety of stochastic dynamic systems under non-adversarial disturbances over a finite-time horizon. In particular, our formulation approximates 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 Conditional Value-at-Risk instead of the expectation operator. This is the first use of (non-neutral) risk measures in the reachability literature to our knowledge. We formally define risk-sensitive safe sets in Sec. II and justify our definition in Sec. III. Sec. IV presents a value-iteration algorithm with theoretical justification to compute tractable approximations of risk-sensitive safe sets. Sec. V provides a numerical example in the domain of stormwater infrastructure design. Sec. VI concludes the paper.

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

[5]The essential supremum of a random variable, Z, is the supremum of the realizations of Z that occur with non-zero probability.

II. Problem Statement

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

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

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 dynamics 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 $\sum_{j=1}^{W} p_j = 1$. 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,

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

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. Let $g : \mathcal{X} \to \mathbb{R}$ satisfy,

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

which is also done by [9] in Equation 2.3. For example, we may choose $g(x) := x - 5$ to characterize $\mathcal{K} := [0, 5\text{ft})$ on the state space, $\mathcal{X} := [0, \infty)$.

A *risk-sensitive safe set* is a set of initial states from which the risk of extreme constraint violation over time can be made small using an admissible control policy, where we quantify risk using the *Conditional Value-at-Risk* measure. We use the term, *risk level*, to mean the allowable level of risk of constraint violation. 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,

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

where

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

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 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: For any initial state, $x \in \mathcal{X}$, 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 are possibly 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^*. ∎

The goal of this paper is to compute a family of risk-sensitive safe sets at different levels of confidence, $\alpha \in (0, 1]$, and risk, $r \in \mathbb{R}$.

III. Rationale

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 [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. In particular, 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,

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

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

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 the risk-sensitive safe sets at risk level, $r := 0$, enjoy probabilistic safety guarantees.

[6]See [6], Equations 11 and 13. Note that ref. [6] considers hybrid dynamic systems that evolve under Markov policies.

Lemma 2: 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 $\text{CVaR}_\alpha[Z_x^\pi] \leq 0 \implies \mathbb{P}[Z_x^\pi \geq 0] \leq \alpha$ (see [11], Sec. 6.2.4). 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 $g(x_k) \geq 0 \iff x_k \notin \mathcal{K}$. ∎

Remark 1: 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 (see [6], Equations 9 and 11).

IV. Computational Method

Computing risk-sensitive safe sets is challenging because the computation involves a maximum of costs (as opposed to a summation of costs) and the Conditional Value-at-Risk measure (as opposed to the expectation operator). Here we provide a value-iteration algorithm with theoretical justification to compute tractable approximations of risk-sensitive safe sets. Specifically, we provide under-approximations of risk-sensitive safe sets and a well-justified algorithm to approximate the under-approximations. The computational method is inspired by Chow et al. [15]. 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 are defined in terms of the dynamics of the augmented state, $(x, y) \in \mathcal{X} \times \mathcal{Y}$. In the following section, we explain these dynamics in detail.

A. Dynamics of the Augmented State

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 4. Note that the augmented state space, $\mathcal{X} \times \mathcal{Y}$, is fully observable. Indeed, the history of states and actions, $(x_0, u_0, \ldots, x_{k-1}, u_{k-1}, x_k)$, is available at time k, according to (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 sets 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, $(x_0, \dots, x_k) \in H_k$. 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.[7]

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, $(y_0, x_0, x_1, \dots, x_k)$. 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 in the following sections.

B. Under-Approximation of Risk-Sensitive Safe Set

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

$$\mathcal{U}^r_\alpha := \{x \in \mathcal{X} \mid J^*_0(x, \alpha) \leq \beta e^{m \cdot r}\}, \tag{10}$$

where

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

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

Lemma 3: 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, $\mathcal{U}^r_\alpha$, as defined in (10), is a subset of $\mathcal{S}^r_\alpha$, as defined in (5). Further, the gap between $\mathcal{U}^r_\alpha$ and $\mathcal{S}^r_\alpha$ can be made smaller by increasing m.

Remark 2: β is included in the definition of c to help counter numerical issues that may arise if m is very large. The proof of Lemma 3 is provided in the Appendix.

C. Temporal Decomposition of Conditional Value-at-Risk

Here we state an existing result that motivates a value-iteration algorithm to compute tractable approximations of the risk-sensitive-safe-set under-approximations, $\{\mathcal{U}^r_\alpha\}$. In particular, the result specifies how the CVaR of a sum of costs can be partitioned over time and how the confidence level evolves over time, which we initially presented in (8).

Lemma 4: Lemma 22 in [19] implies the following CVaR-decomposition for the system (2) at time k, initialized at $x_k \in \mathcal{X}$, with the confidence level, $y_k \in \mathcal{Y}$, under a given policy, $\pi_k := (\mu_k, \pi_{k+1}) \in \bar{\Pi}_k$. Then,

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

where

$$\begin{aligned} &\mathcal{R}(x_k, y_k, \mu_k, \mathbb{P}) \\ &:= \big\{R : D \to (0, \tfrac{1}{y_k}] \;\big|\; \textstyle\sum_{j=1}^W R(w_k)\mathbb{P}[w_k = d_j] = 1\big\}, \\ &Z := \textstyle\sum_{i=k+1}^N c(x_i), \end{aligned} \tag{12b}$$

such that $c : \mathcal{X} \to \mathbb{R}$ is a stage cost. Further, given the current state (x_k, y_k), the current control $u_k := \mu(x_k, y_k)$ 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,

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

Remark 3: 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 y_{k+1} 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 4: The proof of Lemma 4, which we leave out due to lack of space, is an application of Lemma 22 in [19].

Remark 5: $\text{CVaR}_\alpha[Z|x_k, y_k, \pi_k]$ is the risk of the cumulative cost starting at time $k+1$ of the trajectory, $(x_k, \dots, x_N)$, that is initialized at the augmented state, (x_k, y_k), and evolves under the policy, $\pi_k \in \bar{\Pi}_k$.

D. Value-Iteration Algorithm

In this section, we use Lemma 4 to derive a value-iteration dynamic programming algorithm that computes a lower bound, J_0, on J^*_0. We will implement the algorithm with the stage cost, $c(x) := \beta e^{m \cdot g(x)}$, to approximate the set, $\mathcal{U}^r_\alpha$, which is itself an under-approximation of the set, $\mathcal{S}^r_\alpha$, at different levels of confidence, α, and risk, r.

Theorem 1: Define the functions, J_{N-1}, …, J_0, recursively as follows, $\forall z_k := (x_k, y_k) \in \mathcal{X} \times \mathcal{Y}$,

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

for $k = N-1, \dots, 0$, where $J_N(x_k, y_k) := c(x_k)$, $x' := x_{k+1}$ satisfies (2), and $\mathcal{R}(z_k, u, \mathbb{P})$ is defined in (12). Then, for any $(x, \alpha) \in \mathcal{X} \times \mathcal{Y}$,

$$J_0(x, \alpha) \leq J^*_0(x, \alpha). \tag{15}$$

Using the previous dynamic programming recursion, we can compute $\widehat{\mathcal{U}}^r_\alpha := \{x \in \mathcal{X} \mid J_0(x, \alpha) \leq \beta e^{mr}\}$, which is an approximation of $\mathcal{U}^r_\alpha$. The next conjecture is for future work.

Conjecture (C): $\forall (x, \alpha) \in \mathcal{X} \times \mathcal{Y}$, $J_0(x, \alpha) = J^*_0(x, \alpha)$ and $\widehat{\mathcal{U}}^r_\alpha = \mathcal{U}^r_\alpha$.

Remark 6: Chow et al. proposed the recursion (14) and applied it to the infinite-horizon discounted problem [15].

[7] More formally, there exists an injective function, $h_k : \mathcal{Y} \times H_k \to \mathcal{X} \times \mathcal{Y}$, such that $h_k(y_0, x_0, x_1, \dots, x_k) = (x_k, y_k)$; 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 $\bar{\mu}_k(h_k(y_0, x_0, x_1, \dots, x_k))$. Define $\mu_k(x_0, x_1, \dots, x_k) := \bar{\mu}_k(h_k(y_0, x_0, x_1, \dots, x_k))$ 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.

They prove that inequality (15) is actually an exact equality. Their arguments require a strong duality assumption, which we do not assume here.

Remark 7: The policy given by (14) requires that the current state and the current confidence level are available. Constructing a policy that depends on the state history and the initial confidence level is important future work, which may require different arguments than those used by Chow et al. [15] due to the finite-time setting of the current paper.

Our proof of Theorem 1 is provided in the Appendix. The proof idea is to use a sub-optimal value function as machinery to demonstrate that each J_k, as defined recursively in (14), is close to the optimal cost-to-go of the sub-problem that starts at time k,

$$J_k^*(x_k, y_k) := \min_{\pi_k \in \Pi_k} \text{CVaR}_{y_k}\left[\textstyle\sum_{i=k}^{N} c(x_i) \middle| (x_k, y_k), \pi_k\right], \tag{16}$$

via induction. This technique is also used to prove the classic finite-time value-iteration algorithm, where the value function is the expected cumulative cost (see [4], Sec. 1.5).

V. Numerical Example

We demonstrate the utility of computing approximate risk-sensitive safe sets on a practical example: to evaluate the design of a stormwater retention pond. Stormwater management facilities, such as retention ponds, are required to operate safely in the presence of precipitation uncertainty, but must be designed within the scope of public resources (e.g., money, land). Standard design practices assess how empty ponds respond to a given *design storm*, which is a synthetic storm based on historical rainfall. In our prior work, we proposed using reachability analysis to augment existing design practices, as it can assess system behavior from a larger number of initial conditions, but we assumed that the surface runoff due to the design storm was deterministic [20]. Here we consider the first pond from the example in our prior work [20] as a stochastic discrete-time dynamic system,

$$\begin{aligned} x_{k+1} &= x_k + \frac{\triangle t}{A}(w_k - q_p(x_k, u_k)), \ k = 0, \dots, N-1, \\ q_p(x_k, u_k) &:= \begin{cases} C_d \pi r^2 u_k \sqrt{2\eta(x - E)} & \text{if } x_k \geq E \\ 0 & \text{if } x_k < E, \end{cases} \end{aligned} \tag{17}$$

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 $w_k \in D := \{d_1, \dots, d_{10}\}$ is the random surface runoff in feet-cubed-per-second at time k. In Eq. (17), $\eta = 32.2 \frac{\text{ft}}{\text{s}^2}$ is the acceleration due to gravity, $\pi \approx 3.14$ is the usual constant, $r = \frac{1}{3}$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. We estimated a finite probability distribution for w_k using the surface runoff samples that we previously generated from a time-varying design storm [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 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® Core™ i7-4700MQ CPU @ 2.40GHz). Our code is at `https://github.com/chapmanmp/Risk_Sensitive_Reachability_Project/tree/stormwater_example/MATLAB_Code`.

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 computed approximations of $\{\widehat{\mathcal{U}}_y^r\}$ using the value-iteration algorithm (14), and approximations of $\{\mathcal{S}_y^r\}$ using a Monte Carlo procedure. Fig. 1 shows these approximations at different levels of confidence and risk. We computed over a finite grid of states and confidence levels, $G := G_s \times G_c$, where $G_s := \{0, 0.1\text{ft}, \dots, 6.4\text{ft}, 6.5\text{ft}\}$, and $G_c := \{0.999, 0.95, 0.80, \dots, 0.20, 0.05, 0.001\}$. 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.

Value-iteration implementation. The results for $\{\widehat{\mathcal{U}}_y^r\}$ are approximations because the computations were done over a grid using linear interpolations. 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 (LP). 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 $x^i \leq x_{k+1} \leq x^{i+1}$. We chose the stage cost, $c(x) := \beta e^{m \cdot g(x)}$, such that $\beta := 10^{-3}$ and $m := 10$. Fig. 2 shows our approximation of J_0, generated by the value-iteration algorithm (14), over the grid, G, using the interpolations just described. The computation time was roughly 3h 6min, which is fast given that at each time point an LP was solved for each grid point, yielding $N \cdot |G_s| \cdot |G_c| = 48 \cdot 66 \cdot 9 = 28,512$ LP's in total. A more efficient implementation would require solving only $N = 48$ LP's in total, where each LP encodes the entire grid.

Monte Carlo implementation. An optimal control policy

was known *a priori* for our one-pond system, which made a Monte Carlo implementation feasible. (If prior information on an optimal policy was not available, then $|U|^N = 2^{48}$ possible control signals would need to be simulated.) 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, regardless of the state history up to the current time. For each $(x, \alpha) \in G$, we sampled 100,000 trajectories starting from $x_0 := x$, subject to keeping the valve open over time. 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; see [11], Sec. 6.5.1. Since the 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. 3 provides a Monte Carlo estimate of W_0^*, as defined in (5).

100,000 samples per grid point appears to be sufficient. Also using this number of samples, we estimated $\min_{\pi \in \Pi} \text{CVaR}_\alpha\left[Y_x^\pi\right]$ via an analogous Monte Carlo procedure as just described, where Y_x^π is the random sum of stage costs (Fig. 4).[8] See that Fig. 4 and Fig. 2 are comparable in most regions of the grid. However, our Monte Carlo estimate does not provide the higher costs at the smallest confidence level, $\alpha = 0.001$, that are provided by our value-iteration estimate.

VI. Conclusion

In this paper, we propose the novel idea of a risk-sensitive safe set that encodes the safety of a stochastic dynamic system in terms of an allowable level of risk of constraint violation, r, in the α-fraction of the most damaging outcomes. We reduce the computation of risk-sensitive safe sets to the problem of optimizing the Conditional Value-at-Risk of the maximum extent of constraint violation over the state trajectory. Further, we provide a dynamic programming algorithm with theoretical justification to compute tractable approximations of risk-sensitive safe sets.

Risk-sensitive safe sets have 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 formulation on a stormwater 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. 1, $r = 0.25$ plot at $x = 0$).

On the applications side, future steps include: 1) adjust the parameters of the dynamics model (e.g., outlet radius) to

[8]We added zero-mean Gaussian noise with a small standard deviation, $\sigma := 10^{-7}$, to each sample of Y_x^π prior to computing the CVaR. See that the magnitude of J_0 (Fig. 2) is about 10^5-times greater than the magnitude of W_0 (Fig. 3).

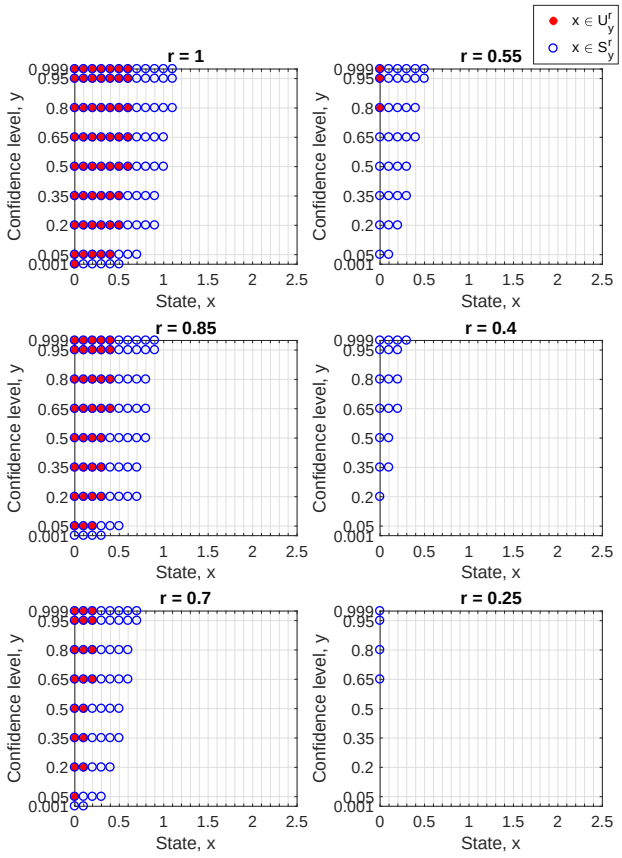

Fig. 1. Approximations of $\{\widehat{\mathcal{U}}_y^r\}$ and $\{\mathcal{S}_y^r\}$ are shown for the pond system at various levels of confidence, y, and risk, r. $\{\widehat{\mathcal{U}}_y^r\}$ were approximated using the value-iteration algorithm (14), and $\{\mathcal{S}_y^r\}$ were approximated via a Monte Carlo procedure.

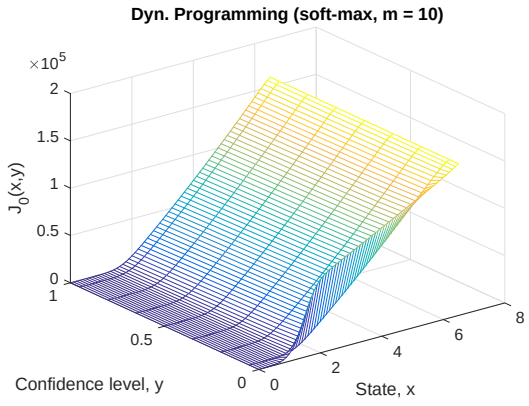

Fig. 2. Our value-iteration estimate of $J_0(x, \alpha)$ versus $(x, \alpha) \in G$ for the pond system, see (14). $c(x) := \beta e^{m \cdot g(x)}$, $\beta := 10^{-3}$, $m := 10$, and $g(x) := x - 5$.

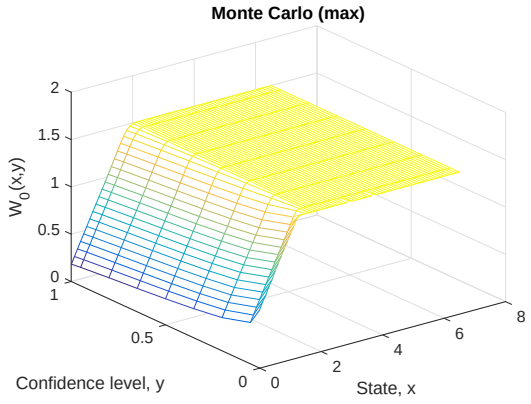

Fig. 3. A Monte Carlo estimate of $W_0^*(x,\alpha)$, as defined in (5), versus $(x,\alpha) \in G$ for the pond system. 100,000 samples were generated per grid point, and $g(x) := x - 5$. The maximum is 1.5ft because the system state was prevented from exceeding 6.5ft.

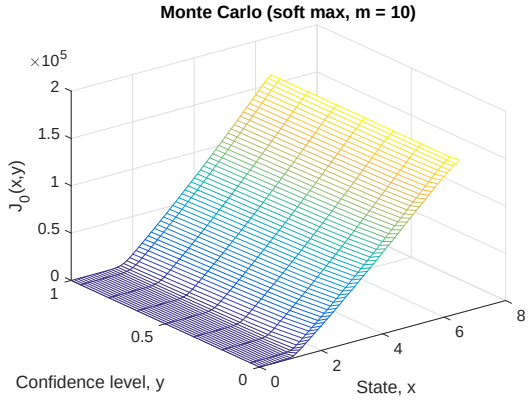

Fig. 4. A Monte Carlo estimate of $\min_{\pi\in\Pi} \text{CVaR}_\alpha\big[Y_x^\pi\big]$ versus $(x,\alpha) \in G$ for the pond system. $Y_x^\pi := \sum_{k=0}^N \beta e^{m\cdot g(x_k)}$, $x_0 := x$, $\beta := 10^{-3}$, $m := 10$, and $g(x) := x - 5$. 100,000 samples were generated per grid point. See also Fig. 2.

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) solve one linear program that encodes the entire grid per time point. On the methods side, we seek to prove our conjecture, possibly using another state augmentation. We are hopeful that with further development the concept of risk-sensitive safe sets will become a valuable tool for the design of safety-critical systems.

ACKNOWLEDGMENTS

We thank Dr. Sumeet Singh, Dr. Mo Chen, and Dr. Murat Arcak 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

We prove Lemma 3 below.

Proof: The proof relies on two facts. The first fact is,

$$\begin{aligned}\max\{y_1,\dots,y_p\} &\le \frac{1}{m}\log(e^{my_1}+\cdots+e^{my_p}) \\ &\le \max\{y_1,\dots,y_p\} + \frac{\log p}{m},\end{aligned} \tag{18a}$$

for any $y \in \mathbb{R}^p$, $m > 0$.[9] So, as $m \to \infty$,

$$\frac{1}{m}\log(e^{my_1}+\cdots+e^{my_p}) \to \max\{y_1,\dots,y_p\}. \tag{18b}$$

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 $\lambda \ge 0$, and *monotonic*, $\text{CVaR}_\alpha[Y] \le \text{CVaR}_\alpha[Z]$, for any random variables, $Y \le Z$.[10] Also, CVaR can be expressed as the supremum expectation over a particular set of probability density functions.[11] Using this property and the fact, $\mathbb{E}[\log(Z)] \le \log\left(\mathbb{E}[Z]\right)$, one can show,

$$\text{CVaR}_\alpha[\log(Z)] \le \log\left(\text{CVaR}_\alpha[Z]\right), \tag{19}$$

for any random variable, Z, with finite expectation.

By monotonicity, positive homogeneity, (18), and (19),

$$\text{CVaR}_\alpha\big[Z_x^\pi\big] \le \tfrac{1}{m}\text{CVaR}_\alpha\big[\log\big(\bar{Y}_x^\pi\big)\big] \le \tfrac{1}{m}\log\big(\text{CVaR}_\alpha\big[\bar{Y}_x^\pi\big]\big), \tag{20}$$

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

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

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

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

where the second inequality holds by (20). So, $x \in \mathcal{S}_\alpha^r$. ■

Lastly, we prove Theorem 1.

Proof: 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,

$$c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}[RJ_{k+1}(x_{k+1}, y_k R)|z_k, \mu_k^\epsilon] \le J_k(z_k) + \epsilon. \tag{21}$$

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

$$J_k^\epsilon(z_k) := \text{CVaR}_{y_k}\big[\textstyle\sum_{i=k}^N c(x_i)\big|z_k, \pi_k^\epsilon\big], \tag{22}$$

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), and J_k^*, as defined in (16). We will show by induction, $\forall z_k := (x_k, y_k) \in \mathcal{X}\times\mathcal{Y}$ and $k = N-1,\dots,0$,

$$J_k^\epsilon(z_k) \le J_k(z_k) + (N-k)\epsilon, \tag{23a}$$

$$J_k^*(z_k) \le J_k^\epsilon(z_k) \le J_k^*(z_k) + (N-k)\epsilon, \tag{23b}$$

$$J_k(z_k) = J_k^*(z_k), \tag{23c}$$

which is the proof technique in [4], Sec. 1.5. One can show (23) for the base case, $k := N-1$, since J_N is known.

[9] Use the log-sum-exp relation stated in [22], Sec. 3.1.5.
[10] See [12], Sec. 2.2
[11] See [11], Eqs. 6.40 and 6.70

Assuming that (23) holds for index $k+1$ (induction hypothesis), we will show that (23) holds for index k (induction step). The key idea is to use the following recursion,

$$J_k^\epsilon(z_k) \geq c(x_k) + \max_{R\in\mathcal{R}(y_k,\mathbb{P})} \mathbb{E}[R \cdot J_{k+1}^\epsilon(x_{k+1}, y_k R)|z_k, \mu_k^\epsilon], \tag{24}$$

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

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

where we use (12), (22), and translation equivariance.[12] We will show (23a) for index k using (24) 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 (21). By (14), sub-optimality of $\mu_k^\epsilon(z_k) \in U$, $J_{k+1} \leq J_{k+1}^\epsilon$, and (24),

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

The induction step for (23a) is complete. Next, we show (23b) for index k. $J_k^* \leq J_k^\epsilon$ by (16) and (22). 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 completes the induction step for (23b).

We have shown that (23a) and (23b) hold for index k, for any $\epsilon > 0$. So, (23c) holds for index k. This completes the proof. ■

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

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] G. Serraino and S. Uryasev, "Conditional Value-at-Risk (CVaR)," in *Encyclopedia of Operations Research and Management Science*. Springer, 2013, pp. 258–266.

[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.