2019 American Control Conference (ACC)
Philadelphia, PA, USA, July 10-12, 2019

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], Kevin M. Smith[5], Victoria Cheng[6], Jaime F. Fisac[1], Susmit Jha[2], 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 a given 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 a set of initial states from which the risk of large constraint violations can be reduced to a required level via a 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 arguments about its correctness. Finally, we present a realistic example inspired from stormwater catchment design to 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 a given 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 usually non-adversarial. Most storms do not cause major floods, and most vehicles are not involved in pursuit-evasion 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 to compute a set of initial states from which the probability of safety of the state trajectory can be increased to a required level by a 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 an expectation of a 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 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 will always occur, which may yield solutions with limited practical utility, especially in appli-

[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 the Computer Science Laboratory at 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.

cations with large uncertainty sets. Stochastic reachability methods [6], [7] do not explicitly account for rare high-consequence events, as costs are evaluated in terms of an expectation.

In contrast, in this paper, we harness *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 time, but they are unlikely to occur all the time. 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 possibility of danger 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], [18]. A risk-sensitive method may provide more practical and protective decision-making machinery (versus stochastic or minimax methods) by encoding a flexible degree of conservativeness.

In this paper, we use a particular risk measure, called *Conditional Value-at-Risk* (CVaR). If Z is a random variable representing 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]

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

CVaR captures a full spectrum of risk assessments from risk-neutral to worst-case, since $\text{CVaR}_\alpha[Z]$ increases from $\mathbb{E}[Z]$ to $\operatorname{ess\,sup} Z$, as α decreases from 1 to 0. CVaR has desirable mathematical properties for optimization [19] and chance-constrained stochastic control [20]. There is a well-established relationship between CVaR and chance constraints that we will use to obtain probabilistic safety guarantees in this paper. Please see [12] and [21] for additional background on CVaR.

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 a set of initial states from which the risk of large constraint violations can be reduced to a required level via a 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. 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 arguments about its correctness. Finally, we present a realistic example inspired from stormwater catchment design to 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].

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],

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

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 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,

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

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 in a stormwater catchment, then x_k may be the water level of the pond in feet at time k, and $\mathcal{K} = [0, 5)$ indicates that the pond overflows if the water level exceeds 5 feet. 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 $\mathcal{K} = [0, 5)$ 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 large constraint violations can be reduced to a required level via a 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,

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

where

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

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 distance to the constraint set $\mathcal{K}$ according to (4). The minimum in W_0^* is attained as stated below.

Lemma 1 (Existence of a minimizer): For any initial state $x_0 \in \mathcal{X}$ and confidence level $\alpha \in (0, 1]$, $\exists\, \pi^* \in \Pi$ 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: Since U and D are finite, the set of policies restricted to realizable histories from x_0 is finite. ∎

C. Discussion

Computing risk-sensitive safe sets, as defined by (5), is well-motivated. 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 $g(x) := \mathbf{1}_{\bar{\mathcal{K}}}(x) - \frac{1}{2}$. ($\mathbf{1}_{\bar{\mathcal{K}}}(x) = 1$ if $x \notin \mathcal{K}$; $\mathbf{1}_{\bar{\mathcal{K}}}(x) = 0$ if $x \in \mathcal{K}$.) Then, the risk-sensitive safe set (5) is equal to

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

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. (Abate et al. considers hybrid systems under Markov policies [6].)

Further, $\mathcal{S}_\alpha^r$ encodes a higher degree of safety as r or α decrease. Formally, for any $r_1 \ge r_2$ and $1 \ge \alpha_1 \ge \alpha_2 > 0$, we have $\mathcal{S}_{\alpha_2}^{r_2} \subseteq \mathcal{S}_{\alpha_1}^{r_1}$. Another useful property is the following relation to probabilistic safety at risk level $r := 0$.

Lemma 2 (Probabilistic safety guarantee): If $x \in \mathcal{S}_\alpha^0$, then the probability that the state trajectory initialized at x exits $\mathcal{K}$ can be reduced to α by a control policy.

Proof: Note that $\text{CVaR}_\alpha[Z_x^\pi] \le 0 \implies \mathbb{P}[Z_x^\pi \ge 0] \le \alpha$ [11, Sec. 6.2.4, pp. 257-258]. The event $Z_x^\pi \ge 0$ is equivalent to the event that there is a state x_k of the associated trajectory that exits $\mathcal{K}$ since $g(x_k) \ge 0 \iff x_k \notin \mathcal{K}$. ∎

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

A key difference between our risk-sensitive safe set (5) and the risk-constrained safe set in [18] is that we specify the CVaR of the worst constraint violation of the state trajectory $(x_0, \ldots, x_N)$ to be below a required threshold, while ref. [18] specifies the CVaR of the constraint violation of x_k to be below a required threshold for each k.

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 (versus a summation) and the Conditional Value-at-Risk measure (versus the Expectation). Here we assert that computing an under-approximation of a risk-sensitive safe set may be reduced to solving a CVaR-MDP [15], [22]. Such a reduction will be used 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 $\mathcal{Y} := (0, 1]$. Under-approximations of risk-sensitive safe sets will be defined in terms of the dynamics of the augmented state $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

Fix the initial condition $(x_0, y_0) := (x, \alpha)$. The augmented state at time $k+1$ 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 $w_k \in D$, $x_{k+1} \in \mathcal{X}$ satisfies the dynamics (2). The next confidence level $y_{k+1} \in \mathcal{Y}$ is given by

$$y_{k+1} = \bar{R}_{x_k, y_k}(w_k) \cdot y_k, \tag{7}$$

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 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,

$$\begin{aligned}&\bar{\Pi}_t := \{(\bar{\mu}_t, \bar{\mu}_{t+1}, \ldots, \bar{\mu}_{N-1}) \mid \bar{\mu}_k : \mathcal{X} \times \mathcal{Y} \to U\},\\ &t = 0, \ldots, N-1.\end{aligned} \tag{8}$$

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 is processed instead of the initial confidence level and the state history up to time k. Second, we are able to define an under-approximation of the risk-sensitive safe set using $\bar{\Pi}_0$ as detailed below.

B. Under-Approximation of Risk-Sensitive Safe Set

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

$$\mathcal{U}_\alpha^r := \{x \in \mathcal{X} \mid J_0^*(x, \alpha) \le \beta e^{m \cdot r}\}, \tag{9}$$

where

$$J_0^*(x, \alpha) := \min_{\pi\in\bar{\Pi}_0} \text{CVaR}_\alpha\big[Y_x^\pi\big], \quad Y_x^\pi := \textstyle\sum_{k=0}^N c(x_k), \tag{10}$$

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 (7) 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 c takes a particular form, then $\mathcal{U}_\alpha^r$ under-approximates 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 (9)

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 reduced by increasing m. ■

The parameter β is included above to address numerical issues that may arise if m is set to a very large number.

IV. Numerical Method

By using Theorem 1, one can apply existing CVaR-MDP algorithms to compute approximations of risk-sensitive safe sets. Here we adapt a value-iteration algorithm from [15] to compute tractable approximations of the risk-sensitive safe set under-approximations $\{\mathcal{U}^r_\alpha\}$. We start by stating an existing result from operations research that will be instrumental for devising the value-iteration algorithm.

A. Temporal Decomposition of Conditional Value-at-Risk

Here we present an existing result (using Lemma 22 in [23]) that specifies how the Conditional Value-at-Risk of a sum of costs can be partitioned over time, which motivates the choice of the update function (7).

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

$$\text{CVaR}_{y_k}[Z|x_k, \pi_k] = \max_{R \in \mathcal{R}(y_k, \mathbb{P})} C(R, Z; x_k, y_k, \pi_k), \quad \text{(11a)}$$

where

$$\mathcal{R}(y_k, \mathbb{P}) := \big\{R : D \to \big(0, \tfrac{1}{y_k}\big] \;\big|\; \mathbb{E}_{w_k \sim \mathbb{P}}\big[R(w_k)\big] = 1\big\},$$
$$Z := \textstyle\sum_{i=k+1}^{N} c(x_i),$$
$$C(R, Z; x_k, y_k, \pi_k) :=$$
$$\mathbb{E}_{w_k \sim \mathbb{P}}\big[R(w_k) \cdot \text{CVaR}_{y_k R(w_k)}[Z|x_{k+1}, \pi_{k+1}]\big|x_k, \mu_k\big], \quad \text{(11b)}$$

and $c : \mathcal{X} \to \mathbb{R}$ 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 (7) $\bar{R}_{x_k, y_k} : D \to (0, \frac{1}{y_k}]$ is defined as

$$\bar{R}_{x_k, y_k}(w_k) := \operatorname*{arg\,max}_{R \in \mathcal{R}(y_k, \mathbb{P})} C(R, Z; x_k, y_k, \pi_k). \quad \text{(12)}$$

Remark 1: The proof of Lemma 3 is a consequence of Lemma 22 in [23], 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 any disturbance that satisfies $x_{k+1} = f(x_k, u_k, w)$.

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 $\mathcal{U}^r_\alpha$ at different levels of confidence α and risk r.

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

$$J_k(z_k)$$
$$:= \min_{u \in U}\Big\{c(x_k) + \max_{R \in \mathcal{R}(y_k, \mathbb{P})} \mathbb{E}_{w_k \sim \mathbb{P}}\big[R J_{k+1}(x', y_k R)\big|z_k, u\big]\Big\},$$
$$\text{for } k = N-1, \ldots, 0, \quad \text{(13)}$$

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 (11).

Then, we approximate the set $\mathcal{U}^r_\alpha$ as $\widehat{\mathcal{U}}^r_\alpha := \{x \in \mathcal{X} \mid J_0(x, \alpha) \le \beta e^{m \cdot r}\}$, where we have replaced J^*_0 in (9) with J_0. The function, J_0, is obtained from the last step of the value iteration (13). We present theoretical arguments inspired by [15] and [4, Sec. 1.5] that justify such an approximation in our extended version [24]. In particular, theoretical evidence for the following conjecture is provided.

Conjecture (C): Assume that the functions $J_{N-1}, \ldots, J_0$ are computed recursively as per (13). Then, for any $(x, \alpha) \in \mathcal{X} \times \mathcal{Y}$, $J_0(x, \alpha) = J^*_0(x, \alpha)$, where J^*_0 is given by (10).

This conjecture is also supported by the next example.

V. Numerical Example

Here we provide empirical results to demonstrate: 1) our value-iteration estimate of J_0 is close to a Monte Carlo estimate of J^*_0, 2) our value-iteration estimate of $\widehat{\mathcal{U}}^r_y$ is an under-approximation of a Monte Carlo estimate of $\mathcal{S}^r_y$, and 3) estimating J_0 (and $\widehat{\mathcal{U}}^r_y$) via the value-iteration algorithm is tractable on a realistic example. In our experiments, we used MATLAB and MOSEK with CVX [25] on a standard laptop (64-bit OS, 16GB RAM, Intel Core i7-4700MQ CPU @ 2.40GHz). Our code is available at `github.com/chapmanmp/ACC_2019_Github`.

We computed approximate risk-sensitive safe sets to evaluate the design of a retention pond in a stormwater catchment system. We adopted an example from our prior work [26] and assumed the following pond dynamics, $x_{k+1} = x_k + \frac{\triangle t}{A}(w_k - q_p(x_k, u_k))$ for $k = 0, \ldots, N-1$, where $x \ge 0$ is the water level (in feet), $u \in U := \{0, 1\}$ is the valve setting, $w \in D := \{d_1, \ldots, d_{10}\}$ is the random surface runoff rate, q_p is the outflow rate, and A is the pond surface area. We estimated a finite probability distribution for w using the *design storm* from [26]. We set $N := 48$, $\mathcal{K} := [0, 5)$ft, and $g(x) := x - 5$. We computed over a grid of states and confidence levels $G := G_s \times G_c$, where $G_s := \{0, 0.1, \ldots, 6.4, 6.5\}$ft and $G_c := \{0.999, 0.95, 0.80, 0.65, 0.5, 0.35, 0.20, 0.05, 0.001\}$. If $x_{k+1} > 6.5$ft, we set $x_{k+1} := 6.5$ft to stay within the grid. We were able to empirically assess the accuracy of our proposed approach because an optimal policy is known *a priori* for the one-pond system. In our setting, $x_{k+1} \ge x_k$ for all k, and the only way to exit $\mathcal{K}$ is if $x_k \ge 5$ft. So, it is optimal to keep the valve open over all time, regardless of the current state, the current confidence level, or the state history up to the current time. Please see [24], [26] for more details on our example.

Our value-iteration estimate of J_0 is shown in Fig. 1, and a Monte Carlo estimate of J^*_0 is shown in Fig. 2. The estimates of J_0 and J^*_0 are similar throughout the grid (except near the smaller confidence levels). Our value-iteration estimate of $\widehat{\mathcal{U}}^r_\alpha$ and a Monte Carlo estimate of $\mathcal{S}^r_y$ are shown in Fig. 3 at various values of y and r. The empirical results indicate that $\widehat{\mathcal{U}}^r_\alpha$ is an under-approximation of $\mathcal{S}^r_y$. We estimated each $\mathcal{S}^r_y$ using a Monte Carlo estimate of W^*_0 (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 entailed a realistically sized grid ($|G_s| \cdot |G_c| = 594$ grid points) and time horizon ($N = 48$), and 3) our implementation is not optimized. Please refer to [24] for additional implementation details, including the interpolation methods [15].

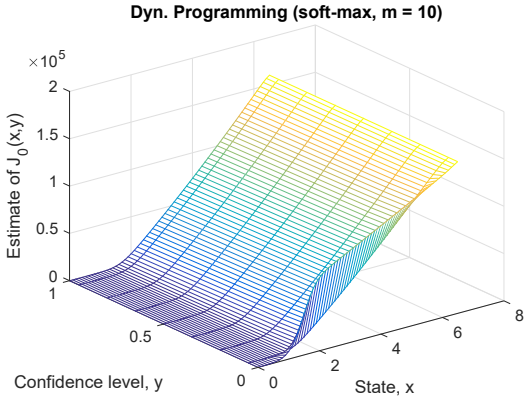

Fig. 1. Our value-iteration estimate of $J_0(x, \alpha)$ versus $(x, \alpha) \in G$ for the pond system, see (13). $c(x) := \beta e^{m \cdot g(x)}$, $\beta := 10^{-3}$, $m := 10$, and $g(x) := x - 5$. The computation time was roughly 3h 6min.

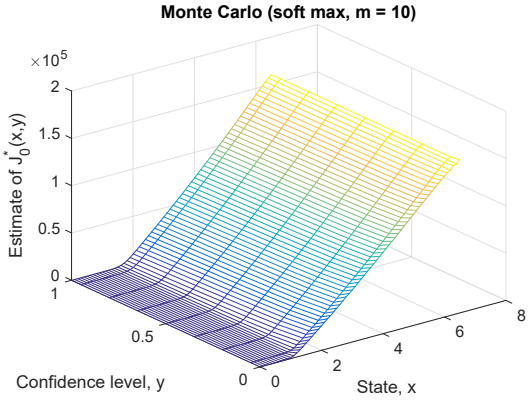

Fig. 2. A Monte Carlo estimate of $J_0^*(x, \alpha)$ versus $(x, \alpha) \in G$ for the pond system. $c(x) := \beta e^{m \cdot g(x)}$, $\beta := 10^{-3}$, $m := 10$, and $g(x) := x - 5$. 100,000 samples were generated per grid point. See also Fig. 1.

VI. Conclusion

In this paper, we proposed the novel idea of a risk-sensitive safe set to quantify safety of a stochastic dynamic system over a spectrum of confidence levels. We showed 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 devised a tractable algorithm to approximate risk-sensitive safe sets and provided empirical justification for the algorithm. Theoretical justification is provided in [24].

Risk-sensitive safe sets may become powerful design tools for safety-critical infrastructure systems by revealing trade-offs between various design choices at different levels of confidence. We illustrated our risk-sensitive reachability approach on a stormwater retention pond that must be designed

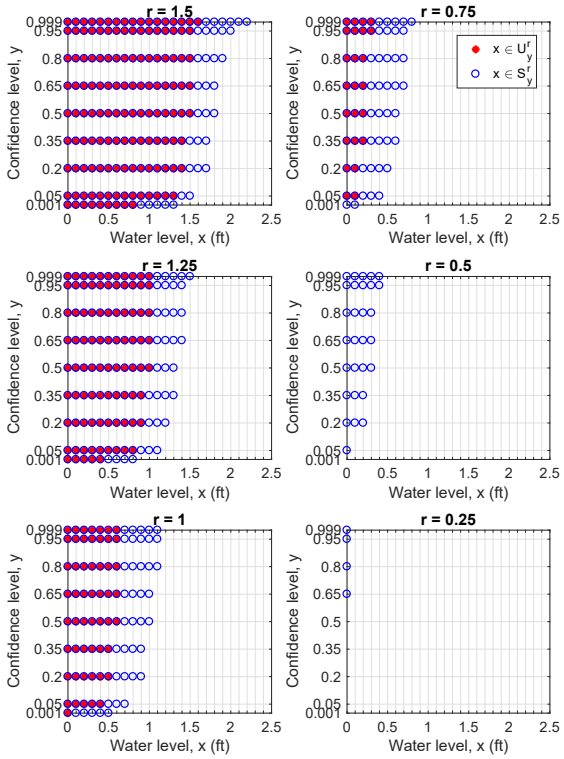

Fig. 3. Approximations of $\{\widehat{\mathcal{U}}_y^r\}$ and $\{\mathcal{S}_y^r\}$ are shown for the pond system at various levels of confidence y and risk r. In the legend, $\widehat{\mathcal{U}}_y^r$ is denoted by U_y^r, and $\mathcal{S}_y^r$ is denoted by S_y^r. Approximations of $\{\widehat{\mathcal{U}}_y^r\}$ were obtained from our 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).

to operate safely in the presence of uncertain rainfall. Our results revealed 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 (see Fig. 3, $r = 0.25$ plot at $x = 0$).

Future steps include: 1) 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. We are hopeful that with further development, 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, Dr. Alessandro Abate, and Dr. David Freyberg for discussions. M.C. and V.C. are supported by NSF GRFP. This work is supported by NSF CPS 1740079, NSF PIRE UNIV59732, and NSF DGE 1633740.

Appendix

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

$$\begin{aligned} \max\{y_1, \ldots, y_p\} &\leq \frac{1}{m} \log(e^{m y_1} + \cdots + e^{m y_p}) \\ &\leq \max\{y_1, \ldots, y_p\} + \frac{\log p}{m}, \end{aligned} \tag{14a}$$

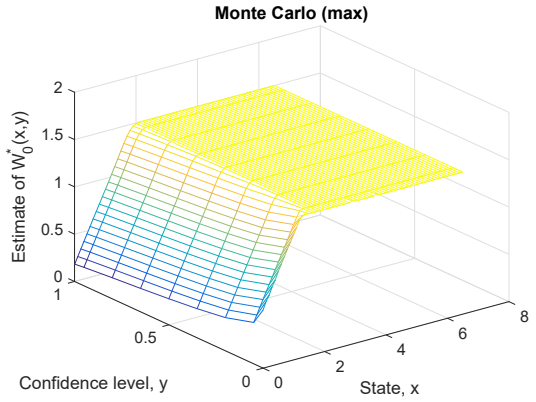

Fig. 4. A Monte Carlo estimate of $W_0^*(x,\alpha)$, as defined in (5), versus $(x,\alpha) \in G$ for the pond system. 100,000 samples were generated per grid point, and $g(x) := x - 5$. The maximum is approximately 1.5ft because the system state was prevented from exceeding 6.5ft.

for any $y \in \mathbb{R}^p$, $m > 0$. (Use the log-sum-exp relation stated in [27, Sec. 3.1.5].) So, as $m \to \infty$,

$$\frac{1}{m}\log(e^{my_1} + \cdots + e^{my_p}) \to \max\{y_1, \ldots, y_p\}. \quad (14b)$$

The second fact is that CVaR is a *coherent risk measure*, so it satisfies certain properties. 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 $\mathbb{E}[\log(Z)] \leq \log(\mathbb{E}[Z])$, one can show,

$$\text{CVaR}_\alpha[\log(Z)] \leq \log\left(\text{CVaR}_\alpha[Z]\right), \quad (15)$$

for any random variable Z with finite expectation.

By monotonicity, positive homogeneity, (14), and (15),

$$\begin{aligned} \text{CVaR}_\alpha\left[Z_x^\pi\right] &\leq \tfrac{1}{m}\text{CVaR}_\alpha\left[\log\left(\bar{Y}_x^\pi\right)\right] \\ &\leq \tfrac{1}{m}\log\left(\text{CVaR}_\alpha\left[\bar{Y}_x^\pi\right]\right), \end{aligned} \quad (16)$$

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

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

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

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

where the second inequality holds by (16). So, $x \in \mathcal{S}_\alpha^r$. ■

REFERENCES

[1] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi Reachability: A Brief Overview and Recent Advances," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 2242–2253.

[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. [Online]. Available: https://www.maths.ed.ac.uk/~prichtar/docs/Kisiala_Dissertation.pdf

[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] S. Samuelson and I. Yang, "Safety-Aware Optimal Control of Stochastic Systems Using Conditional Value-at-Risk," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 6285–6290.

[19] R. T. Rockafellar and S. P. Uryasev, "Optimization of Conditional Value-at-Risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[20] M. P. Vitus, Z. Zhou, and C. J. Tomlin, "Stochastic control with uncertain parameters via chance constrained control," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2892–2905, 2016.

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

[22] 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.

[23] 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.

[24] M. P. Chapman, J. Lacotte, A. Tamar, D. Lee, K. M. Smith, V. Cheng, J. F. Fisac, S. Jha, M. Pavone, and C. J. Tomlin, "A Risk-Sensitive Finite-Time Reachability Approach for Safety of Stochastic Dynamic Systems," *arXiv preprint arXiv:1902.11277*, 2019.

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

[26] 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*, Nov. 2018.

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