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

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

Abstract*— A classic reachability analysis 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 in 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 reachability literature. First, 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). Stochastic reachability does not explicitly account for rare high-consequence events, since the optimal control problem is posed in terms of the expectation operator. Second, our formulation quantifies the distance between the boundary of the constraint set and the state trajectory in a stochastic setting. Stochastic reachability quantifies the probability that the state trajectory stays within the constraint set, and Hamilton-Jacobi reachability quantifies the distance between the boundary of the constraint set and the state trajectory in a deterministic setting. We define a *risk-sensitive safe set* as the set of initial states from which the risk of extreme constraint violation 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 to compute under-approximations for risk-sensitive safe sets and prove the correctness of the algorithm under the assumption of finite probability spaces. Our proof is a key contribution to reinforcement learning literature, as it does not require the assumption of strong duality, which was required in a previous paper. Finally, we demonstrate the utility of risk-sensitive reachability analysis as a design tool for stormwater infrastructure, which is required to operate safely in the presence of rainfall uncertainty.

I. Introduction

Reachability analysis is a formal verification method based on optimal control theory that can be 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 the uncertain disturbances live in given sets, but whose probability distributions are unknown [2], [3].[1] In this context, the problem is solved using a minimax formulation [2], [3], where the disturbances are adversarial and safety is described solely as a binary notion according to set membership.

In practice, however, disturbances do not usually behave adversarially: most storms do not cause major floods; most drivers are not involved in car chases; most drones (outside of war) do not intend to collide with one another or with humans. In applications where assuming adversarial disturbances yields overly conservative solutions, instead one may assume that the disturbances are drawn from a given probability distribution. Abate et al. posed the stochastic reachability problem for safety in terms of maximizing the probability that the (stochastic) state trajectory stays inside a given constraint set over time [5]. Summers and Lygeros extended the formulation of Abate et al. to also encode the goal of the state trajectory reaching a given target set [6].

Similar to the minimax formulation [2], [3], the stochastic formulations [5], [6] describe safety solely as a binary notion according to set membership. In particular, the optimal control problem is expressed as an expectation of the summation (or maximum) of indicator functions, where each indicator encodes whether a given state is inside or outside a particular set [5]. Stochastic formulations [5], [6] do not generalize to quantify the (random) distance between the state trajectory and the boundary of the constraint set, because they rely on indicator functions to encode the (random) position of a state with respect to a particular set. Quantification of the (random) distance between the state trajectory and the boundary of the constraint set may be important in practical applications where the boundary of the constraint set is not known exactly, or where minor constraint violations are inevitable, but extreme constraint violations are particularly damaging.

In this work, we provide a formulation for reachability analysis of stochastic dynamic systems

Stochastic formulations [5], [6] pose the optimal control problem in terms of expectation, which

Typically, the dynamic system has a control signal to be designed to ensure constraint satisfaction and an uncertain disturbance signal that may try to prevent constraint satisfaction (i.e., is *adversarial*). The guarantees enjoyed by the

[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]J.L. is with the Department of Aeronautics and Astronautics, Stanford University, USA.

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

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

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

[6]S.J. is with SRI International, Menlo Park, California, USA.

[1]See also [4], Sec. 3.6.2 "Control within a Target Tube"

states of the safe set are either deterministic or probabilistic in nature, depending on what we assume about the system dynamics.

For example, we may assume that the control and disturbance signals are bounded, and we may not specify their probability distributions; in this case, the dynamics are said to be *nondeterministic* [7]. If the dynamics are nondeterministic and the disturbance signal is adversarial, then the safe set may be defined as the set of states from which the system can start, such that for any disturbance signal, there exists a control signal that ensures constraint satisfaction (see [8], Sec. 2.2.1). The safety guarantee is deterministic in this case, which has been studied extensively and applied mainly to vehicles (see [1] and the references therein).

In contrast, we may assume that the state of the system is a random variable that evolves according to some probability distribution (e.g., see [4], Sec. 1.2); here the dynamics are said to be *stochastic*. If the dynamics are stochastic and the disturbance input is adversarial, then the safe set may be defined as the set of states from which the system can start, such that for any disturbance signal, there exists a control signal that ensures constraint satisfaction with sufficiently high probability (e.g., see [9]). If the dynamics are stochastic and the disturbance input is non-adversarial (i.e., behaves as random noise), then the safe set may be defined as the set of states from which the system can start, such that there exists a control signal that ensures constraint satisfaction with sufficiently high probability (e.g., see [6] and [5]). In these last two examples, the safety guarantees are probabilistic.

Whether the safety guarantee is deterministic or probabilistic in nature, its essential purpose is to inform decision-making in an uncertain world. The key distinction between deterministic and probabilistic safety guarantees is how the uncertainty is quantified, and whether we assume a pessimistic world, where the uncertainty is adversarial, or a realistic world, where the uncertainty behaves as random noise. In this paper, we develop a framework for reachability theory that lies on the spectrum between pessimism and realism, by leveraging the theory of risk from finance and mathematics.

Risk may be defined qualitatively as "danger, or the possibility of danger, defeat, or loss" [10]. To quantify risk, the mathematical concept of a *risk measure* has been developed. A risk measure is a function that maps a random variable, X, representing loss into the real line, according to the risk associated with X (see [11], Sec. 6.3; see [12], Sec. 2.2). Risk-sensitive optimization algorithms, which minimize the risk of predicted losses via a risk measure, have been receiving more attention in the communities of applied mathematics [13], reinforcement learning [14], [15], [16], and optimal control [17]. Optimization programs that appreciate risk are desirable due to the limitations of alternative methods. In particular, formulations that minimize worst-case losses under adversarial disturbances may produce conservative results with limited practical utility, and formulations that minimize expected losses under random disturbances do not account for low-probability extreme events [17], [18]. On the other hand, risk-sensitive formulations have the potential to generate decisions that can be used in practice and that also protect against particularly damaging outcomes [19].

In this paper, we leverage existing computational results for a particular risk measure, called *Conditional Value-at-Risk* (CVaR), to propose a framework for risk-sensitive reachability analysis. CVaR is a well-justified choice for several reasons. CVaR is a *coherent risk measure*, meaning that it satisfies several intuitive axioms, such as *subadditivity*, which can be interpreted as "diversification decreases risk" (see [12], Sec. 2.2). On finite probability spaces, coherent risk measures are expectations that have been maximized over a collection of perturbed probability distributions, or expectations that have been made more robust to large losses [17], [11], [15], [20]. Recent work [15] provides an algorithm to minimize the Conditional Value-at-Risk of total cost over time, which we leverage to compute risk-sensitive safe sets. Further, probabilistic safety guarantees and risk-sensitive safety guarantees are closely related, if the risk measure is CVaR, as we shall explain in Sec. V.

We propose a formulation for risk-sensitive reachability with several desirable attributes. At a fixed confidence level for CVaR, our formulation partitions the state space into regions of varying degrees of safety quantified via the extent of constraint violation likely to be attained by the stochastic dynamic system. Quantification of varying degrees of safety is a feature of safety guarantees for non-deterministic systems (see [8], Eq. 2.3) but not for stochastic systems currently. Existing safety guarantees for stochastic systems are binary, meaning that they encode whether the system is likely to be inside or outside a given set (e.g., see [5], [6], and [9]). Our formulation, however, provides a non-binary quantification of safety for stochastic systems, which is particularly useful when constraint violation is not catastrophic (e.g., routine flooding of a pond after a large storm). Further, our formulation inherits the benefits of risk-sensitive optimization and the benefits of reachability theory. By using a risk measure, our formulation may protect against rare harmful outcomes, which are ignored by reachability formulations that provide safety guarantees in expectation (e.g., [5], [6], and [9]), and may also avoid unnecessary conservatism, which is a common limitation of deterministic safety guarantees (e.g., see [1]). Like existing reachability methods, our formulation provides a comprehensive characterization of the state space in terms of safety. This is not provided by recent work in risk-sensitive optimization, which computes optimal paths emanating from different initial conditions separately (e.g., see [17] and [15]). A comprehensive safety characterization of the state space may be used to inform the cost-effective design of infrastructure that must withstand rare extreme storms, to reduce overly conservative error bounds that arise in safe dynamic motion planning (e.g., [21]), and to increase the amount of time that an autonomous vehicle can operate safely while simultaneously optimizing for performance.

Our formulation also inherits the disadvantages of risk-sensitive optimization and reachability analysis. Since we evoke existing methods for risk-sensitive optimization, we are required to assume finite probability spaces. Since we are not yet learning probability mass functions on-line, we assume that estimates of these functions are available, which is the case for evaluating designs of stormwater infrastructure but not the case for real-time motion planning of a vehicle. Further, like existing methods for risk-sensitive optimization and reachability, our formulation generally requires a dynamic programming algorithm that is computationally expensive.

II. System Model

The system model is a special case of the model given by [4] in Sec. 1.2. We consider a stochastic discrete-time dynamic system over a finite time horizon,

$$x_{k+1} = f_k(x_k, u_k, w_k), \quad k = 0, 1, \dots, T-1, \tag{1}$$

such that $x_k \in S$ is the state of the system at time k, $u_k \in C$ is the control input at time k, and $w_k \in D_k = \{d_1^k, \dots, d_N^k\}$ is the random disturbance input at time k defined over a finite probability space. The control input is not random, but the state generally is random because it depends on random disturbances. The initial condition, x_0, is not random for simplicity. The collection of admissible control policies is,

$$\Pi = \{(\mu_0, \mu_1, \dots, \mu_{T-1}), \text{ such that } \mu_k : S \to C\}. \tag{2}$$

The random disturbance at time k, w_k, is characterized by a time-dependent probability mass function that is independent of any control policy, $\pi \in \Pi$, and other disturbances, $w_{\not k} = (w_0, \dots, w_{k-1}, w_{k+1}, \dots, w_{T-1})$.[2] Formally, we have

$$\begin{aligned} P_k(w_k = d_j^k | x_k) &= p_j^k, \\ \textstyle\sum_{j=1}^N p_j^k &= 1, \; p_j^k \geq 0, \\ P_k(w_k = d_j^k | x_k, \pi, w_{\not k}) &= P_k(w_k = d_j^k | x_k), \end{aligned} \tag{3}$$

for each disturbance sample $j = 1, \dots, N$ and each time point $k = 0, 1, \dots, T-1$. We are given a (non-empty) constraint set, $\mathcal{K} \subset S$, and the safety criterion that the state of the system should stay inside $\mathcal{K}$ over time. For example, if our application is the flow of water through a network of ponds and streams, $\mathcal{K}$ may indicate that the water does not overflow the banks during a storm event.

III. Problem Statement

The goal of this paper is to design an algorithm that computes a *risk-sensitive safe set* for a system of the form specified in Sec. II. A risk-sensitive safe set is, informally, the set of initial conditions of the system, from which there is small risk of large constraint violations over time.

We quantify risk using the well-established risk measure, *Conditional Value-at-Risk* (CVaR), which is equal to,

$$\text{CVaR}_\delta(Z) = \min_{c \in \mathbb{R}} \Big\{ c + \frac{1}{\delta}\mathbb{E}\big[\max\{Z - c, 0\}\big] \Big\}, \tag{4}$$

[2]The probability mass function may be state-dependent as well.

where $\delta \in (0, 1)$, and Z is a random variable representing loss [22].[3] If Z is a continuous random variable, then $\text{CVaR}_\delta(Z)$ is the expected value of Z over large realizations of Z, where the meaning of large is based on δ (Fig. 1).

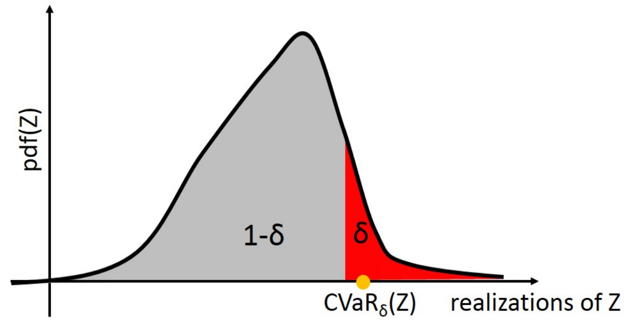

Fig. 1. An illustration of $\text{CVaR}_\delta(Z) \in \mathbb{R}$, if Z is a continuous random variable. The graph shows the probability density function of Z versus the realizations of Z. The area of the right portion under the curve, shown in red, is $\delta \in (0, 1)$. The area of the left portion under the curve, shown in grey, is $1 - \delta$. $\text{CVaR}_\delta(Z)$ is the expectation of the values along the right portion under the curve, indicated by a yellow circle.

We quantify the extent of constraint violation via a surface function that characterizes the constraint set, $\mathcal{K}$. Let $g : S \to \mathbb{R}$ satisfy,

$$x \in \mathcal{K} \iff g(x) < 0, \tag{5}$$

where we adopt the convention provided by [8] in Eq. (2.3). The particular form of g is chosen based on how safety of the system changes with distance to the boundary of $\mathcal{K}$ for the application at hand. For example, if the relationship between safety and distance to the boundary of $\mathcal{K}$ is linear, then the signed distance function for $\mathcal{K}$ is a suitable choice for g (Fig. 2, dotted). However, if the relationship between safety and distance to the boundary of $\mathcal{K}$ is non-linear, then a quadratic function may be more appropriate (Fig. 2, solid).

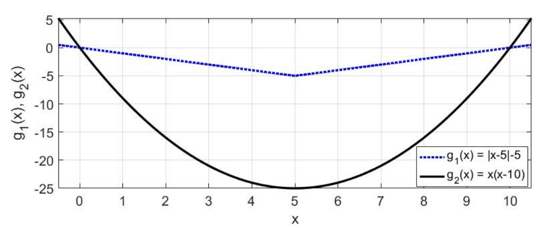

Fig. 2. Different choices for the particular form of g, see (5), for an example constraint set, $\mathcal{K} = (0, 10)$. The relationship between safety and distance to the boundary of $\mathcal{K}$ is linear for $g_1(x) = |x - 5| - 5$ (dotted), and non-linear for $g_2(x) = x(x - 10)$ (solid). For example, the degree of safety at $x = 6$ characterized by the linear relationship is, $g_1(6) = -4$, since the closest distance from $x = 6$ to the boundary of $\mathcal{K}$ is 4, and $x = 6$ is inside $\mathcal{K}$.

We are now ready to define the risk-sensitive safe set formally. Let $\xi_y^\pi(k) \in S$ be the random state of the system at time k that satisfies (1) under a given control policy $\pi \in \Pi$, starting from a given (non-random) state $y \in S$ at time 0.

[3]Definitions of CVaR are presented in various forms. The original paper is [22]. Other references on CVaR include [19] (see Eq. (9)) and [12].

The maximum extent of constraint violation attained by the system under policy $\pi \in \Pi$, starting from initial condition $y \in S$, is given by the random variable,

$$X_y^\pi = \max_{k \in \{0,\ldots,T\}} \Big\{ g\big(\xi_y^\pi(k)\big) \Big\}, \tag{6}$$

where g satisfies (5). For any $\delta \in (0,1)$, the risk-sensitive safe set is,

$$\begin{aligned} \mathcal{U}_\delta :=& \Big\{ y \in S \mid \exists \pi \in \Pi \text{ such that } \mathrm{CVaR}_\delta\left(X_y^\pi\right) < 0 \Big\} \\ =& \Big\{ y \in S \mid \min_{\pi \in \Pi} \left\{ \mathrm{CVaR}_\delta\left(X_y^\pi\right) \right\} < 0 \Big\}, \end{aligned} \tag{7}$$

where the random variable, X_y^π, is defined in (6), and the conditional value-at-risk is taken with respect to the probability distribution of $(w_0, \ldots, w_{T-1})$. To summarize, the risk-sensitive safe set is the set of initial conditions from which the risk of large constraint violations can be made small by an appropriate control policy. The problem addressed in this paper is how to compute (7).

IV. Properties of $\mathcal{U}_\delta$

Here we present two key properties of the risk-sensitive safe set. The first property is that every state in $\mathcal{U}_\delta$ enjoys a probabilistic safety guarantee. To prove this property, we need the following result adopted from [11].
SJ: The simplified proof is using $\delta = 1 - \alpha$. I suggest we use this notation throughout the paper.

Lemma 1: Let $\delta \in (0,1)$, and Z be a random variable. If $\mathrm{CVaR}_\delta(Z) < 0$, then $\mathbb{P}[Z \geq 0] < \delta$ (see [11], Sec. 6.2.4).[4]

Proof: SJ: simplifying the proof below.

$$\begin{aligned} & \mathrm{CVaR}_{1-\alpha}(Z) < 0 \\ \iff & \frac{1}{1-\alpha}\mathbb{E}\big[\max\{Z-c,0\}\big] < -c \\ \iff & \exists c \in \mathbb{R} \;\; c + \frac{1}{1-\alpha}\mathbb{E}\big[\max\{Z-c,0\}\big] < 0 \text{ [using (4)]} \\ \iff & \mathbb{E}\big[\max\{Z-c,0\}\big] < -c(1-\alpha) \end{aligned} \tag{8}$$

Now, the LHS of the inequality, $\mathbb{E}\big[\max\{Z-c,0\}\big] \geq 0$ because the expectation of non-negative values cannot be negative. Consequently, the RHS of the inequality must also be non-negative, that is, $-c(1-\alpha) \geq 0$, that is, $c \leq 0$ since $1-\alpha \geq 0$. So, we can rewrite inequality (8) using $a = -c \geq 0$ as follows:

$$\begin{aligned} & \mathbb{E}\big[\max\{Z+a,0\}\big] < a(1-\alpha), \text{ where } a \geq 0 \\ \iff & \frac{1}{a}\mathbb{E}\big[\max\{Z+a,0\}\big] < 1-\alpha, \text{ where } a \geq 0 \end{aligned} \tag{9}$$

Using Markov's Inequality, $\mathbb{P}\big[\max\{Z+a,0\} \geq a\big] \leq \frac{1}{a}\mathbb{E}\big[\max\{Z+a,0\}\big]$. Combining with inequality (9),

$$\begin{aligned} & \mathbb{P}\big[\max\{Z+a,0\} \geq a\big] \leq \frac{1}{a}\mathbb{E}\big[\max\{Z+a,0\}\big] < 1-\alpha \\ \Rightarrow & \mathbb{P}\big[\max\{Z+a,0\} \geq a\big] < 1-\alpha \end{aligned} \tag{10}$$

Now, $Z \geq 0 \iff Z + a \geq a \iff \max\{Z+a,0\} \geq a$ since $a \geq 0$, and so, $\mathbb{P}\big[Z \geq 0\big] = \mathbb{P}\big[\max\{Z+a,0\} \geq a\big]$. Combining with the inequality (10),

$$\begin{aligned} & \mathbb{P}\big[Z \geq 0\big] = \mathbb{P}\big[\max\{Z+a,0\} \geq a\big] < 1-\alpha \\ \iff & \mathbb{P}\big[Z \geq 0\big] < 1-\alpha \end{aligned}$$

■

The only one-sided implication is in the use of Markov's Inequality to get inequality (9), and this corresponds to the approximation gap in using $\mathrm{CVaR}_{1-\alpha}(Z) < 0$ to approximate $\mathbb{P}[Z \geq 0] < 1-\alpha$.

The next corollary indicates that every state in $\mathcal{U}_\delta$ enjoys a probabilistic safety guarantee.

Corollary 1: $\mathcal{U}_\delta$, as defined in (7), is a subset of $\mathcal{S}_\delta$,

$$\mathcal{S}_\delta := \Big\{ y \in S \mid \exists \pi \in \Pi, \, \mathbb{P}\left[\forall k \in \mathbb{T}, \, \xi_y^\pi(k) \in \mathcal{K}\right] > 1-\delta \Big\}, \tag{11}$$

where $\mathbb{P}$ is the probability measure for the state trajectory, and $\mathbb{T} = \{0, 1, \ldots, T\}$ is the time horizon.

Proof: may want to remove this proof b/c it's not very important? Take $y \in \mathcal{U}_\delta$. Then, there exists $\pi \in \Pi$ such that $\mathrm{CVaR}_\delta(X_y^\pi) < 0$, which implies $\mathbb{P}[X_y^\pi \geq 0] < \delta$ by Lemma 1. After some algrebra using (5) and (6),

$$\mathbb{P}[X_y^\pi \geq 0] = 1 - \mathbb{P}\left[\forall k \in \mathbb{T}, \, \xi_y^\pi(k) \in \mathcal{K}\right]. \tag{12}$$

So, $\exists \pi \in \Pi$ such that $\mathbb{P}\left[\forall k \in \mathbb{T}, \, \xi_y^\pi(k) \in \mathcal{K}\right] > 1-\delta$, implying that $y \in \mathcal{S}_\delta$. ■

The second key property of the risk-sensitive safe set, $\mathcal{U}_\delta$, is, as δ decreases, the states of $\mathcal{U}_\delta$ enjoy a stronger probabilistic safety guarantee, and $\mathcal{U}_\delta$ becomes smaller.

Lemma 2: Let $1 > \delta_1 \geq \delta_2 > 0$. Then, $\mathcal{S}_{\delta_2} \supset \mathcal{S}_{\delta_1}$, and $\mathcal{U}_{\delta_2} \supset \mathcal{U}_{\delta_1}$, where $\mathcal{U}_\delta$ is given by (7) and $\mathcal{S}_\delta$ is given by (11).

Remark 1: Remark

Please see Table I for a summary of relevant notation.

Problem 1. An important problem is to compute the set of initial states for which there exists an admissible control policy that keeps the system inside the constraint set over time with sufficiently high probability. The *safe set* with confidence $1-\delta \in (0,1)$ is defined as,

$$\mathcal{S}(\delta) := \{ x \in S \mid \exists \pi \in \Pi \text{ such that } \mathbb{P}[\forall k \in \mathbb{T}, \, \xi_x^\pi(k) \in \mathcal{K}] > 1-\delta \}. \tag{13}$$

V. Relation between Probablistic Safety and CVaR Safety

VI. Conclusion

Acknowledgment

We thank Mo Chen and Jaime Fisac for discussions. M.C. is supported in part by a NSF Graduate Research Fellowship. This work is supported in part by NSF CPS 1740079.

[4]Ref. [11] indicates that the constraint, $\mathrm{CVaR}_\delta(Z) \leq 0$, gives a conservative approximation of the chance constraint, $\mathbb{P}[Z > 0] \leq \delta$. $\mathrm{CVaR}_\delta(Z) \leq 0$ is written as "$\mathrm{AV@R}_\alpha(Z_x) \leq 0$" in [11], see (6.24). $\mathbb{P}[Z > 0] \leq \delta$ is equivalent to $\mathbb{P}[Z \leq 0] \geq 1-\delta$, which is written as "$\Pr(Z_x \leq 0) \geq 1-\alpha$" in [11], see text below (6.18).

TABLE I

Symbol	Definition	Expression (if applicable)
g	Surface function that characterizes the constraint set, $\mathcal{K}$	$x \in \mathcal{K} \iff g(x) \leq 0$
C	Set of possible values for the control input	
D_k	Sample space for the random disturbance input at time k	$D_k := \{d_1^k, d_2^k, \ldots, d_N^k\}$
S	Set of (continuous) states	$S := \mathbb{R}^n$
$\mathcal{K}$	Constraint set	$\mathcal{K} \subset S$
Π	Set of admissible control policies	$\Pi := \{(\mu_0, \mu_1, \ldots, \mu_{T-1}), \text{ such that } \mu_k : S \to C\}$
$\mathbb{P}$	The probability measure with respect to $(w_0, w_1, \ldots, w_{T-1})$	
$\mathbb{T}$	Finite discrete time horizon	$\mathbb{T} := \{0, 1, \ldots, T\}$
$\xi_x^\pi(k)$	Random state at time k under (fixed) policy π, starting from (fixed) initial condition, $x \in S$, at time 0	

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] 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.

[6] 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.

[7] 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.

[8] 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

[9] M. Kamgarpour, J. Ding, S. Summers, A. Abate, J. Lygeros, and C. Tomlin, "Discrete Time Stochastic Hybrid Dynamical Games: Verification & Controller Synthesis," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, 2011, pp. 6122–6127.

[10] "Risk." [Online]. Available: https://dictionary.cambridge.org/us/dictionary/english/risk

[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] S. Jha, V. Raman, D. Sadigh, and S. A. Seshia, "Safe autonomy under perception uncertainty using chance-constrained temporal logic," *Journal of Automated Reasoning*, vol. 60, no. 1, pp. 43–62, 2018.

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

[20] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[21] 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.

[22] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.