On Optimizing the Conditional Value-at-Risk of a Maximum Cost for Risk-Averse Safety Analysis*

Margaret P. Chapman†, *Member, IEEE*, Michael Faußß‡, *Member, IEEE*, and Kevin M. Smith**

***Abstract*— The popularity of Conditional Value-at-Risk (CVaR), a risk functional from finance, has been growing in the control systems community due to its intuitive interpretation and axiomatic foundation. We consider a nonstandard optimal control problem in which the goal is to minimize the CVaR of a maximum random cost subject to a Borel-space Markov decision process. The objective represents the maximum departure from a desired operating region averaged over a given fraction of the worst cases. This problem provides a safety criterion for a stochastic system that is informed by both the *probability* and *severity* of the potential consequences of the system's behavior. In contrast, existing safety analysis frameworks apply stage-wise risk constraints or assess the probability of constraint violation without quantifying the potential severity of the violation. To the best of our knowledge, the problem of interest has not been solved. To solve the problem, we propose and study a family of stochastic dynamic programs on an augmented state space. We prove that the optimal CVaR of a maximum random cost enjoys an equivalent representation in terms of the solutions to these dynamic programs under appropriate assumptions. For each dynamic program, we show the existence of an optimal policy that depends on the dynamics of an augmented state under the assumptions. In a numerical example, we illustrate how our safety analysis framework is useful for assessing the severity of combined sewer overflows under precipitation uncertainty.**

***Index Terms*— Conditional Value-at-Risk, Risk-averse optimal control, Safety analysis, Markov decision processes.**

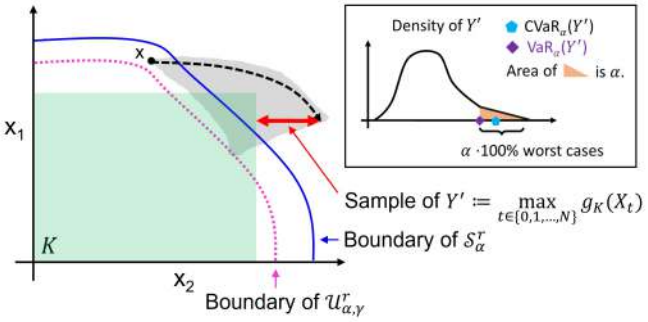

Fig. 1. A risk-averse safe set $\mathcal{S}^r_\alpha$ is the set of initial states from which the Conditional Value-at-Risk (CVaR) at level $\alpha \in (0, 1]$ of a trajectory-wise maximum random cost can be reduced to a threshold $r \in \mathbb{R}$. (A random cost is a random variable in which smaller realizations are preferred.) While we depict a state-dependent maximum random cost Y' in this figure, our theory permits control-dependent random costs as well. Our framework applies to settings in which leaving a desired operating region K may be inevitable, but the extent of a departure should be limited when possible. (K need not be a polytope. However, we require stage and terminal cost functions to be continuous and bounded. In this figure, $g_K(x)$ is a signed distance between a state x and the boundary of K.) In this work, we prove that any collection of $\mathcal{S}^r_\alpha$ is given by the solutions to a family of stochastic dynamic programs under a measurable selection assumption. In a numerical example, we compare this characterization to our underapproximation method from [12]. An underapproximation set $\mathcal{U}^r_{\alpha,\gamma} \subseteq \mathcal{S}^r_\alpha$ depends on a soft-maximum parameter γ that requires tuning [12].

I. Introduction

Control system safety is often assessed through minimax optimal control problems [1]–[4], which assume bounded nonstochastic adversarial disturbances that try to inhibit safe or efficient operation. In cases where disturbances are not well-modeled as bounded inputs (e.g., Gaussian noise), then it is standard to define safety in terms of a stochastic optimal control problem, whose optimal value is a probability of satisfactory operation. This framework, called stochastic safety analysis, can accommodate either adversarial [5], [6] or nonadversarial [7], [8] stochastic disturbances. However, a minimax approach may lead to controllers that are too cautious in practice. On the other hand, a purely probabilistic risk assessment indicates the likelihood of a harmful event but has a limited capacity to quantify the amount of harm the event would cause. These different limitations have motivated a growing body of research that lies in the intersection of formal methods and risk analysis for control systems [9]–[13].

Here, we study a nonstandard safety analysis problem, which concerns the notion of a *risk-averse safe set* $\mathcal{S}^r_\alpha := \{\mathbf{x} \in S : \mathcal{J}^*_\alpha(\mathbf{x}) \leq r\}$. $\mathcal{S}^r_\alpha$ represents the set of initial states from which the maximum distance between the trajectory and a desired operating region averaged over the $\alpha \cdot 100\%$ worst cases can be reduced to a threshold r (Fig. 1). The system of interest is a Markov decision process (MDP) with Borel spaces of states, controls, and disturbances, operating on a discrete-time horizon of length N, a natural number. $\mathcal{J}^*_\alpha(\mathbf{x})$ is the optimal value of a stochastic optimal control problem

This work was supported in part by the Computational Hydraulics International University Grant Program for complementary use of PCSWMM Professional software. K. M. Smith was supported in part by the U.S. National Science Foundation under Grant NSF-NRT 2021874. The work of M. Faußß was supported by the German Research Foundation (DFG) under grant number 424522268. M. P. Chapman acknowledges support from the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, and the Natural Sciences and Engineering Research Council of Canada Discovery Grants Program [RGPIN-2022-04140]. Cette recherche a été financée par le Conseil de Recherches en Sciences Naturelles et en Génie du Canada.

†M. P. Chapman is with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4 Canada (email: mchapman@ece.utoronto.ca).

‡M. Faußß is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, New Jersey 08544 USA (email: mfauss@princeton.edu).

**K. M. Smith is with the Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155 USA and OptiRTC, Inc., Boston, MA 02116 USA (email: kevin.smith@tufts.edu).

*This work solves the risk-averse safety analysis problem. Our prior works [10], [12] offer approximations.

with a Conditional Value-at-Risk maximum cost objective:

$$\mathcal{J}^*_\alpha(\mathbf{x}) \coloneqq \inf_{\pi\in\Pi} \mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y), \tag{1a}$$

$$Y \coloneqq \max_{t\in\{0,1,\ldots,N-1\}} \{c_t(X_t,U_t), c_N(X_N)\}. \tag{1b}$$

The random variable Y depends on stage and terminal cost functions c_t, random states X_t, and random controls U_t. The quantity $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ represents the average value of Y in the $\alpha\cdot 100\%$ worst cases when the initial state is $\mathbf{x}$ and the system uses the control policy π. A control policy provides distributions for the realizations of $U_0, U_1, \ldots, U_{N-1}$. (We will formalize $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ and π in Sec. III-B and Sec. IV-A, respectively.) The setting is fairly general in theory. It permits nonlinear dynamics, nonconvex bounded cost functions, continuous spaces, and non-Gaussian stochastic disturbances. First, we will explain why (1) is an important problem to solve, and then we will explain the novelty of our contribution.

A. Relevance of the CVaR

The CVaR functional, which defines the objective of (1), provides an *intuitive and quantitative interpretation for risk* because it represents the average value of a random variable in a fraction α of the worst cases [24, Th. 6.2]. Other common risk functionals do not have interpretations that are as consistent or clear. Expected utility risk functionals encode risk preferences using utility functions and their parameters [14]–[19]. It is challenging to provide a precise meaning for the parameter of the classical expected exponential utility functional, which limits its applicability to control systems with specific safety or performance requirements [20]. It may be difficult to interpret a recursive risk functional because it takes the form $\rho_1(C_1+\rho_2(C_2+\cdots+\rho_{N-1}(C_{N-1}+\rho_N(C_N))\cdots))$, where C_i is a random variable and ρ_i is a map between spaces of random variables [21]–[23]. A weighted sum of the mean and a moment-based dispersion functional, e.g., variance, standard deviation, and upper-semideviation [24], provides an heuristic for the probability and severity of more rare and harmful outcomes. The CVaR is arguably more intuitive than the broader class of spectral risk functionals, which are "mixtures" of the CVaR_α over the values of α [25, Prop. 2.5]. The Value-at-Risk (VaR) at level α, which is the left-side $(1-\alpha)$-quantile, has a clear quantitative interpretation. However, the VaR's ability to summarize the severity of harmful outcomes is limited because it is insensitive to the shape of the distribution beyond the $(1-\alpha)$-quantile. From a decision-theoretic perspective, the VaR has the disadvantage of lacking a desirable property called subadditivity [26]. Both of these shortcomings are overcome by the CVaR [24], [27].

B. Relevance of the Maximum Cost

We focus on a maximum cost (1b) generated by an MDP rather than a cumulative cost. While a cumulative cost is typical for MDP problems [17], [21], [23], [25], [28]–[30], a maximum cost is typical for robust safety and reachability analysis problems for nonstochastic systems, e.g., see [3], [4], and the references therein. Maximum costs have natural roles in systems theory, beyond robust safety and reachability analysis. The theory of the long-term behavior of normalized maxima of random variables, i.e., extreme value theory, has applications in finance, the study of human longevity, and hydrology [31].

A maximum cost is appropriate for applications in which the *extent* of a constraint violation over a brief time interval is more critical to assess than its accumulation.[1] For example, in stormwater management, the *maximum* water level can be a useful surrogate for the maximum flood extent (in more extreme cases) and the maximum discharge rate (in general). These are instantaneous rather than cumulative properties. For gravity-drained stormwater systems, the instantaneous discharge rate through an uncontrolled outlet into open atmosphere is a function of the water level behind the outlet. Therefore, from water levels, we can estimate instantaneous demands on downstream conveyance infrastructure (i.e., infrastructure to transport water rather than to store it). Designing this infrastructure for the worst maximum discharge rate may be cost-prohibitive. However, assessing the average maximum water level in the worst $\alpha\cdot 100\%$ of cases from historical data would allow designers to estimate downstream conveyance capacity demands along a spectrum of worst cases.

C. Related Literature

The problem of computing risk-averse safe sets $\mathcal{S}^r_\alpha$ is distinct from established problems in the stochastic and risk-averse control theory literature and necessitates different techniques. Classical discrete-time stochastic control theory, e.g., [32], studies the problem of optimizing the expectation of a cumulative cost. In contrast, our focus is optimizing the CVaR of a maximum cost (1). The dynamic programming (DP) proofs from stochastic control theory do not apply to our problem directly. Theoretical challenges arise because the CVaR satisfies only some of the properties that are enjoyed by the expectation. Moreover, while sums and integrals of nonnegative Borel-measurable functions can be interchanged, this is not the case for maxima and integrals in general. Such technical differences between our problem and the scenarios that prevail in the literature make it necessary to build a pathway from measure-theoretic first principles. Doing so enables us to solve for the sets $\mathcal{S}^r_\alpha$ and the associated optimal control policies under appropriate assumptions.

We take inspiration from a technique called *state-space augmentation*, which has been used to solve risk-averse MDP problems with cumulative costs [17], [25], [28]–[30]. The problem of minimizing the expectation of a cumulative cost subject to an upper bound on the CVaR of a cumulative cost has been studied in [29]. The authors propose offline and online algorithms on augmented state spaces to update a Lagrange multiplier and a lower bound on a cumulative cost [29]. Several risk-averse control problems with cumulative costs over an infinite-time horizon have been investigated using infinite-dimensional linear programming and state-space augmentation [30]. Bäuerle and Ott provide a DP solution to the problem of minimizing the CVaR of a cumulative cost [28]. While we also use DP, our approach requires different proof techniques to manage a maximum cost (1b) and to

[1] A constraint violation means that a state or control leaves a desired operating region, and its extent refers to the severity of the violation.

study our proposed algorithm, which we define in terms of dynamics functions $x_{t+1} = f_t(x_t, u_t, w_t)$, stage and terminal cost functions c_t, and disturbance distributions $p_t(\mathrm{d}w_t|x_t, u_t)$.

Most literature about risk-averse MDPs concerns exponential utility, taking inspiration from decision theory in economics and extending from 1972 to present-day [14]–[16], [18], [19], [33]. Bäuerle and Rieder study the problem of optimizing an expected utility for systems on Borel spaces with state-space augmentation, analyzing exponential utility as a special case [17]. Another line of work considers the optimization of recursive risk functionals [21]–[23], [33]; the basic approach is to replace a conditional expectation with a "conditional risk functional" to derive a risk-based Bellman equation. The problem of minimizing an expected cumulative cost subject to a risk constraint has been studied by, e.g., [9], [11], [29], [30], [35], [37]. Linear-quadratic settings have been studied in [11], [35], [37], and a safety analysis problem with CVaR has been proposed by [9]. Our problem (1) assesses the risk of the entire trajectory, whereas the framework in [9] is concerned with the risk of each state in the trajectory separately, i.e., $\mathrm{CVaR}_\alpha(\psi(X_t))$ must be small for every t. An emerging research direction proposes risk-averse signal temporal logic specifications for linear-quadratic model predictive control [11] and for a setting with continuous-time systems of the form $\dot{x} = f(x) + g(x)u$ [13]. We refer the reader to our survey about risk-averse autonomous systems [34] and the references therein for additional literature.

Contributions. We show that any collection of risk-averse safe sets is characterized exactly using the solutions to a family of stochastic dynamic programs on an augmented state space under a measurable selection assumption. We derive this characterization by expressing the minimum CVaR (for a given initial state $\mathbf{x}$ and a given level α) as a nested optimization problem with respect to a control policy and a dual parameter s. We propose a nonstandard stochastic dynamic program that is parametrized by s to assess a maximum random cost. We show that the algorithm returns an optimal s-dependent value function and policy under regularity conditions on the dynamics functions, stage and terminal cost functions, and disturbance distributions. Subsequently, we perform an outer minimization over s to obtain $\mathcal{J}^*_\alpha(\mathbf{x})$ (1). The framework permits nonlinear dynamics, non-Gaussian noise, nonconvex bounded cost functions, and continuous spaces. We solve the risk-averse safety analysis problem, whereas our prior works [10], [12] provide approximations. For detailed derivations of our theory, we refer the interested reader to [36, Appendix].

The numerical tractability of the method is limited due to its reliance on DP and an augmented state space. In this work, we provide a nonlinear two-dimensional example motivated by a stormwater management application and offer a comparison to our underapproximation method from [12]. Our on-going and future work involves developing more scalable approaches using extreme value theory and value function approximations.

Notation. We define $\mathbb{R}^* := \mathbb{R} \cup \{+\infty, -\infty\}$ and $\mathbb{N} := \{1, 2, \dots\}$. Given $N \in \mathbb{N}$, we define $\mathbb{T} := \{0, 1, \dots, N-1\}$ and $\mathbb{T}_N := \mathbb{T} \cup \{N\}$. If $\mathcal{M}$ is a metrizable space, then $\mathcal{B}_\mathcal{M}$ is the Borel sigma algebra on $\mathcal{M}$. If $g : \mathcal{M} \to \mathbb{R}^*$, then $\min_{x\in\mathcal{M}} g(x)$ means that there is a point $x^* \in \mathcal{M}$ such that $g(x^*) = \inf_{x\in\mathcal{M}} g(x)$; i.e., g attains its infimum, and x^* is a minimizer. If $g' : \mathcal{M}' \to \mathcal{M}$, where $\mathcal{M}'$ is a metrizable space, then we define $g \circ g' : \mathcal{M}' \to \mathbb{R}^*$ by $(g \circ g')(y) := g(g'(y))$. If $\mathcal{M}$ is a Borel space, then $\mathcal{P}(\mathcal{M})$ is the space of probability measures on $(\mathcal{M}, \mathcal{B}_\mathcal{M})$ with the weak topology; if $x \in \mathcal{M}$, then δ_x is the Dirac measure in $\mathcal{P}(\mathcal{M})$ that is concentrated at x. We distinguish between random objects and their realizations (i.e., values) using capital letters and lowercase letters, respectively. The abbreviation l.s.c. means lower semi-continuous.

II. Control System Model

We consider a fully observable MDP operating on a finite discrete-time horizon $\mathbb{T}_N$, where $N \in \mathbb{N}$ is given. The state space S, control space C, and disturbance space D are nonempty Borel spaces. X_t, U_t, and W_t are random objects, whose codomains are S, C, and D, respectively.[2] The disturbance process $(W_0, W_1, \dots, W_{N-1})$ satisfies the following property: for every $t \in \mathbb{T}$, given (X_t, U_t), W_t is conditionally independent of W_τ for every $\tau \neq t$. The realizations of X_0 are concentrated at an arbitrary element $\mathbf{x}$ of S. For every $t \in \mathbb{T}$, $p_t(\cdot|\cdot,\cdot)$ is a Borel-measurable stochastic kernel on D given $S \times C$, providing a conditional distribution for the realizations of W_t. For every $t \in \mathbb{T}$, if $(x, u) \in S \times C$ is the realization of (X_t, U_t), then the probability that X_{t+1} is in $\underline{S} \in \mathcal{B}_S$ is defined by

$$q_t(\underline{S}|x, u) := p_t\big(\{w \in D : f_t(x, u, w) \in \underline{S}\}\big|x, u\big), \qquad (2)$$

where $f_t : S \times C \times D \to S$ is a Borel-measurable function for the dynamics. The stage cost function $c_t : S \times C \to \mathbb{R}$ for every $t \in \mathbb{T}$ and the terminal cost function $c_N : S \to \mathbb{R}$ are Borel-measurable.

Assumption 1 (Measurable selection): We assume:

1) There exist $a \in \mathbb{R}$ and $b \in \mathbb{R}$ such that $a \leq c_t \leq b$ for every $t \in \mathbb{T}_N$. (We define $\mathcal{Z} := [a, b]$.)
2) The control space C is compact.
3) For every t, f_t and c_t are continuous functions, and $p_t(\cdot|\cdot,\cdot)$ is a continuous stochastic kernel.

We will show that Assumption 1 guarantees the existence of an optimal policy that depends on the dynamics of a running maximum (Sec. IV). It is standard to impose a measurable selection assumption for stochastic optimal control problems on Borel spaces, e.g., see [32]. As risk-aware MDP problems can pose additional technical challenges, it is common to assume bounded costs, e.g., [17], [28], [30], [33]. We assume continuous cost functions c_t because our cost-update operation is a composition of two functions (rather than a sum). Hence, we replace the typical l.s.c. assumption by a property that is preserved under compositions. In the theoretical sections of this work, we assume that Assumption 1 holds, even without an explicit statement.

III. Risk-Averse Safety Analysis

First, we will present an example of the maximum random cost Y (1b) in terms of a desired operating region K. Then, we will provide measure-theoretic definitions of Y and CVaR to formalize our risk-averse safety specification $\mathcal{S}^r_\alpha$.

[2]The realizations of X_t, U_t, and W_t include the possible states, controls, and disturbances at time t, respectively.

A. Y as a Distance between the State Trajectory and K

Suppose that $K \in \mathcal{B}_S$ is a desired operating region. While we would like the state trajectory to remain inside K always, this may not be possible due to disturbances that may arise. We will explain how one may choose Y (1b) to represent a distance between the state trajectory and K.

Let $g_K : S \to \mathbb{R}$ be bounded and continuous, where $g_K(x)$ quantifies a signed distance between a state x and the boundary of K. For example, if $S \in \mathcal{B}_{\mathbb{R}^2}$ is bounded and $K = [0, k_1] \times [0, k_2] \subset S$ is the set of desired water levels in two storage tanks, then $\max\{x_1 - k_1, x_2 - k_2, 0\}$ or $\max\{x_1 - k_1, x_2 - k_2\}$ are suitable choices for $g_K(x)$ with $x = [x_1, x_2]^T \in S$. More generally, if x is outside K and far from its boundary, then $g_K(x)$ has a large positive value. Otherwise, if x is inside K, then there are two options: 1) $g_K(x)$ equals zero, or 2) $g_K(x)$ equals a more negative value if x is located more deeply inside K. The former applies when there is no preference for certain trajectories inside K. The latter applies when there is a preference for trajectories that are inside K *and* farther from its boundary.

To quantify the extent of the state trajectory's departure relative to K, we can choose the terminal and stage cost functions to be g_K. That is, we can choose $c_N = g_K$ and $c_t(x, u) = g_K(x)$ for every $t \in \mathbb{T}$ and $(x, u) \in S \times C$. In this case, if $(x_0, x_1, \ldots, x_N) \in S^{N+1}$ is the realization of $(X_0, X_1, \ldots, X_N)$, then $y = \max\{g_K(x_t) : t \in \mathbb{T}_N\}$ is the realization of Y (1b). In this example, Y represents the extent of the state trajectory's departure from K, and we use the notation $Y' = Y$ (Fig. 1).

B. A CVaR-based Trajectory-wise Safety Specification

To define risk-averse safe sets formally, we must describe Y (1b) in measure-theoretic terms. Let $\mathbf{x} \in S$ be an initial state and $\pi \in \Pi$ be a control policy. (We will specify the control policy class Π in Sec. IV.) Y is a random variable defined on a probability space $(\Omega, \mathcal{B}_\Omega, P^\pi_{\mathbf{x}})$. The sample space Ω contains all possible trajectories; a trajectory is a tuple of states, maximum stage costs, and controls over time. From Assumption 1, every c_t is bounded below by $a \in \mathbb{R}$. Given $(\mathbf{x}, a)$, π, and the system dynamics, there exists a unique probability measure $P^\pi_{\mathbf{x}} \in \mathcal{P}(\Omega)$ (Ionescu-Tulcea Theorem). We write $P^\pi_{\mathbf{x}}$ instead of $P^\pi_{\mathbf{x},a}$ for brevity. $E^\pi_{\mathbf{x}}(\cdot)$ denotes the expectation operator with respect to $P^\pi_{\mathbf{x}}$. Since the stage and terminal cost functions are bounded (Assumption 1), Y is bounded everywhere. This is one way to ensure that $E^\pi_{\mathbf{x}}(|Y|)$ is finite, which will allow us to define $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$.

As we have mentioned, $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ represents the average value of Y in the $\alpha \cdot 100\%$ worst cases when the initial state is $\mathbf{x}$ and the system uses the control policy π. The meaning of the $\alpha \cdot 100\%$ worst cases is made precise using a quantity called the Value-at-Risk of Y at level α, which we denote by $\mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y)$. Formally, $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ is the expectation of Y conditioned on the event that Y exceeds $\mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y)$, provided that $\alpha \in (0, 1)$ and the distribution function of Y is continuous at $\mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y)$ [24, Th. 6.2]. The Value-at-Risk of Y at level $\alpha \in (0, 1)$ is defined by

$$\mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y) := \inf\{y \in \mathbb{R} : P^\pi_{\mathbf{x}}(\{Y \le y\}) \ge 1 - \alpha\}, \quad (3)$$

where $y \mapsto P^\pi_{\mathbf{x}}(\{Y \le y\})$ is the distribution function of Y. Now, for every $\alpha \in (0, 1]$, we define $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ by

$$\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y) := \inf_{s \in \mathbb{R}} \left(s + \tfrac{1}{\alpha} E^\pi_{\mathbf{x}}(\max\{Y - s, 0\}) \right), \quad (4)$$

following Shapiro et al. [24, Eq. (6.22)]. We call $s \in \mathbb{R}$ a *dual parameter*. Using the derivation from [24, p. 258], one can show that if $\alpha \in (0, 1)$, then $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ equals

$$\mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y) + \tfrac{1}{\alpha} E^\pi_{\mathbf{x}}(\max\{Y - \mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y), 0\}). \quad (5)$$

This relation implies that $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ assesses a probability-weighted average of the realizations of Y above $\mathrm{VaR}^\pi_{\alpha,\mathbf{x}}(Y)$.

CVaR is an attractive choice for defining safety specifications for two reasons. First, the parameter α has a quantitative interpretation as a fraction of the worst cases. Second, CVaR assesses the part of a distribution *above* a particular quantile and therefore is designed to assess more rare and harmful outcomes. We define risk-averse safe sets $\mathcal{S}^r_\alpha$ as the sublevel sets of the optimal CVaR of the maximum random cost Y.

Definition 1 ($\mathcal{S}^r_\alpha$): For every $\alpha \in (0, 1]$ and $r \in \mathbb{R}$, we define the (α, r)-risk-averse safe set by $\mathcal{S}^r_\alpha := \{\mathbf{x} \in S : \mathcal{J}^*_\alpha(\mathbf{x}) \le r\}$ with $\mathcal{J}^*_\alpha(\mathbf{x}) := \inf_{\pi \in \Pi} \mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ (1).

In the next section, we will show that risk-averse safe sets can be characterized exactly using stochastic dynamic programs on an augmented state space.

IV. Characterization of Risk-averse Safe Sets using Stochastic Dynamic Programs

Unlike the minimum expectation of a cumulative cost, $\mathcal{J}^*_\alpha$ cannot be computed using a DP recursion on the state space S alone. Such a recursion holds in special cases due to the structure inherent in certain problems, but it does not hold universally. To alleviate the challenge of optimizing the CVaR of a maximum cost, we will construct an augmented state space to record the running maximum. Recall that $\mathcal{Z} = [a, b]$.

A. Construction of an Augmented State Space

We define the random augmented state by $\mathcal{X}_t := (X_t, Z_t)$ for every $t \in \mathbb{T}_N$. X_t is the original S-valued random state. Z_t is a $\mathcal{Z}$-valued random object that records the maximum stage cost up to time t (to be further described). The realizations of $\mathcal{X}_0 = (X_0, Z_0)$ are concentrated at $(\mathbf{x}, a)$, where we recall that $\mathbf{x} \in S$ is arbitrary. Z_{t+1} depends on X_t, U_t, and Z_t as follows: $Z_{t+1} = \max\{c_t(X_t, U_t), Z_t\}$ for every $t \in \mathbb{T}$. We define $\mathbb{S} := S \times \mathcal{Z}$ for brevity.

$\mathcal{X}_t$ and U_t are functions defined on $\Omega := (\mathbb{S} \times C)^N \times \mathbb{S}$. Every $\omega \in \Omega$ takes the form

$$\omega = (x_0, z_0, u_0, \ldots, x_{N-1}, z_{N-1}, u_{N-1}, x_N, z_N) \quad (6)$$

with $(x_t, z_t) \in \mathbb{S}$ for every $t \in \mathbb{T}_N$ and $u_t \in C$ for every $t \in \mathbb{T}$. We define $\mathcal{X}_t(\omega) := (X_t(\omega), Z_t(\omega)) := (x_t, z_t)$ and $U_t(\omega) := u_t$ for every $\omega \in \Omega$ whose coordinates are specified by (6). It follows that $\mathcal{X}_t$ and U_t are Borel-measurable functions. While these definitions are general enough to capture arbitrary dependencies between the coordinates of ω, we restrict ourselves to particular casual dependencies, which we have discussed and will continue to present. Next, we will define the class Π of control policies using the augmented state space $\mathbb{S}$.

Definition 2 (Π): Every control policy $\pi \in \Pi$ takes the form $\pi = (\pi_0, \pi_1, \dots, \pi_{N-1})$, where $\pi_t(\cdot|\cdot,\cdot)$ is a Borel-measurable stochastic kernel on C given $\mathbb{S}$ for every $t \in \mathbb{T}$.

Remark 1 (Π is history-dependent): Let $\pi \in \Pi$ be given, and suppose that $(x_t, z_t) \in \mathbb{S}$ is the realization of $\mathcal{X}_t = (X_t, Z_t)$. The distribution $\pi_t(\cdot|x_t, z_t) \in \mathcal{P}(C)$ for the realizations of U_t depends on (x_t, z_t), which depends on the previous states and controls.

Remark 2 (A deterministic control law δ_κ): Let $\kappa : \mathbb{S} \to C$ be Borel-measurable. We use the notation δ_κ to denote the following Borel-measurable stochastic kernel on C given $\mathbb{S}$: for every $(x, z) \in \mathbb{S}$, $\delta_{\kappa(x,z)}$ is the Dirac measure in $\mathcal{P}(C)$ that is concentrated at the point $\kappa(x, z) \in C$.

The next remark presents a convenient notation for an element of $\mathbb{S}$ and a transition law for the realizations of $\mathcal{X}_{t+1}$.

Remark 3 (χ_t, $\tilde{q}_t$): The notation $\chi_t = (x_t, z_t)$ denotes an element of $\mathbb{S}$. For every $t \in \mathbb{T}$ and $(\chi_t, u_t) \in \mathbb{S} \times C$, let $\tilde{q}_t(\cdot|\chi_t, u_t)$ be the product measure of $q_t(\cdot|x_t, u_t)$ (2) and $\delta_{\max\{c_t(x_t,u_t),z_t\}}$. $\tilde{q}_t$ is a continuous stochastic kernel on $\mathbb{S}$ given $\mathbb{S} \times C$ by applying Assumption 1 [36, Appendix].

Now, we are ready to formalize the expectation operator $E^\pi_{\mathbf{x}}(\cdot)$. Let $\mathbf{x} \in S$ and $\pi \in \Pi$ be given. If $G : \Omega \to \mathbb{R}^*$ is Borel-measurable and $E^\pi_{\mathbf{x}}(G) \coloneqq \int_\Omega G \, \mathrm{d}P^\pi_{\mathbf{x}}$ exists, then

$$E^\pi_{\mathbf{x}}(G) = \int_{\mathbb{S}} \int_C \cdots \int_{\mathbb{S}} G(\chi_0, u_0, \dots, \chi_N) \, \tilde{q}_{N-1}(\mathrm{d}\chi_N|\chi_{N-1}, u_{N-1}) \cdots \pi_0(\mathrm{d}u_0|\chi_0) \, \delta_{\mathbf{x},a}(\mathrm{d}\chi_0), \tag{7}$$

by applying [32, Prop. 7.28] and Assumption 1 [36, Appendix]. The kernels in (7) describe how an augmented state $\chi_0 = (x_0, z_0)$ may lead to a control u_0, how (χ_0, u_0) may lead to a subsequent augmented state $\chi_1 = (x_1, z_1)$, and so on. The point $(\mathbf{x}, a)$ serves as the initial augmented state.

B. Characterization of Risk-Averse Safe Sets

Here, we show that risk-averse safe sets enjoy an equivalent representation in terms of a family of stochastic dynamic programs on the augmented state space under Assumption 1. For convenience, for every $s \in \mathbb{R}$, we define $h^s : \mathbb{R} \to \mathbb{R}$ by

$$h^s(y) \coloneqq \max\{y - s, 0\}. \tag{8}$$

Let $\mathbf{x} \in S$ and $\alpha \in (0, 1]$ be given. The optimal value $\mathcal{J}^*_\alpha(\mathbf{x})$ (1) can be expressed using the definitions of $\mathrm{CVaR}^\pi_{\alpha,\mathbf{x}}(Y)$ (4) and h^s (8) as follows:

$$\mathcal{J}^*_\alpha(\mathbf{x}) = \inf_{s \in \mathbb{R}} \Big(s + \tfrac{1}{\alpha} \inf_{\pi \in \Pi} E^\pi_{\mathbf{x}}(h^s(Y)) \Big), \tag{9}$$

where we exchange the order of the infima over $\mathbb{R}$ and Π. By the definition of Y (1b) and Assumption 1, we have that $Y(\omega) \in \mathcal{Z}$ for every $\omega \in \Omega$. Consequently, a minimizer in $\mathcal{Z}$ exists for the outer problem of (9) by the next lemma.

Lemma 1 (Existence of a minimizer): Let Assumption 1 hold, $\mathbf{x} \in S$, $\alpha \in (0, 1]$, $G : \Omega \to \mathbb{R}$ be Borel-measurable, and $G(\omega) \in [a, b]$ for every $\omega \in \Omega$. Define $L^\alpha_{\mathbf{x}}(s) \coloneqq s + \frac{1}{\alpha} \inf_{\pi \in \Pi} E^\pi_{\mathbf{x}}(h^s(G))$ for every $s \in \mathbb{R}$. Then, $\inf_{s \in \mathbb{R}} L^\alpha_{\mathbf{x}}(s) = \min_{s \in [a,b]} L^\alpha_{\mathbf{x}}(s)$; i.e., a minimizer $s^*_{\mathbf{x},\alpha} \in [a, b]$ exists.

Proof: Define $\ell \coloneqq \inf_{s \in [a,b]} L^\alpha_{\mathbf{x}}(s)$. Then, for every $s \in [a, b]$, $L^\alpha_{\mathbf{x}}(s) \geq \ell$. Now, if $s \leq a$, then $h^s(G) = G - s$, and hence, $L^\alpha_{\mathbf{x}}(s) \geq L^\alpha_{\mathbf{x}}(a) \geq \ell$. However, if $s \geq b$, then $h^s(G) = 0$, and thus, $L^\alpha_{\mathbf{x}}(s) \geq L^\alpha_{\mathbf{x}}(b) \geq \ell$. Since $L^\alpha_{\mathbf{x}}(s) \geq \ell$ for every $s \in \mathbb{R}$, $\ell = \inf_{s \in \mathbb{R}} L^\alpha_{\mathbf{x}}(s)$ holds. Since $L^\alpha_{\mathbf{x}}(s)$ is continuous in s and $[a, b]$ is compact, the infimum ℓ is attained by a point $s^*_{\mathbf{x},\alpha} \in [a, b]$ [38, Th. A6.3]. ∎

For every $s \in \mathbb{R}$, we define $V^s : S \to \mathbb{R}^*$ by

$$V^s(\mathbf{x}) \coloneqq \inf_{\pi \in \Pi} E^\pi_{\mathbf{x}}(h^s(Y)). \tag{10}$$

By Lemma 1, there exists a point $s^*_{\mathbf{x},\alpha} \in \mathcal{Z}$ such that

$$\mathcal{J}^*_\alpha(\mathbf{x}) = \min_{s \in \mathcal{Z}} \big(s + \tfrac{1}{\alpha} V^s(\mathbf{x}) \big) = s^*_{\mathbf{x},\alpha} + \tfrac{1}{\alpha} V^{s^*_{\mathbf{x},\alpha}}(\mathbf{x}). \tag{11}$$

We will develop a dynamic programming-based solution for V^s to characterize $\mathcal{J}^*_\alpha$. Toward this aim, we define extended random variables that represent costs-to-go. For every $s \in \mathbb{R}$ and $t \in \mathbb{T}_N$, we define $Y^s_t : \Omega \to \mathbb{R}^*$ by

$$Y^s_t \coloneqq \begin{cases} h^s(\max\{c_N(X_N), A_t, Z_t\}), & \text{if } t \in \mathbb{T}, \\ h^s(\max\{c_N(X_N), Z_N\}), & \text{if } t = N, \end{cases} \tag{12}$$

with $A_t : \Omega \to \mathbb{R}$, $A_t \coloneqq \max_{i \in \{t,\dots,N-1\}} c_i(X_i, U_i)$, $t \in \mathbb{T}$. The next theorem specifies some properties of a conditional expectation $\phi^{\pi,s}_t(x, z) = E^\pi(Y^s_t | \mathcal{X}_t = (x, z))$ of Y^s_t given $\mathcal{X}_t$. The theorem is based on the definition of conditional expectation [38, Th. 6.3.3] and a basic change-of-measure theorem [38, Th. 1.6.12]. For brevity, we use the notation $\int_\Omega \varphi \circ \mathcal{X}_t \, \mathrm{d}P^\pi_{\mathbf{x}} \coloneqq \int_\Omega \varphi(\mathcal{X}_t(\omega)) \, \mathrm{d}P^\pi_{\mathbf{x}}(\omega)$, where $\varphi : \mathbb{S} \to \mathbb{R}^*$ is Borel-measurable.

Theorem 1 (Properties of $\phi^{\pi,s}_t$): Let Assumption 1 hold, and let $\mathbf{x} \in S$, $\pi \in \Pi$, and $s \in \mathbb{R}$ be given. Define the function $J^s_N : \mathbb{S} \to \mathbb{R}^*$ by

$$J^s_N(x, z) \coloneqq h^s(\max\{c_N(x), z\}). \tag{13}$$

Then, the following statements hold:

$$E^\pi_{\mathbf{x}}(h^s(Y)) = \int_\Omega \phi^{\pi,s}_0 \circ \mathcal{X}_0 \, \mathrm{d}P^\pi_{\mathbf{x}} = \phi^{\pi,s}_0(\mathbf{x}, a), \tag{14}$$

$$\int_\Omega \phi^{\pi,s}_N \circ \mathcal{X}_N \, \mathrm{d}P^\pi_{\mathbf{x}} = \int_\Omega J^s_N \circ \mathcal{X}_N \, \mathrm{d}P^\pi_{\mathbf{x}}, \tag{15}$$

$$\int_\Omega \phi^{\pi,s}_t \circ \mathcal{X}_t \, \mathrm{d}P^\pi_{\mathbf{x}} = \int_\Omega \phi^{\pi,s}_{t+1} \circ \mathcal{X}_{t+1} \, \mathrm{d}P^\pi_{\mathbf{x}}, \quad t \in \mathbb{T}. \tag{16}$$

Proof: For every $t \in \mathbb{T}_N$, Y^s_t is an extended random variable on $(\Omega, \mathcal{B}_\Omega, P^\pi_{\mathbf{x}})$, $\mathcal{X}_t : \Omega \to \mathbb{S}$ is Borel-measurable, and $\int_\Omega Y^s_t \, \mathrm{d}P^\pi_{\mathbf{x}}$ exists (recall that Y^s_t is nonnegative). The probability measure induced by $\mathcal{X}_t$ is defined by $P^\pi_{\mathbf{x},\mathcal{X}_t}(\underline{\mathbb{S}}) \coloneqq P^\pi_{\mathbf{x}}(\mathcal{X}_t^{-1}(\underline{\mathbb{S}}))$ for every $\underline{\mathbb{S}} \in \mathcal{B}_{\mathbb{S}}$. By the definition of conditional expectation [38, Th. 6.3.3] and the change-of-measure theorem [38, Th. 1.6.12], we have

$$\int_\Omega Y^s_t \, \mathrm{d}P^\pi_{\mathbf{x}} = \int_\Omega \phi^{\pi,s}_t \circ \mathcal{X}_t \, \mathrm{d}P^\pi_{\mathbf{x}}, \quad t \in \mathbb{T}_N, \tag{17}$$

where the integrals exist. Now,

$$\int_\Omega Y^s_t \, \mathrm{d}P^\pi_{\mathbf{x}} = \int_\Omega Y^s_{t+1} \, \mathrm{d}P^\pi_{\mathbf{x}}, \quad t \in \mathbb{T}, \tag{18}$$

as a consequence of $Z_{t+1} = \max\{c_t(X_t, U_t), Z_t\}$. The relations (17)–(18) imply the relation (16). The relation (14) is derived using (7) and (17) with $t = 0$; note that $E^\pi_{\mathbf{x}}(Y^s_0) = E^\pi_{\mathbf{x}}(h^s(Y))$ because $a \leq c_t$ for every $t \in \mathbb{T}_N$ and the realizations of (X_0, Z_0) are concentrated at $(\mathbf{x}, a)$. The relation (15) holds by (17) with $t = N$ and by $Y^s_N = J^s_N \circ \mathcal{X}_N$. ∎

Subsequently, we will use Theorem 1 to derive a DP-based solution for V^s (10), and we will show the existence of a control policy that is optimal for V^s under Assumption 1.

Theorem 2 (DP on $\mathbb{S}$): Let Assumption 1 hold, and let $s \in \mathbb{R}$ be given. Recall the definition of J_N^s (13). For $t = N-1, \dots, 1, 0$, we define $J_t^s : \mathbb{S} \to \mathbb{R}^*$ recursively by

$$J_t^s(x,z) \coloneqq \inf_{u \in C} v_t^s(x,z,u), \tag{19a}$$

where we define $v_t^s : \mathbb{S} \times C \to \mathbb{R}^*$ by $v_t^s(x,z,u) \coloneqq$

$$\int_D J_{t+1}^s\big(f_t(x,u,w), \max\{c_t(x,u), z\}\big)\, p_t(\mathrm{d}w|x,u). \tag{19b}$$

Then, for every $t \in \mathbb{T}_N$, J_t^s is l.s.c. and bounded below by zero. Moreover, for every $t \in \mathbb{T}$, there is a Borel-measurable function $\kappa_t^s : \mathbb{S} \to C$ such that

$$J_t^s(x,z) = v_t^s(x,z,\kappa_t^s(x,z)), \quad (x,z) \in \mathbb{S}. \tag{20}$$

We define $\pi^s \coloneqq (\delta_{\kappa_0^s}, \delta_{\kappa_1^s}, \dots, \delta_{\kappa_{N-1}^s})$, which is an element of Π. Then, for every $\mathbf{x} \in S$, we have

$$J_0^s(\mathbf{x}, a) = V^s(\mathbf{x}) = E_{\mathbf{x}}^{\pi^s}(h^s(Y)). \tag{21}$$

Proof: J_t^s being l.s.c. and bounded below by zero for every $t \in \mathbb{T}_N$ and the existence of a Borel-measurable function $\kappa_t^s : \mathbb{S} \to C$ that satisfies (20) for every $t \in \mathbb{T}$ follow from standard induction arguments. These arguments use Assumption 1, properties that are preserved under integration with respect to a continuous stochastic kernel [32, Prop. 7.30], and a measurable selection result [32, Prop. 7.33].

Next, we prove (21). We work on the probability spaces $\{(\Omega, \mathcal{B}_\Omega, P_{\mathbf{x}}^\pi) : \mathbf{x} \in S,\ \pi \in \Pi\}$. For (21), it suffices to show that for every $t \in \mathbb{T}_N$ and $\mathbf{x} \in S$,

$$\forall \pi \in \Pi, \quad \int_\Omega \phi_t^{\pi,s} \circ \mathcal{X}_t \, \mathrm{d}P_{\mathbf{x}}^\pi \geq \int_\Omega J_t^s \circ \mathcal{X}_t \, \mathrm{d}P_{\mathbf{x}}^\pi, \tag{22a}$$

$$\int_\Omega \phi_t^{\pi^s,s} \circ \mathcal{X}_t \, \mathrm{d}P_{\mathbf{x}}^{\pi^s} = \int_\Omega J_t^s \circ \mathcal{X}_t \, \mathrm{d}P_{\mathbf{x}}^{\pi^s}. \tag{22b}$$

Indeed, if $t = 0$, then the above statement implies that for every $\mathbf{x} \in S$ and $\pi \in \Pi$,

$$E_{\mathbf{x}}^\pi(h^s(Y)) \geq J_0^s(\mathbf{x}, a) = E_{\mathbf{x}}^{\pi^s}(h^s(Y)), \tag{23}$$

using (14) from Theorem 1 and the realizations of $\mathcal{X}_0$ being concentrated at $(\mathbf{x}, a)$ (7). Then, we take the infimum of the expression in (23) with respect to $\pi \in \Pi$ to derive (21). The function $\phi_t^{\pi,s}$ appears inside an integral in (22) because a conditional expectation is not unique everywhere in general [38, Th. 6.3.3]. We proceed by induction to prove (22). The base cases ($t = N$) for (22) hold by (15) from Theorem 1. Now, suppose that for some $t \in \mathbb{T}$, we have: for every $\mathbf{x} \in S$,

$$\forall \pi \in \Pi, \quad \int_\Omega \phi_{t+1}^{\pi,s} \circ \mathcal{X}_{t+1} \, \mathrm{d}P_{\mathbf{x}}^\pi \geq \int_\Omega J_{t+1}^s \circ \mathcal{X}_{t+1} \, \mathrm{d}P_{\mathbf{x}}^\pi. \tag{24}$$

Let $\mathbf{x} \in S$ and $\pi \in \Pi$ be given. To show the induction step for (22a), it suffices to show that

$$\int_\Omega J_{t+1}^s \circ \mathcal{X}_{t+1} \, \mathrm{d}P_{\mathbf{x}}^\pi \geq \int_\Omega J_t^s \circ \mathcal{X}_t \, \mathrm{d}P_{\mathbf{x}}^\pi, \tag{25}$$

by applying (16) from Theorem 1 and the induction hypothesis (24). Noting that $J_{t+1}^s \circ \mathcal{X}_{t+1} : \Omega \to \mathbb{R}^*$ is Borel-measurable and nonnegative, we use (7), the change-of-measure result [38, Th. 1.6.12], and the Fubini Theorem [38, Th. 2.6.6] to derive

$$\int_\Omega J_{t+1}^s \circ \mathcal{X}_{t+1} \, \mathrm{d}P_{\mathbf{x}}^\pi = \int_\Omega v_t^{s,\pi} \circ \mathcal{X}_t \, \mathrm{d}P_{\mathbf{x}}^\pi, \tag{26}$$

where $v_t^{s,\pi} : \mathbb{S} \to \mathbb{R}^*$ is given by

$$v_t^{s,\pi}(x,z) \coloneqq \int_C v_t^s(x,z,u)\, \pi_t(\mathrm{d}u|x,z). \tag{27}$$

Since $v_t^{s,\pi} \circ \mathcal{X}_t : \Omega \to \mathbb{R}^*$ and $J_t^s \circ \mathcal{X}_t : \Omega \to \mathbb{R}^*$ are Borel-measurable and satisfy $v_t^{s,\pi} \circ \mathcal{X}_t \geq J_t^s \circ \mathcal{X}_t \geq 0$ and (26) holds, the relation (25) follows. An induction argument for (22b) is similar. A key step is using (20) to find that $v_t^{s,\pi^s} = J_t^s$. ■

In particular, by letting any $s \in \mathbb{R}$ be the dual parameter's value and any $\mathbf{x} \in S$ be the initial state, we have shown that (21) holds under Assumption 1. Therefore, under Assumption 1, we conclude that for every $s \in \mathbb{R}$ and $\mathbf{x} \in S$, $J_0^s(\mathbf{x}, a) = V^s(\mathbf{x})$. This conclusion permits a useful characterization of risk-averse safe sets (Def. 1) in terms of the family $\{J_0^s : s \in \mathcal{Z}\}$ under Assumption 1:

$$\mathcal{S}_\alpha^r = \Big\{\mathbf{x} \in S : \min_{s \in \mathcal{Z}} \big(\, s + \tfrac{1}{\alpha} J_0^s(\mathbf{x}, a)\,\big) \leq r\Big\}. \tag{28}$$

To derive (28), we use (11) as well. Since $\{J_0^s : s \in \mathcal{Z}\}$ does not depend on α or r, the family $\{J_0^s : s \in \mathcal{Z}\}$ characterizes any collection of risk-averse safe sets $\{\mathcal{S}_\alpha^r : \alpha \in \Lambda, r \in R\}$, where Λ is a subset of $(0,1]$ and R is a subset of $\mathbb{R}$.

The results in this section provide a nonunique optimal policy on the augmented state space $\pi^{s^*_{\mathbf{x},\alpha}} \in \Pi$ under Assumption 1. Policies on augmented state spaces have also been developed by, e.g., [17], [25], [28], [30]. Nonunique optimal policies are typical in stochastic optimal nonlinear control.

Remark 4 (Policy deployment): Let $\alpha \in (0,1]$ and $\mathbf{x} \in S$ be given. Let $\pi^{s^*_{\mathbf{x},\alpha}} \in \Pi$ satisfy (21), where $s^*_{\mathbf{x},\alpha} \in \mathcal{Z}$ satisfies (11). Let $\kappa_t^{s^*_{\mathbf{x},\alpha}}$ be the control law for time $t \in \mathbb{T}$ associated with $\pi^{s^*_{\mathbf{x},\alpha}}$. Let $(x_0, z_0) = (\mathbf{x}, a)$ and $t = 0$. For $t = 0, 1, \dots, N-1$, repeat the following four steps: 1) choose $u_t = \kappa_t^{s^*_{\mathbf{x},\alpha}}(x_t, z_t)$; 2) nature provides a realization w_t of W_t according to the distribution $p_t(\cdot|x_t, u_t)$; 3) the realization (x_{t+1}, z_{t+1}) of (X_{t+1}, Z_{t+1}) is $(f_t(x_t, u_t, w_t), \max\{c_t(x_t, u_t), z_t\})$; 4) t updates by 1.

V. Numerical Example

Risk-averse safety analysis, as presented here, suffers from the curse of dimensionality inherent to DP and requires an augmented state space. Despite these computational challenges, risk-averse safety analysis may be a useful tool for designing control systems. At the design stage, large-scale offline simulations may be commonplace, and designers may be required to assess multiple alternatives in light of uncertainty.

We consider the task of modifying the design of an urban stormwater system. We refer the reader to [36] for more details about the simulation setting. The system consists of two tanks (tank 1, tank 2) connected by an automated valve. Water flows by gravity between the tanks based on the relative difference in water levels and the position of the valve. Water enters the system through a random process of surface runoff. Water exits the system through a *storm sewer* drain that is connected to tank 2 or through outlets that lead to a *combined sewer*. The storm sewer directs stormwater to a nearby water body, which occurs without penalty. However, the storm sewer's capacity is limited. When water levels are too great, excess flows are directed to a combined sewer, which can release a mixture of stormwater and untreated wastewater into a local waterway. This event is called a *combined sewer overflow* (CSO), which can disturb local ecosystems. We apply risk-averse safety analysis to examine how design modifications may reduce the

risk of CSOs by managing the system's maximum water levels. The designs are: (a) baseline, (b) replace the valve with a controllable bidirectional pump, and (c) retrofit tank 1 with an outlet that drains to a storm sewer without penalty.

$X_t = [X_{t1}, X_{t2}]^T$ is the vector of random water levels in tank 1 and tank 2 at time t. The state space is $S = [0, \bar{k}_1] \times [0, \bar{k}_2]$ ft^2, where $\bar{k}_i = k_i + 2$ ft and k_i is the maximum water level that tank i can hold without releasing water into the combined sewer. The desired operating region is $K = [0, k_1] \times [0, k_2]$. We choose $c_t = g_K$ for every $t \in \mathbb{T}_N$, where $g_K(x) = \max\{x_1 - k_1, x_2 - k_2, 0\}$ quantifies the maximum water elevation outside K. In the baseline design, the control input is the valve position at time t, and the control space is $C = [0, 1]$ (closed to open, unitless). The tuple $(W_0, W_1, \ldots, W_{N-1})$ is a random process of surface runoff that arises due to precipitation uncertainty (units: ft^3/s). We select a disturbance distribution to reflect simulations from our prior work [39]. We use Newtonian physics and a mass balance to form a dynamics function ($f = f_t \ \forall t \in \mathbb{T}$) for each design. We consider $N = 20$; $[t, t+1)$ represents a duration of 3 minutes. $\hat{\mathcal{S}}$ denotes a computation of a set $\mathcal{S}$.

Fig. 2 presents computations of $\mathcal{S}^r_\alpha$ (28). For comparison, Fig. 2 also presents computations of the underapproximations $\mathcal{U}^r_{\alpha,\gamma}$ using the method from [12]. As the risk-aversion level α becomes smaller (more pessimistic), the contours of $\hat{\mathcal{S}}^r_\alpha$ and $\hat{\mathcal{U}}^r_{\alpha,\gamma}$ contract, as we expect, while the qualitative features are preserved. The contours for the pump design (b) are more rectangular than those for the baseline design (a). The contours for the outlet design (c) are stretched farther along the x_1-axis compared to those for the baseline design; i.e., tank 1's effective capacity increases under the outlet design.

The underapproximation method [12] requires the manual tuning of a soft-max parameter γ that impacts the fidelity of the underapproximation set $\mathcal{U}^r_{\alpha,\gamma}$ at different levels of α. The current method provides $\mathcal{S}^r_\alpha$ exactly in principle and does not involve the tuning of an additional parameter, but it does require significantly more resources.[3] This is due to the augmented state space $\mathbb{S} = S \times \mathcal{Z}$ and solving dynamic programs for different values of the dual parameter s. In contrast, for a fixed γ, solving one MDP problem on S is required to compute $\mathcal{U}^r_{\alpha,\gamma}$ for every α and r of interest [12]. Consequently, the underapproximation method provides a *preliminary screening* tool to identify more promising designs from a collection of candidate designs. The current method provides a tool for *in-depth analysis* of a small number of promising designs that have been selected through preliminary screening. The risk-aversion level α allows one to specify a degree of pessimism in terms of a fraction of the worst cases. The interpretability and flexibility of α may be useful for assessing trade-offs between system performance and financial considerations in practice, especially given the limited budgets afforded to "ordinary" rather than "safety-critical" infrastructure.

[3]We provide a rough comparison of resources; we have made no attempt to optimize efficiency beyond parallelizing the operations in a given DP recursion. One can run the underapproximation method on a standard laptop (2-4 CPU cores) in approximately 10 minutes for a fixed γ and a fixed design. However, the current method takes about 13.5 hours and 30 CPU cores for a fixed design. We used the Tufts Linux Research Cluster (Medford, MA) running MATLAB (The Mathworks, Inc.). Our code is available from https://github.com/risk-sensitive-reachability/RSSAVSA-2021.

VI. Conclusions

By overcoming theoretical challenges attributed to optimizing the CVaR of a trajectory-wise maximum cost, we have shown that risk-averse safe sets enjoy an equivalent representation in terms of the solutions to a family of stochastic dynamic programs. We are investigating extensions to higher-dimensional systems in the finite-time case using extreme value theory [31] and in the infinite-time case using value function approximations. In the future, we hope to study new problems that combine performance and risk-averse safety criteria, such as optimizing a utility functional subject to a constraint on the CVaR of a maximum cost.

Acknowledgement

The authors would like to thank Dr. H. Vincent Poor and Mr. Chuanning Wei for discussions.

References

[1] 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.

[2] K. Margellos and J. Lygeros, "Hamilton-Jacobi formulation for reach-avoid differential games," *IEEE Trans. Autom. Control*, vol. 56, no. 8, pp. 1849–1861, 2011.

[3] M. Chen and C. J. Tomlin, "Hamilton-Jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management," *Annu. Rev. Control Rob. Auton. Syst.*, vol. 1, no. 1, pp. 333–358, 2018.

[4] S. L. Herbert, Safe Real-World Autonomy in Uncertain and Unstructured Environments (Doctoral dissertation). Technical Report No. UCB/EECS–2020–147. University of California Berkeley, 2020.

[5] J. Ding, M. Kamgarpour, S. Summers, A. Abate, J. Lygeros, and C. Tomlin, "A stochastic games framework for verification and control of discrete time stochastic hybrid systems," *Automatica*, vol. 49, pp. 2665–2674, 2013.

[6] I. Yang, "A dynamic game approach to distributionally robust safety specifications for stochastic systems," *Automatica*, vol. 94, pp. 94–101, 2018.

[7] 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.

[8] 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.

[9] S. Samuelson and I. Yang, "Safety-aware optimal control of stochastic systems using Conditional Value-at-Risk," in *Proc. Am. Control Conf.*, pp. 6285–6290, 2018.

[10] 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," in *Proc. Am. Control Conf.*, pp. 2958–2963, 2019.

[11] S. Safaoui, L. Lindemann, D. V. Dimarogonas, I. Shames, and T. H. Summers, "Control design for risk-based signal temporal logic specifications," *IEEE Control Syst. Lett.*, vol. 4, no. 4, pp. 1000–1005, 2020.

[12] M. P. Chapman, R. Bonalli, K. Smith, I. Yang, M. Pavone, and Claire J. Tomlin, "Risk-sensitive safety analysis using Conditional Value-at-Risk," *IEEE Trans. Autom. Control*, 2022, doi: 10.1109/TAC.2021.3131149.

[13] L. Lindemann, G. J. Pappas, and D. V. Dimarogonas, "Reactive and risk-aware control for signal temporal logic," *IEEE Trans. Autom. Control*, 2022, doi: 10.1109/TAC.2021.3120681.

[14] R. A. Howard and J. E. Matheson, "Risk-sensitive Markov decision processes," *Manage. Sci.*, vol. 18, no. 7, pp. 356–369, 1972.

[15] D. H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Trans. Autom. Control*, vol. 18, no. 2, pp. 124–131, 1973.

[16] P. Whittle, "Risk-sensitive linear/quadratic/Gaussian control," *Adv. Appl. Probab.*, vol. 13, no. 4, pp. 764–777, 1981.

[17] N. Bäuerle and U. Rieder, "More risk-sensitive Markov decision processes," *Math. Oper. Res.*, vol. 39, no. 1, pp. 105–120, 2014.

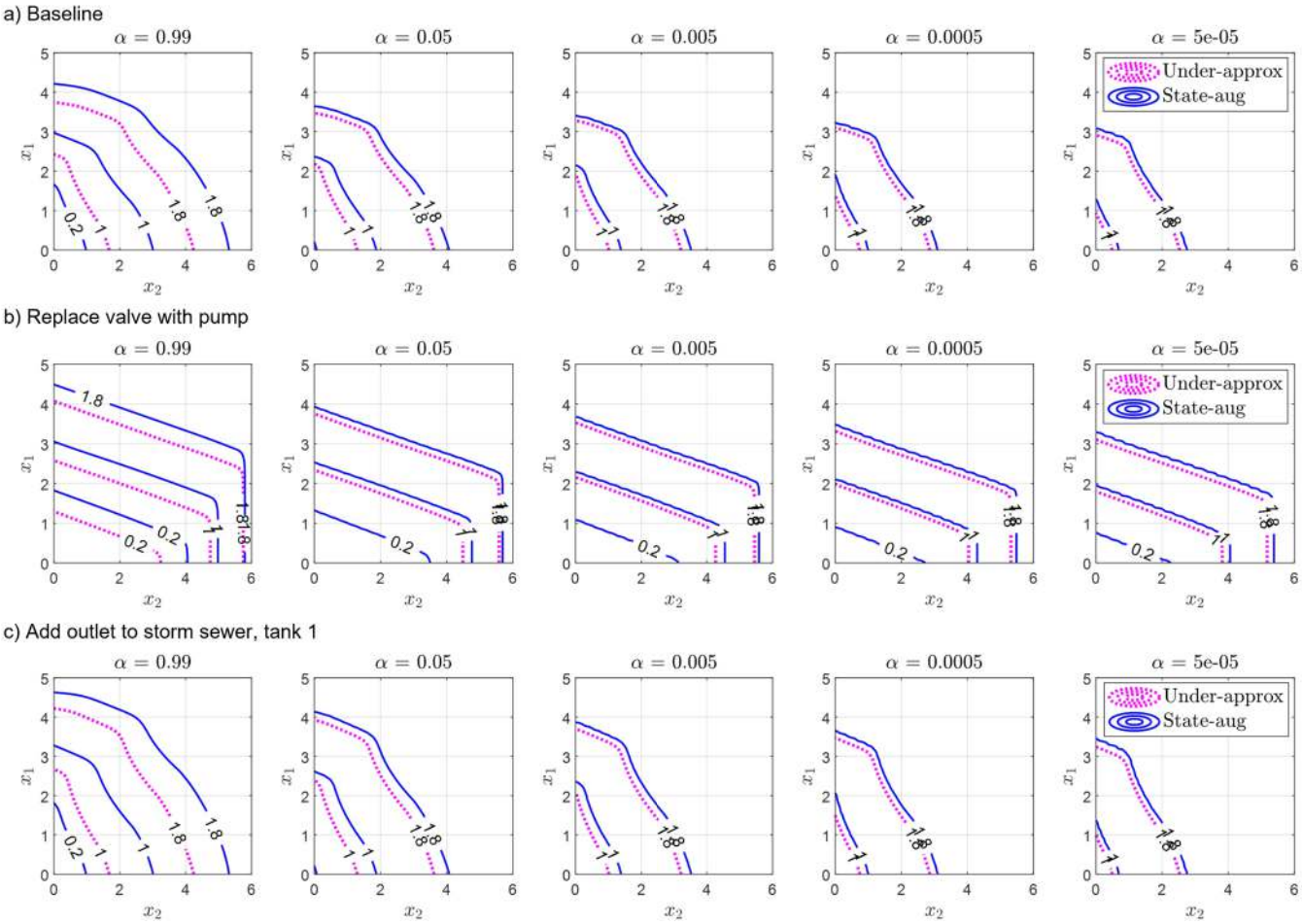

Fig. 2. Contours of computations of risk-averse safe sets for $\alpha \in \{0.99, 0.05, 0.005, 0.0005, 5 \cdot 10^{-5}\}$ and $r \in \{0.2, 1, 1.8\}$. Each row pertains to a particular design. Solid blue lines show the numerical results for $\mathcal{S}_\alpha^r$ using (28). Pink dotted lines show the numerical results for $\mathcal{U}_{\alpha,\gamma}^r$ ($\gamma = 20$) using the underapproximation method from [12]. We have explored values of γ between 10 and 120 in increments of roughly 10. We have chosen $\gamma = 20$ because this value provides relatively large estimates of $\mathcal{U}_{\alpha,\gamma}^r$ for more risk-averse values of α. We present the differences between four distinct designs qualitatively and quantitatively in the extended version [36] of this work.

[18] N. Saldi, T. Başar, and M. Raginsky, "Approximate Markov-Nash equilibria for discrete-time risk-sensitive mean-field games," *Math. Oper. Res*, vol. 45, no. 4, pp. 1596–1620, 2020.

[19] M. P. Chapman and K. M. Smith, "Classical risk-averse control for a finite-horizon Borel model," *IEEE Control Syst. Lett.*, vol. 6, pp. 1525–1530, 2021, https://arxiv.org/abs/2107.13981.

[20] K. M. Smith and M. P. Chapman, "On Exponential Utility and Conditional Value-at-Risk as risk-averse performance criteria," under review, arXiv preprint, https://arxiv.org/abs/2108.01771.

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

[22] S. Singh, Y. Chow, A. Majumdar, M. Pavone, "A framework for time-consistent, risk-sensitive model predictive control: Theory and algorithms," *IEEE Trans. Autom. Control*, vol. 64, no. 7, pp. 2905–2912, 2018.

[23] N. Bäuerle and A. Glauner, "Markov decision processes with recursive risk measures," *Eur. J. Oper. Res.*, vol. 296, no. 3, pp. 953–966, 2022.

[24] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, Philadelphia: MPS-SIAM, 2009.

[25] N. Bäuerle and A. Glauner, "Minimizing spectral risk measures applied to Markov decision processes," *Math. Methods Oper. Res.*, vol. 94, no. 1, pp. 35–69, 2021.

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

[27] R. T. Rockafellar and S. Uryasev, "Conditional Value-at-Risk for general loss distributions," *J. Bank. Financ.*, vol. 26, no. 7, pp. 1443–1471, 2002.

[28] N. Bäuerle and J. Ott, "Markov decision processes with Average-Value-at-Risk criteria," *Math. Methods Oper. Res.*, vol. 74, no. 3, pp. 361–379, 2011.

[29] V. Borkar and R. Jain, "Risk-constrained Markov decision processes," *IEEE Trans. Autom. Control*, vol. 59, no. 9, pp. 2574–2579, 2014.

[30] W. B. Haskell and R. Jain, "A convex analytic approach to risk-aware Markov decision processes," *SIAM J. Control Optim.*, vol. 53, no. 3, pp. 1569–1598, 2015.

[31] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*, New York: Springer, 2006.

[32] D. P. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*, Belmont: Athena Scientific, 1996.

[33] H. Asienkiewicz and A. Jaśkiewicz, "A note on a new class of recursive utilities in Markov decision processes," *Applicationes Mathematicae*, vol. 44, no. 2, pp. 149–161, 2017.

[34] Y. Wang and M. P. Chapman, "Risk-averse autonomous systems: A brief history and recent developments from the perspective of optimal control," Art. No. 103743, *J. Artif. Intell.*, 2022.

[35] A. Tsiamis, D. S. Kalogerias, L. F. Chamon, A. Ribeiro, and G. J. Pappas, "Risk-constrained linear-quadratic regulators," in *Proc. IEEE Conf. Decis. Control*, pp. 3040–3047, 2020.

[36] M. P. Chapman, M. Fauß, and K. M. Smith, "On optimizing the Conditional Value-at-Risk of a maximum cost for risk-averse safety analysis," arXiv preprint, https://arxiv.org/abs/2106.00776.

[37] B. P. G. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, "Distributionally robust control of constrained stochastic systems," *IEEE Trans. Automat. Control*, vol. 61, no. 2, pp. 430–442, 2015.

[38] R. B. Ash, *Real Analysis and Probability*, New York: Academic Press, 1972.

[39] M. P. Chapman, K. M. Smith, V. Cheng, D. L. Freyberg, and C. J. Tomlin, "Reachability analysis as a design tool for stormwater systems," in *Proc. IEEE Conf. Technol. Sustain.*, pp. 1–8, 2018.