On Exponential Utility and Conditional Value-at-Risk as Risk-Averse Performance Criteria

Kevin M. Smith[ID] and Margaret P. Chapman[ID], *Member, IEEE*

Abstract— **The standard approach to risk-averse control is to use the *exponential utility* (EU) functional, which has been studied for several decades. Like other risk-averse utility functionals, EU encodes risk aversion through an increasing convex mapping φ of objective costs to subjective costs. An objective cost is a realization y of a random variable Y. In contrast, a subjective cost is a realization $\varphi(y)$ of a random variable $\varphi(Y)$ that has been transformed to measure preferences about the outcomes. For EU, the transformation is $\varphi(y) = \exp((-\theta/2)y)$, and under certain conditions, the quantity $\varphi^{-1}(E(\varphi(Y)))$ can be approximated by a linear combination of the mean and variance of Y. More recently, there has been growing interest in risk-averse control using the *conditional value-at-risk* (CVaR) functional. In contrast to the EU functional, the CVaR of a random variable Y concerns a fraction of its possible realizations. If Y is a continuous random variable with finite $E(|Y|)$, then the CVaR of Y at level α is the expectation of Y in the $\alpha \cdot 100\%$ worst cases. Here, we study the applications of risk-averse functionals to controller synthesis and safety analysis through the development of numerical examples, with an emphasis on EU and CVaR. Our contribution is to examine the decision-theoretic, mathematical, and computational tradeoffs that arise when using EU and CVaR for optimal control and safety analysis. We are hopeful that this work will advance the interpretability of risk-averse control technology and elucidate its potential benefits.**

Index Terms— **Conditional value-at-risk (CVaR), exponential utility (EU), risk aversion, safety analysis, stochastic systems.**

I. Introduction

WHILE there is no universal definition of risk, there is growing recognition that measures of risk should be informed by both the *probability* and *severity* of harmful

Kevin M. Smith is with the Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155 USA, and also with OptiRTC Inc., Boston, MA 02114 USA (e-mail: kevin.smith@tufts.edu).

Margaret P. Chapman is with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: mchapman@ece.utoronto.ca).

events.[1] This represents a departure from traditional definitions of risk in many engineering disciplines, where risk has often been expressed in terms of probability alone, e.g., see [1]. The expanded definition of risk raises important questions about how to combine the probability and severity of outcomes into a useful risk measure. One approach is to simply measure risk as the expected cost, that is, the probability-weighted average of the outcomes. Such a measure is said to be *risk-neutral* because it is insensitive to the characteristics of the outcome distribution (e.g., spread, higher-order moments, and so on) around the expected cost. In contrast, we use the term *risk-sensitive* to describe risk measures that are responsive to these characteristics.

Here, we are mainly concerned with *risk aversion*, a type of risk sensitivity that generally prefers outcome distributions with smaller spreads and tail-costs (for the same expected cost). Incorporating risk aversion into the analysis and synthesis of control systems provides a potentially useful alternative to risk-neutral or worst-case methods. Indeed, summarizing a random outcome in terms of its expectation neglects other characteristics of its distribution that may have practical importance. For example, average performance measures can mask the presence of rare outcomes that would be considered ruinous and unacceptable. On the other hand, approaches that focus solely on the worst-case outcome may be too sensitive to the estimate of that outcome, or lead to designs that are unnecessarily conservative or too expensive to implement in practice. Altogether, the abovementioned limitations motivate the investigation of risk-averse methods that allow a decision-maker some flexibility between these two extremes.

Since the 1970s, algorithms have been developed to minimize a random cost incurred by a control system, where the cost is assessed in terms of a *risk-averse functional* [13], [18], [36], [38], [39], [41], [42], [43], [44], [45], [46], [48], [49]. Recent work has also sought to incorporate insights from prospect theory into control systems by making them risk-seeking toward some outcomes while being risk-averse toward others [12].[2] A survey of approaches to risk-sensitive

---

[1]This recognition is evident in the changing definitions of risk codified by the International Organization for Standardization. Compare the definition of risk as the "probability of loss or injury from a hazard" in [3] to a more contemporary standard that accounts for "consequences" as well as their "likelihood," e.g., see [2].

[2]Risk-seeking control is the other form of risk-sensitive control. Under risk-seeking control, larger spreads in the outcomes are assumed to represent opportunities rather than liabilities, and tail-rewards are considered to be more important than tail-costs.

control from an optimal control perspective can be found in our recent work [15].

Nonetheless, the question of when a particular risk-averse functional may be more suitable for the analysis or synthesis of a control system is not well-understood. We seek to shed light on this question by focusing our study on optimal control using exponential utility (EU) and conditional value-at-risk (CVaR), which are arguably the two most popular and well-established risk-averse functionals.

EU is the classical risk-averse functional in control engineering. A value iteration algorithm, where the domain of the value functions is the state space of a control system, can be derived to solve an EU-optimal control problem, e.g., see [18], [46], [48], [78]. Like other utility functionals, EU maps objective costs to subjective costs. To encode "rational" risk aversion, risk-averse utility functionals transform objective costs using an increasing and convex mapping φ [8, Ch. 1]. EU is of particular interest because under certain conditions, the quantity $\varphi^{-1}(E(\varphi(Y)))$ can be approximated by a linear combination of the mean and variance of a random cost Y. This provides a concise interpretation of risk-averse EU-optimal control as an approximation for a multiobjective mean-variance minimization problem. However, outside of the limited set of conditions where this interpretation is valid, EU-optimal control can only be said to minimize an infinite linear combination of moments. This latter case arguably makes it EU harder to interpret, parametrize, and identify practical situations in which its use is appropriate.

In financial portfolio optimization, it is conventional to approximate the first and second moments of random returns of financial assets directly rather than compute an expected utility. In this application, which typically involves the optimization of one decision at one time, approximating the first two moments may be preferred because choosing a utility function and knowing distributions exactly are not required, e.g., see [21] and the references therein. This perspective is called the *Markowitz model*, which is founded on the seminal work by Markowitz from 1952 [22].

The optimal control problem of minimizing variance subject to an equality constraint on the mean can be solved efficiently in a linear–quadratic setting via Riccati equations [24], [25]. More generally, mean-variance optimal control is computationally expensive, as it cannot be solved using dynamic programming (DP) on the original state space. Miller and Yang [23], for instance, develop an interesting bilevel optimization approach. The investigation of EU-optimal control has been motivated by its theoretical connections to mean-variance and its computational simplicity.

EU-optimal control was first studied in the context of finite state spaces in 1972 by Howard and Matheson [36]. In the following year, EU was applied to optimal control of linear systems with quadratic costs on continuous state spaces by Jacobson [13]. This theory was extended over the next two decades by Whittle [16], [18], and [19] and Whittle and Kuhn [17], in particular, to the setting of partially observable systems. In the context of linear systems with quadratic costs subject to Gaussian noise, EU-optimal control is often called LEQR or LEQG. The minimum entropy $\mathcal{H}_\infty$ controller and the infinite-time LEQR controller are equivalent [26], [27]. Connections between minimax model predictive control and model predictive control with an EU objective have been studied for linear systems with quadratic costs [70, Ch. 8.3]. EU-optimal control is a special case of mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control synthesis in the linear–quadratic case (see [28] and the references therein), cost-cumulant control [24], [31], [32], and optimizing an expected utility [46]. The EU functional has been applied to, for example, missile guidance [29], inventory control [30], control of active suspensions on vehicles [33], and control of satellite attitude [24, Sec. 5.6.1].

While expected utility optimization dates back to the 1950s, if not earlier, the optimization of *CVaR* was not studied until the early 2000s [50], [51], [52]. This functional has been studied primarily by the operations research and financial engineering communities. The CVaR of a random cost represents the expected cost in a given fraction of the worst outcomes. Unlike EU-optimal control, CVaR-optimal control problems do not satisfy Bellman's principle of optimality on the original state space, in general, [34]; the term *time-inconsistent* is used to describe such problems [35]. One approach for overcoming this issue is to utilize state-space augmentation.[3] This approach involves specifying the dynamics of a system on an enlarged state space so that Bellman's principle is, in fact, satisfied for value functions that are defined appropriately on the augmented space. One can guarantee the existence of an optimal policy that depends on the augmented state dynamics under a measurable selection condition [45], [46], [47]; such a policy may be called an optimal precommitment policy to emphasize its extra dependencies. A state-augmentation approach has been used to solve a CVaR-optimal control problem exactly in [45] and approximately in [41]. A related approach for solving a CVaR-optimal control problem (on an infinite time horizon) is to pose an infinite-dimensional linear program (LP) in occupation measures on an augmented state space [43].

Another line of research has focused on minimizing an expected cumulative cost incurred by a stochastic system subject to a CVaR constraint, e.g., see [42], [44], [49]. In particular, Samuelson and Yang [44] used this formulation to define a stagewise safety specification, where the CVaR of the stage cost at time t must be sufficiently small for each t.

More broadly, research on safety analysis for control systems has been active since at least the 1970s. In 1971, Bertsekas and Rhodes [53] proposed a safety analysis method for discrete-time systems using robust (minimax) optimal control. A safety analysis method for continuous-time systems with bounded disturbances via Hamilton–Jacobi (HJ) equations was introduced in the mid-2000s [54]. This method, which is called HJ reachability analysis, has been further developed theoretically and in applications over the last decade, see [55], [56], [57], and the references therein. The abovementioned two methods allow the computation of a set of initial conditions from which an uncertain system reaches a target set or avoids an unsafe region when subject to bounded adversarial

---

[3]While we focus on discrete-time systems with stage and terminal costs, we note that CVaR-optimal control problems for continuous-time systems with terminal costs have been studied without state-space augmentation [23].

disturbances. In general, these disturbances lack probabilistic descriptions and are assumed to realize their most detrimental values.

In 2008, Abate et al. [58] proposed a less conservative safety analysis method for stochastic systems. Stochastic safety analysis allows one to compute a set of initial conditions from which a system's probability of avoiding an unsafe region is sufficiently large [58]. This method has been extended to reach-avoid and distributionally robust settings [59], [60], [61].

Using risk-averse functionals to define safety specifications for control systems is a relatively new idea and is motivated by the practical importance of quantifying both the probability and severity of harmful outcomes. Examples from the literature include [44] and our prior work [9], [10], [11]. Samuelson and Yang [44] proposed a stagewise, risk-averse safety specification (mentioned previously), whereas we proposed a trajectorywise, risk-averse safety specification [9], [10], [11]. In this prior work, we considered the problem of minimizing the CVaR of a maximum cost of the state trajectory. We defined risk-sensitive safe sets as level sets of the optimal value function, and we derived methods for their estimation [9], [10] and computation [11].

Here, our contribution is to examine the decision-theoretic, mathematical, and computational tradeoffs that arise when using EU and CVaR for optimal control and safety analysis. We illustrate such tradeoffs by developing numerical examples of a thermostatic regulator and a stormwater system with a cumulative cost. In particular, we study how the empirical statistics of an optimal cost distribution (for a given control system and risk-averse functional) vary as the level of risk aversion varies.[4] We investigate the degree to which risk-averse EU-optimal control provides a useful approximation to mean-variance multiobjective optimization.

Notation: If S is a metrizable space, $\mathcal{B}_S$ is the Borel sigma algebra on S.[5] $\mathbb{R}_+^n := \{x \in \mathbb{R}^n : x_i \geq 0, \ i = 1, 2, \ldots, n\}$ is the non-negative orthant in $\mathbb{R}^n$. We define $\mathbb{T} := \{0, 1, \ldots, N - 1\}$ and $\mathbb{T}' := \{0, 1, \ldots, N\}$, where $N \in \mathbb{N}$ is given. We use the abbreviations: w.r.t. = with respect to, s.t. = such that, a.e. = almost everywhere or almost every, l.s.c. = lower semicontinuous, and cfs = cubic feet per second.

Organization: Section II studies how EU and CVaR encode risk aversion, and Section III presents algorithms for EU- and CVaR-optimal control. Section IV provides models of a thermostatic regulator and a stormwater system. Section V develops numerical examples of optimal control. Section VI focuses on safety analysis, and we provide concluding remarks in Section VII.

II. Quantification of Risk Aversion

Consider a random variable Z, representing a cost, that arises as a control system operates over time. Z representing a

---

[4]The use of a Pareto frontier to assess tradeoffs between competing objectives is common, for example, in reservoir management and financial portfolio optimization [21], [22], [65], [66], [67].

[5]$\mathbb{R}^n$ with the Euclidean metric, and more generally any metric space, is a metrizable space. Informally, $\mathcal{B}_S$ is a large collection of subsets of S that are "regular enough" to be measured. For a formal definition of $\mathcal{B}_S$ and further details about measure-theoretic concepts, please refer to [37], for example.
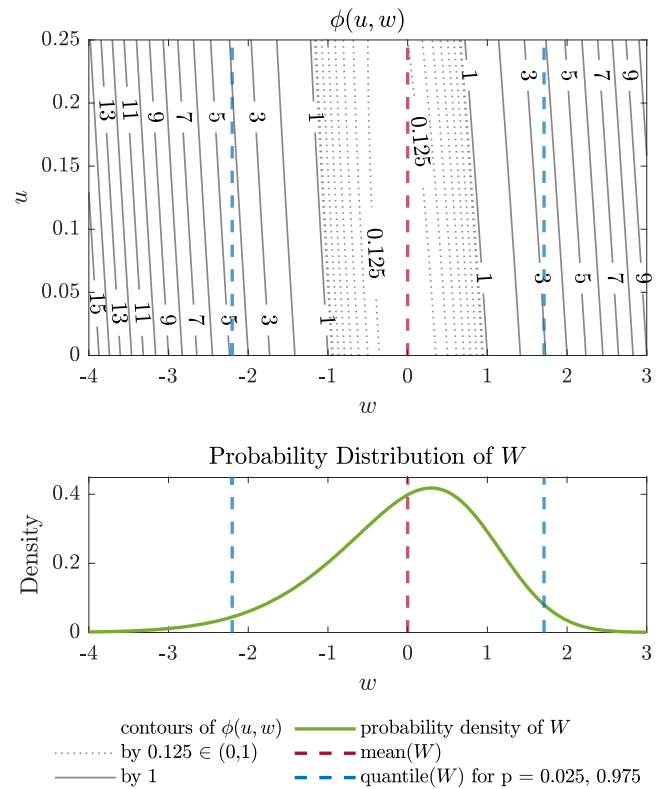

Fig. 1. **Top:** Contour plot of ϕ with vertical reference lines drawn at the 2.5%, 50%, and 97.5% quantiles of W. Note how the values of $\phi(u, W)$ change along these reference lines. The value of $\phi(u, W)$ increases slightly with increasing u when W takes on its mean value of 0. However, the value of $\phi(u, W)$ decreases (increases) more dramatically with increasing u at the 2.5% (97.5%) quantiles of W. **Bottom:** Probability plot of W with the same quantiles was marked.

cost means that smaller realizations of Z correspond to better outcomes in the real world, whereas larger realizations of Z correspond to worse outcomes. A standard stochastic control problem is to minimize the expectation of Z, subject to a given dynamics model, over a class of control policies (e.g., deterministic Markov). Different distributions can have the same expectation, but such distributions appear equivalent in the context of this problem. Focusing solely on the expectation of Z ignores other characteristics of Z (e.g., spread and higher order moments) that may have practical significance. In this sense, minimizing the expectation of Z is considered to be *risk neutral*. In contrast, we use the term *risk sensitive* to describe control problems that are aware of these characteristics. Here, we are mainly concerned with *risk-averse control*, a form of risk-sensitive control that penalizes outcome distributions with larger spreads and tail-costs (for the same expected cost). Specifically, we focus on formulating risk-averse control problems using the EU and CVaR functionals.

A. Pedagogical Example

Before presenting these functionals formally, we provide an example. Suppose that we would like to choose an input $u \in \mathbb{R}$ to minimize a quadratic cost, $\phi(u, w) := u^2 + (w + u)^2$, where $w \in \mathbb{R}$ is an unknown value of a random disturbance W with a mean of zero. If our preferences are risk neutral, the optimal choice for minimizing the expectation of $\phi(u, W)$ is $u = 0$,
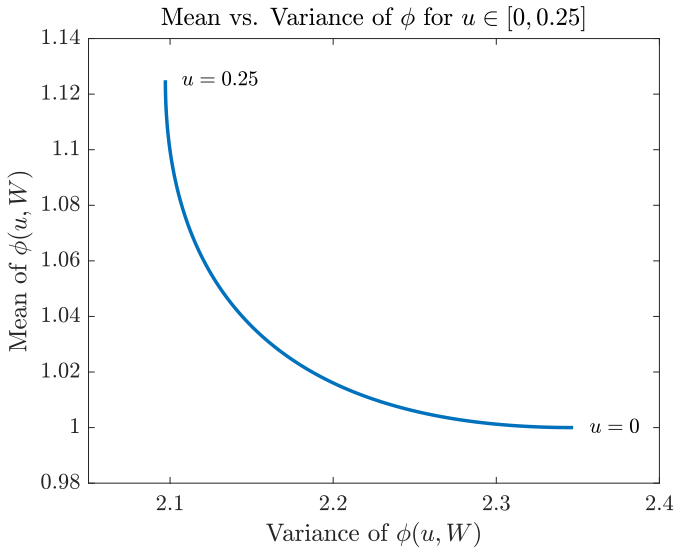
Fig. 2. This plot presents the tradeoff between the mean and variance of $\phi(u, W)$ for values of u in the range [0, 0.25].
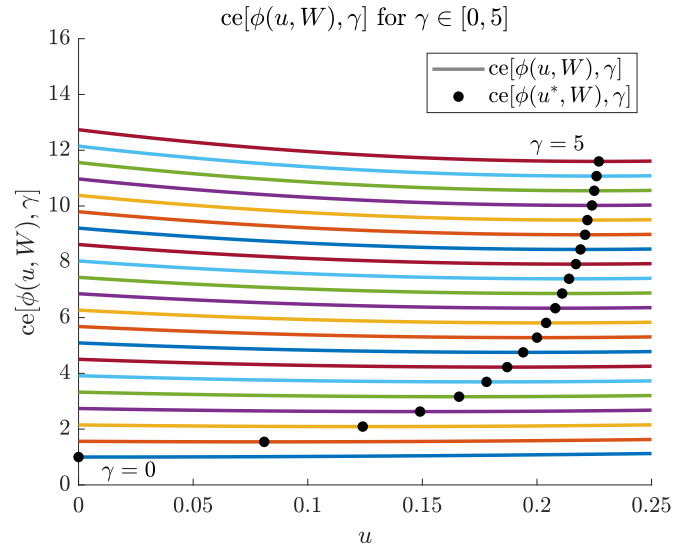

Fig. 3. These curves show the certainty equivalent values of $\phi(u, W)$ for different values of γ. The certainty equivalence relation is defined in (1), and γ reflects the price one is willing to pay to reduce the variance of $\phi(u, W)$.

regardless of the characteristics of W other than its mean. For example, assume that W follows a zero-mean skew normal distribution with unit variance and skewness of -0.5. This distribution is shown in Fig. 1, below a contour plot of ϕ. The intersections of the vertical reference lines with the ϕ-contours show how the value of $\phi(u, W)$ changes with u and how this change depends on the quantile of W. From a risk-neutral perspective, the effect of u on the quantiles of $\phi(u, W)$ is not important because only the expectation of $\phi(u, W)$ is of interest, and the expectation is minimized when $u = 0$.

However, from a risk-averse perspective, this effect is worth investigating, as there may be potential to reduce undesirable features (e.g., variance, average tail risk, and so on) of the distribution of ϕ at the expense of increases in the expectation. For example, if $u \in [0, 0.25]$, then there is a tradeoff where larger values of u can reduce the variance of $\phi(u, W)$ at the expense of increasing its expected value (Fig. 2). The amount of increase in the expectation that we are willing to endure for a reduction in the variance depends on our *risk preferences*. For example, there may be a constant price γ that we are willing to pay. Then, we can define a *certainty equivalent* objective function to minimize

$$\mathrm{ce}_\gamma(\phi(u, W)) := E(\phi(u, W)) + \gamma \cdot \mathrm{var}(\phi(u, W)). \quad (1)$$

The abovementioned relation describes a risk-averse actor's indifference between incurring a certain cost $\mathrm{ce}_\gamma(\phi(u, W))$ or an uncertain cost $\phi(u, W)$. The left-hand side of (1) is called the *certainty equivalent*. This value reflects the maximum price a rational risk-averse actor is willing to pay for insurance to avoid an uncertain outcome. In this sense, the certainty equivalent is a measure of the perceived risk of $\phi(u, W)$, and a risk-averse actor may choose a value u to minimize this risk. Fig. 3 shows plots of (1) as a function of u for a fixed $\gamma \in [0, 5]$ and identifies the optimal values of u that minimize the certainty equivalent $\mathrm{ce}_\gamma(\phi(u, W))$.

However, risk-averse actors can perceive risk in features of $\phi(u, W)$ other than its mean and variance. For example, a risk-
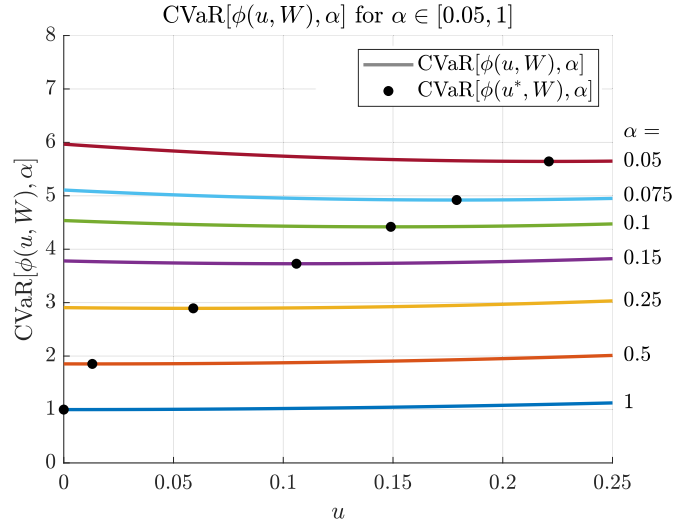

Fig. 4. These curves depict the CVaR of $\phi(u, W)$ for different values of α. CVaR is given by (2), and α reflects the fraction of the worst outcomes considered.

averse actor may express risk in terms of expected costs in the upper tail of $\phi(u, W)$. One such measure is CVaR. Let F be the cumulative distribution function (cdf) of $\phi(u, W)$ and F^{-1} its generalized inverse, the quantile function of $\phi(u, W)$. (F and F^{-1} depend on u and W, which we do not write for brevity.) Then, the CVaR of $\phi(u, W)$ at level $\alpha \in (0, 1]$ is given by

$$\mathrm{CVaR}_\alpha(\phi(u, W)) = \frac{1}{\alpha} \int_{1-\alpha}^{1} F^{-1}(\ell) \, \mathrm{d}\ell. \quad (2)$$

CVaR encodes risk by assessing the expected costs in the worst $\alpha \cdot 100\%$ of values of $\phi(u, W)$. By choosing to minimize CVaR for a particular α, one expresses a desire to minimize the expected cost in one fraction of the worst values, even if doing so increases the expected costs in other fractions. In this way, α captures a preference for a desired level of risk aversion.

Let $u \in [0, 0.25]$ be given. As Fig. 4 demonstrates, one can adopt a risk-neutral perspective and minimize the expected cost over all possible outcomes of $\phi(u, W)$ (i.e., $\alpha = 1$) by selecting $u = 0$. Alternatively, one can adopt a risk-averse perspective and focus only on minimizing the expected cost in a smaller fraction ($\alpha < 1$) of the worst outcomes by selecting a larger value of u. It is important to note that the choice of a risk-aversion level often involves a tradeoff. In this example, any u that minimizes CVaR for a particular choice of α leads to a nonoptimal CVaR at all other levels of α.

This example demonstrates two distinct ways in which risk preferences can be incorporated into an objective function: 1) through a weighted sum of expectation and variance or 2) through an average of a fraction of the largest costs. In certain circumstances, the first approach corresponds to minimizing EU, while the second approach corresponds to minimizing CVaR. As we explore in this article, EU- and CVaR-optimal control approaches offer significant differences in their interpretability and computational efficiency when applied to more complicated examples.

B. Control System Model

Now, we consider the case in which Z is a random variable, representing a cost, that arises as a control system operates over time. In particular, we consider a system on a discrete, finite time horizon of length $N \in \mathbb{N}$ of the form $x_{t+1} = f(x_t, u_t, w_t)$ for $t = 0, 1, \ldots, N-1$ where $x_t \in S$, $u_t \in A$, and $w_t \in D$ are realizations (i.e., values) of the random state X_t, the random control U_t, and the random disturbance W_t, respectively. The state space S, the control space A, and the disturbance space D are Borel spaces; e.g., $\mathbb{R}^n$ and $B \in \mathcal{B}_{\mathbb{R}^n}$ are Borel spaces [37, Definition 7.7, p. 118, Proportion 7.12, p. 119]. The initial state X_0 is fixed at an arbitrary $x \in S$. The dynamics function $f : S \times A \times D \rightarrow S$ is Borel measurable. Given (X_t, U_t), the disturbance W_t is conditionally independent of W_s for all $s \neq t$, and the distribution of W_t is known.[6] If $(x_t, u_t) \in S \times A$ is the realization of (X_t, U_t), then the distribution of W_t is $p(\cdot|x_t, u_t)$.[7]

The random cost $Z : \Omega \rightarrow \mathbb{R}$ is a Borel-measurable function, whose domain is a sample space $\Omega := (S \times A)^N \times S$. Any $\omega = (x_0, u_0, \ldots, x_{N-1}, u_{N-1}, x_N) \in \Omega$ is a realization of the random trajectory $(X_0, U_0, \ldots, X_{N-1}, U_{N-1}, X_N)$. In particular, Z takes the form, for any $\omega = (x_0, u_0, \ldots, x_{N-1}, u_{N-1}, x_N) \in \Omega$

$$Z(\omega) := c_N(x_N) + \sum_{t=0}^{N-1} c(x_t, u_t) \geq \underline{b}. \tag{3}$$

[6]Precise knowledge of the disturbance distributions and the dynamics function f are limitations of this standard formulation. If these assumptions are too strong for one's application of interest, then it may be appropriate to consider a *distributionally robust* formulation, for instance, see [49], which studies a risk-sensitive linear–quadratic setting. Online estimation of f is a growing research area, see [71], [72], [73] for some recent works.

[7]$p(\cdot|\cdot, \cdot)$ is a *Borel-measurable stochastic kernel* on D given $S \times A$ [37, Definition 7.12, p. 134], whose meaning we explain next. Let $\mathcal{P}(D)$ denote the space of probability measures on $(D, \mathcal{B}_D)$ with the weak topology [37, p. 127]. The function $\psi : S \times A \rightarrow \mathcal{P}(D)$ such that $\psi(x, u) := p(\cdot|x, u)$ is Borel measurable, i.e., measurable relative to $\mathcal{B}_{S \times A}$ and $\mathcal{B}_{\mathcal{P}(D)}$. Later, to guarantee the existence of an optimal policy, we assume that ψ is continuous, which holds if $p(\cdot|x, u)$ is constant in (x, u), for example.

The stage cost $c : S \times A \rightarrow \mathbb{R}$ and the terminal cost $c_N : S \rightarrow \mathbb{R}$ are Borel measurable and bounded below by $\underline{d} \in \mathbb{R}$, and Z is bounded below by $\underline{b} := (N+1)\underline{d}$. For convenience, we define a non-negative random cost Z'

$$Z'(\omega) := c'_N(x_N) + \sum_{t=0}^{N-1} c'(x_t, u_t) = Z(\omega) - \underline{b} \tag{4}$$

where $c' := c - \underline{d}$ and $c'_N := c_N - \underline{d}$ are translated versions of c and c_N, respectively.

We consider two classes of control policies: Π is a class of history-dependent policies, and Π' is the class of deterministic Markov policies. In particular, any $\pi \in \Pi'$ takes the form $\pi = (\mu_0, \mu_1, \ldots, \mu_{N-1})$, where $\mu_t : S \rightarrow A$ is Borel measurable for all $t \in \mathbb{T}$. We present Π in detail in Section III-B2.

The risk-neutral approach for managing the uncertainty in Z is to minimize the expectation $E_x^\pi(Z)$ over the class of policies Π'. If $G : \Omega \rightarrow \mathbb{R}$ is Borel measurable, then $E_x^\pi(G) := \int_\Omega G \, dP_x^\pi$ is the expectation of G with respect to P_x^π, a probability measure on $(\Omega, \mathcal{B}_\Omega)$. P_x^π depends on an initial condition x and a policy π and provides the probabilities of the states and controls being in Borel-measurable subsets of S and A, respectively [37, pp. 190–191], [62, p. 16].

In contrast to the risk-neutral approach, we study two risk-averse approaches, where we aim to minimize the EU of Z or the CVaR of Z. We require some conditions to ensure that the optimal values of the problems of interest are finite.

Assumption 1: We assume the following conditions.
1) Let $\Theta \subseteq (-\infty, 0)$ be nonempty. For all $\theta \in \Theta$, there is a $\pi_\theta \in \Pi'$ s.t. $E_x^{\pi_\theta}(e^{(-\theta/2)Z'}) < +\infty$ for all $x \in S$.
2) There is a $\pi \in \Pi$, s.t., $E_x^\pi(|Z|) < +\infty$ for all $x \in S$.

For example, if c and c_N are bounded, then Assumption 1 holds. Next, we introduce the EU and CVaR functionals.

C. Exponential Utility

The EU functional assesses larger values of a random cost through an exponential transformation that depends on a parameter $\theta \in \Theta$. For any $x \in S$ and $\pi \in \Pi'$, the EU of Z at level $\theta \in \Theta$ is given by

$$\rho_{\theta,x}^\pi(Z) := \underline{b} + \frac{-2}{\theta} \log E_x^\pi\left(e^{\frac{-\theta}{2}Z'}\right) = \frac{-2}{\theta} \log E_x^\pi\left(e^{\frac{-\theta}{2}Z}\right). \tag{5}$$

We define $\rho_{\theta,x}^\pi(Z)$ so that $(-\theta/2) > 0$ multiplies Z', which is a non-negative random variable. If θ is more negative, then larger values of Z' are considered to be more harmful, and therefore, more critical to assess when synthesizing a policy. Hence, a more negative value of θ represents a higher degree of risk aversion. The equality in (5) holds as a consequence of $Z = \underline{b} + Z'$. It can be shown under certain conditions that $\lim_{\theta \to 0} \rho_{\theta,x}^\pi(Z) = \underline{b} + E_x^\pi(Z')$. If $|\theta|$ is sufficiently small and if $\{E_x^\pi((Z')^n) : n \in \mathbb{N}\}$ is bounded, then the EU of Z approximates a weighted sum of the expectation $E_x^\pi(Z)$ and variance $\text{var}_x^\pi(Z)$

$$\rho_{\theta,x}^\pi(Z) \approx \underline{b} + E_x^\pi(Z') - \frac{\theta}{4}\text{var}_x^\pi(Z') = E_x^\pi(Z) - \frac{\theta}{4}\text{var}_x^\pi(Z). \tag{6}$$

For details regarding the limit result or (6), please refer to [16, p. 765] or the supplementary material.

D. Conditional Value-at-Risk

CVaR uses quantiles rather than a transformation to assess larger realizations of a random cost. Let $x \in S$ and $\pi \in \Pi$ be given. We denote the CVaR of Z with respect to P_x^π at level $\alpha \in (0, 1]$ by $\text{CVaR}_{\alpha,x}^\pi(Z)$.[8] $\text{CVaR}_{\alpha,x}^\pi(Z)$ can be written in terms of the left-side $(1 - \alpha)$-quantile of the distribution of Z. This quantile is called the *value-at-risk* of Z at level $\alpha \in (0, 1)$, which is defined by

$$\text{VaR}_{\alpha,x}^\pi(Z) := \inf \left\{ z \in \mathbb{R} : P_x^\pi\big(\{Z \leq z\}\big) \geq 1 - \alpha \right\} \quad (7)$$

where $F_{Z,x}^\pi(z) := P_x^\pi(\{Z \leq z\})$ is the cdf of Z for the initial condition x and policy π.[9] $\text{CVaR}_{\alpha,x}^\pi(Z)$ is the expectation of Z conditioned on the event $\{Z \geq \text{VaR}_{\alpha,x}^\pi(Z)\}$, if $\alpha \in (0, 1)$, $F_{Z,x}^\pi$ is continuous at $z = \text{VaR}_{\alpha,x}^\pi(Z)$, and $E_x^\pi(|Z|)$ is finite [68, Th. 6.2]. This fact motivates the name *conditional value at risk*. In this setting, $\text{CVaR}_{\alpha,x}^\pi(Z)$ is the expectation of Z in the $\alpha \cdot 100\%$ worst cases. The parameter α is a risk-aversion level that represents a fraction of the largest values of Z that are of particular concern. Another name for CVaR is *average value-at-risk* because

$$\text{CVaR}_{\alpha,x}^\pi(Z) = \frac{1}{\alpha} \int_{1-\alpha}^1 \text{VaR}_{1-\ell,x}^\pi(Z) \, d\ell \quad (8)$$

provided that $E_x^\pi(|Z|) < +\infty$ and $\alpha \in (0, 1]$; see [68, Th. 6.2] for a proof in the case of $\alpha \in (0, 1)$.

As outlined earlier, various representations for CVaR are used in the literature. The next representation, which often serves as the definition for CVaR [69], is convenient for optimal control problems in particular, e.g., see [11], [23], [45]. For any $\alpha \in (0, 1]$, if $E_x^\pi(|Z|) < +\infty$, then the $\text{CVaR}_{\alpha,x}^\pi(Z)$ is defined by

$$\text{CVaR}_{\alpha,x}^\pi(Z) := \inf_{s \in \mathbb{R}} \left(s + \frac{1}{\alpha} E_x^\pi(\max\{Z - s, 0\}) \right). \quad (9)$$

If $E_x^\pi(|Z|) = +\infty$, then $\text{CVaR}_{\alpha,x}^\pi(Z) := +\infty$. Note that $\text{CVaR}_{1,x}^\pi(Z)$ equals $E_x^\pi(Z)$. Equation (9) allows us to write $\text{CVaR}_{\alpha,x}^\pi(Z)$ as a weighted sum of $\text{VaR}_{\alpha,x}^\pi(Z)$ and the expectation of how much Z exceeds $\text{VaR}_{\alpha,x}^\pi(Z)$. Precisely, if $E_x^\pi(|Z|) < +\infty$ and $\alpha \in (0, 1)$, then a minimizer of the objective function in (9) is $\text{VaR}_{\alpha,x}^\pi(Z)$ [69], and thus,

$$\text{CVaR}_{\alpha,x}^\pi(Z) = \text{VaR}_{\alpha,x}^\pi(Z) + \frac{1}{\alpha} E_x^\pi\big(\max\{Z - \text{VaR}_{\alpha,x}^\pi(Z), 0\}\big). \quad (10)$$

E. Comparison Between EU and CVaR

Table I summarizes how EU and CVaR encode risk aversion differently. The approximation (6) suggests that EU encodes risk aversion in terms of the spread of the distribution of Z relative to the moment $E_x^\pi(Z)$. Equation (10) indicates that CVaR encodes risk aversion in terms of the expected exceedance of Z relative to the quantile $\text{VaR}_{\alpha,x}^\pi(Z)$.

---

[8]For the definition of P_x^π in this context, please refer to the supplementary material (p. 4).

[9]Another name for $\text{VaR}_{\alpha,x}^\pi(Z)$ is the generalized inverse cdf of Z at level $1 - \alpha$.

TABLE I
HOW EU AND CVAR ENCODE RISK AVERSION

	Exponential Utility (EU) (5)	Conditional Value-at-Risk (CVaR) (9)
Approach	Quantifies a distribution in terms of its moments.	Quantifies a distribution in terms of its quantiles.
Interpretation	Approximates a weighted sum of mean and variance if $|\theta|$ is small.	Approximates an expectation in a fraction of worst cases.
Parameter	$\theta \in \Theta \subseteq (-\infty, 0)$	$\alpha \in (0, 1]$
More risk-averse	θ is more negative.	α is near 0.
Less risk-averse	θ is near 0.	α is near 1.

In Section III, we present optimal control problems, in which we define the objective functions in terms of EU and CVaR. Subsequently, we solve these problems numerically for a thermostatic regulator and for a stormwater system. Then, we simulate trajectories under a policy that has been optimized numerically with respect to EU or CVaR. We use such simulations to estimate an optimal distribution of Z and to examine tradeoffs that arise in the empirical statistics of Z, as the risk aversion level varies.

III. ALGORITHMS FOR OPTIMAL CONTROL WITH EU AND CVAR OBJECTIVES

We present two distinct approaches for risk-averse optimal control.

Problem 1 (EU-Optimal Control): Consider an optimal control problem in which the EU at level $\theta \in \Theta$ is used to assess Z, $V_\theta^*(x) := \inf_{\pi \in \Pi'} \rho_{\theta,x}^\pi(Z)$ for all $x \in S$. If there is a $\pi_\theta^* \in \Pi'$ such that $V_\theta^*(x) = \rho_{\theta,x}^{\pi_\theta^*}(Z)$ for all $x \in S$, π_θ^* is said to be *optimal for V_θ^*.*

Problem 2 (CVaR-Optimal Control): Consider an optimal control problem in which the CVaR at level $\alpha \in (0, 1]$ is used to assess Z, $J_\alpha^*(x) := \inf_{\pi \in \Pi} \text{CVaR}_{\alpha,x}^\pi(Z)$ for all $x \in S$. Π is a class of history-dependent policies, which we specify formally later in this section. If there is a $\pi_\alpha^* \in \Pi$ such that $J_\alpha^*(x) = \text{CVaR}_{\alpha,x}^{\pi_\alpha^*}(Z)$ for all $x \in S$, π_α^* is said to be *optimal for J_α^*.*

Remark 1 (Finiteness of Optimal Value Functions): The first (second) condition of Assumption 1 and Z being bounded below by $\underline{b} \in \mathbb{R}$ imply that V_θ^* (J_α^*) is finite.

A DP algorithm on the state space S can be defined to compute V_θ^* under some conditions (to be presented). An algorithm is provided by [20, Problem 7(b), p. 66] in a setting without θ, for example. However, CVaR does not satisfy a DP recursion on S, and thus, computing J_α^* requires more complicated algorithms. Next, we present a method to compute V_θ^* exactly in principle. Then, we present different approaches to estimate J_α^* or compute J_α^* exactly in principle.

A. DP for EU-Optimal Control

The following assumption permits the exact computation of V_θ^* in principle and guarantees the existence of a policy that is optimal for V_θ^*.

Assumption 2 (Measurable Selection Condition):

1) The distribution of W_t, $p(\cdot|\cdot, \cdot)$, is a continuous stochastic kernel on D given $S \times A$ (Footnote 7).
2) The dynamics function f is continuous. c and c_N are l.s.c. and bounded; i.e., $\underline{d} \leq c \leq \bar{d}$ and $\underline{d} \leq c_N \leq \bar{d}$.
3) The set of controls A is compact.

Remark 2 (Justification of Assumption 2): Assumption 2 is a measurable selection condition. Such conditions are used for stochastic control problems on continuous spaces, in which the costs are nonquadratic or the dynamics function is nonlinear. The conditions guarantee the existence of an optimal policy. For additional examples, please see [46, Sec. 2, p. 106], [37, Definition 8.7, pp. 208–209], and [62, Sec. 3.3, pp. 27–29].

Next, we provide a DP algorithm for V_θ^*.

Algorithm 1 (Exact DP for V_θ^*): For any $\theta \in \Theta$, define the functions $V_N^\theta, \ldots, V_1^\theta, V_0^\theta$ on S recursively as follows: for all $x \in S$, $V_N^\theta(x) := c_N(x)$, and for $t = N-1, \ldots, 1, 0$, $V_t^\theta(x) := \inf_{u \in A} v_{t+1}^\theta(x, u)$, where $v_{t+1}^\theta(x, u) := c(x, u) + (-2/\theta) \log(\int_D \exp((-\theta/2)V_{t+1}^\theta(f(x, u, w)))p(dw|x, u))$ for all $(x, u) \in S \times A$.

We have studied Algorithm 1 formally in [78], and we summarize the analysis next. Under Assumption 2, for all $t \in \mathbb{T}'$, V_t^θ is l.s.c. and bounded, and for all $t \in \mathbb{T}$, there is a Borel-measurable function $\mu_t^\theta : S \to A$ such that $V_t^\theta(x) = v_{t+1}^\theta(x, \mu_t^\theta(x))$ for all $x \in S$. The policy $\pi_\theta^* := (\mu_0^\theta, \mu_1^\theta, \ldots, \mu_{N-1}^\theta)$ satisfies $V_\theta^*(x) = \rho_{\theta,x}^{\pi_\theta^*}(Z) = V_0^\theta(x)$ for all $x \in S$, and in particular, π_θ^* is optimal for V_θ^*.

Remark 3 (Algorithm 1 Restricted to Non-Negative Costs): It is common to define the DP iterates $V_N'^\theta, \ldots, V_1'^\theta, V_0'^\theta$ in terms of the non-negative costs c_N' and c'. Similarly in this setting, under Assumption 2, there is a policy $\pi_\theta^* \in \Pi'$ such that $V_\theta^*(x) = \rho_{\theta,x}^{\pi_\theta^*}(Z) = \underline{b} + V_0'^\theta(x)$ for all $x \in S$.

In this article, we numerically investigate Algorithm 1 with and without the restriction to non-negative costs. In Section V, we show that the unrestricted form of Algorithm 1 permits a larger range of numerically stable values of θ.

In the absence of quadratic costs, linear dynamics, and an analytical expression for the disturbance kernel $p(\cdot|\cdot, \cdot)$, one typically implements a DP algorithm numerically. While the details of such implementations are problem-dependent, a standard implementation involves discretizing continuous spaces and interpolating values that are computed on discrete spaces. That is, the state space S and the control space A may be replaced by nonempty subsets $S_G \subseteq S$ and $A_G \subseteq A$, respectively, containing finitely many elements. A discrete approximate distribution for the disturbance may be estimated from observations. To improve the efficiency of DP algorithms, approximate methods are being studied, including stochastic rollout, e.g., see [76], [77], and the references therein for details regarding the state-of-the-art. A related line of research has demonstrated the efficacy of using approximate, efficient "warm-start" computations to estimate high-fidelity, computationally expensive computations; e.g., see [79] for a robust setting with continuous-time nonstochastic systems and our work [11, Sec. VI] for a risk-averse setting with discrete-time stochastic systems. Next, we present different approaches for solving or approximating Problem 2.

B. Approaches for CVaR-Optimal Control

1) Brute Force Simulation: A brute force approach to estimate $J_\alpha^*(x) := \inf_{\pi \in \Pi} \text{CVaR}_{\alpha,x}^\pi(Z)$ is to estimate a distribution for Z for a given policy π and initial condition x via Monte Carlo simulations. Then, one can use a CVaR estimator, e.g., [68, p. 300], to estimate $\text{CVaR}_{\alpha,x}^\pi(Z)$. One can repeat this procedure for many policies and any initial conditions of interest. However, it is generally not clear which policies one should simulate, and simulating all policies is seldom feasible. Fortunately, alternate approaches are available.

2) Exact Method: Bäuerle and Ott [45] developed a method to compute J_α^* exactly in principle. The key machinery in [45] is to define the dynamics of an extra state that records a cumulative cost from time zero to the current time. A DP algorithm is defined on the augmented state space, and a deterministic Markov policy on the augmented state space is shown to be optimal under a measurable selection condition [45]. In the setting of a finite time horizon, the range of the extra state depends on the largest value of the cumulative cost, which limits computational tractability. Our numerical examples (to be shown in Section V) demonstrate this reduction in tractability compared to EU-optimal control, which does not require an extra state. Next, we present the method of [45] to compute J_α^* for the control system model of Section II-B.

1) *Define the Dynamics of the Extra State S_t:* The augmented state is (X_t, S_t). The augmented state space is $S \times \mathcal{Z}$, where $\mathcal{Z} := [-\bar{a}, \bar{a}] \subseteq \mathbb{R}$ and $0 \leq Z' \leq \bar{a} := (\bar{d} - \underline{d})(N+1)$ everywhere. For any $t \in \mathbb{T}$, if $(x_t, s_t, u_t) \in S \times \mathcal{Z} \times A$ is the realization of (X_t, S_t, U_t), then the realization of S_{t+1} is $s_{t+1} = s_t - c'(x_t, u_t)$.
2) *Define a Class of Policies Π That Are History-Dependent Through (X_t, S_t):* Any $\pi \in \Pi$ takes the form $\pi = (\pi_0, \pi_1, \ldots, \pi_{N-1})$, where $\pi_t(\cdot|\cdot, \cdot)$ is a Borel-measurable stochastic kernel on A given $S \times \mathcal{Z}$.[10]
3) *Define a CVaR-Optimal Control Problem for Z':* Define $W_\alpha^*(x) := \inf_{\pi \in \Pi} \text{CVaR}_{\alpha,x}^\pi(Z')$ for all $x \in S$. Since $Z' = Z - \underline{b}$ and Z is bounded everywhere, it holds that $W_\alpha^* = J_\alpha^* - \underline{b}$ for any $\alpha \in (0, 1]$.[11]
4) *Re-Express W_α^* Into a Useful Form for Computation:* By the definition of CVaR (9) and $0 \leq Z' \leq \bar{a}$ everywhere, it holds that, for all $x \in S$ and $\alpha \in (0, 1]$

$$W_\alpha^*(x) = \inf_{s \in \mathbb{R}} \left(s + \frac{1}{\alpha} \inf_{\pi \in \Pi} E_x^\pi(\max\{Z' - s, 0\}) \right)$$
$$= \min_{s \in [0, \bar{a}]} \left(s + \frac{1}{\alpha} \inf_{\pi \in \Pi} E_x^\pi(\max\{Z' - s, 0\}) \right)$$

where a minimizer $s_{\alpha,x}^* \in [0, \bar{a}]$ exists for the outer optimization problem [11, Lemma 1]. We define the

---

[10]Given a policy $\pi \in \Pi$, if $(x_t, s_t) \in S \times \mathcal{Z}$ is the realization of (X_t, S_t), then $\pi_t(\cdot|x_t, s_t) \in \mathcal{P}(A)$ is the distribution of U_t, where $\mathcal{P}(A)$ is the collection of probability measures on $(A, \mathcal{B}_A)$. The map $\psi : S \times \mathcal{Z} \to \mathcal{P}(A)$ such that $\psi(x_t, s_t) := \pi_t(\cdot|x_t, s_t)$ is Borel measurable.

[11]We use translation-equivariance of CVaR. That is, if Y is a random variable with finite expectation, $a \in \mathbb{R}$, and $\alpha \in (0, 1]$, then $\text{CVaR}_\alpha(Y + a) = \text{CVaR}_\alpha(Y) + a$.

inner optimization problem for any $(x, s) \in S \times \mathbb{R}$ by

$$J^*(x, s) := \inf_{\pi \in \Pi} E_x^\pi(\max\{Z' - s, 0\}) \tag{11}$$

and simplify W_α^* as follows:

$$W_\alpha^*(x) = \min_{s \in [0, \bar{a}]} \left(s + \frac{1}{\alpha} J^*(x, s) \right) = s_{\alpha,x}^* + \frac{1}{\alpha} J^*\left(x, s_{\alpha,x}^*\right). \tag{12}$$

5) *Use DP on $S \times \mathcal{Z}$ to compute J^*.* Algorithm 2, to be presented, defines a function J_t on $S \times \mathcal{Z}$ recursively for $t = N, \ldots, 1, 0$, and under appropriate conditions, it holds that $J_0 = J^*$ (Th. 1, to be presented).
6) *Use J^* to compute J_α^*.* For any $\alpha \in (0, 1]$ and $x \in S$, compute $W_\alpha^*(x)$ using (12). Then, using step 3, compute $J_\alpha^* = W_\alpha^* + \underline{b}$.

Next, we present Algorithm 2 and Theorem 1.

Algorithm 2 (Exact DP for J^*): Define the functions $J_N, \ldots, J_1, J_0$ on $S \times \mathcal{Z}$ recursively as follows: for any $(x, s) \in S \times \mathcal{Z}$, $J_N(x, s) := \max\{c_N'(x) - s, 0\}$, and for $t = N - 1, \ldots, 1, 0$, $J_t(x, s) := \inf_{u \in A} v_{t+1}(x, s, u)$, where v_{t+1} is defined by

$$v_{t+1}(x, s, u) := \int_D J_{t+1}(f(x, u, w), s - c'(x, u)) \, p(\mathrm{d}w | x, u)$$

with $(x, s, u) \in S \times \mathcal{Z} \times A$.

The following theorem provides an analysis of Algorithm 2. In the interest of space, technical details and the proof are omitted. Similar proof techniques can be found in [11], [37], [45], and [62].

Theorem 1 (Analysis of Algorithm 2): Let Assumption 2 hold, and assume that c and c_N are continuous. Then, J_t is bounded and continuous for each $t \in \mathbb{T}'$. For each $t \in \mathbb{T}$, there is a Borel-measurable function $\kappa_t : S \times \mathcal{Z} \to A$ such that $J_t(x, s) = v_{t+1}(x, s, \kappa_t(x, s))$ for all $(x, s) \in S \times \mathcal{Z}$. Define $\pi^* := (\kappa_0, \kappa_1, \ldots, \kappa_{N-1})$. It holds that $J_0(x, s) = J^*(x, s) = E_x^{\pi^*}(\max\{Z' - s, 0\})$ for all $(x, s) \in S \times \mathcal{Z}$.

Remark 4 (Policy Deployment on $S \times \mathcal{Z}$): Theorem 1 guarantees the existence of a policy $\pi^* \in \Pi$ such that for all $x \in S$ and $\alpha \in (0, 1]$

$$J_\alpha^*(x) = \underline{b} + \inf_{s \in \mathbb{R}} \left(s + \frac{1}{\alpha} E_x^{\pi^*}(\max\{Z' - s, 0\}) \right) \tag{13}$$

$$= \underline{b} + \mathrm{CVaR}_{\alpha,x}^{\pi^*}(Z') \tag{14}$$

$$= \mathrm{CVaR}_{\alpha,x}^{\pi^*}(Z) \tag{15}$$

and thus, π^* is optimal for J_α^*. Next, we explain how to deploy $\pi^* = (\kappa_0, \kappa_1, \ldots, \kappa_{N-1})$. Let $x \in S$ and $\alpha \in (0, 1]$ be given, and let $s_{\alpha,x}^* \in [0, \bar{a}]$ satisfy (12). The realization of (X_0, S_0) is $(x_0, s_0) = (x, s_{\alpha,x}^*)$. For $t = 0, 1, \ldots, N - 1$, repeat the following steps.

1) Choose the control $u_t = \kappa_t(x_t, s_t)$.
2) A realization w_t of W_t occurs according to the distribution $p(\cdot | x_t, u_t)$.
3) Then, the realization of (X_{t+1}, S_{t+1}) is given by $(x_{t+1}, s_{t+1}) = (f(x_t, u_t, w_t), s_t - c'(x_t, u_t))$.
4) Update t by 1, and proceed to step 1 if $t < N$.

Subsequently, we describe alternative approaches for estimating J_α^*.

3) Upper-Bound Method: We derive an upper bound for J_α^* by adopting techniques from our prior work [10].[12] The upper bound can be computed by DP on S and therefore has the benefit of providing a theoretically guaranteed estimate of J_α^* without state-space augmentation.

Assume that c and c_N are l.s.c. and bounded below by $\underline{d}$. Recall that $Z' = Z - \underline{b} \geq 0$ is given by (4), and Π' is the class of deterministic Markov policies on S. Let $x \in S$, $\alpha \in (0, 1]$, and $\pi \in \Pi'$ be given. First, suppose that $E_x^\pi(Z')$ is finite. Then, $\mathrm{CVaR}_{\alpha,x}^\pi(Z') \leq (1/\alpha)E_x^\pi(Z')$ by [10, Lemma 2] and $\mathrm{CVaR}_{\alpha,x}^\pi(Z') = \mathrm{CVaR}_{\alpha,x}^\pi(Z) - \underline{b}$ by translation-equivariance (Footnote 11). It follows that $\mathrm{CVaR}_{\alpha,x}^\pi(Z) \leq \underline{b} + (1/\alpha)E_x^\pi(Z')$. Otherwise, if $E_x^\pi(Z') = +\infty$, then $E_x^\pi(|Z|) = +\infty$, and then $\mathrm{CVaR}_{\alpha,x}^\pi(Z) = +\infty$ by definition (Section II-D). Thus, regardless of whether $E_x^\pi(Z')$ is finite, it holds that $\mathrm{CVaR}_{\alpha,x}^\pi(Z) \leq \underline{b} + (1/\alpha)E_x^\pi(Z') \; \forall x \in S \; \forall \pi \in \Pi' \; \forall \alpha \in (0, 1]$, and hence,

$$J_\alpha^*(x) \leq \inf_{\pi \in \Pi'} \mathrm{CVaR}_{\alpha,x}^\pi(Z) \leq \underline{b} + \frac{1}{\alpha} \inf_{\pi \in \Pi'} E_x^\pi(Z') \tag{16}$$

for all $x \in S$ and $\alpha \in (0, 1]$. One may compute $J'(x) := \inf_{\pi \in \Pi'} E_x^\pi(Z')$ for all $x \in S$ exactly in principle under appropriate conditions, e.g., see [37], [62], and then use J' to upper-bound J_α^* for any $\alpha \in (0, 1]$ of interest (16).

4) Approximate Method: Chow et al. [41] proposed a method to approximate J_α^* using linear programming and an extra state that takes values in $(0, 1]$. The extra state may be interpreted as a time-varying risk-aversion level. The method in [41] assumes that the CVaR of a random cumulative cost initialized at time t is well-approximated by the optimal value of an LP. The LP is inspired by a history-dependent temporal decomposition for CVaR [39, Th. 6] [40, Th. 21, Lemma 22]. The decomposition is related to a representation for CVaR, which takes the form of a distributionally robust expectation [69, eqs. (3.5) and (3.13)]. The method in [41] defines the iterates of a DP recursion to resemble the decomposition but to be history-dependent only through the current augmented state. While an upper or lower bound to a CVaR-optimal control problem is not guaranteed by [41], the approach is intriguing, as the additional state has a small range. However, numerically solving the required LPs may be difficult in practice because the constraints may span several orders of magnitude. Solving many small LPs in parallel can be used to overcome the difficulty of differently scaled constraints that can arise in large LPs.

IV. System Models

We consider two control systems, a thermostatic regulator and a stormwater system. For each system, we choose a cost function $c = c_N$ that depends on the current state so that Z represents a cumulative deviation of the state trajectory relative to a particular subset K of S.[13] The desired outcome is that the state trajectory remains inside of K over time. However,

---

[12]In this prior work, we derived an upper bound for the CVaR of a maximum cost to approximate a safety analysis problem [10]. However, in the current work, we consider a cumulative cost of Z.

[13]We formalize the meaning of deviation later in this section. One may choose c to depend on the current state and control, in which case, Z would have a different interpretation.
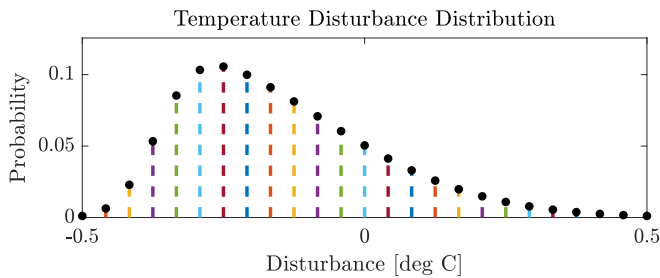
Fig. 5. Right-skewed distribution for the disturbances in the thermostatic regulator model.

TABLE II
THERMOSTATIC REGULATOR PARAMETERS

Symbol	Description	Value
a	Time delay	$e^{\frac{-\triangle\tau}{CR}}$ (no units)
b	Temperature shift	32 °C
C	Thermal capacitance	$2\ \frac{\text{kWh}}{°C}$
η	Control efficiency	0.7 (no units)
K	Constraint set	[20, 21] °C
P	Power	14 kW
R	Thermal resistance	$2\ \frac{°C}{\text{kW}}$
$\triangle\tau$	Duration of $[t, t+1]$	$\frac{5}{60}$ h
N	Number of time points	12 (= 1 h)
A_G	Grid of controls	$\{0, 0.1, \ldots, 1\}$ (no units)
S_G	Grid of states	$\{18, 18.1, \ldots, 23\}$ °C

h = hours, kW = kilowatts, °C = degrees Celsius.

this may not be possible due to disturbances that arise in the environment. Thus, we permit departures from K, and we choose c to quantify the magnitude of departure. That is, K is a soft safety constraint set that characterizes a desired operating region.

A. Thermostatic Regulator

The temperature of a room equipped with a heater that is regulated by a thermostat may be modeled as follows [61], [64]: for $t = 0, 1, \ldots, N - 1$, $x_{t+1} = ax_t + (1 - a)(b - \eta RP u_t) + w_t$. $x_t \in \mathbb{R}$ (°C) is a value of the random temperature X_t. $u_t \in [0, 1]$ is a value of U_t, representing an amount of power supplied to the system. $w_t \in \mathbb{R}$ is a value of W_t, which represents environmental uncertainties. Table II lists the model parameters. We consider a temperature disturbance distribution with positive (i.e., right) skew (Fig. 5).

We define c to quantify a deviation relative to a desired temperature range $K = [20, 21]$ °C as follows: $c(x, u) := c_N(x) := \max\{x - 21, 20 - x\}$ for any $(x, u) \in \mathbb{R} \times [0, 1]$. The random cumulative cost Z (3) is the total deviation of the random state trajectory relative to the desired temperature range.

B. Stormwater System

We consider a stormwater system with two tanks that are connected by an automated pump (Fig. 6). Water enters the system due to a stochastic process of surface runoff, and water exits the system by discharging to a storm sewer or, if the water level is too high, to a combined sewer. *Combined sewers* are present in older cities (e.g., Toronto and San Francisco) and permit untreated wastewater to discharge into natural waterways, if necessary, due to limited sewer capacity.

We aim to quantify and minimize the risk of combined sewer overflows, first in simulation and ultimately in practice. Large infrequent combined sewer overflows are believed to cause disproportionate ecological harm compared to smaller more frequent overflows [4]. Adopting a risk-averse approach in this context may allow decision-makers to calibrate their control systems to minimize ecological harm more effectively.

We model the system as $x_{t+1} = x_t + F(x_t, u_t, w_t) \cdot \triangle\tau$ for $t = 0, 1, \ldots, N - 1$ such that if $x_{t+1,i} > \bar{k}_i$, then we redefine $x_{t+1,i} := \bar{k}_i$. $x_t \in \mathbb{R}^2_+$ is a realization of X_t, and the ith entry $x_{it} \in [0, \bar{k}_i]$ (ft) is a water level of tank i. $u_t \in [-1, 1]$ is a realization of U_t, representing a pump setting. $w_t \in \mathbb{R}$ (cfs) is a realization of W_t, representing surface runoff that enters the system during a storm. $\triangle\tau$ is the duration of $[t, t+1]$, and $N \in \mathbb{N}$ represents the duration of the storm. The function F is a simplified physics-based model, $F(x, u, w) := [F_1(x, u, w), F_2(x, u, w)]^T \in \mathbb{R}^2$, where

$$F_1(x, u, w) := (w - q_{\text{cso},1}(x_1) + q_{\text{pump}}(x, u))/a_1$$
$$F_2(x, u, w) := (w - q_{\text{cso},2}(x_2) - q_{\text{pump}}(x, u) - q_{\text{storm}}(x_2))/a_2. \quad (17)$$

We define $q_{\text{cso},i}$, q_{pump}, and q_{storm} subsequently. Please refer to Table III for the model parameters.

The combined sewer outlets are equipped with flow regulators, where each regulator produces an outflow rate $q_{\text{lin},i}$. The flow rate into the combined sewer from tank i, $q_{\text{cso},i}$, is expressed in terms of $q_{\text{lin},i}$ and the number of combined sewer outlets $N_{\text{cs},i}$ as follows:

$$q_{\text{cso},i}(x_i) := q_{\text{lin},i}(x_i) \cdot N_{\text{cs},i}$$
$$q_{\text{lin},i}(x_i) := \bar{q}_{\text{cs},i} - \frac{\bar{q}_{\text{cs},i}}{\bar{k}_i - z_{\text{cs},i}} \min\{\bar{k}_i - x_i, \bar{k}_i - z_{\text{cs},i}\} \quad (18)$$

where $\bar{q}_{\text{cs},i} := c_d \tilde{\pi} r^2_{\text{cs},i} (2\tilde{g}(\bar{k}_i - z_{\text{cs},i}))^{1/2}$ is tank i's maximum outflow rate to the combined sewer from an outlet with radius $r_{\text{cs},i}$ and elevation $z_{\text{cs},i}$.

We define q_{pump} so that the pumping rate is proportional to u when possible, and we specify intermediary cases to permit continuity and a startup phase. $q_{\text{pump}}(x, u)$ is given by

$$q_{\text{pump}}(x, u) := \begin{cases} 0, & \text{if } \mathcal{I}_1(x_1, u) \text{ or } \mathcal{I}_2(x_2, u) \\ \ell(x_1, u), & \text{if } x_1 \in [z_p - \epsilon, z_p + \epsilon] \text{ and } u < 0 \\ \ell(x_2, u), & \text{if } x_2 \in [z_p - \epsilon, z_p + \epsilon] \text{ and } u \geq 0 \\ u \cdot \bar{q}_p, & \text{otherwise} \end{cases} \quad (19a)$$

where $\bar{q}_p$ (cfs) is the maximum desired pumping rate, $\ell(x_i, u)$ models a startup phase, z_p is a threshold elevation, and $\epsilon \ll z_p$ is a positive number. $\mathcal{I}_i(x_i, u)$ is true or false, depending on whether a water level is high enough for pumping

$$\mathcal{I}_1(x_1, u) := x_1 < z_p - \epsilon \text{ and } u < 0$$
$$\mathcal{I}_2(x_2, u) := x_2 < z_p - \epsilon \text{ and } u \geq 0. \quad (19b)$$

For example, $\mathcal{I}_2(x_2, u)$ is true if and only if the pump attempts to push water from tank 2 to tank 1 ($u \geq 0$), but the water level in tank 2 is not high enough for pumping to occur. We define the startup phase function by $\ell(x_i, u) := ((u \cdot \bar{q}_p)/2\epsilon)(x_i + \epsilon - z_p)$.
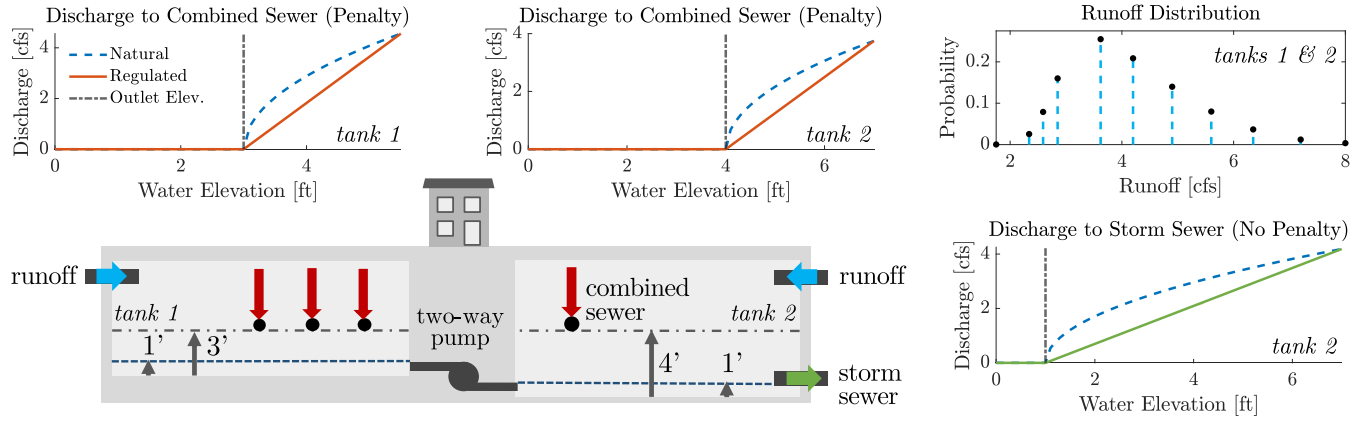
Fig. 6.   Stormwater system with two tanks connected by an automated pump. The pump can be operated at a maximum flow rate of 10 cfs in either direction, as long as the water level is above the 1-ft sump. When water levels are higher than the sump, tank 2 discharges stormwater passively into the storm sewer. At higher water levels, both tanks discharge stormwater passively into a combined sewer. Mechanical outlet regulators throttle discharge so it is linearly increasing with the water level. Combined sewers carry a mixture of stormwater and untreated wastewater that can overflow into natural waterways during rain events. We penalize discharge into combined sewers in our examples to demonstrate control strategies that aim to minimize the social and ecological impacts of combined sewer overflows. We present a discrete, positively skewed distribution for the random surface runoff W_t. The first three moments are approximately 4.0 cfs (mean), 1.2 cfs^2 (variance), and 0.72 (skewness).

TABLE III

STORMWATER SYSTEM PARAMETERS

Symbol	Description	Value
a_1	Surface area of tank 1	30000 ft^2
a_2	Surface area of tank 2	10000 ft^2
A_G	Control space grid	$\{-1, 0, 1\}$
c_d	Discharge coefficient	0.61
$\triangle\tau$	Duration of $[t, t+1)$	5 min
ϵ	Positive number much less than pumping threshold elevation	$\frac{1}{12}$ ft
$\tilde{g}$	Acceleration due to gravity	32.2 $\frac{\text{ft}}{\text{s}^2}$
k_1	Combined sewer (CS) outlet elevation, tank 1	3 ft
k_2	CS outlet elevation, tank 2	4 ft
$\overline{k}_1$	Max value of x_1 in state space grid	5.5 ft
$\overline{k}_2$	Max value of x_2 in state space grid	7 ft
N	Length of discrete time horizon	48 (= 4 h)
$N_{cs,1}$	Number of CS outlets in tank 1	3 outlets
$N_{cs,2}$	Number of CS outlets in tank 2	1 outlets
$\tilde{\pi}$	Circumference-to-diameter ratio	≈ 3.14
$\overline{q}_p$	Maximum desired pumping rate	10 cfs
$r_{cs,1}$	CS outlet radius, tank 1	$\frac{1}{4}$ ft
$r_{cs,2}$	CS outlet radius, tank 2	$\frac{3}{8}$ ft
r_s	Storm sewer outlet radius	$\frac{1}{3}$ ft
S_G	State space grid	$\{0, 0.1, \ldots, \overline{k}_1\}$ ft$\times$ $\{0, 0.1, \ldots, \overline{k}_2\}$ ft
z_p	Pumping threshold elevation	1 ft
$z_{cs,1}$	CS outlet elevation, tank 1	3 ft
$z_{cs,2}$	CS outlet elevation, tank 2	4 ft
z_s	Storm sewer outlet elevation	1 ft

CS = combined sewer, cfs = cubic feet per second, ft = feet, s = seconds, min = minutes, h = hours.

The definition of q_{storm}, which resembles $q_{\text{lin},i}$ (18), is $q_{\text{storm}}(x_2) := \overline{q}_s - (\overline{q}_s)/(\overline{k}_2 - z_s) \min\{\overline{k}_2 - x_2, \overline{k}_2 - z_s\}$, where $\overline{q}_s := c_d \tilde{\pi} r_s^2 (2\tilde{g}(\overline{k}_2 - z_s))^{1/2}$ is tank 2's max outflow rate to the storm sewer from an outlet with radius r_s and elevation z_s.

We use a discrete, positively skewed distribution for W_t (Fig. 6). In previous work, we simulated a design storm in PCSWMM (Computational Hydraulics International) [74]. PCSWMM is an extension of the U.S. Environmental

TABLE IV

COMPUTATIONAL RESOURCES

	Thermostatic Regulator		Stormwater System	
	Grid Size	Runtime	Grid Size	Runtime
Problem 1 (EU)	51	2 min	3976	10 min
Problem 2 (CVaR)	51 x 67	25 min	3976 x 491	136 h

To emphasize the additional complexity of Problem 2, in the last row, we have listed the size of the computational grid as the product of the cardinality of S_G and the cardinality of the discretization of $\mathcal{Z}$. Thermostatic regulator runs utilized 4 cores, and stormwater system runs utilized 30 cores in a multi-tenant cluster environment. Our analysis code is written in MATLAB (The Mathworks, Inc.) and is available from https://github.com/risk-sensitive-reachability/IEEE-TCST-2021.

Protection Agency's Stormwater Management Model [75], an industry-standard software package for the design of stormwater systems. We obtained samples of surface runoff from these simulations, and the empirical distribution had positive skew, which is reflected in the current distribution.

The cost function c quantifies the water volume contributed to the combined sewer in hundreds of cubic feet,

$$c(x, u) = c_N(x) = q_{\text{cso}}(x) \cdot \Delta\tau \cdot 0.01. \qquad (20)$$

The term $q_{\text{cso}}(x)$ is the total discharge rate (cfs) to the combined sewer, $q_{\text{cso}}(x) := \sum_{i=1}^2 q_{\text{cso},i}(x_i)$, where $q_{\text{cso},i}$ is given by (18). $q_{\text{cso},i}(x_i)$ is the discharge rate (cfs) to the combined sewer from tank i when the water level of tank i is x_i (ft). The definition of q_{cso} assumes a constant discharge rate on each time interval $[t - 1, t)$. The factor of 0.01 in (20) is used so that the resulting units are in hundreds of cubic feet. The cumulative random cost Z represents the total water volume (hundreds of ft^3) that is discharged to the combined sewer during a 4-h storm.

V. NUMERICAL RESULTS: OPTIMAL CONTROL

Using the models from the previous section, we have solved Problem 1 (EU) and Problem 2 (CVaR) numerically by implementing Algorithms 1 and 2, respectively. As anticipated, *significantly* reduced computational resources are required for
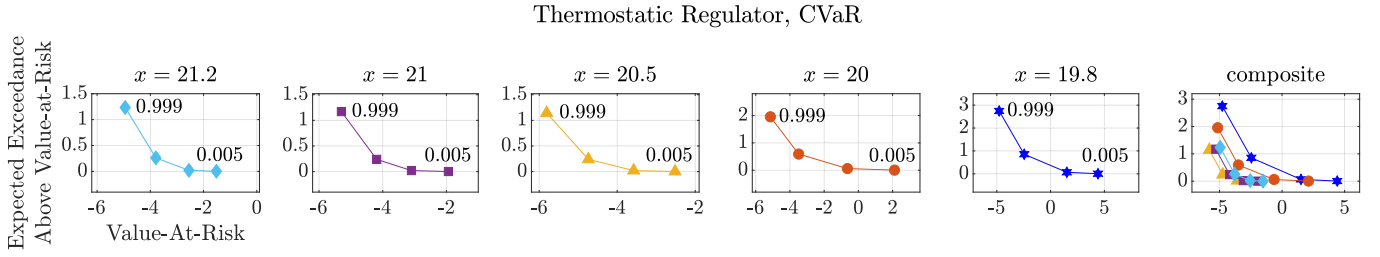
Fig. 7. Empirical estimates of the tradeoffs between the value-at-risk and the expected exceedance above the value-at-risk of a cumulative cost for the thermostatic regulator under CVaR-optimal control. Plots are shown for each initial condition $x \in \{19.8, 20, 20.5, 21, 21.2\}$ °C and $\alpha \in \{0.999, 0.5, 0.05, 0.005\}$. Each point represents the result of sampling 10^7 trajectories for a given x and α. According to (10), nearly risk-neutral preferences (e.g., $\alpha = 0.999$) lead to a nearly equal weighting between the value-at-risk and the expected exceedance beyond the value-at-risk. As risk aversion increases (α decreases), greater emphasis is placed on minimizing the expectation of the outcomes worse than the value-at-risk.
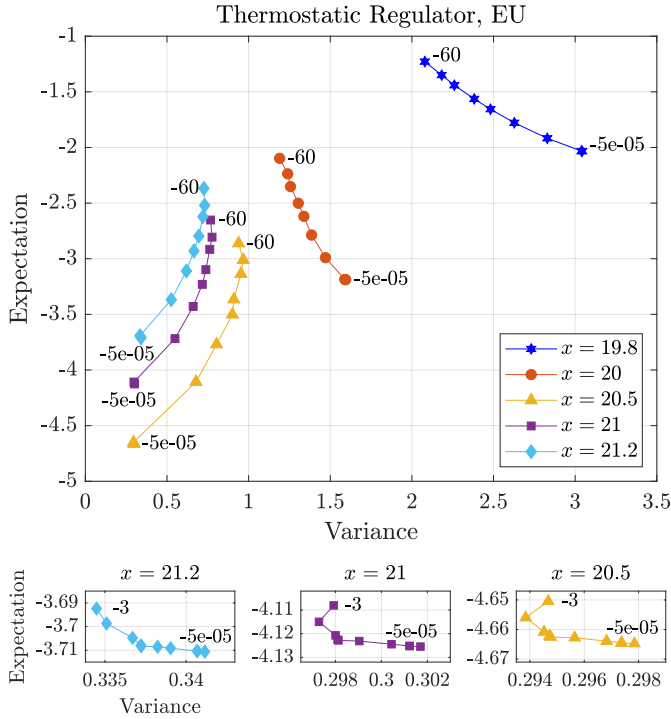

Fig. 8. Estimates of $E_x^{\hat{\pi}_\theta^*}(Z)$ versus $\mathrm{var}_x^{\hat{\pi}_\theta^*}(Z)$ for the thermostatic regulator under EU-optimal control after 10^7 samples for each x and θ. Plots are shown for each initial condition $x \in \{19.8, 20, 20.5, 21, 21.2\}$ °C, and $\theta \in \{-5e-5, -3, -9, -12, -15, -18, -24, -30, -60\}$ (**top**). In the top plot, the points that correspond to $\theta \in \{-5e-5, -3\}$ overlap for each x. There is a neighborhood $\mathcal{N}_x$ of zero in which making θ more negative leads to a reduction in the variance at the expense of an increase in the mean. For $x = 19.8$ and $x = 20$, the policy $\hat{\pi}_\theta^*$ with $\theta = -60$ leads to a distribution with a small variance at the expense of having a large expectation. For $x = 20.5$, $x = 21$, and $x = 21.2$, the policy $\hat{\pi}_\theta^*$ with $\theta = -60$ yields a distribution with a large expectation and a large variance. However, these latter initial conditions achieve a trivial mean-variance tradeoff if $|\theta|$ is sufficiently small (**bottom**). The bottom plots show $\theta \in \{-5e-5, -0.5, -0.75, -1, -1.25, -1.5, -1.75, -2, -2.5, -3\}$. These plots indicate that if $x = 20.5$ or $x = 21$, $\mathcal{N}_x$ is approximately $(-2.5, 0)$; if $x = 21.2$, $\mathcal{N}_x$ is roughly $(-3, 0)$.

Fig. 9. Estimates of $E_x^{\hat{\pi}_\theta^*}(Z')$ versus $\mathrm{var}_x^{\hat{\pi}_\theta^*}(Z')$ for the thermostatic regulator, where we have implemented Algorithm 1 with the restriction to non-negative costs (Remark 3). These empirical moments were estimated from 10^7 samples for each $x \in \{19.8, 20, 20.5, 21, 21.2\}$ °C and $\theta \in \{-5e-5, -6, -7, -8\}$.

Problem 1 versus Problem 2. Recall that solving Problem 2 requires an augmented state space $S \times \mathcal{Z}$, where $\mathcal{Z} := [-\bar{a}, \bar{a}] \subseteq \mathbb{R}$ and $0 \leq Z' \leq \bar{a} := (\bar{d} - \underline{d})(N + 1)$. Hence, the range of $\mathcal{Z}$ depends on the length of the interval $[\underline{d}, \bar{d}]$ formed by the lower and upper bounds of the stage cost and the length of the time horizon N. The discretization of $\mathcal{Z}$ also depends on the desired precision of the cumulative costs. For
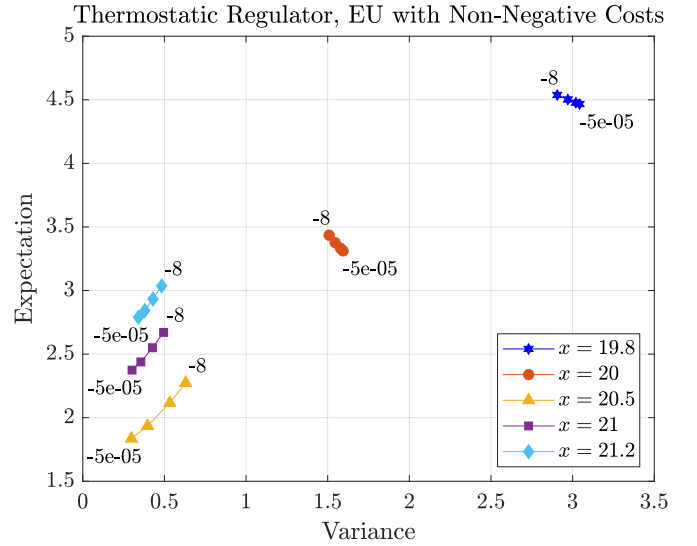
example, we have chosen a precision of approximately 1 °C for the thermostatic regulator and 500 ft^3 for the stormwater system. Covering the range of $\mathcal{Z}$ at this precision increases the cardinality of the overall computational grid by 67 times and 491 times. Table IV outlines the resources utilized in our unoptimized implementation. We have made no attempt to improve efficiency, except for parallelizing operations in a given DP recursion.

To distinguish between an exact solution and a numerical solution returned by a computer, we introduce some notation. $\hat{V}_\theta^*$ ($\hat{J}_\alpha^*$) indicates a computation of V_θ^* (J_α^*), and $\hat{\pi}_\theta^*$ ($\hat{\pi}_\alpha^*$) indicates a computation of a policy that is optimal for V_θ^* (J_α^*). Using $\hat{\pi}_\theta^*$ or $\hat{\pi}_\alpha^*$ and an initial condition x, we have simulated 10 million trajectories to estimate a distribution of Z that is optimal w.r.t. EU or CVaR, respectively. We use these simulations to study the trends between particular empirical statistics of Z, which we have selected based on how EU or CVaR encode risk aversion. In the EU case, we study how the expectation $E_x^{\hat{\pi}_\theta^*}(Z)$ and the variance $\mathrm{var}_x^{\hat{\pi}_\theta^*}(Z)$ vary with θ (6). In the CVaR case, we study how the quantile $\mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z)$
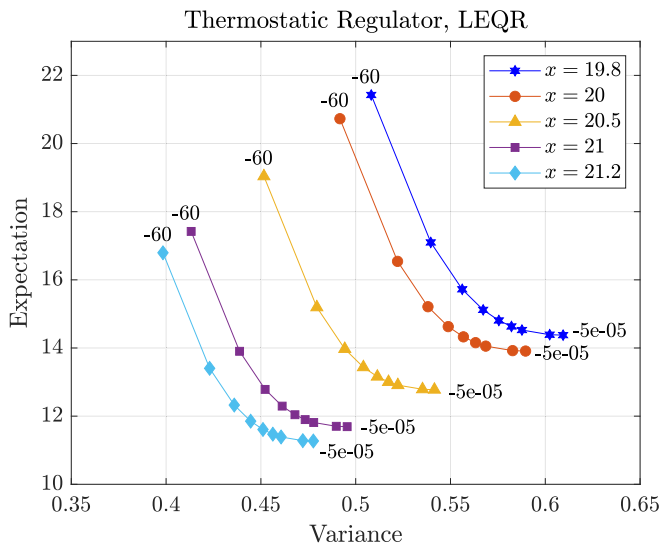
Fig. 10. Empirical mean-variance curves for the thermostatic regulator under a LEQG controller [18, Th. 3], $x \in \{19.8, 20, 20.5, 21, 21.2\}$ °C, and $\theta \in \{-5e-5, -12, -30, -35, -40, -45, -50, -55, -60\}$. Each point represents the outcome of 10^7 sampled trajectories for a given x and θ. The state vector is $\tilde{x}_t := x_t - b$, and the dynamics equation is $\tilde{x}_{t+1} = A\tilde{x}_t + Bu_t + w_t$, where $A = a$ and $B = (a-1)\eta RP$ (Table II). The disturbance process is i.i.d. zero-mean Gaussian noise with variance $\sigma^2 = 0.03$. The initial condition is a Gaussian random variable with mean $x - b$ and variance $\sigma^2/100$. The random cost Z is quadratic, and a realization of Z takes the form, $z = \tilde{q}\tilde{x}_N^2 + \sum_{t=0}^{N-1} \tilde{q}\tilde{x}_t^2 + ru_t^2$, where $\tilde{q} = 0.01$ and $r = 1$.

and the expected exceedance above the quantile $E_x^{\hat{\pi}_\alpha^*}(\max\{Z - \mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z), 0\})$ vary with α (10).

1) Thermostatic Regulator: We present results for the thermostatic regulator in which the disturbance has a right-skewed distribution (Fig. 5). First, we consider the CVaR setting. By (10), minimizing the CVaR of Z at level $\alpha \in (0, 1)$ is equivalent to minimizing a linear combination of the value-at-risk at level α, $\mathrm{VaR}_{\alpha,x}^\pi(Z)$, and the expected exceedance above the $\mathrm{VaR}_{\alpha,x}^\pi(Z)$, which is $E_x^\pi(\max\{Z - \mathrm{VaR}_{\alpha,x}^\pi(Z), 0\})$. The value-at-risk at level α is a quantile that represents the best outcome among the $\alpha \cdot 100\%$ of the worst outcomes. All else being equal, it is desirable to keep the value-at-risk for a given α as small as possible. However, the value-at-risk is not sensitive to the distribution of the values that exceed it. CVaR compensates for this limitation by incorporating a measure of the expected exceedance beyond the value-at-risk. We provide empirical estimates of the Pareto-efficient tradeoffs between these two quantities with $\pi = \hat{\pi}_\alpha^*$, that is, $E_x^{\hat{\pi}_\alpha^*}(\max\{Z - \mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z), 0\})$ versus $\mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z)$, for the thermostatic regulator in Fig. 7.

Now, we consider the EU setting, in which the mean-variance trend is of interest. When the mean-variance approximation (6) is valid, the magnitude of θ represents the amount a decision-maker is willing to increase the expectation for a unit reduction in the variance. Therefore, making θ more negative should prioritize a reduction in the variance $\mathrm{var}_x^\pi(Z)$ rather than the expectation $E_x^\pi(Z)$. We show plots of the empirical estimates of the mean and variance for the thermostatic regulator under EU-optimal control in Fig. 8.

For each initial condition x, there is a small neighborhood $\mathcal{N}_x$ of zero such that varying θ in $\mathcal{N}_x$ leads to a mean-variance

tradeoff. The size of this neighborhood depends on the initial condition. The top portion of Fig. 8 shows that if $x = 19.8$ or $x = 20$ (°C), the variance is reduced at the expense of the mean, as θ decreases from $-5e-5$ to -60. However, the bottom portion of Fig. 8 shows that this trend only exists for a narrow range of θ if $x = 20.5$, $x = 21$, or $x = 21.2$. For these latter initial conditions, varying θ in $\mathcal{N}_x$ leads to a tradeoff that is practically trivial, and increasing the magnitude of θ outside of $\mathcal{N}_x$ leads to increases in both the mean and variance (see Fig. 8, top).

We find that similar trends occur, see Fig. 9, when we implement Algorithm 1 with non-negative costs (Remark 3). Here, the most negative value of θ is -8 because the algorithm suffers from numerical instabilities for more negative values of θ. Notably, in the classical LEQG setting, we find consistent mean-variance tradeoffs for a wide range of θ (Fig. 10, see caption for details regarding the simulation setting). The classical LEQG controller is linear state feedback, and a Riccati recursion provides the optimal control gains [18, Th. 3]. The recursion is well-defined for negative values of θ that satisfy a condition that depends on the Riccati matrices and the noise covariance [18, Th. 3].

2) Stormwater System: When EU-optimal control is applied to the stormwater system, we find that the mean-variance interpretation is again dependent on the initial conditions. In Fig. 11, we show the empirical mean-variance curves from select initial water levels $x \in \mathbb{R}_+^2$ and values of θ. The abovementioned examples demonstrate that special care must be taken when using EU-optimal control. It is known that the mean-variance approximation for EU, $\rho_{\theta,x}^\pi(Z) \approx E_x^\pi(Z) - (\theta/4)\mathrm{var}_x^\pi(Z)$ (6), is valid under a restricted set of conditions. One of our contributions is to demonstrate that such theoretical requirements cannot be ignored in practical applications of EU-optimal control. Indeed, for a given system model and initial condition, the size of $\mathcal{N}_x$ is not known a priori and may be trivial in practice. More importantly, we have shown that inadvertently making θ too negative can lead to a distribution with a higher mean and a higher variance.

On the other hand, CVaR-optimal control provides a consistent tradeoff between the value-at-risk, $\mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z)$, and the expected exceedance above the value-at-risk, $E_x^{\hat{\pi}_\alpha^*}(\max\{Z - \mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z), 0\})$, as α becomes more risk averse (closer to zero). Intuitively, this tradeoff is not surprising because for any random variable Y with a fixed distribution, $E(\max\{Y - \mathrm{VaR}_\alpha(Y), 0\})$ versus $\mathrm{VaR}_\alpha(Y)$ forms a nonincreasing trend, as α becomes more risk averse. While related, the setting of CVaR-optimal control is distinct because the distribution of Z, $P_x^{\hat{\pi}_\alpha^*}$, varies with α. The consistent tradeoff that arises from CVaR-optimal control is useful for analyzing the performance of a control system with respect to competing objectives and varying degrees of pessimism. Fig. 12 shows a summary of performance for one particular design. However, such curves could be estimated for multiple candidate designs, and overlaying these curves could provide a concise visual comparison of performance. The current computational requirements of CVaR-optimal control restrict such comparisons to systems with low-dimensional models.
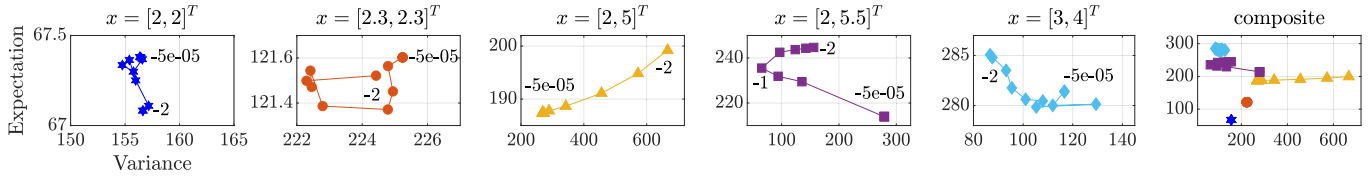
Fig. 11. These plots show estimates of $E_x^{\hat{\pi}_\theta^*}(Z)$ versus $\mathrm{var}_x^{\hat{\pi}_\theta^*}(Z)$ for the stormwater system under EU-optimal control. Z is a non-negative cumulative cost, expressed in hundreds of ft^3 of water. Each point represents the result of sampling 10^7 trajectories for a given x and θ. For this system, we have found the numerically stable range for θ to be approximately $-5e-5$ to -2, and the plots show $\theta \in \{-5e-5, -5e-4, -0.005, -0.05, -0.5, -1, -1.25, -1.5, -1.75, -2\}$ for select initial water levels. For $x = [2, 2]^T$ and $x = [2.3, 2.3]^T$, no consistent mean-variance trends are present, and from the perspective of the composite plot (far right), one may consider the changes in the mean and variance to be negligible. If the system starts from $x = [2, 5]^T$, the mean and variance increase as θ becomes more negative. If $x = [2, 5.5]^T$, there is a mean-variance tradeoff if $|\theta|$ is sufficiently small, but the mean and variance increase as θ decreases from -1 to -2. Trajectories starting from $x = [3, 4]^T$ exhibit a mean-variance tradeoff as θ varies across some, but not all, subintervals of $[-2, 0]$.
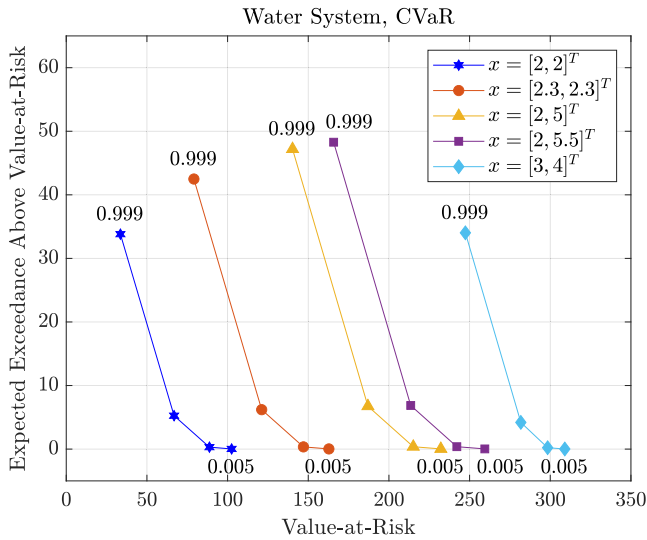

Fig. 12. These plots provide estimates of the tradeoffs between the value-at-risk $\mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z)$ and the expected exceedance above the value-at-risk $E_x^{\hat{\pi}_\alpha^*}(\max\{Z - \mathrm{VaR}_{\alpha,x}^{\hat{\pi}_\alpha^*}(Z), 0\})$ of a cumulative cost Z for the stormwater system under CVaR-optimal control. Plots are shown for different initial conditions x and risk-aversion levels $\alpha \in \{0.999, 0.5, 0.05, 0.005\}$. Each point represents the result of sampling 10^7 trajectories for a given x and α.

Fig. 13. Estimates of EU-safe sets $\mathcal{E}_\theta^r := \{x \in S : V_\theta^*(x) \le r\}$ (**top**) and CVaR-safe sets $\mathcal{C}_\alpha^r := \{x \in S : J_\alpha^*(x) \le r\}$ (**bottom**) for the stormwater system, where $x = [x_1, x_2]^T \in S$. Water levels outside of K (shown in black) cause discharges to a combined sewer. We show $r \in \{100, 200, \ldots, 500\}$ (hundreds of ft^3 of water discharged to a combined sewer) for a nearly risk-neutral setting (θ near zero, α near 1) and a risk-averse setting ($\theta = -2$, $\alpha = 0.005$).

Thus far, we have analyzed system behavior and the distribution of outcomes under such behavior from select initial conditions. In addition, it may be useful to assess the performance of a system on its state space as a whole. An approach for this task is to use the notion of a safe set, which is the topic of the following section.

VI. SAFETY ANALYSIS VIA EU AND CVaR: INTERPRETATIONS AND EXAMPLES

One can assess the performance of a control system by studying the level sets of an optimal value function. The level sets may be called safe sets, reachable sets, or invariant sets, for example, where the precise name is chosen according to the value function's interpretation. Developing algorithms to compute these sets and the associated optimal policies is the core aim of Hamilton–Jacobi reachability analysis [54], [57], minimax safety analysis [53], and stochastic safety analysis [58], [59], [60], [61] (recall our introduction). By formulating and solving robust or stochastic optimal control problems,
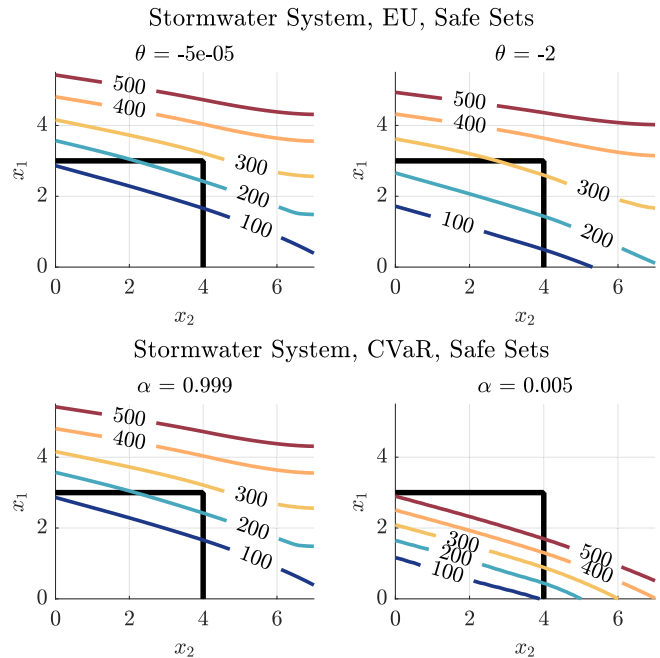
these methods provide principled alternatives to Monte Carlo simulation to synthesize policies and assess whether a system can operate well in an uncertain environment.

Here, we examine the use of risk-averse optimal control for this assessment. We define *exponential-utility-safe sets* in terms of V_θ^*, $\mathcal{E}_\theta^r := \{x \in S : V_\theta^*(x) \le r\}$, and *CVaR-safe sets* in terms of J_α^*, $\mathcal{C}_\alpha^r := \{x \in S : J_\alpha^*(x) \le r\}$, where $r \in \mathbb{R}$. We use the term safe sets because V_θ^* and J_α^* represent an optimal deviation between the state trajectory and a desired operating region. An EU-safe set $\mathcal{E}_\theta^r$ is the set of initial conditions $x \in S$ from which $V_\theta^*(x)$ is no more than r. If a decision-maker has a fixed price θ they are willing to pay to reduce variance and if (6) is valid, then $V_\theta^*(x)$ is the optimal certainty equivalent. The CVaR-safe set $\mathcal{C}_\alpha^r$ represents the set of initial conditions from which the optimal expected value of Z in the $\alpha \cdot 100\%$ worst cases is no more than r. This interpretation is exact if

Z is a continuous random variable for all $x \in S$ and $\pi \in \Pi$, for example.

EU-safe sets and CVaR-safe sets for the stormwater system are shown in Fig. 13. As anticipated, the safe sets in the nearly risk-neutral setting (θ near zero, α near one) are indistinguishable (Fig. 13, left). The EU-safe sets with $\theta = -2$, the most negative value that is numerically stable, are noticeably larger than the CVaR-safe sets with $\alpha = 0.005$ (Fig. 13, right). It is easier to observe the contraction in the contours of the CVaR-safe sets as α becomes more risk averse, suggesting that CVaR may be more suitable for visually conveying the effects of varying degrees of risk aversion.

VII. CONCLUSION

In this article, we have studied the use of EU and CVaR as risk-averse performance criteria for control systems. The optimization of EU is considerably simpler in theory and in practice. However, we have demonstrated that a nontrivial mean-variance tradeoff need not occur, and making θ more negative can yield a higher variance and a higher mean. Therefore, EU-optimal control must be used cautiously. In addition to the abovementioned concerns, the choice of a more appropriate risk-averse functional may depend on several factors, which we highlight below.

1) *Utility Functions:* EU-optimal control assumes a utility function of the form $v_\theta(y) := e^{(-\theta/2)y}$. In contrast, CVaR-optimal control does not require a utility function. There are many utility functions available, and it may be difficult to choose one that describes the desired preferences for all possible outcomes for a particular application. Some utility functions transform the costs in ways that are inappropriate for the application.

2) *Parameter Interpretations:* The parameter θ of EU represents an exponential aversion to larger values of Z' in general (5) and a linear aversion to the variance when (6) is valid. In contrast, the risk-sensitivity level α that parametrizes CVaR corresponds to a fraction of the largest values of Z (9). Parameters with intuitive and precise interpretations, such as α, may be particularly useful for applications that require the development and satisfaction of safety or performance specifications.

3) *Risk Interpretations:* Recall that the EU of Z encodes risk using a subjective utility function v_θ, whereas the CVaR of Z encodes risk in terms of the expected amount that Z exceeds a quantile (10). In EU, the utility function is applied to all possible outcomes, while CVaR focuses exclusively on outcomes in the upper tail. We have provided numerical examples of a stormwater system and a thermostatic regulator to highlight how different ways to quantify risk may be more suitable for different applications.

Regulations for stormwater systems often specify design criteria in terms of quantiles, such as adequate performance under the 10-percentile (10-year) and 1-percentile (100-year) storm events. In these assessments, standard practices measure performance in terms of a random volume of overflow (Z) [5]. CVaR may be more appropriate than EU in these

circumstances because it encodes risk in terms of quantiles and does not distort the units of Z (a physical quantity) through a utility function. However, the computational resources that are required for CVaR-optimal control are not economically feasible for higher dimensional stormwater systems in practice. While EU-optimal control has the benefit of significantly reduced computational requirements, for the previously mentioned reasons, it is unlikely to be useful for stormwater management, or more broadly, for applications with objective performance criteria.

EU-optimal control is better suited for applications with subjective performance criteria. For example, it is natural to measure the performance of indoor heating and cooling systems in terms of perceived comfort. Z may reflect a random deviation from the desired air temperature, and extreme values of Z may be inconvenient but not safety-critical. Office buildings can simply close during rare catastrophic cooling failures. To maximize comfort in normal circumstances, Z should be small on average and have low variance [6]. Moreover, occupant satisfaction with heating and cooling systems is related to the level of perceived control over the indoor environment [7]. In addition to its computational advantages, EU may be more appropriate than CVaR in this setting due to its connection to mean-variance control for values of θ in a neighborhood of zero. Unfortunately, our simple thermostatic regulator example has cast doubt on whether θ provides satisfactory mean-variance control in practice.

While risk-averse optimal control is theoretically attractive in principle, the limitations of current approaches inhibit their adoption in practice. The gap between theory and practice motivates several exciting avenues for future investigations. In particular, we see value in the following developments.

1) New methods to identify the classes of control systems in which EU provides nontrivial mean-variance trade-offs.

2) Scalable approximations for CVaR-optimal control by leveraging stochastic rollout or other grid-free policy improvement techniques.

3) Efficient strategies for assessing the degree to which a given system could benefit from a risk-averse approach relative to a conventional one.

Advances in these areas are needed to fully develop the potential advantages of risk-averse control as a decision-making framework that can accommodate competing objectives and varying degrees of pessimism about an uncertain future.

REFERENCES

[1] L. K. Read and R. M. Vogel, "Reliability, return periods, and risk under nonstationarity," *Water Resour. Res.*, vol. 51, no. 8, pp. 6381–6398, Aug. 2015.

[2] (2018). *Risk Management—Guidelines*. International Organization for Standardization (ISO). [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en

[3] (2007). *Space Systems—Oxygen Safety—Part 4: Hazards Analyses for Oxygen Systems and Components*. International Organization for Standardization (ISO). [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:22538:-4:ed-1:v1:en

[4] W. C. Lucas and D. J. Sample, "Reducing combined sewer overflows by using outlet controls for green stormwater infrastructure: Case study in Richmond, Virginia," *J. Hydrol.*, vol. 520, pp. 473–488, Jan. 2015.

[5] United States Environmental Protection Agency. (2016). *State Post Construction Stormwater Standards*. [Online]. Available: https://www.epa.gov/sites/default/files/2016-08/documents/swstdsummary_7-13-16_508.pdf

[6] M. A. Humphreys, "The variation of comfortable temperatures," *Int. J. Energy Res.*, vol. 3, no. 1, pp. 13–18, 1979.

[7] R. T. Hellwig, "Perceived control in indoor environments: A conceptual approach," *Building Res. Inf.*, vol. 43, no. 3, pp. 302–315, May 2015.

[8] L. Eeckhoudt, C. Gollier, and H. Schlesinger, *Economic and Financial Decisions under Risk*. Princeton, NJ, USA: Princeton Univ. Press, 2005.

[9] M. P. Chapman et al., "A risk-sensitive finite-time reachability approach for safety of stochastic dynamic systems," in *Proc. American Control Conf.*, 2019, pp. 2958–2963.

[10] M. P. Chapman, R. Bonalli, K. M. Smith, I. Yang, M. Pavone, and C. J. Tomlin, "Risk-sensitive safety analysis using conditional value-at-risk," *IEEE Trans. Autom. Control*, vol. 67, no. 12, pp. 6521–6536, Dec. 2022.

[11] M. P. Chapman, M. Faus, and K. M. Smith, "On optimizing the conditional value-at-risk of a maximum cost for risk-averse safety analysis," *IEEE Trans. Autom. Control*, early access, Aug. 1, 2022, doi: 10.1109/TAC.2022.3195381.

[12] L. J. Ratliff and E. Mazumdar, "Inverse risk-sensitive reinforcement learning," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 1256–1263, Mar. 2020.

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

[14] S. Singh, Y. Chow, A. Majumdar, and 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, Jul. 2019.

[15] Y. Wang and M. P. Chapman, "Risk-averse autonomous systems: A brief history and recent developments from the perspective of optimal control," *Artif. Intell.*, vol. 311, Oct. 2022, Art. no. 103743.

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

[17] P. Whittle and J. Kuhn, "A Hamiltonian formulation of risk-sensitive linear/quadratic/Gaussian control," *Int. J. Control*, vol. 43, no. 1, pp. 1–12, Jan. 1986.

[18] P. Whittle, *Risk-Sensitive Optimal Control*. Chichester, U.K.: Wiley, 1990.

[19] P. Whittle, "A risk-sensitive maximum principle: The case of imperfect state observation," *IEEE Trans. Autom. Control*, vol. 36, no. 7, pp. 793–801, Jul. 1991.

[20] D. P. Bertsekas, *Dynamic Programming and Stochastic Control*. New York, NY, USA: : Academic Press, 1976.

[21] H. Markowitz, "Mean-variance approximations to expected utility," *Eur. J. Oper. Res.*, vol. 234, no. 2, pp. 346–355, 2014.

[22] H. Markowitz, "Portfolio selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[23] C. W. Miller and I. Yang, "Optimal control of conditional value-at-risk in continuous time," *SIAM J. Control Optim.*, vol. 55, no. 2, pp. 856–884, Jan. 2017.

[24] C.-H. Won, "Cost-cumulants and risk-sensitive control," in *The Electrical Engineering Handbook*, W. K. Chen, Ed. New York, NY, USA: Academic, 2005, pp. 1061–1068.

[25] C.-H. Won, M. K. Sain, and S. R. Liberty, "Full-state-feedback minimal cost variance control on an infinite time horizon: The risk-sensitive approach," in *Proc. 40th IEEE Conf. Decis. Control*, 2001, pp. 819–824.

[26] K. Glover and J. C. Doyle, "State-space formulae for all stabilizing controllers that satisfy an H_∞-norm bound and relations to relations to risk sensitivity," *Syst. Control Lett.*, vol. 11, no. 3, pp. 167–172, Sep. 1988.

[27] D. Mustafa and K. Glover, *Minimum Entropy H_∞ Control* in (Lecture Notes in Control and Information Sciences). Springer, 1990, ch. 6.

[28] K. Zhang, B. Hu, and T. Basar, "Policy optimization for H_2 linear control with H_∞ robustness guarantee: Implicit regularization and global convergence," in *Proc. 2nd Conf. Learn. Dyn. Control*, 2020, pp. 179–190.

[29] J. Speyer, "An adaptive terminal guidance scheme based on an exponential cost criterion with application to homing missile guidance," *IEEE Trans. Autom. Control*, vol. AC-21, no. 3, pp. 371–375, Jun. 1976.

[30] G. Avila-Godoy and E. Fernández-Gaucherand, "Controlled Markov chains with exponential risk-sensitive criteria: Modularity, structured policies and applications," in *Proc. IEEE Conf. Decis. Control*, Dec. 1998, pp. 778–783.

[31] L. Cosenza, M. K. Sain, R. W. Diersing, and C.-H. Won, "Cumulant control systems: The cost-variance, discrete-time case," in *Advances in Statistical Control, Algebraic Systems Theory, and Dynamic Systems Characteristics*. Basel, Switzerland: Birkhauser, 2008, pp. 29–41.

[32] K. D. Pham, M. K. Sain, and S. R. Liberty, "Cost cumulant control: State-feedback, finite-horizon paradigm with application to seismic protection," *J. Optim. Theory Appl.*, vol. 115, no. 3, pp. 685–710, Dec. 2002.

[33] P. Brezas and M. C. Smith, "Linear quadratic optimal and risk-sensitive control for vehicle active suspensions," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 543–556, Mar. 2014.

[34] K. Boda and J. A. Filar, "Time consistent dynamic risk measures," *Math. Methods Oper. Res.*, vol. 63, no. 1, pp. 169–186, Feb. 2006.

[35] A. Shapiro, "On a time consistency concept in risk averse multistage stochastic programming," *Oper. Res. Lett.*, vol. 37, no. 3, pp. 143–147, May 2009.

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

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

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

[39] G. C. Pflug and A. Pichler, "Time-inconsistent multistage stochastic programs: Martingale bounds," *Eur. J. Oper. Res.*, vol. 249, no. 1, pp. 155–163, Feb. 2016.

[40] G. C. Pflug and A. Pichler, "Time-consistent decisions and temporal decomposition of coherent risk functionals," *Math. Oper. Res.*, vol. 41, no. 2, pp. 682–699, May 2016.

[41] Y. Chow, A. Tamar, S. Mannor, and M. Pavone, "Risk-sensitive and robust decision-making: A CVaR optimization approach," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1522–1530.

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

[43] 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, Jan. 2015.

[44] S. Samuelson and I. Yang, "Safety-aware optimal control of stochastic systems using conditional value-at-risk," in *Proc. Annu. Amer. Control Conf. (ACC)*, Jun. 2018, pp. 6285–6290.

[45] 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, Dec. 2011.

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

[47] 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, Aug. 2021.

[48] G. B. Di Masi and L. Stettner, "Risk-sensitive control of discrete-time Markov processes with infinite horizon," *SIAM J. Control Optim.*, vol. 38, no. 1, pp. 61–78, Jan. 1999.

[49] B. P. G. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, "Distributionally robust control of constrained stochastic systems," *IEEE Trans. Autom. Control*, vol. 61, no. 2, pp. 430–442, Feb. 2016.

[50] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, pp. 21–42, Feb. 2000.

[51] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *J. Banking Finance*, vol. 26, no. 7, pp. 1443–1471, Jul. 2002.

[52] C. Acerbi and D. Tasche, "On the coherence of expected shortfall," *J. Banking Finance*, vol. 26, no. 7, pp. 1487–1503, Jul. 2002.

[53] 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, Mar. 1971.

[54] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin, "A time-dependent Hamilton–Jacobi formulation of reachable sets for continuous dynamic games," *IEEE Trans. Autom. Control*, vol. 50, no. 7, pp. 947–957, Jul. 2005.

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

[56] M. Chen, S. L. Herbert, M. S. Vashishtha, S. Bansal, and C. J. Tomlin, "Decomposition of reachable sets and tubes for a class of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 63, no. 11, pp. 3675–3688, Nov. 2018.

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

[58] 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, Nov. 2008.

[59] 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, Dec. 2010.

[60] 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, no. 9, pp. 2665–2674, Sep. 2013.

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

[62] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. New York, NY, USA: Springer, 1996.

[63] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. New York, NY, USA: John, 1999.

[64] R. E. Mortensen and K. P. Haggerty, "A stochastic computer model for heating and cooling loads," *IEEE Trans. Power Syst.*, vol. PS-3, no. 3, pp. 1213–1219, Aug. 1988.

[65] A. Castelletti, F. Pianosi, and R. Soncini-Sessa, "Water reservoir control under economic, social and environmental constraints," *Automatica*, vol. 44, no. 6, pp. 1595–1607, Jun. 2008.

[66] A. Lotov, V. Bushenkov, G. Kamenev, A. Camara, and D. Loucks, "Water resource conflict resolution based on interactive tradeoffs display," in *Rehabilitation of Degraded Rivers: Challenges, Issues and Experiences*, D. P. Loucks, Ed. Boston, MA, USA: Kluwer Academic Publishers, 1998, pp. 447–470.

[67] A. V. Lotov, L. V. Bourmistrova, R. V. Efremov, V. A. Bushenkov, A. L. Buber, and N. A. Brainin, "Experience of model integration and Pareto frontier visualization in the search for preferable water quality strategies," *Environ. Model. Softw.*, vol. 20, no. 2, pp. 243–260, Feb. 2005.

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

[69] A. Shapiro, "Minimax and risk averse multistage stochastic programming," *Eur. J. Oper. Res.*, vol. 219, no. 3, pp. 719–726, Jun. 2012.

[70] J. Löfberg, "Minimax approaches to robust model predictive control," Ph.D. thesis, Dept. Elect. Eng., Linköping Univ., Linköping, Sweden, 2003.

[71] H. Nishimura, N. Mehr, A. Gaidon, and M. Schwager, "RAT iLQR: A risk auto-tuning controller to optimally account for stochastic model mismatch," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 763–770, Apr. 2021.

[72] G. Yang, H. Wang, and J. Chen, "Disturbance compensation based asymptotic tracking control for nonlinear systems with mismatched modeling uncertainties," *Int. J. Robust Nonlinear Control*, vol. 31, no. 8, pp. 2993–3010, May 2021.

[73] J. Harrison et al., "AdaPT: Zero-shot adaptive policy transfer for stochastic dynamical systems," in *Robotics Research*. Cham, Switzerland: Springer, 2020, pp. 437–453.

[74] 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. Sustainability (SusTech)*, Nov. 2018, pp. 1–8.

[75] L. A. Rossman, *Storm Water Management Model User's Manual, Version 5.0*. Cincinnati, OH, USA: National Risk Management Research Laboratory, Office of Research and Development, 2010.

[76] P. M. Esfahani, T. Sutter, D. Kuhn, and J. Lygeros, "From infinite to finite programs: Explicit error bounds with applications to approximate dynamic programming," *SIAM J. Optim.*, vol. 28, no. 3, pp. 1968–1998, Jan. 2018.

[77] D. P. Bertsekas, *Reinforcement Learning and Optimal Control*. Belmont, CA, USA: Athena Scientific, 2019.

[78] 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, 2022.

[79] A. Akametalu, "A learning-based approach to safety for uncertain robotic systems," Ph.D. thesis, Elect. Eng. Comput. Sci., Univ. California Berkeley, Berkeley, CA, USA, 2018.

Kevin M. Smith received the B.A. degree in environmental studies from Oberlin College, Oberlin, OH, USA, in 2013, and the B.S. degree in earth and environmental engineering from Columbia University, NY, USA, in 2013. He is currently pursuing the Ph.D. degree in environmental and water resources engineering at Tufts University, Medford, MA, USA.

He is the Director of Products at OptiRTC Inc., Boston, MA, USA, where he is responsible for developing flexible real-time systems for the continuous monitoring and adaptive control of stormwater infrastructure. His research seeks to understand the opportunities and risks associated with semiautonomous civil infrastructure when considered as a technology for mediating environmental conflicts.

Mr. Smith was a recipient of the NSF Integrative Graduate Education and Research Traineeship (IGERT) on Water and Diplomacy and the NSF Research Traineeship (NRT) on Data-Driven Decision Making to Address Complex Resource Problems.

Margaret P. Chapman (Member, IEEE) is an Assistant Professor with the Edward S. Rogers Sr. Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada. Her research focuses on risk-aware and stochastic control, with emphasis on safety analysis and applications in healthcare and sustainable cities.

Dr. Chapman was a recipient of the 2021 Leon O. Chua Award for achievement in nonlinear science (Electrical Engineering and Computer Sciences, University of California at Berkeley). She has also received the U.S. National Science Foundation Graduate Research Fellowship, the Berkeley Fellowship for Graduate Study, and the Stanford University Terman Engineering Scholastic Award.