Risk-Averse Semi-Autonomous Systems: A Brief History and Recent Developments from the Perspective of Optimal Control

Margaret P. Chapman

Robotics Science and Systems
Risk Aware Decision Making: From Optimal Control to RL
June 27, 2022

This talk is based on Y. Wang & M. P. Chapman, *J. Artif. Intell.*, in press, 2022.

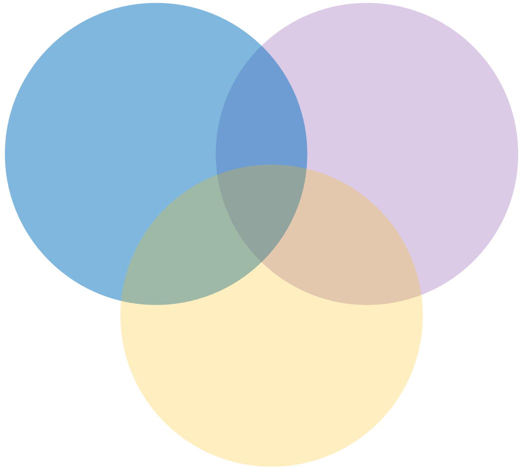
The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

https://www.control.utoronto.ca/~mchapman/


Yuheng Wang


Toronto waterfront, summer 2021

Research Interests



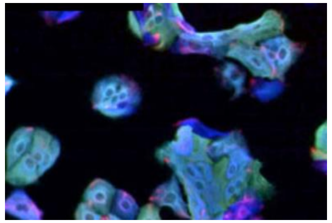
Applications
Urban water infrastructure
Cancer biology

Mathematical Foundations
Measure theory
Systems & control theory
Stochastic processes

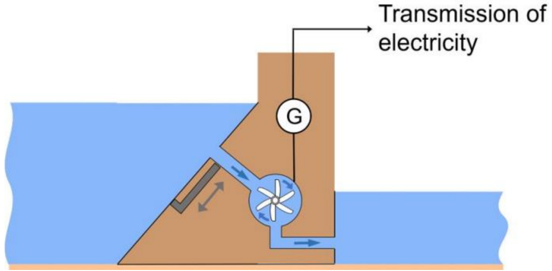
Concepts of
Risk, Safety, and Uncertainty

Safety-Relevant Stochastic Systems

Urban water



Cancer



Hydroelectric dam

Learning Dynamical Models

Risom+, Nature Communications, 2018
Chapman+, PLoS Comp. Biol., 2019

Chapman+, ACC, 2019
Chapman & Smith, L-CSS, 2021
Chapman & Lessard, L-CSS, 2021
Chapman, Bonalli+, TAC, 2022
Chapman, Fauss, Smith, TAC (conditional accept)
Wei, Fauss, Chapman, ACC, 2022
Arsenault, Wang, Chapman, L-CSS, 2022 (in press)

Risk Analysis or
Synthesis of Risk-Averse Policies

Numerical
Experiments

Ultimately deploy &
incorporate into design
practices

UNIVERSITY OF TORONTO

Photo credits: toronto.citynews.ca, Risom+ 2018, Wang & Chapman 2022

ACC: *American Control Conference*
L-CSS: *IEEE Control Systems Letters*
TAC: *IEEE Transactions on Automatic Control*

What is risk?

- There are different definitions for risk, which may be interpreted qualitatively or quantitatively based on application-specific needs.
- Risk is the “possibility of loss or injury” (Merriam-Webster).
- Risk is an entity that “creates or suggests a hazard” (Merriam-Webster).
- Risk is “the effect of uncertainty on objectives” (International Organization for Standardization).

The purpose of this talk is to examine and present the core connections between the analysis of risk and the control of semi-autonomous systems.

Talk Outline

1. Present the two mainstream paradigms for managing systems under uncertainty
2. Motivate the risk-averse paradigm
3. Present three concepts that are helpful for quantifying risk
4. Present risk functionals & optimal control problems
5. Discuss adaptability and scalability challenges
6. Propose future directions

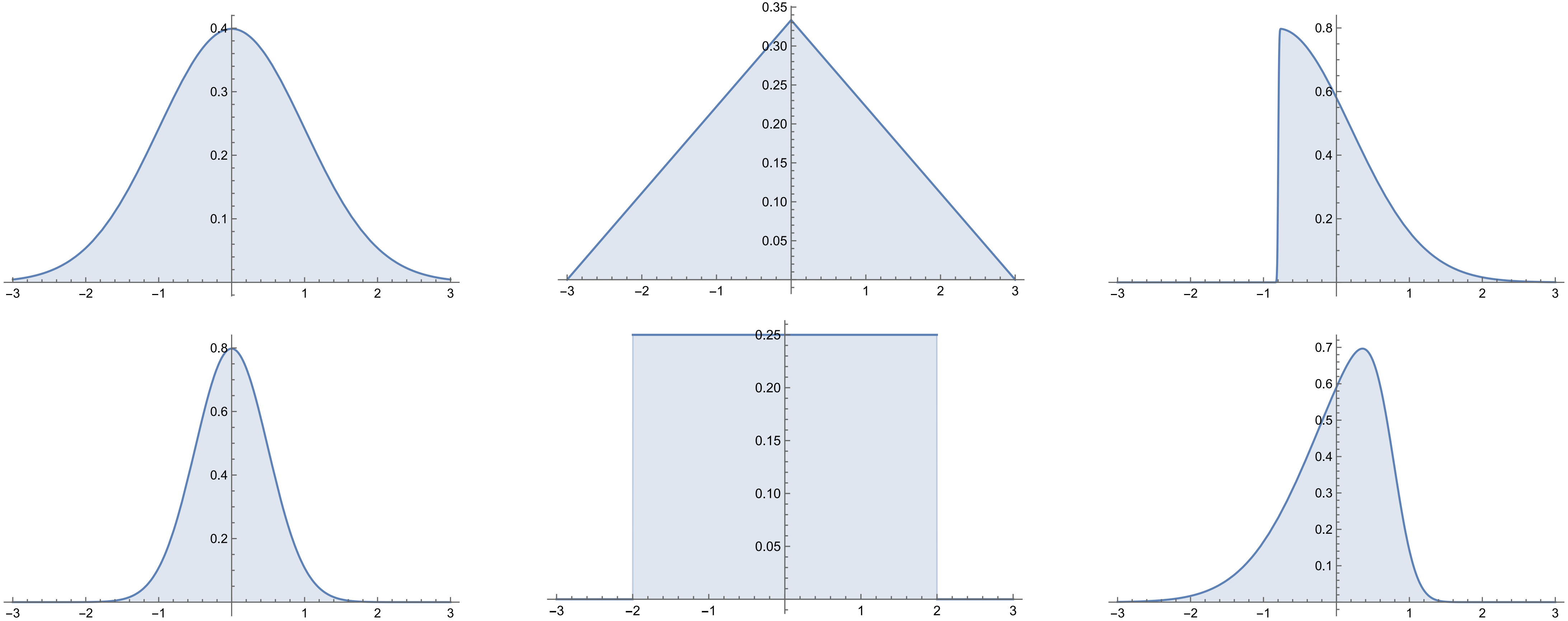
There are two mainstream paradigms for quantifying & managing the potential consequences of a system’s behaviour.

	Robust Paradigm (worst-case)	Stochastic Paradigm (risk-neutral)
Disturbances are modelled as…	nonstochastic bounded adversarial inputs	random variables (may be adversarial)
Quantifies safety or performance…	assuming the “worst” circumstances	in probability or on average
	Bertsekas & Rhodes 1971, Heger 1994, Coraluppi & Marcus 1999, Mitchell+ 2005, Margellos & Lygeros 2011, Chen & Tomlin 2018, Huang+ 2019, Ivanov+ 2020,…	Geibel 2001, Geibel & Wysotzki 2005, Abate+ 2008, Summers & Lygeros 2010, Forejt+ 2011, García & Fernández 2012, Moldovan & Abbeel 2012, Ding+ 2013, Schildbach+ 2014, Yang 2020,…

Includes literature from the machine learning, robust control, stochastic control, and formal verification communities.

More literature can be found in Wang & Chapman 2022.

An issue with assessing safety or performance on average alone is that distributions with diverse characteristics can have the same expectation, e.g.,

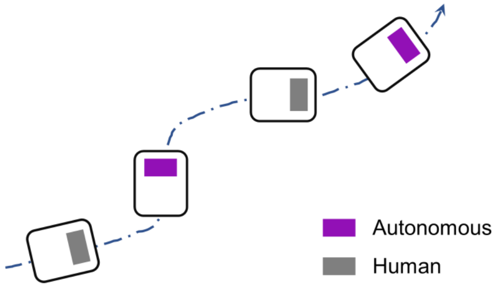


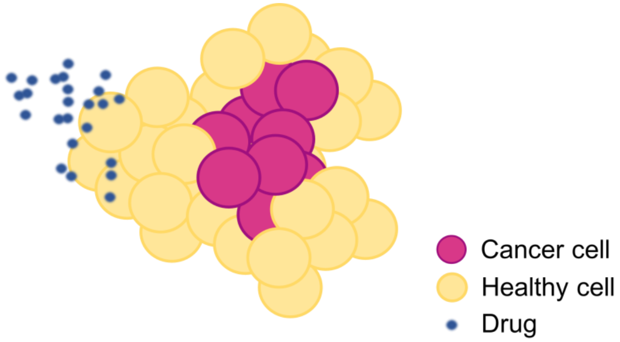
If the outcomes represent potential costs incurred by a system, we may prefer some of these distributions over others.

Courtesy of Kevin Smith

Motivation for Risk-Averse Systems Theory

- The expectation is not designed to assess rare harmful outcomes.
- Assuming bounded disturbances excludes common noise models.
- Systems often require an awareness of rare harmful outcomes without being too cautious.
- Broad applications: civil & environmental infrastructure under weather uncertainty, population growth, medical applications under patient-to-patient variability, human-robot interaction,…





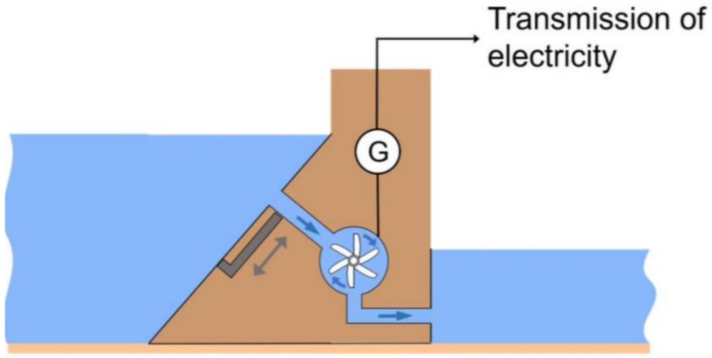


Figure credit: Wang & Chapman 2022

What does risk-averse mean?

- Colloquially, risk-averse describes people or algorithms that prefer outcomes with reduced uncertainty.

- We will propose a definition that emphasizes systems.

We propose a definition for risk-averse that emphasizes systems.

- For a given system, we assume that there is a random variable Z, whose realizations describe the consequences that may arise from the system’s behaviour.
- Risk-averse describes people or algorithms that prefer distributions for Z with specific characteristics, where the characteristics reflect a desire to reduce harm.
- To quantify risk means to summarize (numerically) the potential consequences of the system’s behaviour.

Three concepts that are helpful for quantifying the risk of a system:

1. Probabilities of harmful events

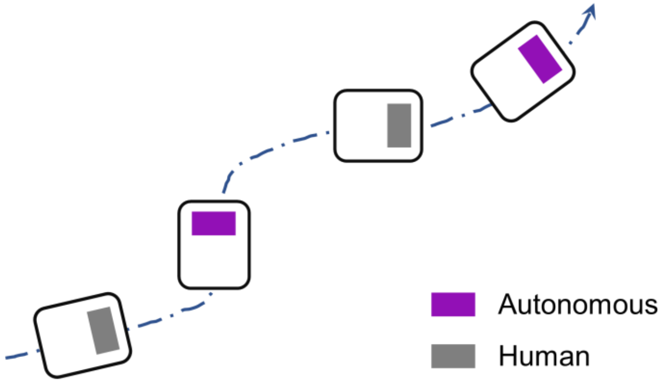
2. Temporal logic (TL) specifications

3. Risk functionals of random variables

Temporal logic (TL) is a mathematical language for describing relationships between different events in time.
TL specifications can be deterministic or probabilistic.

Risk functionals and risk measures are synonyms.

Example: A Network of Autonomous and Human Drivers

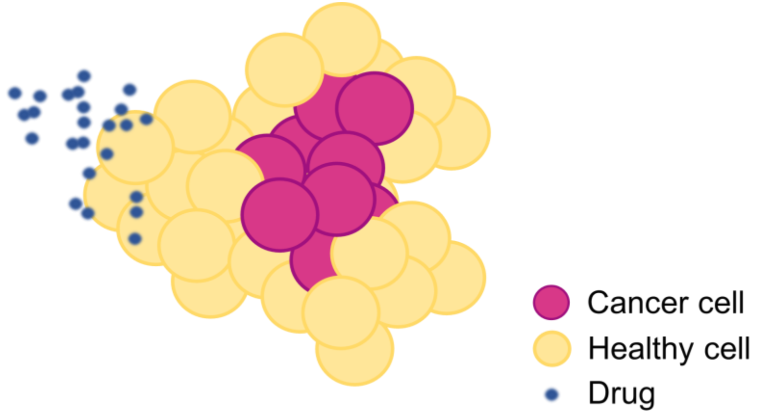


- Z_t : distance between an autonomous car and a human-driven car at time t (larger realizations are preferred)
- z : smallest “allowable” distance

1. Probability of a harmful event: probability that $Z_t \leq z$ for some time t (i.e., probability of a collision)
2. Temporal logic (TL) specification: $Z_t \geq z$ for every time t with probability 1
3. Risk functional of a random variable: average value of $Z_t + (Z_t – [\text{desired distance}])^2$

Figure credit: Wang & Chapman 2022

Example: A Cancer Cell Population



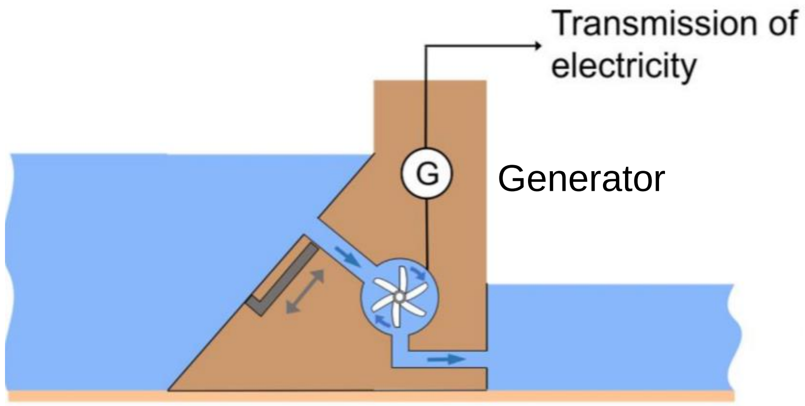
- Z : [# healthy cells]/[desired # healthy cells] at the end of a treatment cycle (larger realizations are preferred)
- z : smallest “allowable” value of the ratio, estimated from a doctor’s expertise
- A distribution for Z can be estimated using a dynamical model with varying parameter estimates.

1. Probability of a harmful event: probability that $Z \leq z$ (e.g., treatment is too toxic, drug resistance develops,…)
2. TL specification: For every cycle after the 5th cycle, $Z \geq 1$ with probability 0.9.
3. Risk functional of a random variable: average value of Z in the 1% worst cases

TL = temporal logic

Figure credit: Wang & Chapman 2022

Example: A Hydroelectric Dam



- Z : volume of water that discharges into the emergency spillway due to a storm (smaller realizations are preferred)
- z : “allowable” discharge volume based on local regulations
- A distribution for Z can be estimated using a model for the dam and historical precipitation data.

1. Probability of a harmful event: probability that Z ≥ z (i.e., probability of an overflow)
2. TL specification: Until the third storm of the current month, the probability that Z ≥ z is less than 0.05.
3. Risk functional of a random variable: average value of max{Z – z, 0}

Figure credit: Wang & Chapman 2022

Talk Outline

1. Present the two mainstream paradigms for managing systems under uncertainty

2. Motivate the risk-averse paradigm

3. Present three concepts that are helpful for quantifying risk

4. Present risk functionals & optimal control problems:
    a. Examples of risk functionals & standard form of a risk-averse optimal control problem
    b. Pros/cons of exponential utility optimal control
    c. Two additional methods

5. Discuss adaptability and scalability challenges

6. Propose future directions

We will focus on risk functionals.

A risk functional ρ is a map from $\mathbb{Z}$, a space of random variables, to $\mathbb{R} \cup \{+\infty\}$.

$$\rho : \mathbb{Z} \longrightarrow \mathbb{R} \cup \{+\infty\}$$

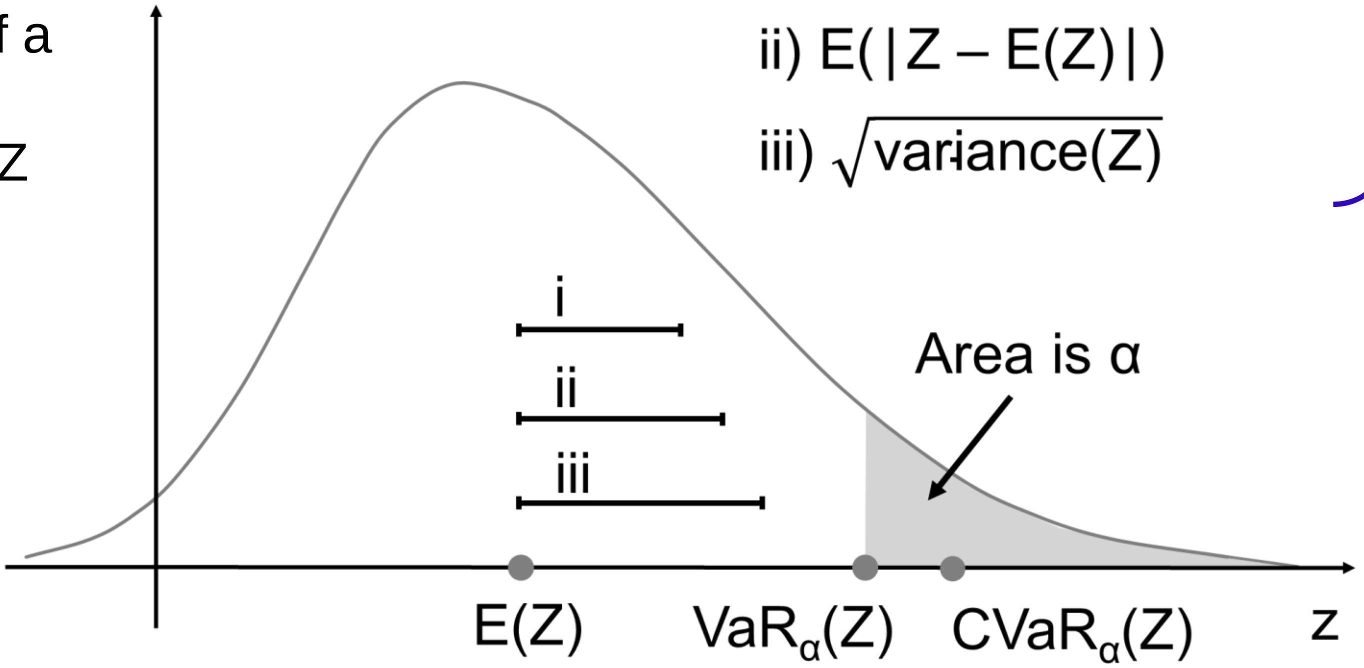
We assume that smaller realizations of $Z \in \mathbb{Z}$ are preferred.

- Typically, one selects $\mathbb{Z}$ to be an L^p-space with $p \in [1, +\infty]$ chosen so that $\rho(Z) \in \mathbb{R}$ for every $Z \in \mathbb{Z}$.
- The term risk *functional* emphasizes that the domain of ρ is a space of functions.

Shapiro, Dentcheva, and Ruszczyński 2009

Risk functionals can quantify heavy-tailed distributions or higher-order moments of random variables.

Density of a random variable Z



i) $E(\max\{Z - E(Z), 0\})$

ii) $E(|Z - E(Z)|)$

iii) $\sqrt{\text{variance}(Z)}$

Different ways to quantify “dispersion” from the mean

Value-at-Risk

Conditional Value-at-Risk

Assumptions: Z has a density; α is sufficiently small so that $E(Z) + \sqrt{\text{variance}(Z)} < \mathrm{VaR}_\alpha(Z)$.

Additional Common Risk Functionals

Compositional: $\rho(Z_1 + Z_2 + Z_3 + Z_4) \coloneqq \rho_1\Big(Z_1 + \rho_2\Big(Z_2 + \rho_3\big(Z_3 + \rho_4(Z_4)\big)\Big)\Big)$

Expected Utility: $\rho(Z) \coloneqq E\big(h(Z)\big)$

$$\rho(Z) \coloneqq h^{-1}\Big(E\big(h(Z)\big)\Big)$$

h is a (sufficiently regular) utility function

- One must choose the domain $\mathbb{Z}$ appropriately so that ρ is well-defined.
- Compositional risk functionals are also called *nested* or *recursive* risk functionals; we show 4 stages for simplicity, but more stages can be chosen.

Standard Form of a Risk-Averse Optimal Control Problem

$$\inf_{\pi\in\Pi} \rho^{\pi}(Z)$$

Risk-averse objective

Π: particular policy class

Subject to: System dynamics,

Explicit equations may not be available.

$$\rho_i^{\pi}(Z_i) \in K_i,\ i \in \mathcal{J}.$$

Risk-averse constraints

- ρ^{π}, ρ_i^{π}: risk functionals
- $Z,\ Z_i$: random variables whose distributions depend on π (and typically also an initial state)

Most research concerns exponential utility…

UNIVERSITY OF TORONTO

Risk Functionals & Systems (not exhaustive)

Exponential Utility $\frac{-2}{\theta}\log E\left(\exp\left(\frac{-\theta}{2}Z\right)\right)$ with $\theta \neq 0$ is viewed as a mean-variance approximation.

- Howard & Matheson 1972: finite state spaces

Linear **E**xponential **Q**uadratic **G**aussian

- Jacobson 1973: Euclidean spaces, linear systems, quadratic costs, Gaussian noise (“LEQG”)
- Whittle+ 1980-1990s: LEQG with partially observable states
- Glover & Doyle 1988: Analyzed relations between controllers satisfying an $\mathcal{H}_\infty$-norm bound & the infinite-time LEQG controller
- di Masi & Stettner 1999: Borel spaces, nonlinear systems, infinite time
- Bäuerle & Rieder 2014: Borel spaces, nonlinear systems, expected utility, studied exponential utility as a special case
- Saldi, Başar, Raginsky 2020: exponential utility & mean-field games
- Chapman & Smith 2021: Borel spaces, nonlinear systems, finite time, optimality from first principles

UNIVERSITY OF TORONTO

More literature can be found in Wang & Chapman 2022.

Classical Example: Finite-Time Exponential Utility Optimal Control

Risk-aversion parameter ($\theta < 0$)

$$\inf_{\pi \in \Pi} \frac{-2}{\theta} \log E_x^{\pi}\left(\exp\left(\frac{-\theta}{2} Z\right)\right)$$

Class of deterministic Markov policies

Initial state

Cumulative cost, whose distribution depends on (π, x)

Subject to:

$$x_{t+1} = f_t(x_t, u_t, w_t), \qquad u_t \in C, \qquad w_t \sim \underbrace{p_t(\,\cdot\,| x_t, u_t)}_{\text{Disturbance distribution}},$$

Dynamics function

Action space

where the disturbances are conditionally independent.

Under appropriate conditions (e.g., θ is “near” zero and others…),

$$\frac{-2}{\theta} \log E_x^{\pi}\left(\exp\left(\frac{-\theta}{2} Z\right)\right) \approx E_x^{\pi}(Z) - \frac{\theta}{4} \mathrm{V}_x^{\pi}(Z).$$

variance

Pros & Cons of Exponential Utility Optimal Control

Pros:

- Admits a dynamic program on the state space
- Represents a user’s subjective preferences (utility model)
- Provides a mean-variance approximation under appropriate conditions

Challenges:

- Can we consistently produce a mean-variance tradeoff by making θ more negative?
- Can we provide a precise quantitative interpretation of Exponential Utility and its parameter θ?

We will provide a simple example to illustrate how different values of θ can affect the “optimal” distribution of Z.

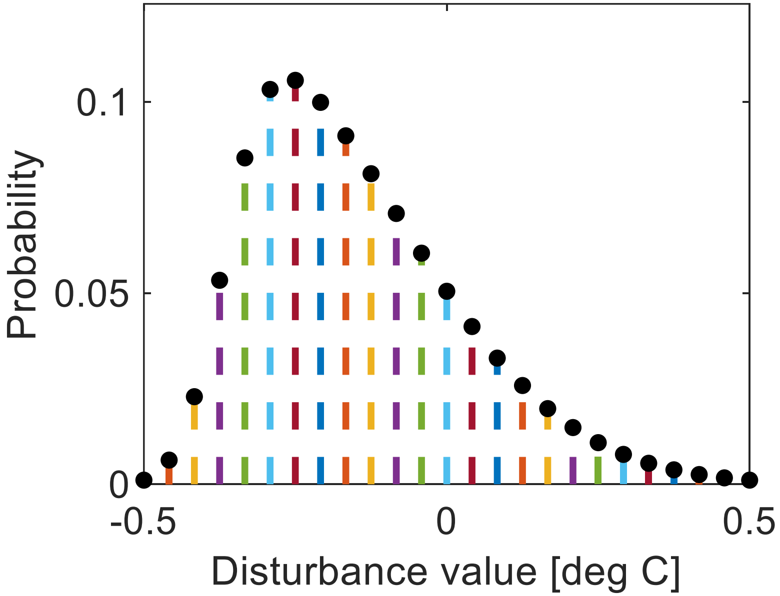
- We consider a thermostatic regulator:

$$x_{t+1} = ax_t + (1-a)(b - \eta RPu_t) + w_t,$$

$$t = 0,1,\ldots,N-1.$$

- A realization of Z takes the form $\sum_{t=0}^{N} c(x_t)$.

- $c(x_t)$ represents the “distance” between a state x_t and a desired range [20, 21] °C.



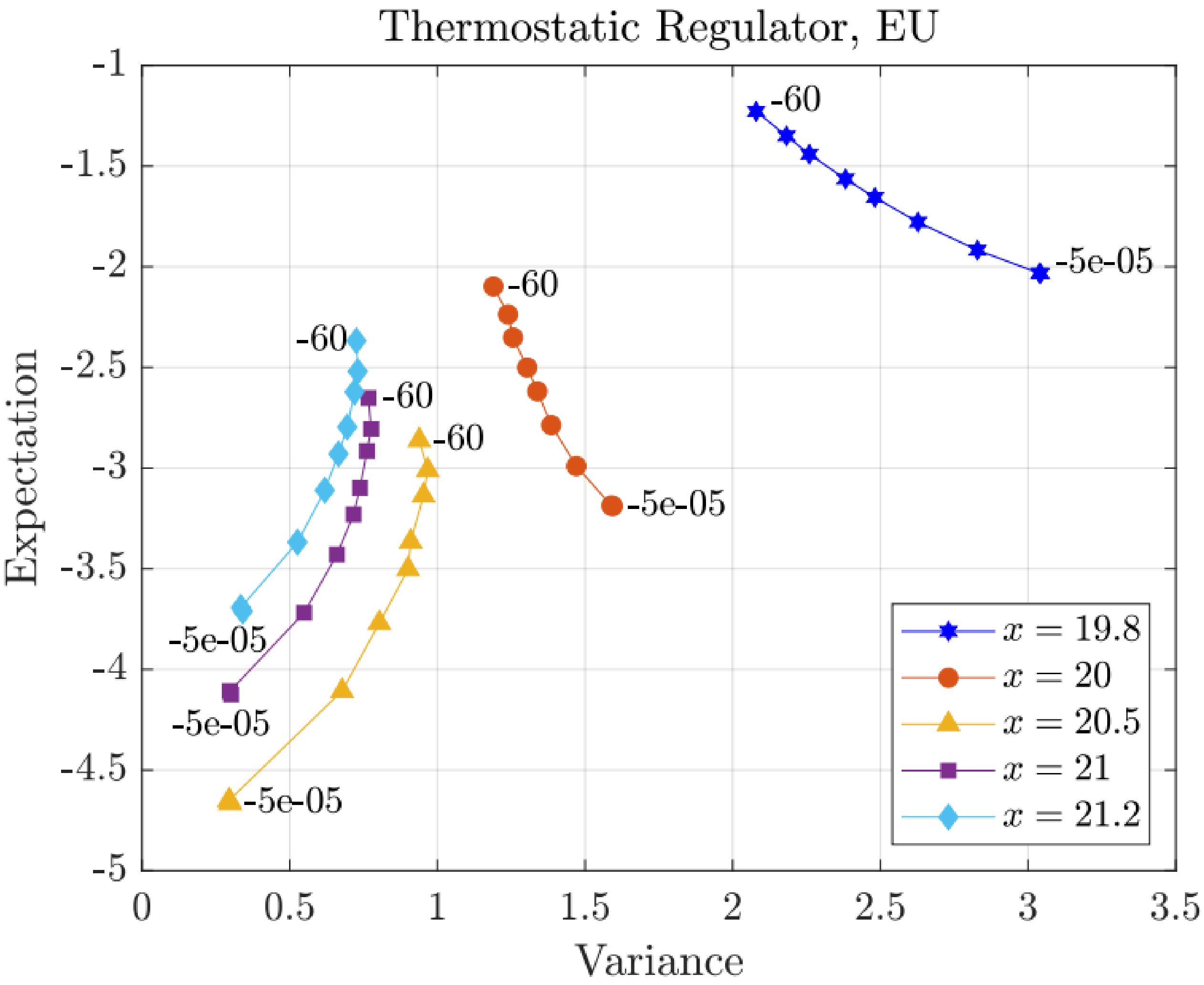
The assumed disturbance distribution is discrete and does not depend on (x_t, u_t).

Model adopted from Yang 2018, originally from Mortensen & Haggerty 1988;

Smith & Chapman, under review

The empirical mean *and* variance may increase, as θ becomes more "risk-averse."

Nearly risk-neutral
$\theta = -5 \cdot 10^{-5}$



$$c(x_t) = \max\{x_t - 21, 20 - x_t\}$$

Kevin Smith

A realization of Z takes the form $\sum_{t=0}^{N} c(x_t)$.

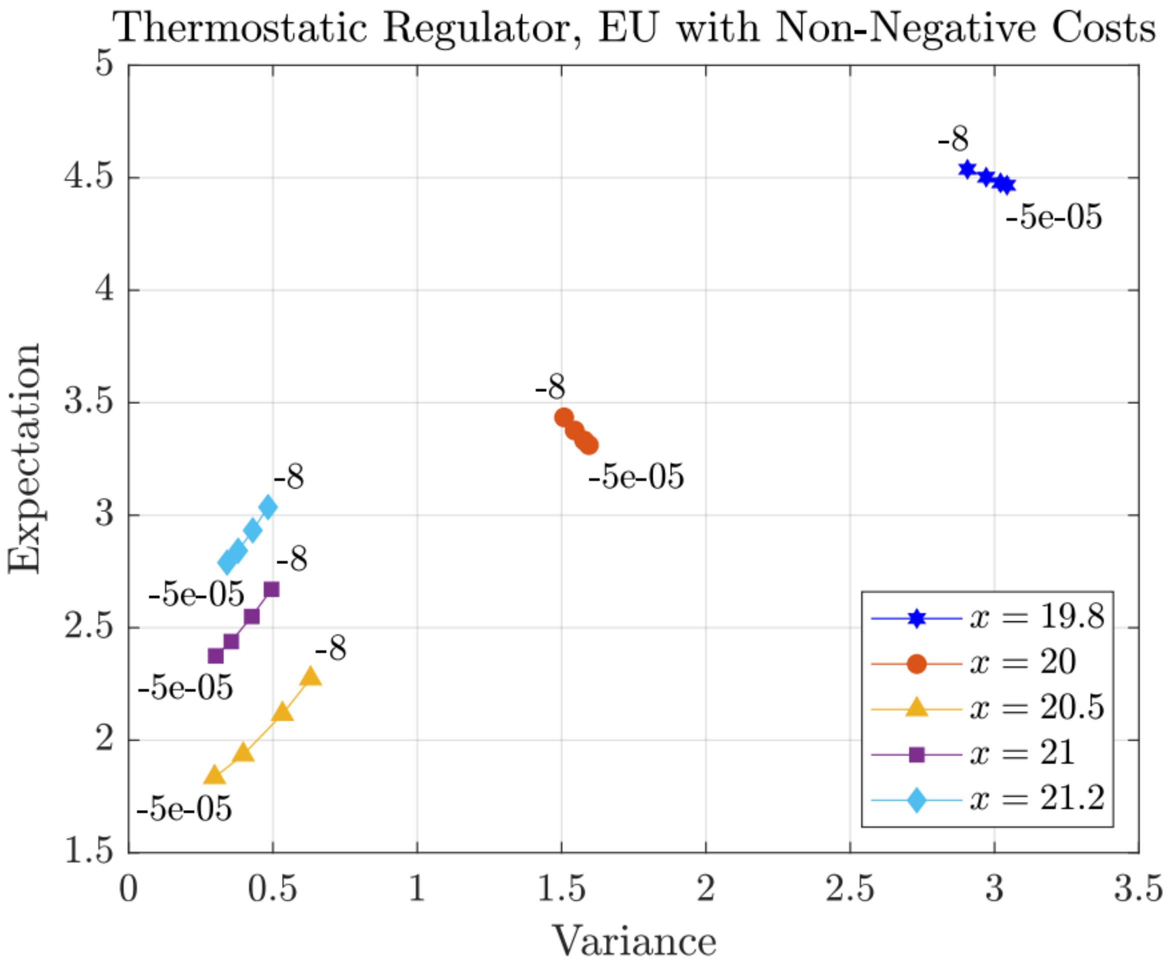
10 million samples of Z per point

Smith & Chapman, under review

The empirical mean *and* variance may increase, as θ becomes more "risk-averse."

Nearly risk-neutral
$\theta = -5 \cdot 10^{-5}$

Kevin Smith



Thermostatic Regulator, EU with Non-Negative Costs

$c(x_t) = \max\{x_t - 21{,}20 - x_t, 0\}$

Non-negative

A realization of Z takes the form $\sum_{t=0}^{N} c(x_t)$.

10 million samples of Z per point

UNIVERSITY OF TORONTO

Smith & Chapman, under review

Recall that there are different types of risk functionals, e.g.,

Mean-dispersion: $E(Z) + \lambda\eta(Z)$, where $\lambda > 0$;

Expected Utility: $h^{-1}(E(h(Z)))$, where h is a sufficiently regular utility function;

Compositional: $\rho_1(Z_1 + \rho_2(Z_2 + \rho_3(Z_3 + \rho_4(Z_4))))$;

Quantile-based: $\mathrm{VaR}_\alpha(Z)$, $\mathrm{CVaR}_\alpha(Z)$.

We will present related literature next.

Risk Functionals & Systems (not exhaustive)

Compositional: $\rho(Z_1 + Z_2 + Z_3 + Z_4) := \rho_1\left(Z_1 + \rho_2\left(Z_2 + \rho_3\left(Z_3 + \rho_4(Z_4)\right)\right)\right)$

- Ruszczyński 2010; Ruszczyński 2014; Asienkiewicz & Jáskiewicz 2017; Singh+ 2018; Bäuerle & Glauner 2021 (Eur. J. Oper. Res.); Köse & Ruszczyński 2021, Ahmadi+ 2021,…

Quantile-based (e.g., CVaR, VaR, Spectral):

- Bäuerle & Ott 2011: CVaR, cumulative cost, Borel spaces, DP
- Borkar & Jain 2014: CVaR constraint on a cumulative cost
- Haskell & Jain 2015: CVaR, cumulative cost, infinite time, infinite-dimensional LP*
- Miller & Yang 2018: CVaR, terminal cost, continuous time*
- Chapman, Lacotte+ 2019, Chapman, Bonalli+ 2022: CVaR, maximum cost, safety analysis, DP
- Bäuerle & Glauner 2021 (Math. Method. Oper. Res.): spectral risk functionals, Borel spaces, DP
- Lindemann+ 2021: CVaR, VaR, signal temporal logic*
- Barbosa+ 2021: CVaR applied to robot motion planning

UNIVERSITY OF TORONTO

*Considers additional risk functionals.

LP: linear programming, DP: dynamic programming.

Talk Outline

1. Present the two mainstream paradigms for managing systems under uncertainty

2. Motivate the risk-averse paradigm

3. Present three concepts that are helpful for quantifying risk

4. Present risk functionals & optimal control problems:
   a. Examples of risk functionals & standard form of a risk-averse optimal control problem
   b. Pros/cons of exponential utility optimal control
   c. Two additional methods:
      i. Reduction to a family of nonstandard risk-neutral problems
      ii. Dynamic programming for compositional risk functionals

5. Discuss adaptability and scalability challenges

6. Propose future directions

UNIVERSITY OF TORONTO

Example Method: Reduction to risk-neutral problems (with state-space augmentation)

- We aim to minimize the CVaR of a maximum random variable $Z : \Omega \to \mathbb{R}$

Trajectory space

State space

$$V_\alpha^*(x) := \inf_{\pi \in \Pi} \mathrm{CVaR}_{\alpha,x}^{\pi}(Z), \qquad \alpha \in (0,1], \qquad x \in S,$$

$$Z := \max\{c_N(X_N), c_t(X_t, U_t) : t \in \{0,1,\dots,N-1\}\}, \qquad c_t \in [a,b] \subset \mathbb{R} \text{ for every } t.$$

Random max. distance between the system's trajectory and a desired operating region

Cost function

- This problem provides a safety criterion for a stochastic system that is informed by *both* the probability and severity of the potential consequences of the system's behaviour.

- $\mathrm{CVaR}_{\alpha,x}^{\pi}(Z)$ may be interpreted as the average value of Z in the $\alpha \cdot 100\%$ worst cases when the initial state is x and the system uses the policy π.

- Π is a particular class of history-dependent policies; π_t in $\pi = (\pi_0, \pi_1, \dots, \pi_{N-1}) \in \Pi$ depends on the state at time t and the running maximum cost up to time t.

High-level description of Chapman, Fauss, Smith, conditionally accepted by *IEEE Trans. Autom. Control*, May 2022.
More conditions (not stated on this slide) are required to ensure that quantities are well-defined.

Example Method: Reduction to risk-neutral problems (with state-space augmentation)

- A key idea: Identify a family of risk-neutral problems within a risk-averse problem.[1]

- $\mathrm{CVaR}_{\alpha,x}^{\pi}(Z)$ is defined by

$$\mathrm{CVaR}_{\alpha,x}^{\pi}(Z) := \inf_{s\in\mathbb{R}} \left(s + \tfrac{1}{\alpha} E_x^{\pi}(\max\{Z - s, 0\}) \right), \qquad \alpha \in (0,1], \qquad x \in S, \qquad \pi \in \Pi.$$

Risk-aversion level

- Therefore, we have

$$\begin{aligned} V_\alpha^*(x) &:= \inf_{\pi\in\Pi} \mathrm{CVaR}_{\alpha,x}^{\pi}(Z), \\ &= \inf_{\pi\in\Pi} \inf_{s\in\mathbb{R}} \left(s + \tfrac{1}{\alpha} E_x^{\pi}(\max\{Z - s, 0\}) \right), \\ &= \inf_{s\in\mathbb{R}} \left(s + \tfrac{1}{\alpha} \underbrace{\inf_{\pi\in\Pi} E_x^{\pi}(\max\{Z - s, 0\})}_{V^s(x)} \right). \end{aligned}$$

- Now, the task is to solve for $\{V^s(x) : s \in \mathbb{R}, x \in S\}$.

High-level description of Chapman, Fauss, Smith, conditionally accepted by *IEEE Trans. Autom. Control*, May 2022. More conditions (not stated on this slide) are required to ensure that quantities are well-defined.

[1] Additional papers use similar reduction techniques (slide 28).

Example Method: Reduction to risk-neutral problems (with state-space augmentation)

- Recall: $V^*_\alpha(x) := \inf_{\pi\in\Pi} \mathrm{CVaR}^\pi_{\alpha,x}(Z) = \inf_{s\in\mathbb{R}}\left(s + \frac{1}{\alpha} V^s(x)\right).$

- To solve for $V^s(x) := \inf_{\pi\in\Pi} E^\pi_x(\max\{\underset{\substack{\uparrow \\ \text{Maximum random variable}}}{Z} - s, 0\})$, we define an extra random state $Y_{t+1} := \max\{\underset{\substack{\uparrow \\ \text{Stage cost function}}}{c_t}(X_t, U_t), Y_t\}$, initialized at a lower bound a for c_t.

- Under appropriate conditions, we show that $V^s(x) = J^s_0(x,a)$, where J^s_0 is defined recursively by

$$J^s_N(x,y) := \max\{\max\{\overset{\substack{\text{Terminal cost function} \\ \downarrow}}{c_N}(x), y\} - s, 0\}, \qquad (x,y) \in S \times [a,b],$$

$$J^s_t(x,y) := \inf_{\substack{u\in C \\ \uparrow \\ \text{Action space}}} \int_D J^s_{t+1}(\underset{\substack{\uparrow \\ \text{Dynamics function}}}{f_t}(x,u,w), \max\{c_t(x,u), y\})\, \underbrace{p_t(\mathrm{d}w|x,u)}_{\text{Disturbance distribution}}, \qquad t = N-1, \dots, 1, 0, \qquad (x,y) \in \underbrace{S \times [a,b]}_{\text{Augmented state space}},$$

and there is a deterministic policy π^s that is optimal for V^s.

High-level description of Chapman, Fauss, Smith, conditionally accepted by *IEEE Trans. Autom. Control*, May 2022.
More conditions (not stated on this slide) are required to ensure that quantities are well-defined.

Talk Outline

1. Present two mainstream paradigms for managing systems under uncertainty

2. Motivate the risk-averse paradigm & define risk-related terms for systems

3. Present three concepts that are helpful for quantifying risk with examples

4. Discuss risk functionals & optimal control problems:
    a. Examples of risk functionals & standard form of a risk-averse optimal control problem
    b. Pros/cons of exponential utility optimal control
    c. Two additional methods:
        i. Reduction to a family of nonstandard risk-neutral problems
        ii. Dynamic programming for compositional risk functionals

5. Adaptability and scalability challenges

6. Proposed future directions

Example Method: Dynamic Programming for Compositional Risk Functionals

- High-level idea: Replace a conditional expectation with a risk-averse analog.

Terminal cost function

Dynamics function

Random disturbance

State space

- For $\pi \in \Pi$, define $J_N^\pi \coloneqq c_N$, and for $t = N-1, \dots, 1, 0$, define

Class of deterministic Markov policies

$$J_t^\pi(x) \coloneqq c_t\big(x, \pi_t(x)\big) + v_t\Big(J_{t+1}^\pi\big(f_t(x, \pi_t(x), W_t)\big)\Big), \qquad x \in S,$$

$$J_t^*(x) \coloneqq \inf_{\pi \in \Pi} J_t^\pi(x), \qquad x \in S.$$

a risk functional

- A special case is $$v_t\Big(J_{t+1}^\pi\big(f_t(x,u,W_t)\big)\Big) = \int_D J_{t+1}^\pi\big(f_t(x,u,w)\big)\, \underbrace{p_t(\mathrm{d}w|x,u)}_{\text{Disturbance distribution}}, \qquad (x,u) \in S \times C.$$

Space of disturbances

Action space

- Under appropriate conditions, J_t^* satisfies

$$J_t^*(x) = \inf_{u \in C}\bigg(c_t(x,u) + v_t\Big(J_{t+1}^*\big(f_t(x,u,W_t)\big)\Big)\bigg), \qquad x \in S.$$

High-level description of a piece of Bäuerle & Glauner 2021 (Eur. J. Oper. Res.). Additional papers about compositional risk functionals use related techniques (slide 28). More conditions (not stated on this slide) are required to ensure that quantities are well-defined.

Many problem-solving methods are based on reformulating a risk-neutral algorithm. We have seen two examples:

1. Reduce a risk-averse problem to a family of nonstandard risk-neutral problems (e.g., when the objective is defined using CVaR, a spectral risk functional, or an extremal risk functional).

2. Replace a conditional expectation with a risk-averse analog (e.g., when the objective is defined using a compositional risk functional).

→ Such a method *inherits adaptability and scalability issues* of the original risk-neutral algorithm.

A risk-neutral algorithm: An algorithm that aims to minimize an expected cumulative cost subject to a (partially observable) Markov decision process (e.g., dynamic programming, Q-learning, temporal difference learning).

Some Adaptability Issues & Related Questions

- A *representative* family of disturbance distributions or an ability to generate *representative* data samples (e.g., by simulating transitions) is required to "optimize" a system's behaviour.

How can we design off-line or on-line experiments to...

…understand and model the sources of uncertainty and estimate their (potentially) time-varying effects on the system *without being overly conservative*?

…generate samples with desired statistical properties at runtime?

…accomplish the above when sampling is expensive?

Some Scalability Issues & Related Questions

- Typically, dynamic programming, Q-learning, and temporal difference learning cannot scale to high-dimensional state spaces without value function or policy approximations.

  ➡ For a risk-averse optimal control problem of interest, what theoretical conditions justify the use of such approximations?

- Training value function or policy approximations (e.g., using polynomial basis functions, neural networks, etc.) may require large data sets.

  ➡ How can existing statistical methods (e.g., importance sampling, extreme value theory) be applied to model rare high-consequence events in *data-sparse* applications?

Extreme Value Theory (EVT) may be useful for risk estimation in data-sparse applications.

- EVT is the study of the long-term behaviour of normalized maxima of random variables (de Haan & Ferreira 2006).
- Recently, we proposed an EVT-based estimator for the upper semi-deviation $E(\max\{Z - E(Z), 0\})$ in a fraction α of the largest realizations of Z.
- We showed that the estimator enjoys a closed-form representation in terms of CVaR.
- In experiments, we illustrated the extrapolation power of the estimator using a small number of i.i.d. samples.

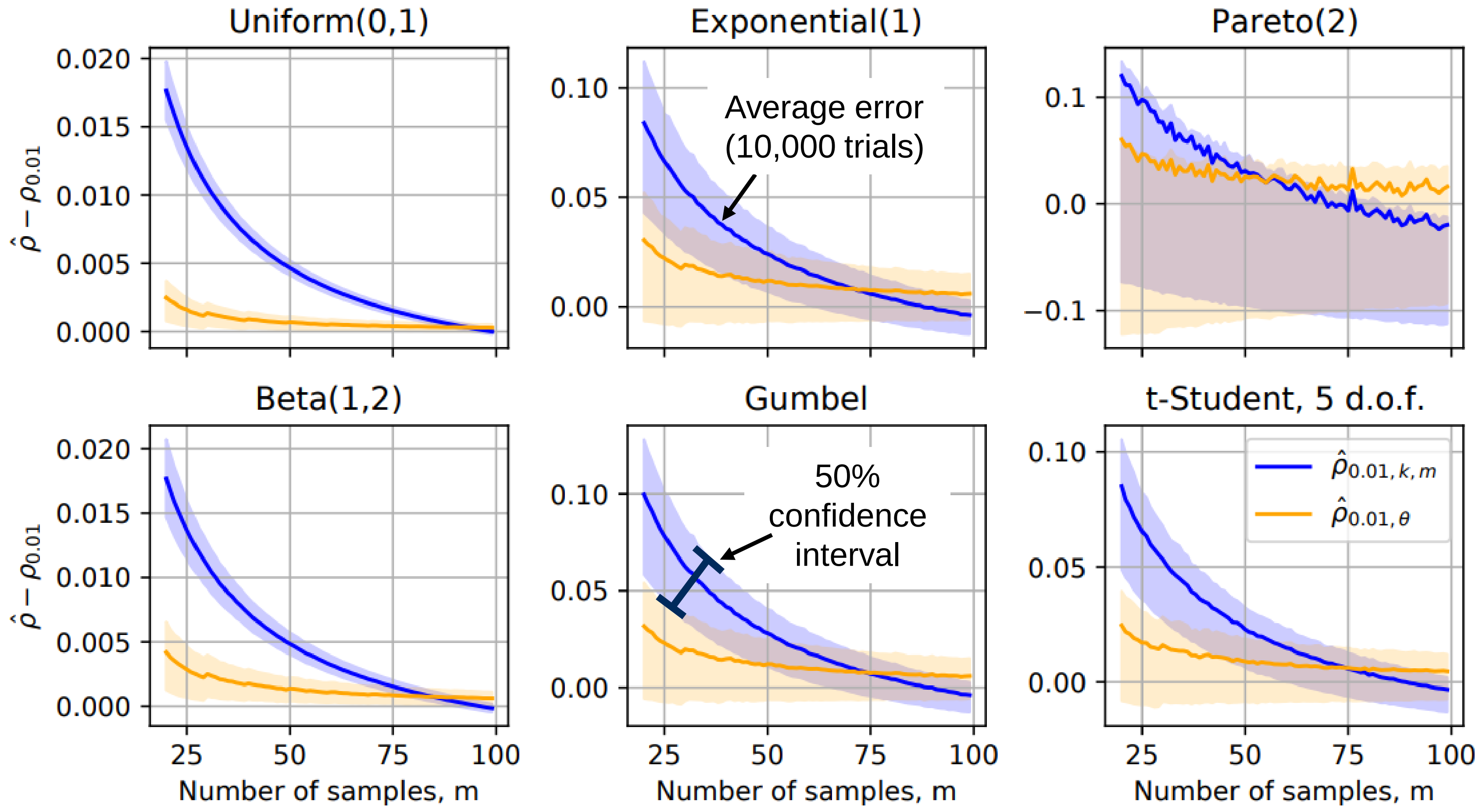

Yuheng Wang


Evan Arsenault

UNIVERSITY OF TORONTO

Arsenault, Wang, Chapman, *IEEE Control Systems Letters*, 2022 (in press)

The EVT-based estimator outperforms the typical estimator for smaller numbers of samples.



Average error (10,000 trials)

50% confidence interval

$\rho_{0.01}$: Monte Carlo estimate with ≥4 million samples

Typical estimator: based on a sample average

Blue: Typical estimator – $\rho_{0.01}$

Orange: EVT-based estimator – $\rho_{0.01}$

Yuheng Wang

Evan Arsenault

Arsenault, Wang, Chapman, *IEEE Control Systems Letters*, 2022 (in press)

Proposed Future Directions

- Use of risk functionals and data samples to *adapt* estimates of system models or “safe” regions
    - Köse and Ruszczyński 2021 provides empirical evidence that risk functionals can protect against modelling (specifically, value function approximation) errors.

- Further studies about risk-averse model-free methods from a *nonasymptotic* viewpoint
    - Finite-time horizons and finitely many samples are used in practice.
    - Huang & Haskell 2021 offers a nonasymptotic analysis of a risk-averse Q-learning algorithm.

- Additional investigations about which risk functional(s) may be more appropriate for a particular application

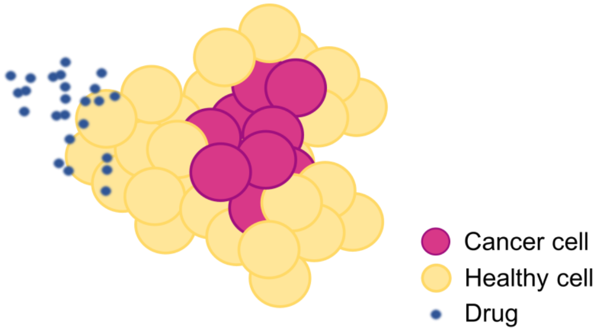
Proposed Future Directions

- Studies that develop theoretical risk-averse optimal control methods with *both* model-free and model-based aspects

- Different variations of such methods are needed to accommodate diverse applications. E.g.,

    - To improve a cancer patient’s outcomes, consider a blend of the oncologist’s expertise, the patient’s recent and historical data, and biological and chemical models.

    - To support environmental health (e.g., related to hydroelectricity), combine advice from water & energy experts, precipitation data, weather forecasts, and models for the flow and quality of water.



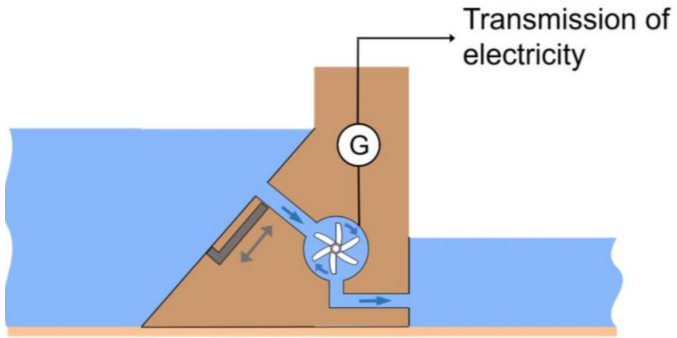


Figure credit: Wang & Chapman 2022

Acknowledgements

The Edward S. Rogers Sr. Department
of Electrical & Computer Engineering
UNIVERSITY OF TORONTO

- My research group: Evan Arsenault, Yuheng Wang, Chuanning Wei
- My collaborators: Kevin Smith (Tufts, OptiRTC) & Michael Fauss (Princeton)
- I thank Peter Caines (McGill) & Yuxi Han (Wisconsin-Madison) for fruitful discussions.
- I appreciate research funding provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants program [RGPIN-2022-04140], NSERC Undergraduate Summer Research Award program, and Edward S. Rogers Sr. Department of Electrical & Computer Engineering, University of Toronto.


Evan Arsenault


Yuheng Wang


Chuanning Wei


Kevin Smith


Michael Fauss

UNIVERSITY OF TORONTO

Natural Sciences and Engineering Research Council of Canada

Conseil de recherches en sciences naturelles et en génie du Canada

Canada

References

A. Abate, M. Prandini, J. Lygeros, and S. Sastry, Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems, *Automatica*, vol. 44, no. 11, pp. 2724–2734, 2008.

M. Ahmadi, U. Rosolia, M. D. Ingham, R. M. Murray, and A. D. Ames, Risk-averse decision making under uncertainty, arXiv preprint arXiv:2109.04082, 2021.

E. Arsenault, Y. Wang, and M. P. Chapman, Towards scalable risk analysis for stochastic systems using Extreme Value Theory, *IEEE Control Systems Letters*, accepted June 15, 2022.

H. Asienkiewicz and A. Jáskiewicz, A note on a new class of recursive utilities in Markov decision processes, *Applicationes Mathematicae* 44 (2017) 149–161.

F. S. Barbosa, B. Lacerda, P. Duckworth, J. Tumova, and N. Hawes, Risk-aware motion planning in partially known environments, arXiv preprint arXiv:2109.11287, 2021.

N. Bäuerle and A. Glauner, Markov decision processes with recursive risk measures, *European Journal of Operational Research*, 2021.

N. Bäuerle and A. Glauner, Minimizing spectral risk measures applied to Markov decision processes, *Mathematical Methods of Operations Research*, 2021.

References

N. Bäuerle and J. Ott, Markov decision processes with Average-Value-at-Risk criteria, *Mathematical Methods of Operations Research*, vol. 74, no. 3, pp. 361–379, 2011.

N. Bäuerle and U. Rieder, More risk-sensitive Markov decision processes*, Mathematics of Operations Research*, vol. 39, no. 1, pp. 105–120, 2014.

D. P. Bertsekas and I. B. Rhodes, On the minimax reachability of target sets and target tubes, *Automatica*, 7 (2) (1971) 233–247.

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

V. Borkar and R. Jain, Risk-constrained Markov decision processes, *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2574–2579, 2014.

M. P. Chapman, R. Bonalli, K. M. Smith, I. Yang, M. Pavone, C. J. Tomlin, Risk-sensitive safety analysis using Conditional Value-at-Risk, *IEEE Transactions on Automatic Control* (2022) in press.

M. P. Chapman, M. Fauss, K. M. Smith, On optimizing the conditional value-at-risk of a maximum cost for risk-averse safety analysis, *IEEE Transactions on Automatic Control*, accepted conditionally (May 2022).

M. P. Chapman, J. Lacotte, A. Tamar, D. Lee, K. M. Smith, V. Cheng, J. F. Fisac, S. Jha, M. Pavone, C. J. Tomlin, A risk-sensitive finite-time reachability approach for safety of stochastic dynamic systems, in: Proceedings of the *American Control Conference*, IEEE, 2019, pp. 2958–2963.

References

M. P. Chapman and L. Lessard, Toward a scalable upper bound for a CVaR-LQ problem, *IEEE Control Systems Letters*, 2021.

M. P. Chapman, T. Risom, A. Aswani, E. Langer, R. Sears, and C. J. Tomlin, Modeling differentiation-state transitions linked to therapeutic escape in triple-negative breast cancer, *PLoS Computational Biology*, 2019.

M. P. Chapman and K. M. Smith, Classical risk-averse control for finite-horizon Borel models, *IEEE Control Systems Letter*s, 2021.

M. Chen and C. J. Tomlin, Hamilton–Jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management, *Annual Review of Control, Robotics, and Autonomous Systems* 1 (2018) 333–358.

S. P. Coraluppi and S. I. Marcus, Risk-sensitive and minimax control of discrete-time, finite-state Markov decision processes, *Automatica* 35 (2) (1999) 301–309.

L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*. New York, NY: Springer, 2006, vol. 21.

G. B. di Masi and L. Stettner, Risk-sensitive control of discrete-time Markov processes with infinite horizon, *SIAM Journal on Control and Optimization*, vol. 38, no. 1, pp. 61–78, 1999.

References

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

Flooding from storm turns Toronto streets into rivers, *City News*, August 7, 2018, https://toronto.citynews.ca/2018/08/07/flooding-storm-turns-toronto-streets-rivers/

V. Forejt, M. Kwiatkowska, G. Norman, D. Parker, Automated verification techniques for probabilistic systems, in: *International School on Formal Methods for the Design of Computer, Communication and Software Systems*, Springer, 2011, pp. 53–113.

J. García and F. Fernández, Safe exploration of state and action spaces in reinforcement learning, *Journal of Artificial Intelligence Research* 45 (2012) 515–564.

P. Geibel, Reinforcement learning with bounded risk, in: Proceedings of the *International Conference on Machine Learning*, 2001, pp. 162–169.

P. Geibel, F. Wysotzki, Risk-sensitive reinforcement learning applied to control under constraints, *Journal of Artificial Intelligence Research* 24 (2005) 81–108.

K. Glover and J. C. Doyle, State-space formulae for all stabilizing controllers that satisfy an H∞-norm bound and relations to risk sensitivity, *Systems & Control Letters*, vol. 11, no. 3, pp. 167–172, 1988.

References

W. B. Haskell and R. Jain, A convex analytic approach to risk-aware Markov decision processes, *SIAM Journal on Control and Optimization*, vol. 53, no. 3, pp. 1569–1598, 2015.

M. Heger, Consideration of risk in reinforcement learning, in: *Proceedings of the International Machine Learning Conference*, 1994, pp. 105–111.

R. A. Howard and J. E. Matheson, Risk-sensitive Markov decision processes, *Management Science*, vol. 18, no. 7, pp. 356–369, 1972.

C. Huang, J. Fan, W. Li, X. Chen, Q. Zhu, ReachNN: Reachability analysis of neural-network controlled systems, *ACM Transactions on Embedded Computing Systems* 18 (5s) (2019) 1–22.

W. Huang, W. B. Haskell, Stochastic approximation for risk-aware Markov decision processes, *IEEE Transactions on Automatic Control* 66 (3) (2021) 1314–1320.

R. Ivanov, T. J. Carpenter, J. Weimer, R. Alur, G. J. Pappas, I. Lee, Verifying the safety of autonomous systems with neural network controllers, *ACM Transactions on Embedded Computing Systems* 20 (1) (2020) 1–26.

D. H. Jacobson, Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games, *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, 1973.

References

U. Köse and A. Ruszczyński, Risk-averse learning by temporal difference methods with Markov risk measures, *Journal of Machine Learning Research* 22 (2021) 1–34.

L. Lindemann, G. J. Pappas, D. V. Dimarogonas, Reactive and risk-aware control for signal temporal logic, *IEEE Transactions on Automatic Control* (2021).

K. Margellos, J. Lygeros, Hamilton–Jacobi formulation for reach–avoid differential games, *IEEE Transactions on Automatic Control* 56 (8) (2011) 1849–1861.

C. W. Miller and I. Yang, Optimal control of Conditional Value-at-Risk in continuous time, *SIAM Journal on Control and Optimization*, vol. 55, no. 2, pp. 856–884, 2018.

I. M. Mitchell, A. M. Bayen, C. J. Tomlin, A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games, *IEEE Transactions on Automatic Control* 50 (7) (2005) 947–957.

T. M. Moldovan, P. Abbeel, Safe exploration in Markov decision processes, in: *Proceedings of the International Conference on Machine Learning*, 2012.

R. E. Mortensen and K. P. Haggerty, A stochastic computer model for heating and cooling loads, *IEEE Transactions on Power Systems*, vol. 3, no. 3, pp. 1213-1219, 1988.

References

Risk, in *Merriam-Webster.com Dictionary*, https://www.merriam-webster.com/dictionary/risk, accessed December 26, 2021 (2021).

Risk management–Guidelines, *International Organization for Standardization*, https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en, accessed July 11, 2021 (2018).

Risom, Langer, Chapman#, Rantala#, Fields, Boniface, Alvarez, Kendsersky, Pelz, Johnson-Camacho, Dobrolecki, Chin, Aswani, Wang, Califano, Lewis, Tomlin, Spellman, Adey, Gray, and Sears, Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer, *Nature Communications*, 2018. #equal

A. Ruszczyński, Risk-averse dynamic programming for Markov decision processes, *Mathematical Programming*, vol. 125, no. 2, pp. 235– 261, 2010.

A. Ruszczyński, Erratum to: Risk-averse dynamic programming for Markov decision processes, *Mathematical Programming* 145 (1) (2014) 601–604.

N. Saldi, T. Başar, M. Raginsky, Approximate Markov-Nash equilibria for discrete-time risk-sensitive mean-field games, *Mathematics of Operations Research* 45 (4) (2020) 1596–1620.

G. Schildbach, L. Fagiano, C. Frei, M. Morari, The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations, *Automatica* 50 (12) (2014) 3009–3018.

References

A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial and Applied Mathematics, Mathematical Programming Society, 2009.

S. Singh, Y. Chow, A. Majumdar, M. Pavone, A framework for time-consistent, risk-sensitive model predictive control: Theory and algorithms, *IEEE Transactions on Automatic Control* 64 (7) (2018) 2905–2912.

K. Smith & M. P. Chapman, On Exponential Utility and Conditional Value-at-Risk as risk-averse performance criteria, under review for *IEEE Transactions on Control Systems Technology*.

S. Summers and J. Lygeros, Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem, *Automatica*, vol. 46, no. 12, pp. 1951–1961, 2010.

Y. Wang and M. P. Chapman, Risk-Averse autonomous systems: A brief history and recent developments from the perspective of optimal control, *Journal of Artificial Intelligence*, in press, 2022.

C. Wei, M. Fauss, and M. P. Chapman, CVaR-based safety analysis for the infinite time setting, *American Control Conference*, in press, 2022.

P. Whittle, Risk-Sensitive Linear/Quadratic/Gaussian Control, *Advances in Applied Probability*, vol. 13, no. 4, pp. 764–777, 1981.

References

P. Whittle, *Risk-sensitive Optimal Control*, Chichester: Wiley, 1990.

I. Yang, A dynamic game approach to distributionally robust safety specifications for stochastic systems, *Automatica*, vol. 94, pp. 94-101, 2018.

I. Yang, Wasserstein distributionally robust stochastic control: A data-driven approach, *IEEE Transactions on Automatic Control* 66 (8) (2020) pp. 3863-3870.

Systems & Risk