Risk Quantification and Control Challenges and Opportunities

Andrzej Ruszczyński

RUTGERS

Systems Analysis 2015

IIASA

Methodological Challenges in Modern System Analysis

Complexity

- Large dimension
- Distributed information and control
- Multiple agents

Dynamics

- Random evolution
- Learning
- Multiple time scales

Risk

- How to quantify risk?
- Local *versus* systemic risk
- Dynamics of risk

New concepts, problem formulations, and methods are needed

Uncertain Outcomes and Risk

Why Probabilistic Models?

- Wealth of results of probability theory
- Connection to real data via statistics
- Universal language (engineering, economics, medicine, ...)

- Probability space $(\Omega, \mathcal{F}, P)$
- Decision space $\mathcal{X}$
- Random outcome (*e.g.*, cost) $Z_x(\omega)$, $\quad Z : \mathcal{X} \times \Omega \to \mathbb{R}$

Expected Value Model

$$\min_x \ \mathbb{E}[Z_x] = \int_\Omega Z_x(\omega)\, P(d\omega)$$

It optimizes the outcome on average (Law of Large Numbers?)

What is Risk?

Existence of unlikely and undesirable outcomes - high $Z_x(\omega)$ for some ω

Classical Utility Models

Expected Utility Models [von Neumann and Morgenstern, 1944]

$$\min_{x \in X} \; \mathbb{E}\left[u(Z_x)\right] \qquad \left(= \int_{\Omega} u\big(Z_x(\omega)\big)\, dP(\omega) \right)$$

$u : \mathbb{R} \to \mathbb{R}$ is a nondecreasing disutility function

Rank Dependent Utility (Distortion) Models [Quiggin, 1982; Yaari, 1987]

$$\min_{x \in X} \int_0^1 F_{Z_x}^{-1}(p)\, dw(p) \qquad F_{Z_x}^{-1}(\cdot) \text{ - quantile function}$$

$w : [0,1] \to \mathbb{R}$ is a nondecreasing rank dependent utility function

Existence of utility functions is derived from systems of axioms, but in practice they are difficult to elicit

Chance Constraints (Miller-Wagner, 1965; Prékopa, 1970)

Z_x - random outcome (*e.g.*, cost) or a vector of outcomes, y - fixed target

The Model

$$\begin{aligned} &\min f(x) \\ &\text{subject to } \mathbb{P}\left[Z_x \leq y\right] \geq 1-\alpha \\ &x \in X \end{aligned}$$

- Individual constraints: separately for each component of Z_x
- Joint constraints: the entire vector Z_x is below target with high probability

Application: Service level in a supply network, with random demands and delivery times (with M. Lejeune, 2007)

Joint constraints model systemic risk but are very hard

Chance Constraints (Miller-Wagner, 1965; Prékopa, 1970)

Z_x - random outcome (*e.g.*, cost) or a vector of outcomes, y - fixed target

The Model

$$\begin{aligned} \min\ & f(x) \\ \text{subject to}\ & \mathbb{P}\big[\,(Z_x)_j \le y_j\,\big] \ge 1-\alpha,\ j=1,\dots,m \\ & x \in X \end{aligned}$$

- Individual constraints: separately for each component of Z_x
- Joint constraints: the entire vector Z_x is below target with high probability

Application: Service level in a supply network, with random demands and delivery times (with M. Lejeune, 2007)

Joint constraints model systemic risk but are very hard

Chance Constraints (Miller-Wagner, 1965; Prékopa, 1970)

Z_x - random outcome (*e.g.*, cost) or a vector of outcomes, y - fixed target

The Model

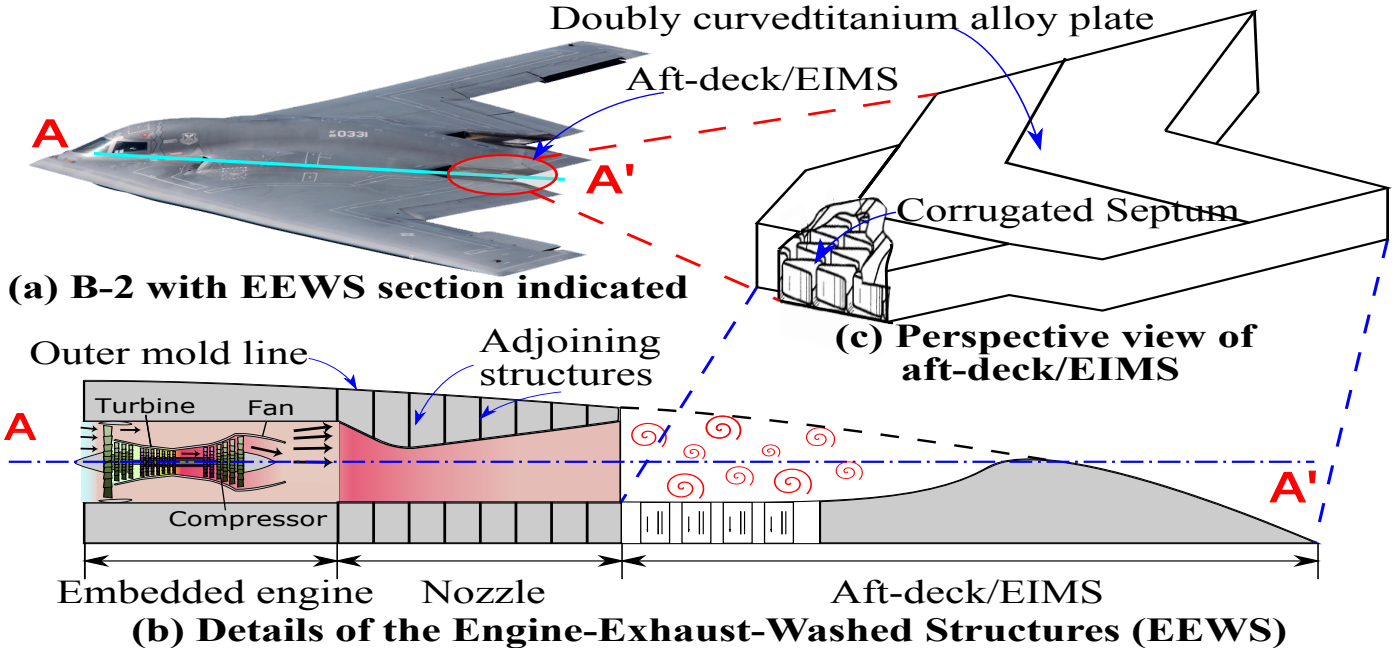
$$\begin{aligned} \min\ & f(x) \\ \text{subject to}\ & \mathbb{P}\big[\,(Z_x)_j \le y_j,\ j = 1, \ldots, m\,\big] \ge 1 - \alpha \\ & x \in X \end{aligned}$$

- Individual constraints: separately for each component of Z_x
- Joint constraints: the entire vector Z_x is below target with high probability

Application: Service level in a supply network, with random demands and delivery times (with M. Lejeune, 2007)

Joint constraints model systemic risk but are very hard

Example: Topology Optimization (with a Virginia Tech team, 2015–)



Optimize the shape of the exhaust structure, subject to structural equations, fluid and solid thermal equations (system of PDEs). The fluid thermal boundary conditions and the material properties are random.

- Individual Risk Constraints: At each point x the probability that the stress exceeds the critical stress (both are random) is below α
- Systemic Risk Constraint: The probability that anywhere in the structure the stress exceeds the critical stress is below α

Stochastic Dominance Constraints (with Darinka Dentcheva, 2003–)

Z_x - random outcome (*e.g.*, cost, or a vector of outcomes)
Y - benchmark random outcome, e.g. $Y(\omega) = Z_{\bar{x}}(\omega)$ for some $\bar{x} \in X$

$$\begin{aligned} &\min\ f(x) \\ &\text{subject to } Z_x \preceq_{\mathrm{SD}} Y \qquad \text{(stochastic dominance constraint)} \\ &\qquad\qquad x \in X \end{aligned}$$

First Order: $\mathbb{P}\big[Z_x \geq \eta\big] \leq \mathbb{P}\big[Y \geq \eta\big]$ for all $\eta_{\min} \leq \eta \leq \eta_{\max}$

Second Order: $\mathbb{E}\big[(Z_x - \eta)_+\big] \leq \mathbb{E}\big[(Y - \eta)_+\big]$ for all $\eta_{\min} \leq \eta \leq \eta_{\max}$

Stochastic dominance constraints consider the entire distribution of the outcome Z_x, rather than the probability of one event.

- Optimality conditions and solutions methods have been developed
- Current research: multivariate (vector) constraints

Mean–Risk Models

Two Objectives (for scalar Z_x)

- Minimize the expected outcome, the mean $\mathbb{E}[Z_x]$
- Minimize a scalar measure of uncertainty of Z_x, the risk $r[Z_x]$

$$r[Z] = \mathbb{V}\mathrm{ar}[Z] \qquad \text{(Markowitz' model)}$$

$$\sigma_p^+[Z] = \left(\mathbb{E}[(Z - \mathbb{E}\,Z)_+^p]\right)^{1/p} \qquad \text{(semideviation)}$$

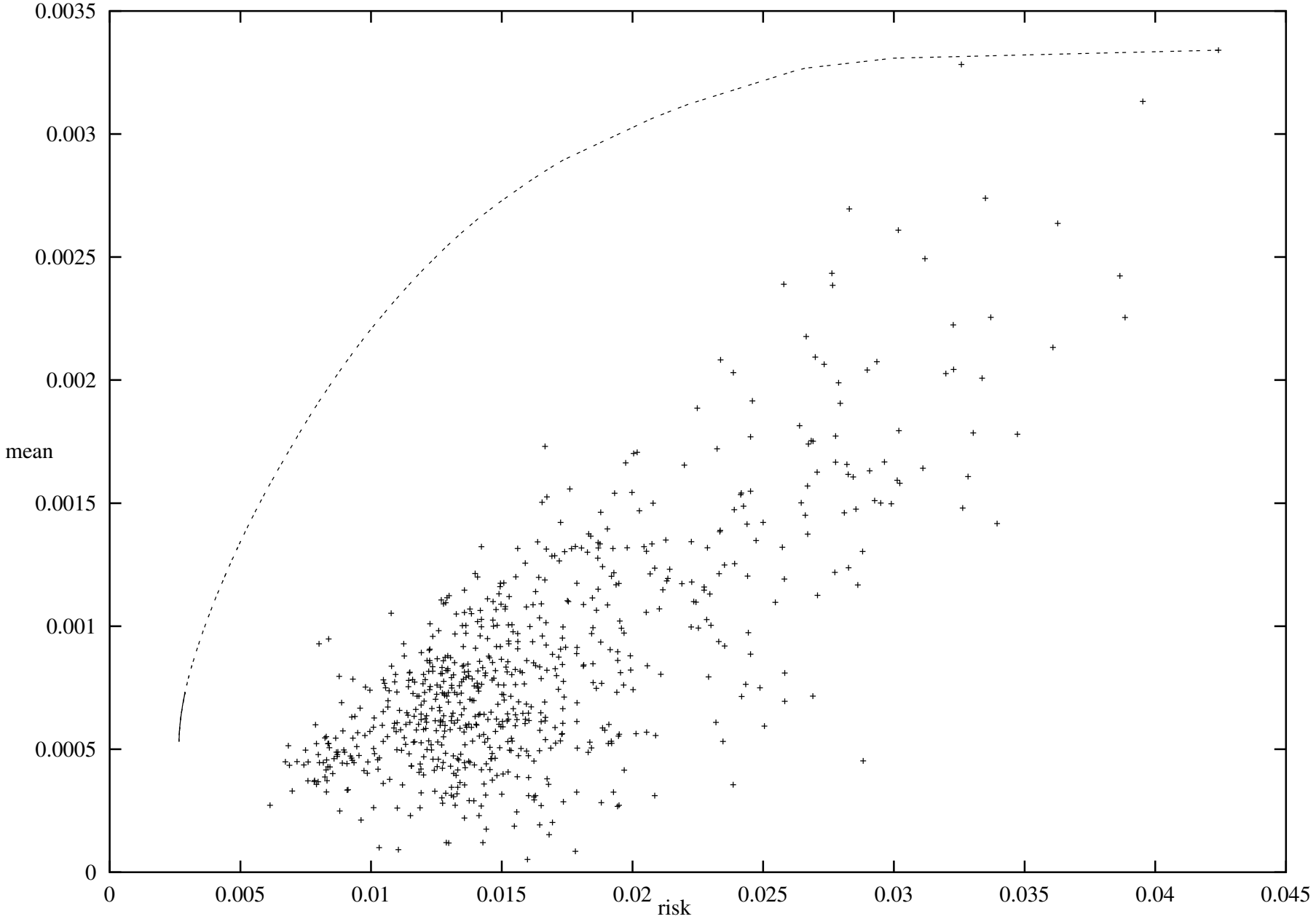
$$\delta_\alpha^+[Z] = \min_\eta \mathbb{E}\left[\max\left(\eta - Z, \frac{\alpha}{1-\alpha}(Z-\eta)\right)\right] \qquad \text{(deviation from quantile)}$$

Mean–Risk Optimization

$$\min_{x \in X} \ \rho[Z_x] = \mathbb{E}[Z_x] + \kappa\, r[Z_x], \qquad 0 \le \kappa \le \kappa_{\max}$$

$r[Z_x]$ is nonlinear w.r.t. probability and possibly nonconvex in x

Mean–Semideviation Model (719 stocks) (with R. Vanderbei, 2002)



Coherent Risk Measures

Space of uncertain outcomes $\mathcal{Z}$ (example: $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$)

A functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is a coherent risk measure if it satisfies the following axioms

- Convexity: $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$ $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}$
- Monotonicity: If $Z \leq W$ then $\rho(Z) \leq \rho(W)$, $\quad \forall\ Z, W \in \mathcal{Z}$
- Translation Equivariance: $\rho(Z + a) = \rho(Z) + a$, $\quad \forall\ Z \in \mathcal{Z},\ a \in \mathbb{R}$
- Positive Homogeneity: $\rho(\gamma Z) = \gamma\rho(Z)$, $\quad \forall\ Z \in \mathcal{Z},\ \gamma \geq 0$

Kijima-Ohnishi (1993) – no monotonicity
Artzner-Delbaen-Eber-Heath (1999–) - space $\mathcal{L}_\infty$
Ruszczyński-Shapiro (2005) – spaces $\mathcal{L}_p$, ...

Coherent Risk Measures

Space of uncertain outcomes $\mathcal{Z}$ (example: $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$)

A functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is a coherent risk measure if it satisfies the following axioms

- Convexity: $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$ $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}$
- Monotonicity: If $Z \leq W$ then $\rho(Z) \leq \rho(W)$, $\quad \forall\ Z, W \in \mathcal{Z}$
- Translation Equivariance: $\rho(Z + a) = \rho(Z) + a$, $\quad \forall\ Z \in \mathcal{Z},\ a \in \mathbb{R}$
- Positive Homogeneity: $\rho(\gamma Z) = \gamma\rho(Z)$, $\quad \forall\ Z \in \mathcal{Z},\ \gamma \geq 0$

Kijima-Ohnishi (1993) – no monotonicity
Artzner-Delbaen-Eber-Heath (1999–) - space $\mathcal{L}_\infty$
Ruszczyński-Shapiro (2005) – spaces $\mathcal{L}_p$, ...

Coherent Risk Measures

Space of uncertain outcomes $\mathcal{Z}$ (example: $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$)

A functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is a coherent risk measure if it satisfies the following axioms

- Convexity: $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$
  $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}$
- Monotonicity: If $Z \leq W$ then $\rho(Z) \leq \rho(W), \quad \forall\ Z, W \in \mathcal{Z}$
- Translation Equivariance: $\rho(Z + a) = \rho(Z) + a, \quad \forall\ Z \in \mathcal{Z},\ a \in \mathbb{R}$
- Positive Homogeneity: $\rho(\gamma Z) = \gamma\rho(Z), \quad \forall\ Z \in \mathcal{Z},\ \gamma \geq 0$

Kijima-Ohnishi (1993) – no monotonicity
Artzner-Delbaen-Eber-Heath (1999–) - space $\mathcal{L}_\infty$
Ruszczyński-Shapiro (2005) – spaces $\mathcal{L}_p$, ...

Coherent Risk Measures

Space of uncertain outcomes $\mathcal{Z}$ (example: $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$)

A functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is a coherent risk measure if it satisfies the following axioms

- Convexity: $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$
  $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}$
- Monotonicity: If $Z \leq W$ then $\rho(Z) \leq \rho(W), \quad \forall\ Z, W \in \mathcal{Z}$
- Translation Equivariance: $\rho(Z + a) = \rho(Z) + a, \quad \forall\ Z \in \mathcal{Z},\ a \in \mathbb{R}$
- Positive Homogeneity: $\rho(\gamma Z) = \gamma\rho(Z), \quad \forall\ Z \in \mathcal{Z},\ \gamma \geq 0$

Kijima-Ohnishi (1993) – no monotonicity
Artzner-Delbaen-Eber-Heath (1999–) - space $\mathcal{L}_\infty$
Ruszczyński-Shapiro (2005) – spaces $\mathcal{L}_p$, ...

Coherent Risk Measures

Space of uncertain outcomes $\mathcal{Z}$ (example: $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$)

A functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is a coherent risk measure if it satisfies the following axioms

- Convexity: $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$
  $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}$
- Monotonicity: If $Z \leq W$ then $\rho(Z) \leq \rho(W), \quad \forall\ Z, W \in \mathcal{Z}$
- Translation Equivariance: $\rho(Z + a) = \rho(Z) + a, \quad \forall\ Z \in \mathcal{Z},\ a \in \mathbb{R}$
- Positive Homogeneity: $\rho(\gamma Z) = \gamma\rho(Z), \quad \forall\ Z \in \mathcal{Z},\ \gamma \geq 0$

Kijima-Ohnishi (1993) – no monotonicity
Artzner-Delbaen-Eber-Heath (1999–) - space $\mathcal{L}_\infty$
Ruszczyński-Shapiro (2005) – spaces $\mathcal{L}_p$, ...

Coherent Risk Measures

Space of uncertain outcomes $\mathcal{Z}$ (example: $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$)

A functional $\rho : \mathcal{Z} \to \overline{\mathbb{R}}$ is a coherent risk measure if it satisfies the following axioms

- Convexity: $\rho(\lambda Z + (1-\lambda)W) \leq \lambda\rho(Z) + (1-\lambda)\rho(W)$
  $\forall\ \lambda \in (0,1),\ Z, W \in \mathcal{Z}$
- Monotonicity: If $Z \leq W$ then $\rho(Z) \leq \rho(W), \quad \forall\ Z, W \in \mathcal{Z}$
- Translation Equivariance: $\rho(Z + a) = \rho(Z) + a, \quad \forall\ Z \in \mathcal{Z},\ a \in \mathbb{R}$
- Positive Homogeneity: $\rho(\gamma Z) = \gamma\rho(Z), \quad \forall\ Z \in \mathcal{Z},\ \gamma \geq 0$

Kijima-Ohnishi (1993) – no monotonicity
Artzner-Delbaen-Eber-Heath (1999–) - space $\mathcal{L}_\infty$
Ruszczyński-Shapiro (2005) – spaces $\mathcal{L}_p$, ...

Examples of Risk Measures

Mean–Semideviation (coherent): [Ogryczak–Ruszczyński, 1997]

$$\rho(Z) \triangleq \mathbb{E}[Z] + \kappa \left(\mathbb{E}\Big[\big((Z - \mathbb{E}[Z])_+ \big)^p \Big] \right)^{1/p}, \qquad \kappa \in [0,1], \quad p \geq 1$$

Average Value at Risk (coherent): [Rockafellar–Uryasev, 2001]

$$\begin{aligned} \mathrm{AV@R}^+_\alpha(Z) &\triangleq \frac{1}{\alpha} \int_0^\alpha F_Z^{-1}(1-\beta)\, \mathrm{d}\beta \\ &= \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{\alpha} \mathbb{E}\big[(Z-\eta)_+ \big] \right\} \qquad \alpha \in (0,1] \end{aligned}$$

Entropic Risk Measure (convex, not coherent):

$$\rho(Z) \triangleq \frac{1}{\gamma} \ln \Big(\mathbb{E}\big[e^{\gamma Z} \big] \Big), \qquad \gamma > 0$$

Optimization of Risk Measures (with Alexander Shapiro, 2006–)

"Minimize" over $x \in X$ a random outcome $Z_x(\omega) = f(x, \omega)$, $\omega \in \Omega$

Composite Optimization Problem

$$\min_{x \in X} \rho(Z_x) \qquad \text{(P)}$$

A convex set $\mathcal{A}$ of probability measures exists, such that

$$\rho(Z_x) = \max_{\mu \in \mathcal{A}} \mathbb{E}_\mu[Z_x]$$

If $x \mapsto Z_x(\omega)$ is convex and $\hat{x} \in X$ is an optimal solution of (P), then a probability measure $\hat{\mu} \in \partial\rho(Z_{\hat{x}}) \subseteq \mathcal{A}$ exists such that

$$\min_{x \in X} \rho(Z_x) = \min_{x \in X} \mathbb{E}_{\hat{\mu}}[Z_x] = \min_{x \in X} \max_{\mu \in \mathcal{A}} \mathbb{E}_\mu[Z_x] = \max_{\mu \in \mathcal{A}} \inf_{x \in X} \mathbb{E}_\mu[Z_x]$$

Intriguing game against an opponent choosing the distribution $\mu \in \mathcal{A}$

How to Measure Risk of Sequences?

Probability space $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T \subset \mathcal{F}$
Adapted sequence of random variables (costs) $Z_1, Z_2, \ldots, Z_T$
Spaces: $\mathcal{Z}_t = \mathcal{L}_p(\Omega, \mathcal{F}_t, P)$, $p \in [1, \infty]$, and $\mathcal{Z}_{t,T} = \mathcal{Z}_t \times \cdots \times \mathcal{Z}_T$

Conditional Risk Measure

A mapping $\rho_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{Z}_t$ satisfying the monotonicity condition:

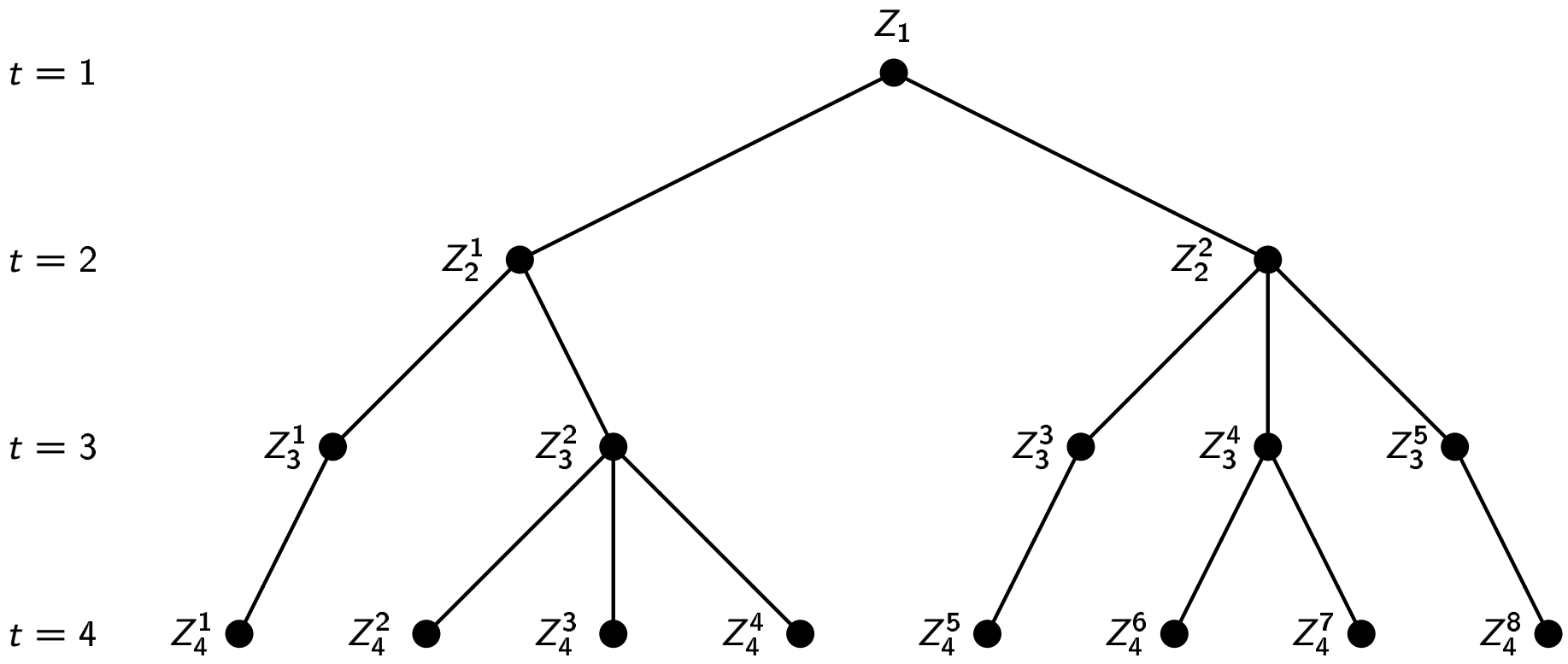
$$\rho_{t,T}(Z) \leq \rho_{t,T}(W) \text{ for all } Z, W \in \mathcal{Z}_{t,T} \text{ such that } Z \leq W$$

Dynamic Risk Measure

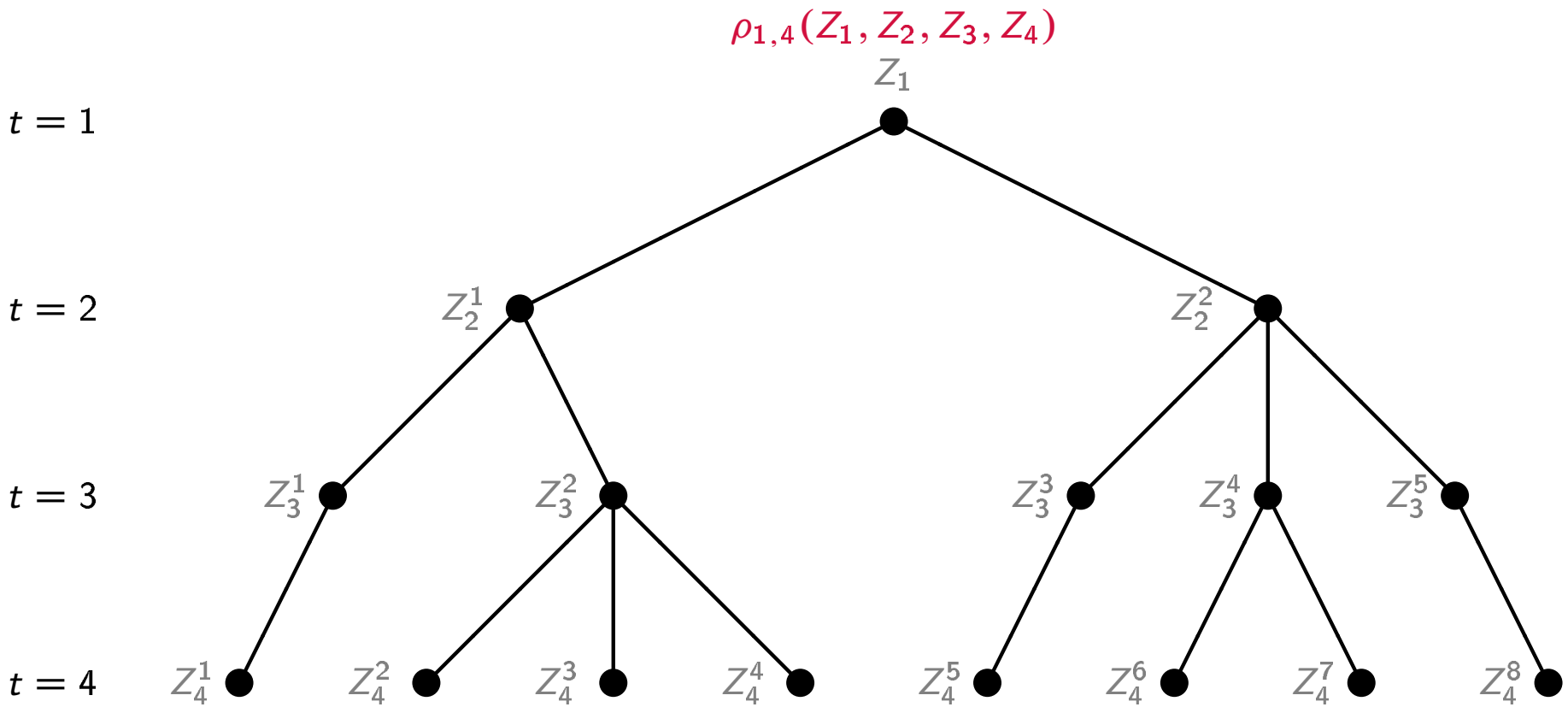
A sequence of conditional risk measures $\rho_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{Z}_t$, $t = 1, \ldots, T$

$$\begin{aligned}
\rho_{1,T}(Z_1, Z_2, Z_3, \ldots, Z_T) &\in \mathcal{Z}_1 = \mathbb{R} \\
\rho_{2,T}(Z_2, Z_3, \ldots, Z_T) &\in \mathcal{Z}_2 \\
\rho_{3,T}(Z_3, \ldots, Z_T) &\in \mathcal{Z}_3 \\
&\vdots
\end{aligned}$$

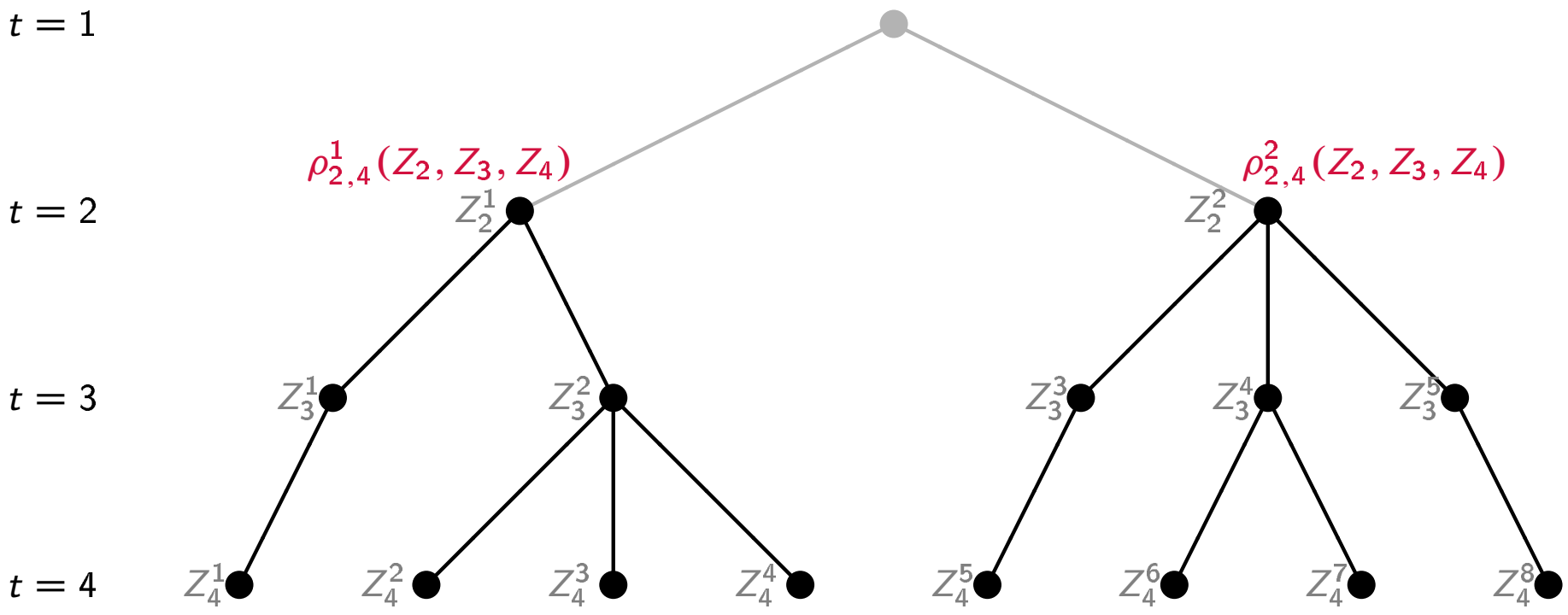
Evaluating Risk on a Scenario Tree



Evaluating Risk on a Scenario Tree



Evaluating Risk on a Scenario Tree



Time Consistency of Dynamic Risk Measures

A dynamic risk measure $\{\rho_{t,T}\}_{t=1}^{T}$ is time-consistent if for all $\tau < \theta$

$$Z_k = W_k, \;\; k = \tau, \ldots, \theta - 1 \quad \text{and} \quad \rho_{\theta,T}(Z_\theta, \ldots, Z_T) \leq \rho_{\theta,T}(W_\theta, \ldots, W_T)$$

imply that $\quad \rho_{\tau,T}(Z_\tau, \ldots, Z_T) \leq \rho_{\tau,T}(W_\tau, \ldots, W_T)$

Define $\quad \rho_t(Z_{t+1}) = \rho_{t,T}(0, Z_{t+1}, 0, \ldots, 0)$

Nested Decomposition Theorem

Suppose a dynamic risk measure $\{\rho_{t,T}\}_{t=1}^{T}$ is time-consistent

$$\rho_{t,T}(Z_t, Z_{t+1}, \ldots, Z_T) = Z_t + \rho_{t,T}(0, Z_{t+1}, \ldots, Z_T) \quad \text{and} \quad \rho_{t,T}(0, \ldots, 0) =$$

Then for all t we have the representation

$$\rho_{t,T}(Z_t, \ldots, Z_T) = Z_t + \rho_t\Big(Z_{t+1} + \rho_{t+1}\Big(Z_{t+2} + \cdots + \rho_{T-1}(Z_T)\Big)\cdots\Big)$$

Using Dynamic Risk Measures for Markov Decision Processes

- Controlled Markov process x_t, $t = 1, \ldots, T$, with kernels $Q_t(x_t, u_t)$
- Policy $\Pi = \{\pi_1, \pi_2, \ldots, \pi_T\}$ with $u_t = \pi_t(x_t)$ implies measure P^Π
- Cost sequence $Z_t^\Pi = c(x_t, \pi_t(x_t))$ (bounded), $t = 1, \ldots, T$,
- Dynamic time-consistent risk measure:

$$J_T(\Pi) = Z_1^\Pi + \rho_1^\Pi\big(Z_2^\Pi + \cdots + \rho_{T-1}^\Pi(Z_T^\Pi)\cdots\big)\Big)$$

- Risk-averse optimal control problem: $\min_\Pi J_T(\Pi)$

Difficulties

- Probability measure P^Π, processes x_t and Z_t^Π depend on policy Π
- The risk measures $\rho_t^\Pi(\cdot)$ depend on Π and may depend on history; no Markov policies

Idea

We only need to measure risk of random sequences that may occur

Risk-Averse Dynamic Programming

Foundation: (with Jingnan Fan)

- New refined concept of time-consistency
- Markov risk measures

Dynamic Programming Equations

$$v_{T+1}(x) = c_{T+1}(x), \quad x \in \mathcal{X}$$

$$v_t(x) = \min_{u \in U(x)} \Big\{ c_t(x,u) + \sigma_t\big(x, Q_t(x,u), v_{t+1}\big) \Big\}, \quad x \in \mathcal{X}, \quad t = T, \ldots, 1.$$

Optimal Markov policy $\hat{\Pi} = \{\hat{\pi}_1, \ldots, \hat{\pi}_T\}$ - the minimizers above

Further Advances:

- Infinite horizon models (discounted or absorbing)
- Solution methods
- Special cases, applications (with Özlem Çavuş)

Risk-Averse Dynamic Programming

Foundation: (with Jingnan Fan)

- New refined concept of time-consistency
- Markov risk measures

Dynamic Programming Equations

$$v_{T+1}(x) = c_{T+1}(x), \quad x \in \mathcal{X}$$

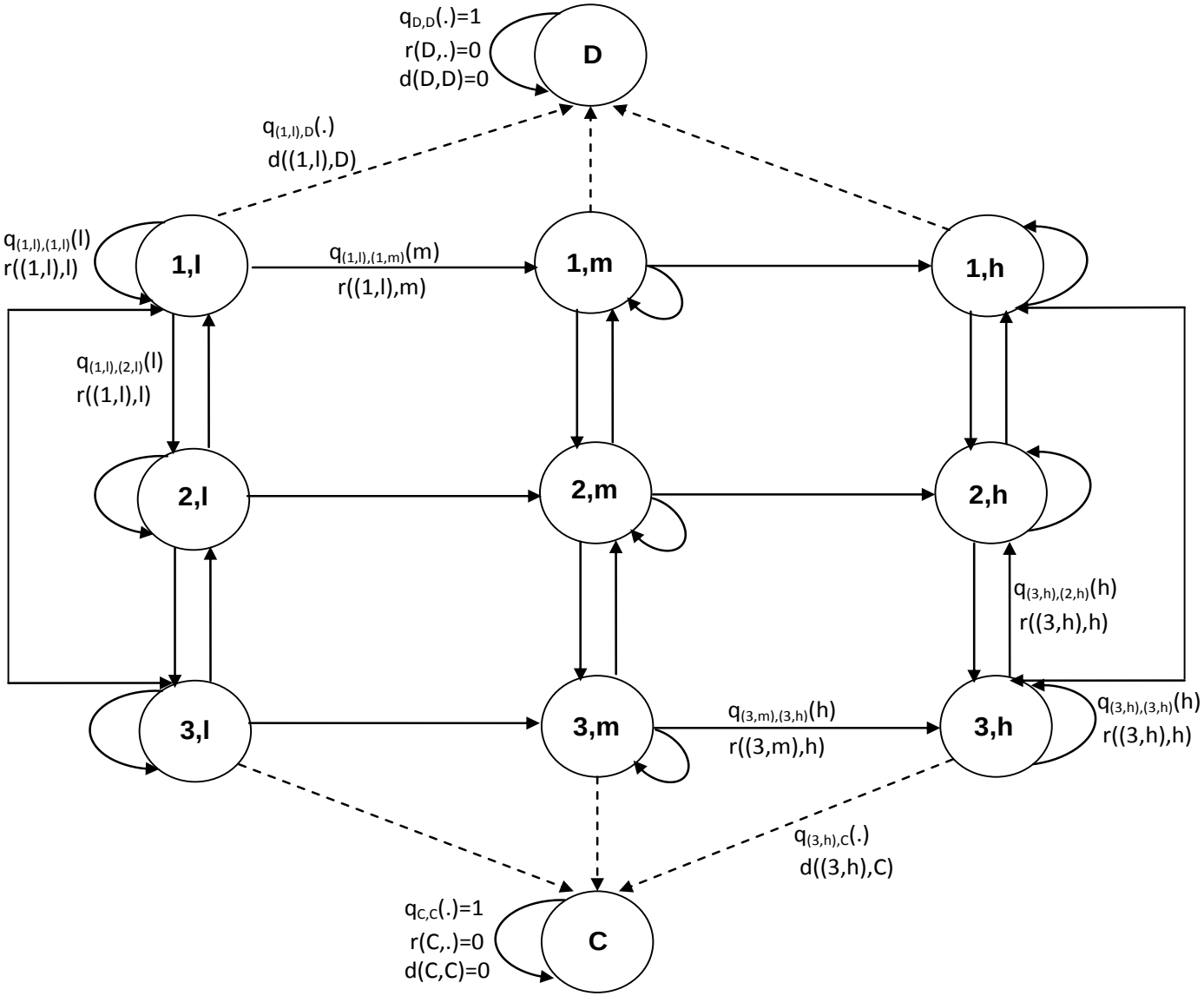
$$v_t(x) = \min_{u \in U(x)} \left\{ c_t(x,u) + \max_{\mu \in \mathcal{A}_t(x, Q_t(x,u))} \mathbb{E}_\mu \big[v_{t+1} \big] \right\}, \quad x \in \mathcal{X}, \quad t = T, \ldots,$$

Optimal Markov policy $\hat{\Pi} = \{\hat{\pi}_1, \ldots, \hat{\pi}_T\}$ - the minimizers above

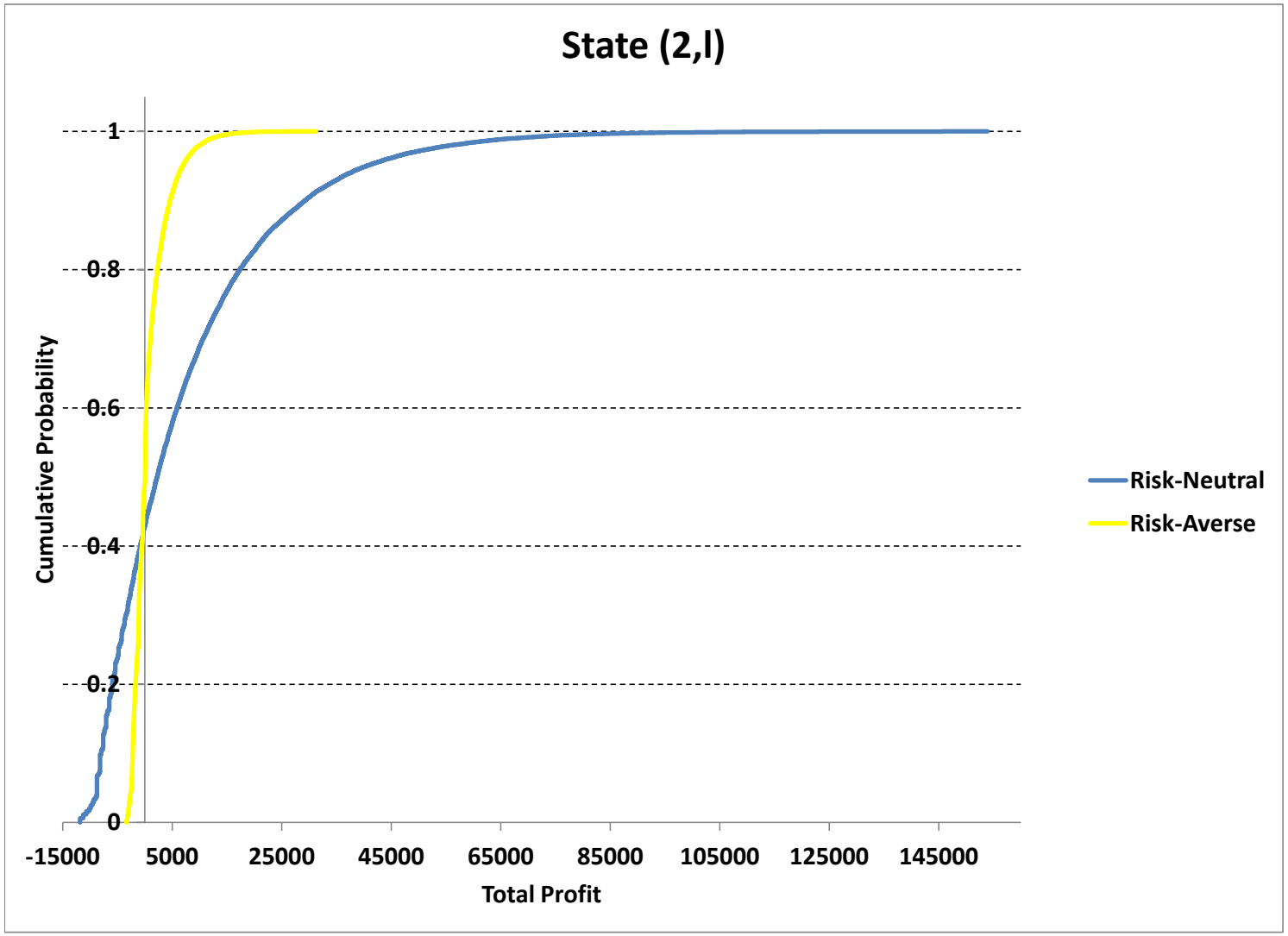
Further Advances:

- Infinite horizon models (discounted or absorbing)
- Solution methods
- Special cases, applications (with Özlem Çavuş)

Credit Card Example (with Özlem Çavuş)



Profit Distribution (with Özlem Çavuş)



Clinical Trials for Cytotoxic Agents

Main Challenges:

- Severe toxic response (e.g. death) is possible, even likely
- Trial subjects are patients, with a dire need for the potential treatment

Treatment *vs.* Experimentation Dilemma

- Dose escalation at sub- and/or supra-therapeutic doses potentially incurs large costs to individual patients
- Offering a poor estimation of the safe and effective (i.e. optimal) dose incurs systemic and future costs

Medical ethics naturally motivates risk-averse optimization

Risk-Averse Clinical Trials (with D. Dentcheva and Curtis McGinity)

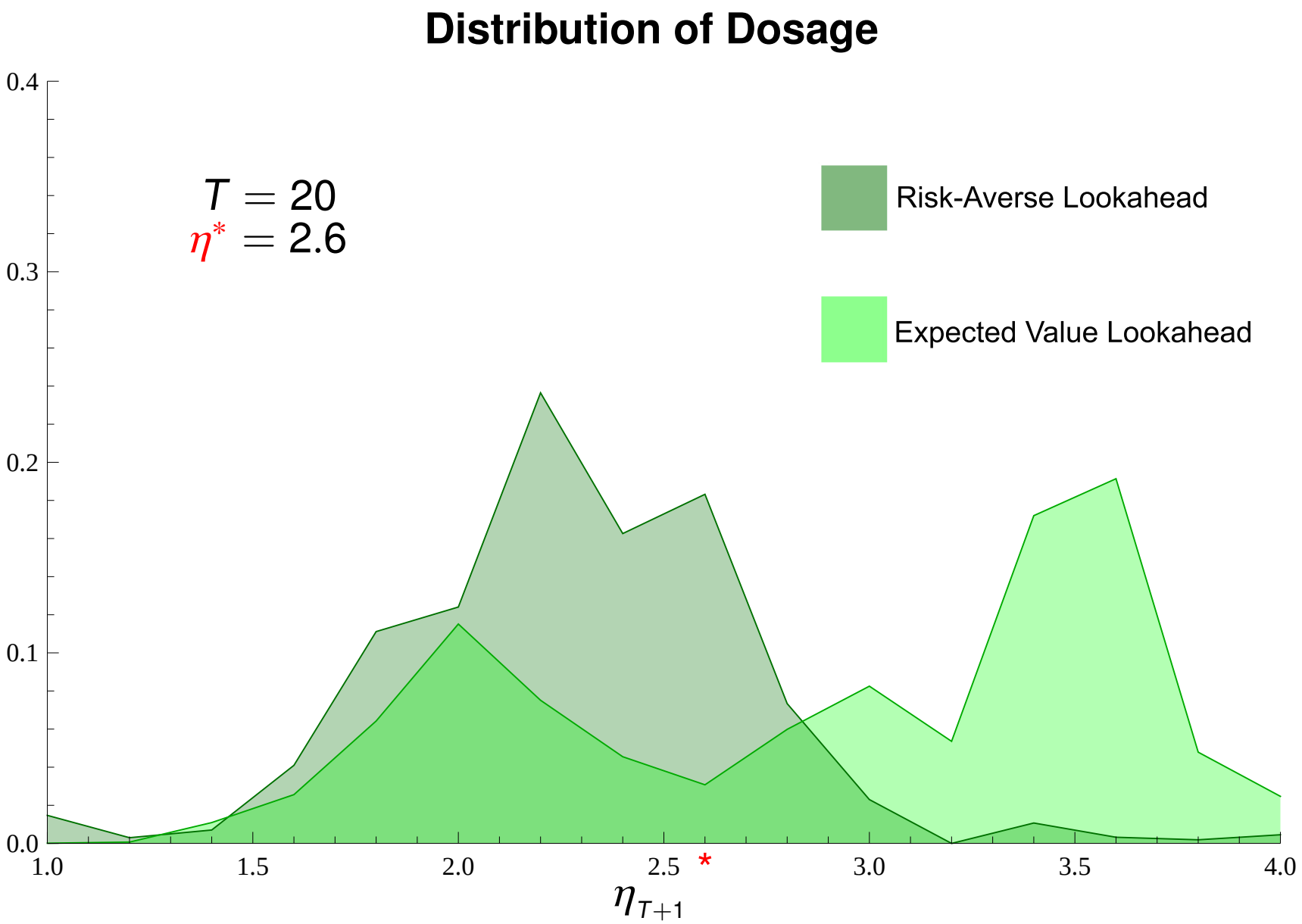
The Model

- Probability of toxic response depends on an unknown parameter (optimal dose)
- "Belief state" - probability distribution of the parameter - is the current state of the system
- The state evolves according to Bayes operator, depending on the response of the patient

Risk-Averse Dynamic Optimization

- Two sources of risk: unknown optimal dose, and
  random response of the patient
- Special Markov risk measures
- Approximate dynamic programming
- Simulation studies

Simulation Results for Expected Value and Risk-Averse Policies



Opportunities

Multiple research directions

- Risk-averse equilibria (*e.g.*, in networks, energy systems, etc.)
- Risk aversion in healthcare
- Risk-averse supply networks
- ...

Interdisciplinary cooperation

- Business: Finance, supply chains, and business analytics
- Engineering: Energy production and distribution, reliability
- Computer Science and Statistics: Big data, learning, artificial intelligence