Control by Coin Flips: Randomized Control for Risk Aware Human Cooperative Autonomy

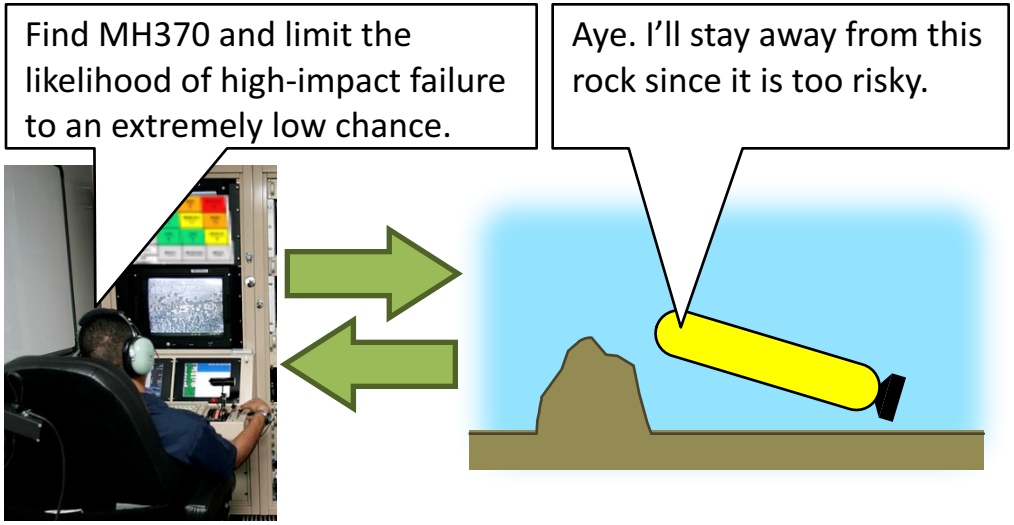
Masahiro Ono (JPL, Caltech)
Ufuk Topcu (UT Austin)
Behçet Açıkmeşe (UT Austin)
Missy Cummings (Duke University)

Funded by the ONR Science of Autonomy Program
Office of Naval Research Grant N00014-15-IP-00052

Copyright 2015. All rights reserved. U.S. Government sponsorship acknowledged.

Project Objective

Example: MH370 search and rescue



Objective: Realize *risk-aware, human-cooperative autonomy*

Risk-aware

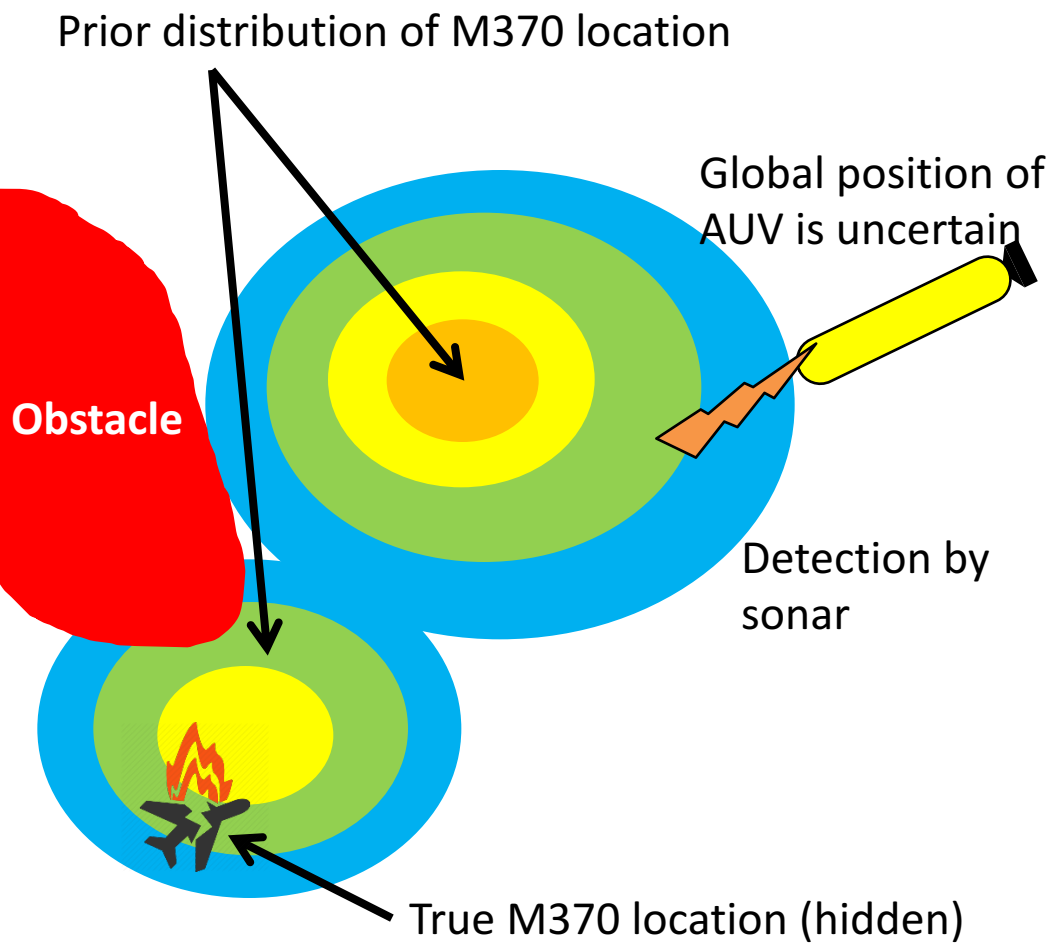
- Makes optimal decisions based on risk
- Manages risks at multiple levels of abstraction

Human-cooperative

- Shares the responsibility of managing risks
- Combines the strengths of each other synergistically

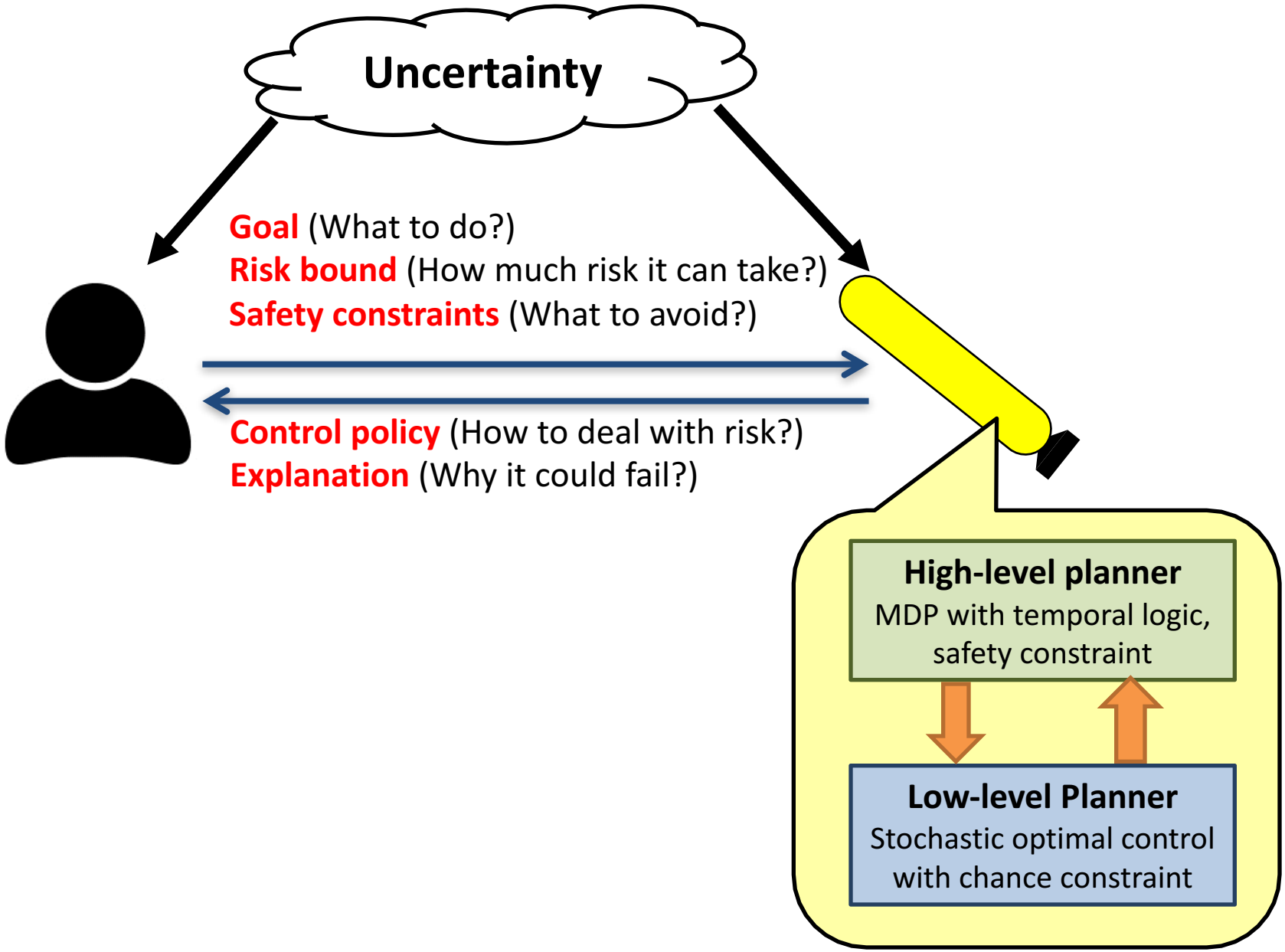
Sample Scenario: Search for M370

Problem: Maximize the chance of detecting the lost airplane while limiting the chance of the loss of vehicle

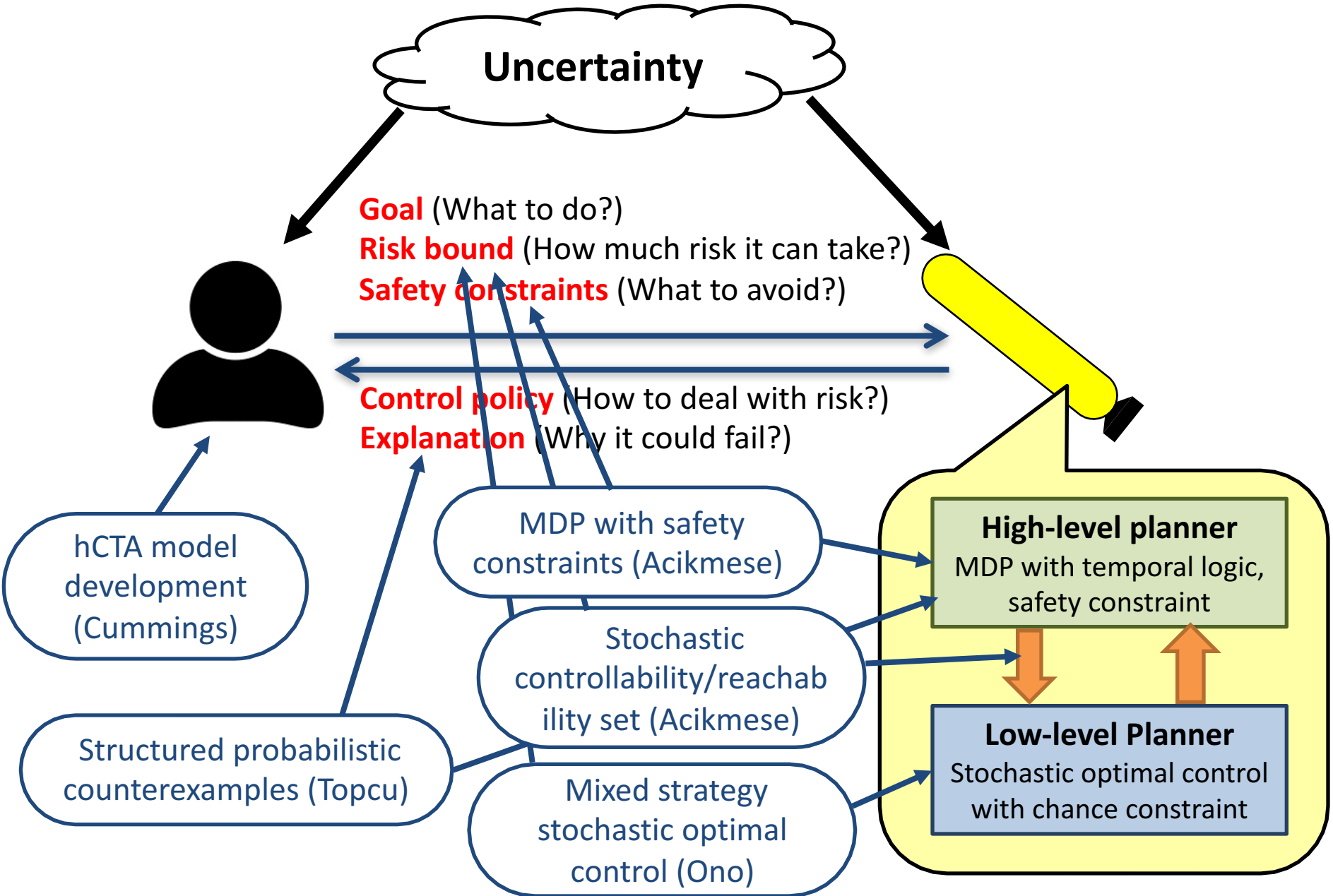


- Guesses of mine locations available
  - Represented by probability distribution
- AUV's global position has uncertainty
- Probability of detection is a decreasing function of the distance between AUV and M370
- Once detected, position of the airplane is specified *in relative to* AUV
  - Global position still uncertain
  - Prior distribution updated
- Safety constraints:
  - Limit the chance of vehicle loss due to collision with obstacles/other vessels
  - Avoid no-go zones

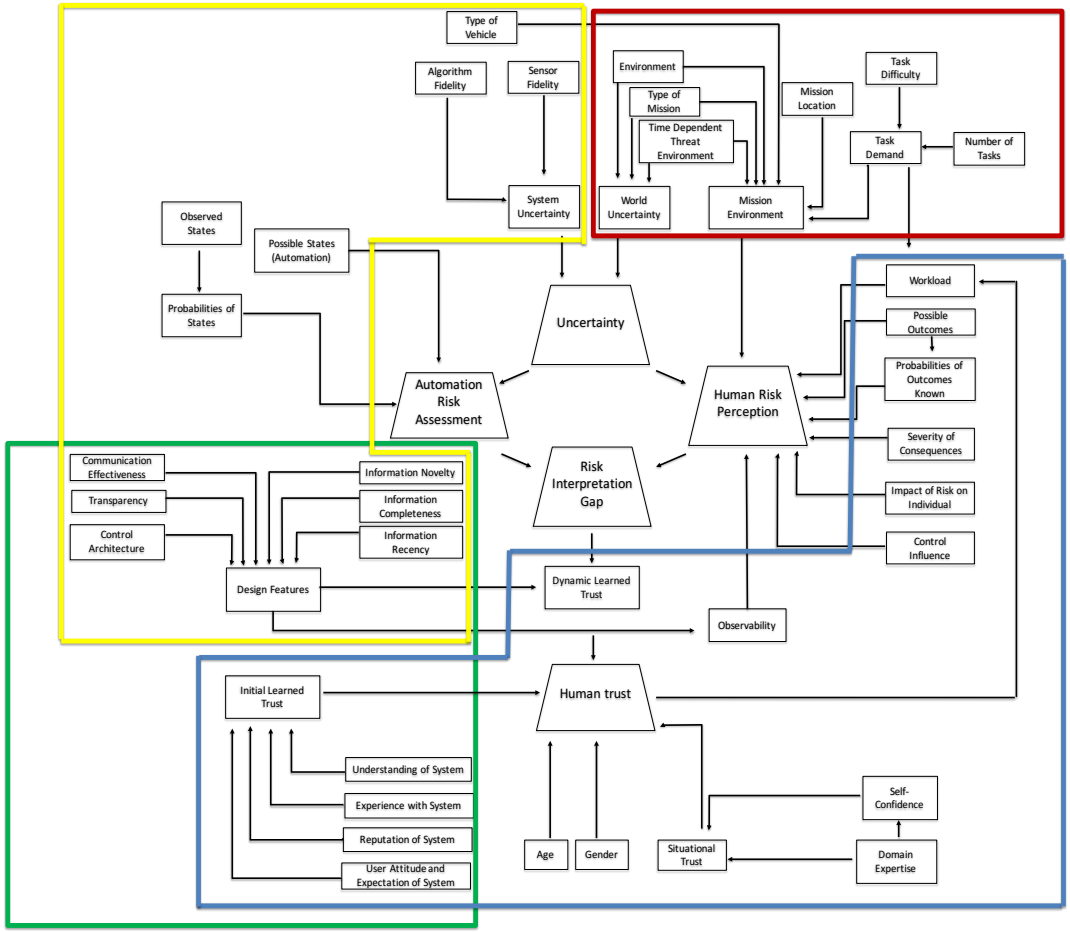
Approach


Uncertainty
Goal (What to do?)
Risk bound (How much risk it can take?)
Safety constraints (What to avoid?)
Control policy (How to deal with risk?)
Explanation (Why it could fail?)
High-level planner
MDP with temporal logic, safety constraint
Low-level Planner
Stochastic optimal control with chance constraint

FY15Q4-FY16Q3 Accomplishments



Task1: Human-machine interface for risk-aware planning



Influence diagram that identifies the key factors and variables associated with uncertainty, automated risk assessment, human risk perception, and human trust in automation

- Hybrid-Cognitive Task Analysis (hCTA) process helped to elucidate the factors crucial for design features for both human interfaces and algorithms
- Factors grouped together by color
  - Red: Environment
  - Yellow: Autonomy
  - Blue: Human
  - Green: Human-machine Interaction
- Human-in-the-loop experiments in the coming months to better understand:
  - The influence of the algorithm's time horizon on risk perception
  - Replanning frequency under risk
  - How decisions are influenced by risk budgeting

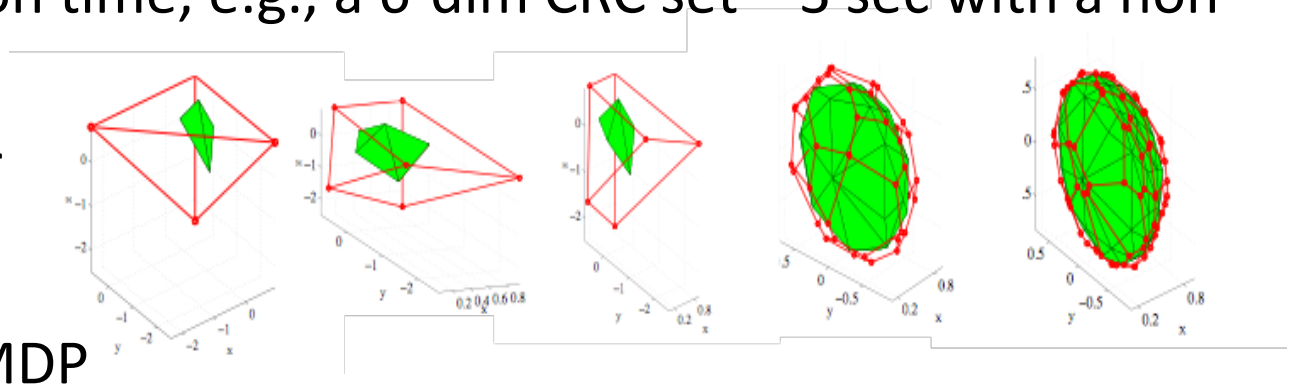
Task 2: Chance-constrained temporal logic planning and informative feedback to human operators

- Automated construction of structured counterexample in Markov decision processes. Counterexamples describe the core reasons for conflicts between the models, specifications and operator requirements and preferences.
- Controller synthesis for autonomous systems interacting with human operators
- Risk-averse control of Markov decision processes with ω-regular objectives

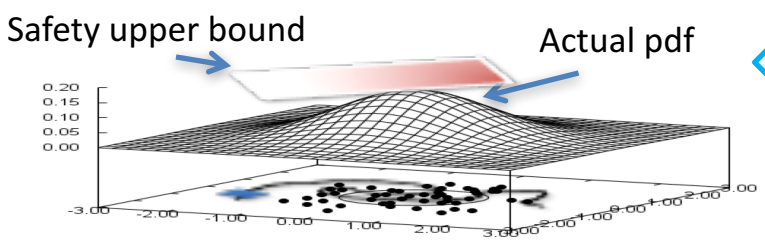
Image source: AFRL

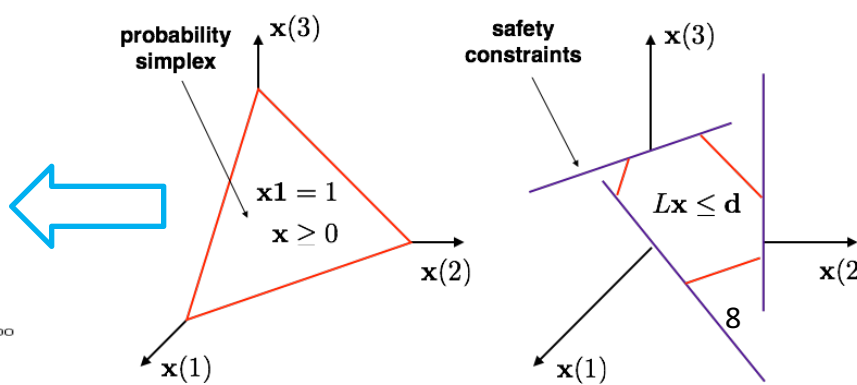
Task 3: Stochastic reachability/controllability, MDP with safety constraints

- CRC sets
  - Hausdorff distance based inner outer approximations via using methods from convex optimization and computational geometry
  - Fast computation time, e.g., a 6-dim CRC set ~3 sec with a non-optimized code
  - ECC 2016 paper



- Finite horizon SC-MDP
  - LP based solution methods utilizing duality theory and DP to compute feasible policies
  - ACC 2016 papers, IJCAI paper



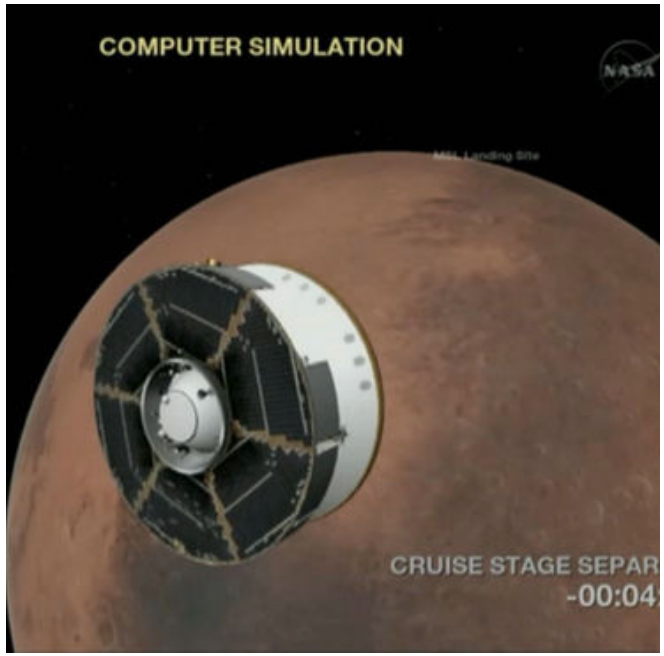


Task 4: Mixed Strategy Stochastic Optimal Control

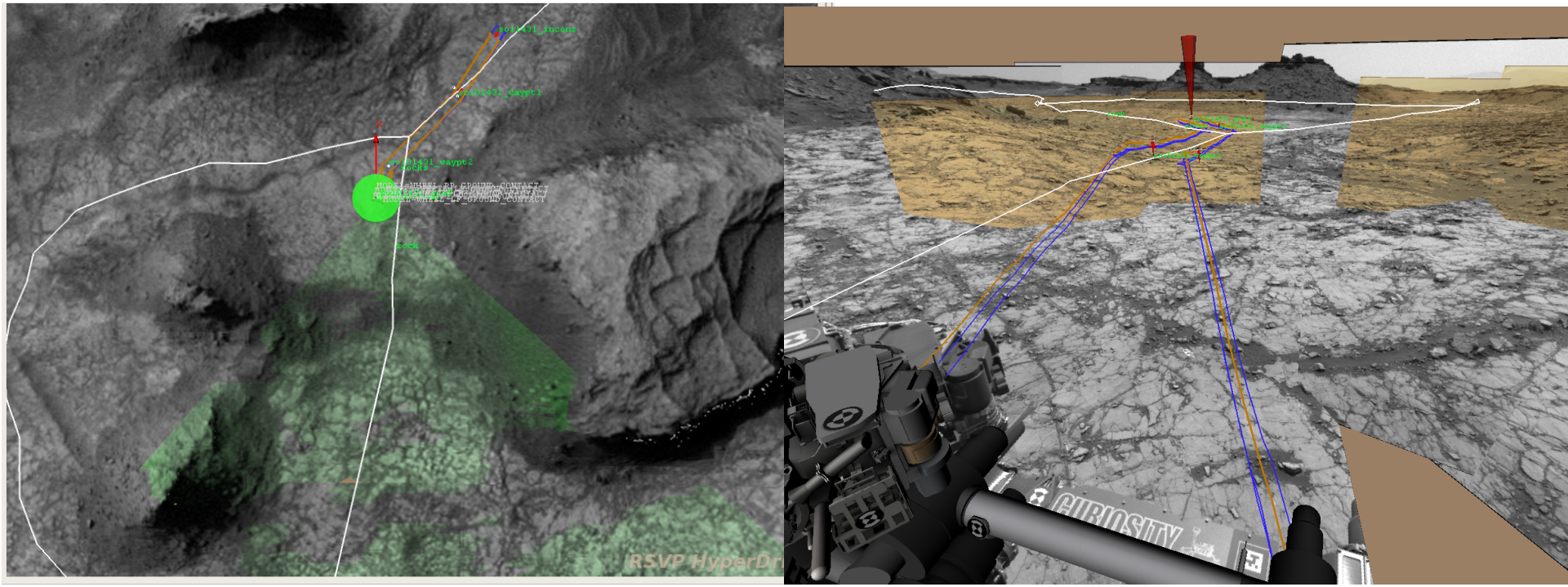
Chance Constrained Decision Making

Minimize cost, s.t. Prob. of failure $\leq V$

- Can change the level of conservatism by changing V
- Can avoid the excessive conservatism of robust decision making
  - Should we consider one-in-billion chance?
  - The nature does not play a game
- Example:
  - Mars Exploration Rovers: perform trajectory correction maneuvers if the landing failure probability is more than:
    - 9% for Spirit
    - 4% for Opportunity
  - Mars Pathfinder: the probability of the unsterilized launch vehicle upper stage impacting the surface of Mars must be less than 0.0001



Chance-constrained Decision Making for Mars Rover



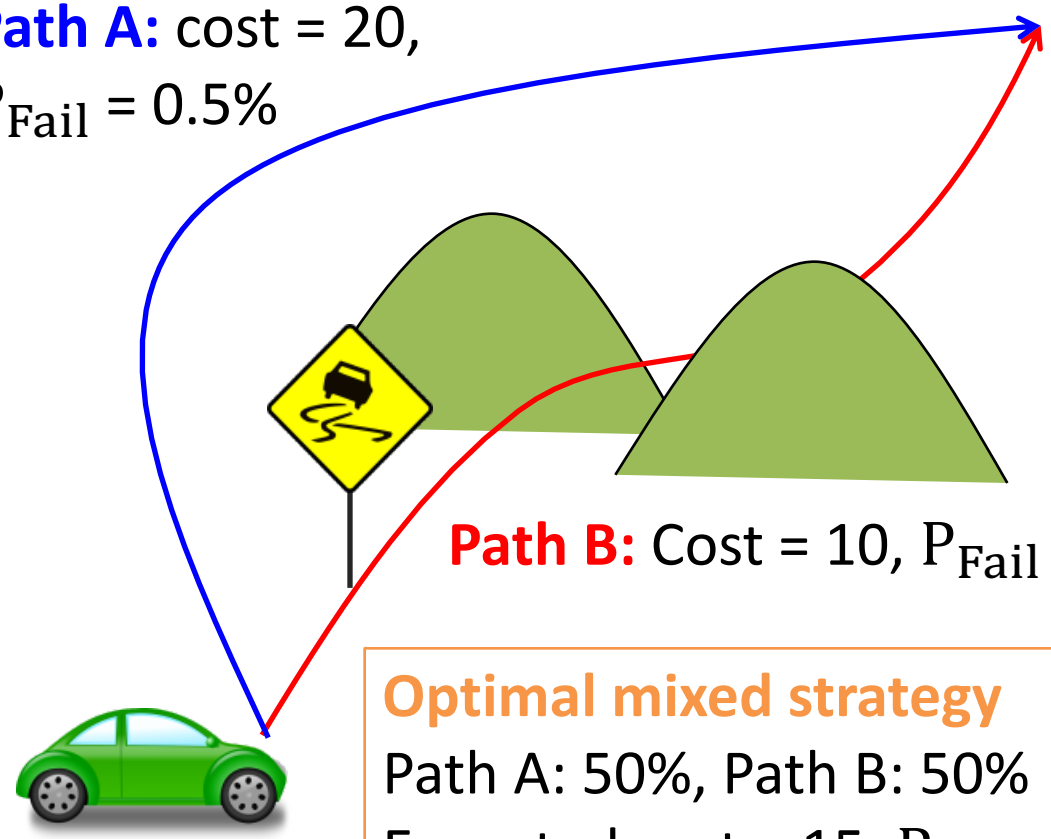
MSL, Sol1431: The rover is at the red dot at top center, heading down. The central route is shorter in distance, but blocked by a potentially hazardous sand. The left route is safer but has a longer distance. The rover drives up to the green circle, get a better view of the sand patch, and decides if it is “safe enough” (risk-aware decision making!!) to go over it. Otherwise it will turn around and take the longer route.

- Choosing control inputs randomly (e.g., by coin flip) can result in a reduced expected cost in optimal control problems with stochastic constraints

Intuition: Toy Example

Chance constraint: $P_{Fail} \leq 1\%$

Path A: cost = 20, P_{Fail} = 0.5%



Path B: Cost = 10, P_{Fail} = 1.5%

Optimal mixed strategy

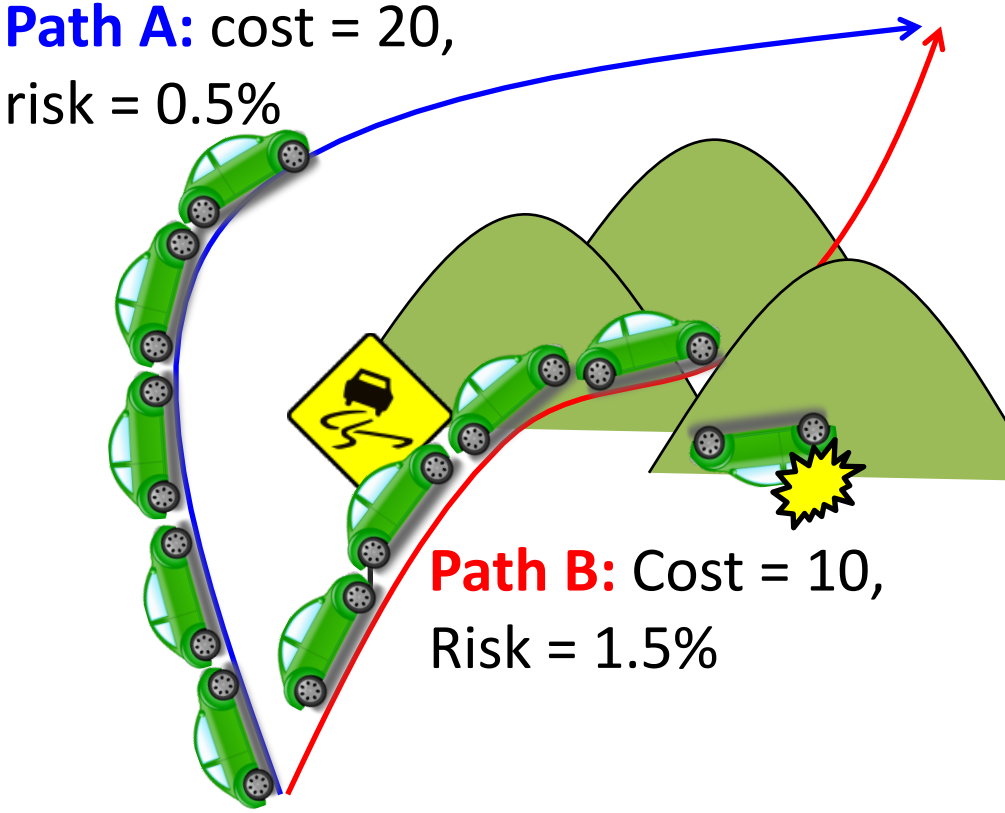
Path A: 50%, Path B: 50%

Expected cost = 15, P_{Fail} = 1%

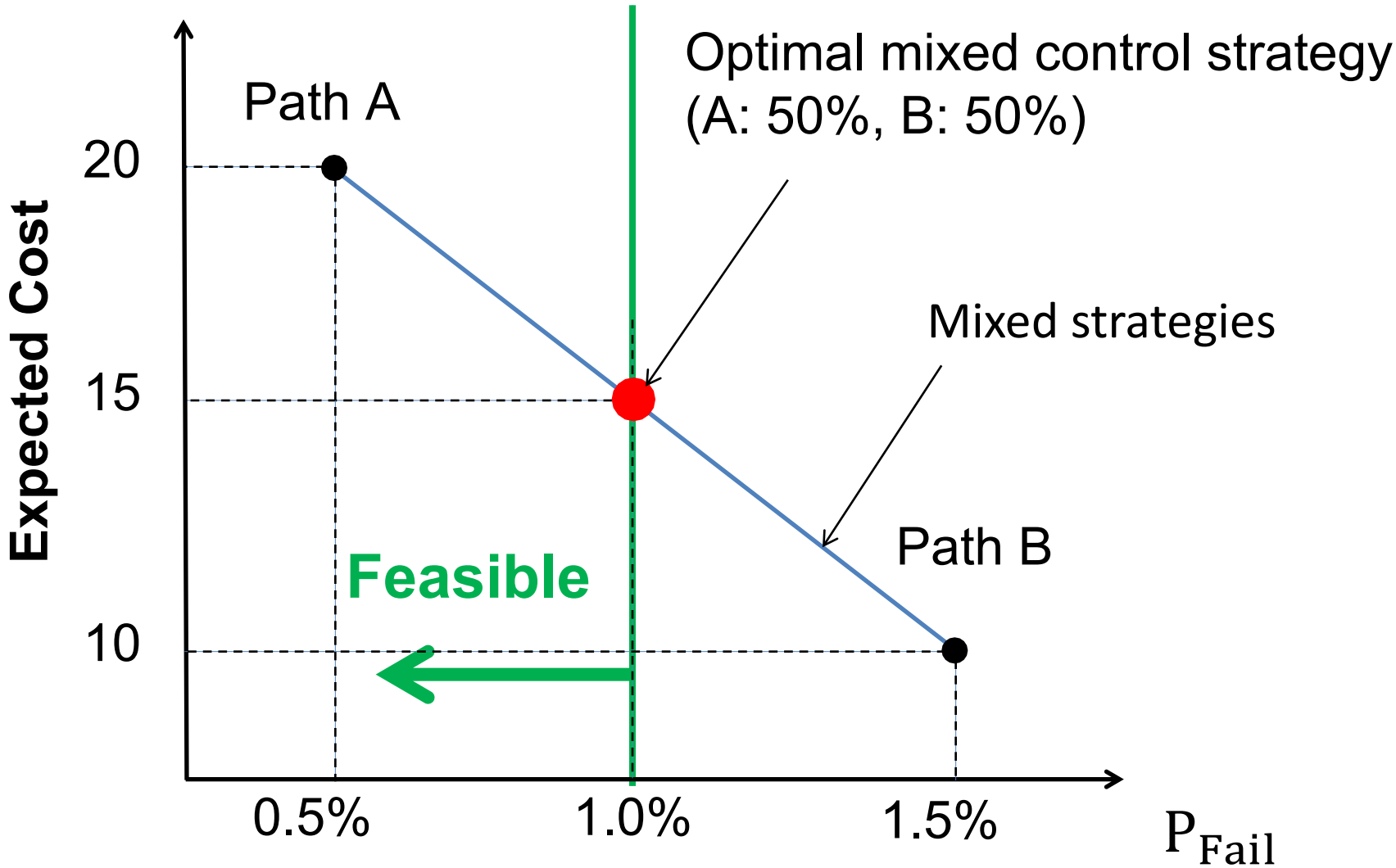
Multi-agent Interpretation

- If there are 100 homogeneous agents
  - Minimize ***total*** cost while ensuring that 99 of them are expected to be safe
- Optimal strategy: send 50 to Path A, send 50 to Path B.

Optimal mixed strategy
Path A: 50%, Path B: 50%
Expected cost = 15, P_{Fail} = 1%



Intuition


Optimal mixed control strategy
(A: 50%, B: 50%)
Path A
Mixed strategies
Path B
Feasible
Expected Cost
20
15
10
0.5%
1.0%
1.5%
P_{Fail}

Relation to Existing Work

- Scenario-based MPC[1]
  - Obtains deterministic control inputs by a randomized method
  - This work obtains randomized control inputs by a deterministic method
- Constrained MDP[2]
  - Randomized policy improves *constrained* MDP
  - E.g., MDP with safety constraint
  - CMDP on continuous, uncountable state/control not studied
- Game theory
  - Mixed strategy is discussed in adversary games
  - In our work, objective function and constraints act like adversary players

1. Bernardini and Bemporad (2009), Calafiore , Dabbene, and Tempo (2011), etc
2. Altman (1999) *Constrained Markov decision processes*

Problem Formulation (Pure SOC)

PSOC (Pure-strategy Stochastic Optimal Control

$$c_{\mathrm{P}}^{\star} = \min_{\substack{\mu \in \mathcal{M}^T \\ g(x,\mu,w)=0}} \mathbb{E}\left[f_0(x,\mu)\right]$$

$$s.t.\ \mathbb{E}\left[f_i(x,\mu)\right] \leq V_i, \quad i = 1 \cdots K.$$

f_0 : Objective function

$f_1 \cdots f_K$: Constraint functions

$x := \{x_1, x_2, \cdots, x_T\} \in \mathcal{X}^T$	State sequence
$u := \{u_1, u_2, \cdots, u_T\} \in \mathcal{U}^T$	Control sequence
$\mu_k : \mathcal{X} \mapsto \mathcal{U}$	Policy
$\mathcal{M}$	Set of policy to be considered
$\mu := \{\mu_1, \mu_2, \cdots, \mu_T\} \in \mathcal{M}^T$	Sequence of policy

Problem Formulation (Mixed SOC)

MSOC^N (Mixed-strategy Stochastic Optimal Control)

$$c_{\mathrm{M}}^{\star N} = \min_{\substack{\mu^1 \cdots \mu^{N+1} \in \mathcal{M}^T \\ \sum_{j=1}^{N} p^j = 1,\ p^j \geq 0 \\ g(x,\mu,w)=0}} \sum_{j=1}^{N+1} p^j \mathbb{E}\left[f_0(x,\mu^j)\right]$$

$$s.t. \sum_{j=1}^{N+1} p^j \mathbb{E}\left[f_i(x,\mu^j)\right] \leq V_i$$

$$i = 1 \cdots K$$

$N+1$: Number of policy sequences to be mixed

μ^j: j-th policy sequence

p^j: Probability that j-th policy sequence is chosen

Question 1: What *N* is sufficient in order to achieve optimality?

Theorem 1: Sufficient Degree of Ranzomization

What N is sufficient in order to achieve optimality?

- If there are K stochastic constraints, $N=K$ (having $K+1$ options) is sufficient for optimality (called *K-randomization*)

Alternative Representation of the Problems

Let $c = (c_0, c_1, \cdots c_K)$ where c_i is the i-th cost value:

$$c_i = \mathbb{E}\left[f_i(x, \mu)\right].$$

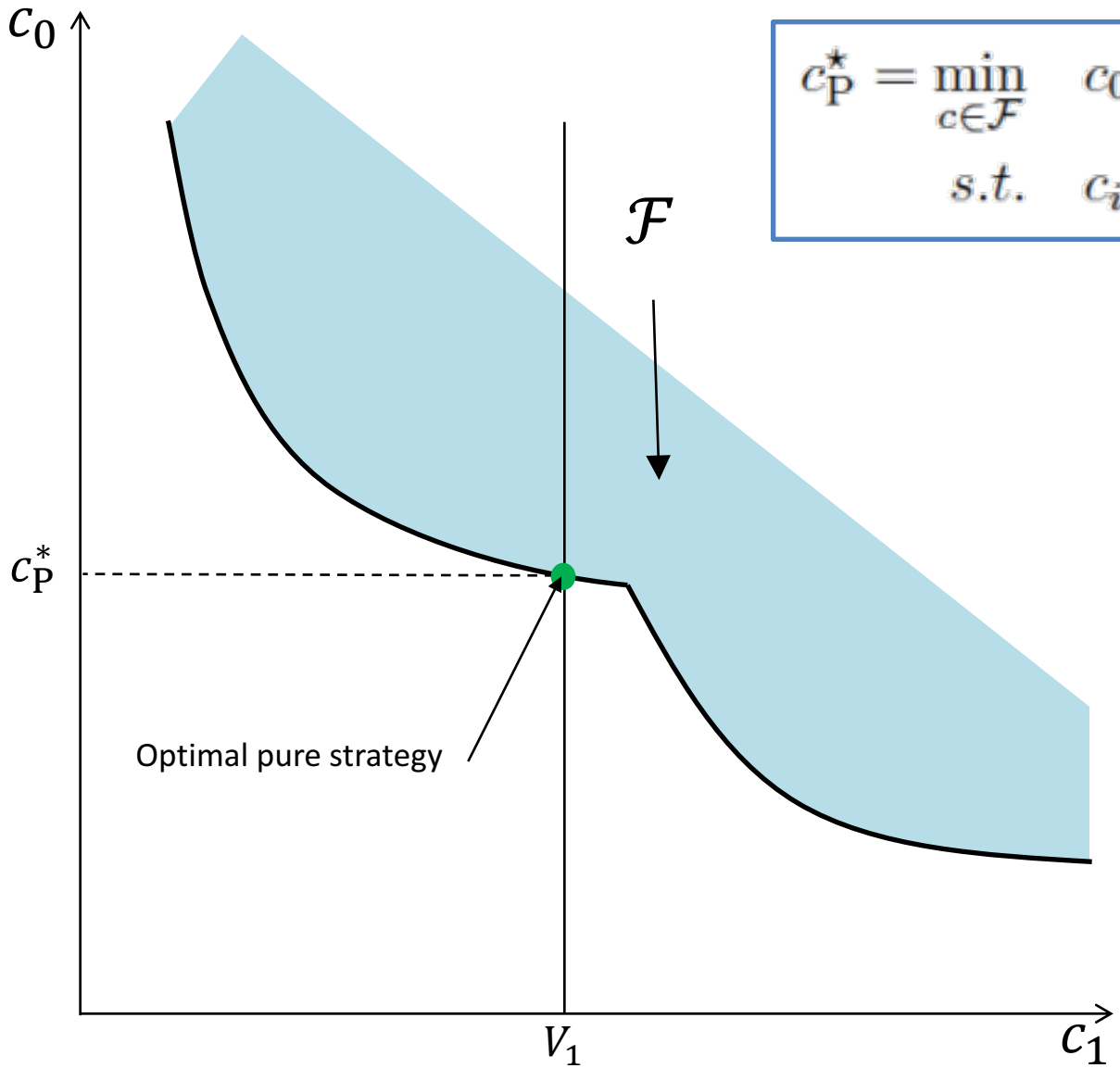
We denote by $\mathcal{F} \subset \mathbb{R}^{K+1}$ the feasible set of the costs of the original problem, that is,

$$\mathcal{F} := \{c \mid \mu \in \mathcal{M} \wedge g(x, \mu, w) = 0\}. \tag{6}$$

PSOC':

$$\begin{aligned} c_{\mathrm{P}}^{\star} = \min_{c \in \mathcal{F}} \quad & c_0 \\ s.t. \quad & c_i \leq V_i, \quad i = 1 \cdots K. \end{aligned}$$

Graphical Interpretation



$$c_{\mathrm{P}}^\star = \min_{c \in \mathcal{F}} \quad c_0$$
$$s.t. \quad c_i \leq V_i, \quad i = 1 \cdots K.$$

Alternative Representation of the Problems

MSOC^N':

$$\min_{\substack{c^1 \cdots c^{N+1} \in \mathcal{F} \\ \sum_{j=1}^{N+1} p^j = 1,\ p^j \geq 0}} \sum_{j=1}^{N+1} p^j c_0^j$$

$$s.t. \quad \sum_{j=1}^{N+1} p^j c_i^j \leq V_i,$$

Yet Another Alternative Representation

MSOC^N":

$$c_{\mathrm{M}}^{\star N} = \min_{c \in \mathcal{F}_{\mathrm{M}}^N} \quad c_0$$

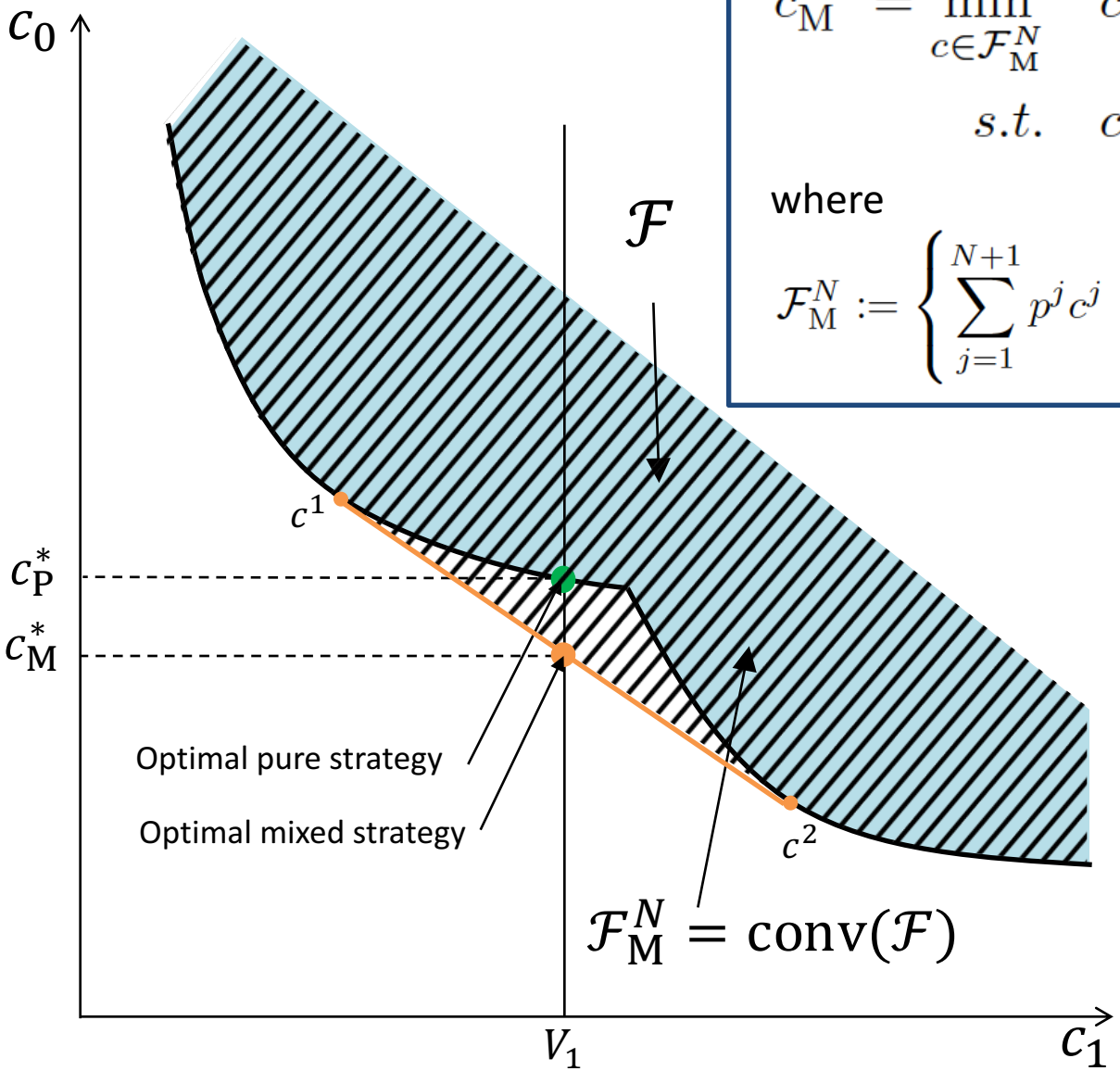
$$s.t. \quad c_i \leq V_i, \quad i = 1 \cdots K.$$

where

$$\mathcal{F}_{\mathrm{M}}^N := \left\{ \sum_{j=1}^{N+1} p^j c^j \mid c^j \in \mathcal{F}, 0 \leq p^j, \sum_{j=1}^{N+1} p^j = 1 \right\}$$

Graphical Interpretation

K=1, N=1



$$c_{\mathrm{M}}^{\star N} = \min_{c \in \mathcal{F}_{\mathrm{M}}^N} \quad c_0$$

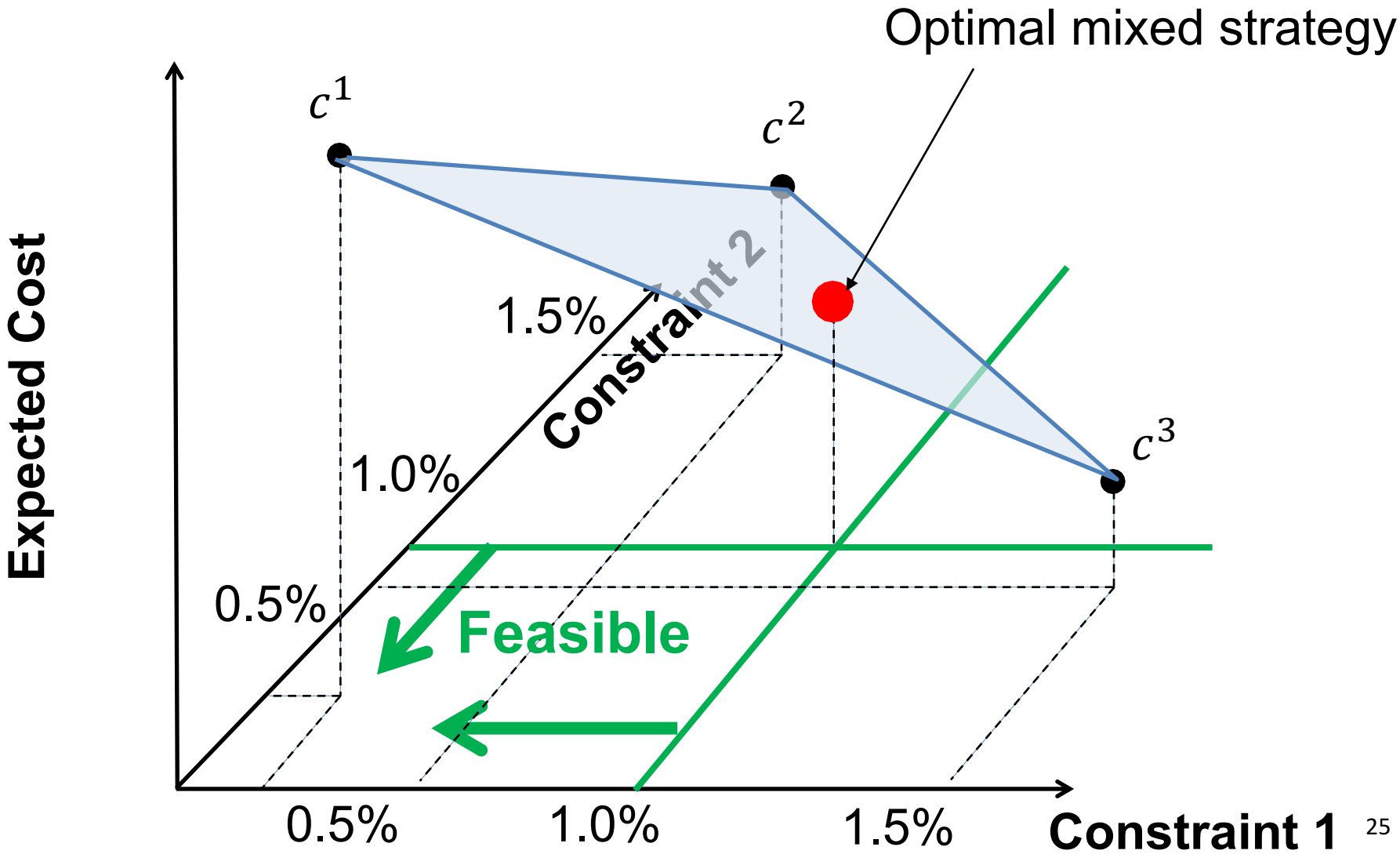
$$s.t. \quad c_i \leq V_i, \quad i = 1 \cdots K.$$

where

$$\mathcal{F}_{\mathrm{M}}^N := \left\{ \sum_{j=1}^{N+1} p^j c^j \mid c^j \in \mathcal{F}, 0 \leq p^j, \sum_{j=1}^{N+1} p^j = 1 \right\}$$

Theorem 1: Intuition

2 constraints => 3D cost space => need 3 points to cover the convex hull



Question 2: How much can we gain by mixed strategy?

Theorem 2

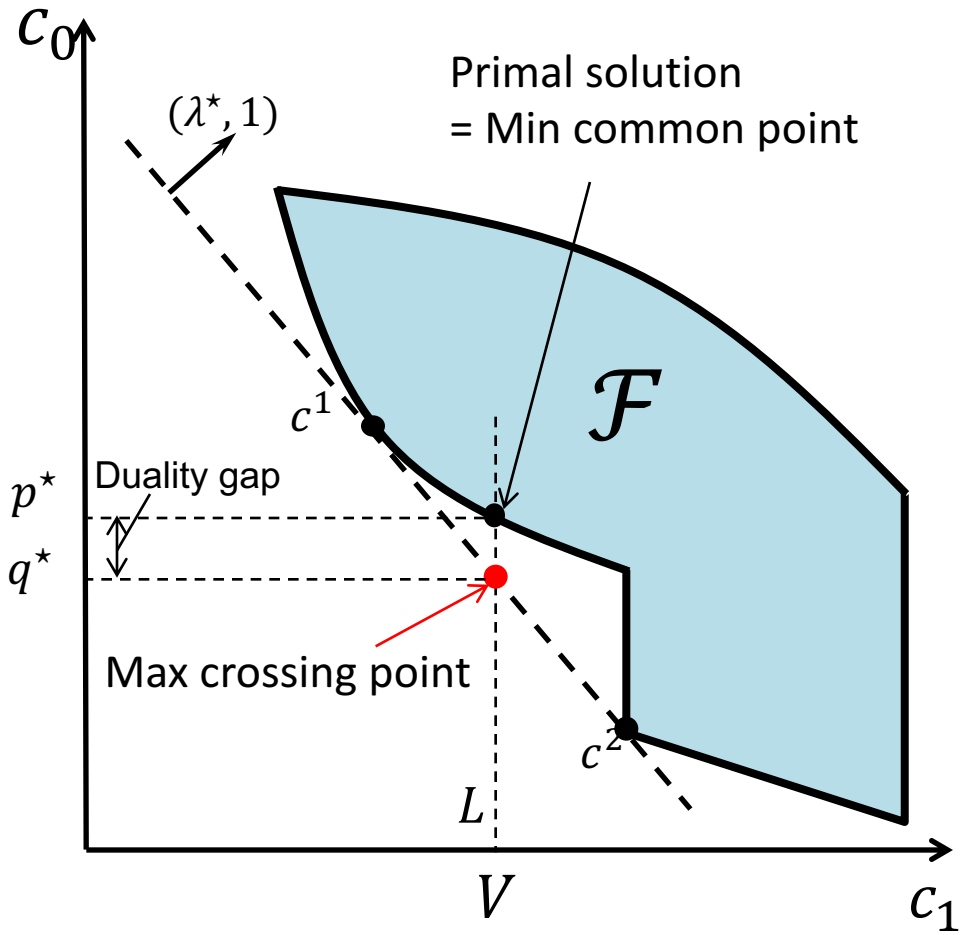
$$c_{\mathrm{M}}^{\star} = c_{\mathrm{P}}^{\star} - \Delta$$

where Δ is the duality gap of PSOC

Min Common/Max Crossing (MCMC) Framework

Bertsekas 2009

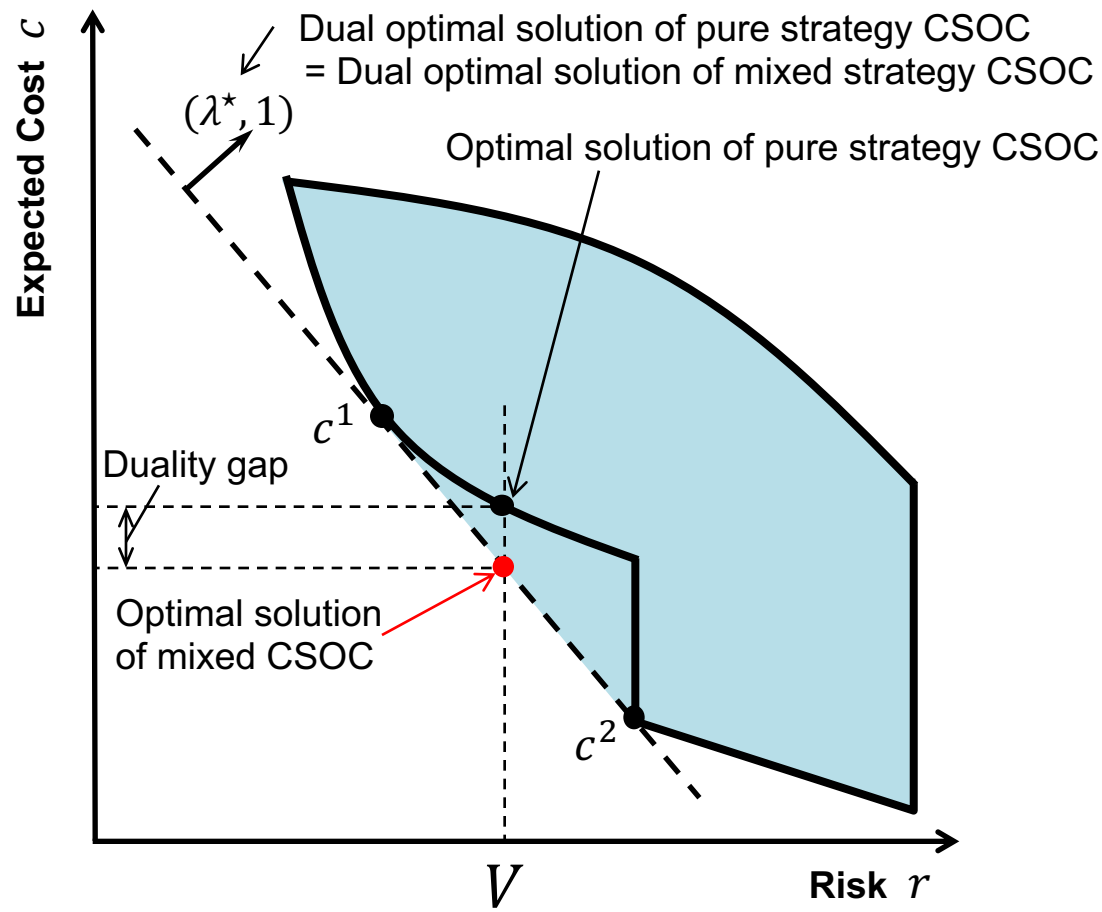
- **Dual optimal solution = Max crossing point**: the point with the greatest c_0 among the intersection of L and hyperplanes that contain $\mathcal{F}$ in their upper closed halfspace



$$\mathcal{F} = \{(c_0, c_1) | x \in X\}$$
$$L = \{(c_0, c_1) | c_1 = V\}$$

Strong Duality of MSOC

- MSOC is better than PSOC only when PSOC has a duality gap (≈nonconvex)
- MSOC does not have a duality gap



Solution Approach

- A naïve approach (for N-randomization): co-optimize $u^1 \cdots u^{N+1}$ and $p^1 \cdots p^{N+1}$
  - Difficult to solve, if not impossible…

$$c_{\mathrm{M}}^{\star N} = \min_{\substack{u^1 \cdots u^{N+1} \in \mathcal{U}^T \\ \sum_{j=1}^{N} p^j = 1,\ p^j \geq 0 \\ g(x,u,w)=0}} \sum_{j=1}^{N+1} p^j \mathbb{E}\left[f_0(x, u^j)\right]$$

$$s.t. \sum_{j=1}^{N+1} p^j \mathbb{E}\left[f_i(x, u^j)\right] \leq V_i$$

$$i = 1 \cdots K$$

Our Solution Approach: Basic Idea

1. Solve the dual of PSOC
   - Because the dual of the pure and mixed problems are the same!
2. Reconstruct the primal solution of MSOC from the dual solution

Dual of PSOC

DPSOC (Dual of PSOC)

$$p_{\mathrm{P}}^{\star} = \max_{\lambda \geq 0} \min_{c \in \mathcal{F}} \quad c_0 + \lambda(c_{1:K} - V)$$

And let

$$C(\lambda) = \arg\min_{c \in \mathcal{F}} \quad c_0 + \lambda(c_{1:K} - V)$$

Theorem 3: Condition for Optimality

- $c^0 \cdots c^K, p^0 \cdots p^K$ is an optimal solution to **DSOC^K** if and only if there exists $\lambda \geq 0$ such that:

$$
\begin{aligned}
&a)\ c^i \in C(\lambda) \ \vee \ p^i = 0, \ \forall i = 1 \cdots K+1 \\
&b)\ \lambda \left(\sum_{i=1}^{K+1} p^i c^i_{1:K} - V \right) = 0, \\
&c)\ \sum_{i=1}^{K+1} p^i = 1, \\
&d)\ p^i \geq 0 \ \forall i = 1 \cdots K+1, \\
&e)\ \sum_{i=1}^{K+1} p^i c^i_{1:K} \leq V, and \\
&f)\ c^i \in \mathcal{F}, \ \forall i = 1 \cdots K+1.
\end{aligned}
$$

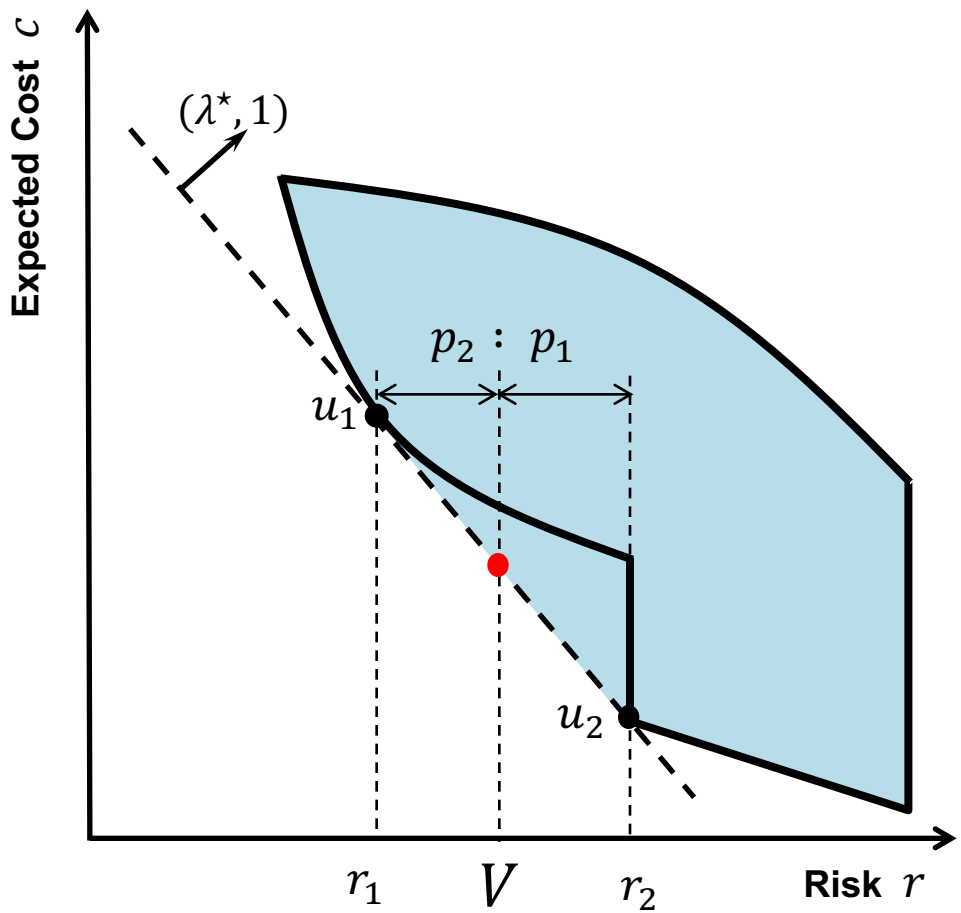
$c^0 \cdots c^K$ are the optimal solution to the inner optimization of DPSOC

KKT conditions

*Details are in the arXiv paper

Computing probability distribution

Get $p^1 \cdots p^{N+1}$ that makes the equality in the stochastic constraints hold



Solution Approach

1. Solve the dual of the *original* (non-randomized) problem
   - Because the dual of the original and mixed problems are the same!
2. If $\lambda^* = 0$ (stochastic constraints are inactive)
   - The optimal solution to the original problem is also optimal for the mixed problem
3. If $\lambda^* > 0$
   - Get the primal solutions $u_1 \cdots u_{K+1}$ that minimizes the Lagrangian with λ^*
   - Get $p^1 \cdots p^{N+1}$ that makes the equality in the stochastic constraints hold:

$$\sum_{j=1}^{N+1} p^j \mathbb{E}\left[f_i(x, u^j)\right] = V_i \qquad \text{and} \qquad \sum_{j=1}^{N+1} p^j = 1$$

$$i = 1 \cdots K$$

Special Case: K=1

- If there is only one stochastic constraint, the dual optimization can be solved efficiently by **root finding**
  - E.g., Joint chance-constrained optimal control

Special case with K=1

Dual Problem

$$\max_{\lambda} \min_{u \in \mathcal{U}} \mathbb{E}[f_0(x,u)] + \lambda(\mathbb{E}[f_1(x,u)] - \mathrm{V})$$

$$\|$$

$$q(\lambda)$$

Primal Problem

$$\min_{u \in \mathcal{U}} \mathbb{E}[f_0(x,u)]$$

$$\text{s.t. } \mathbb{E}[f_1(x,u)] \leq V$$

Special case with K=1

Dual Problem

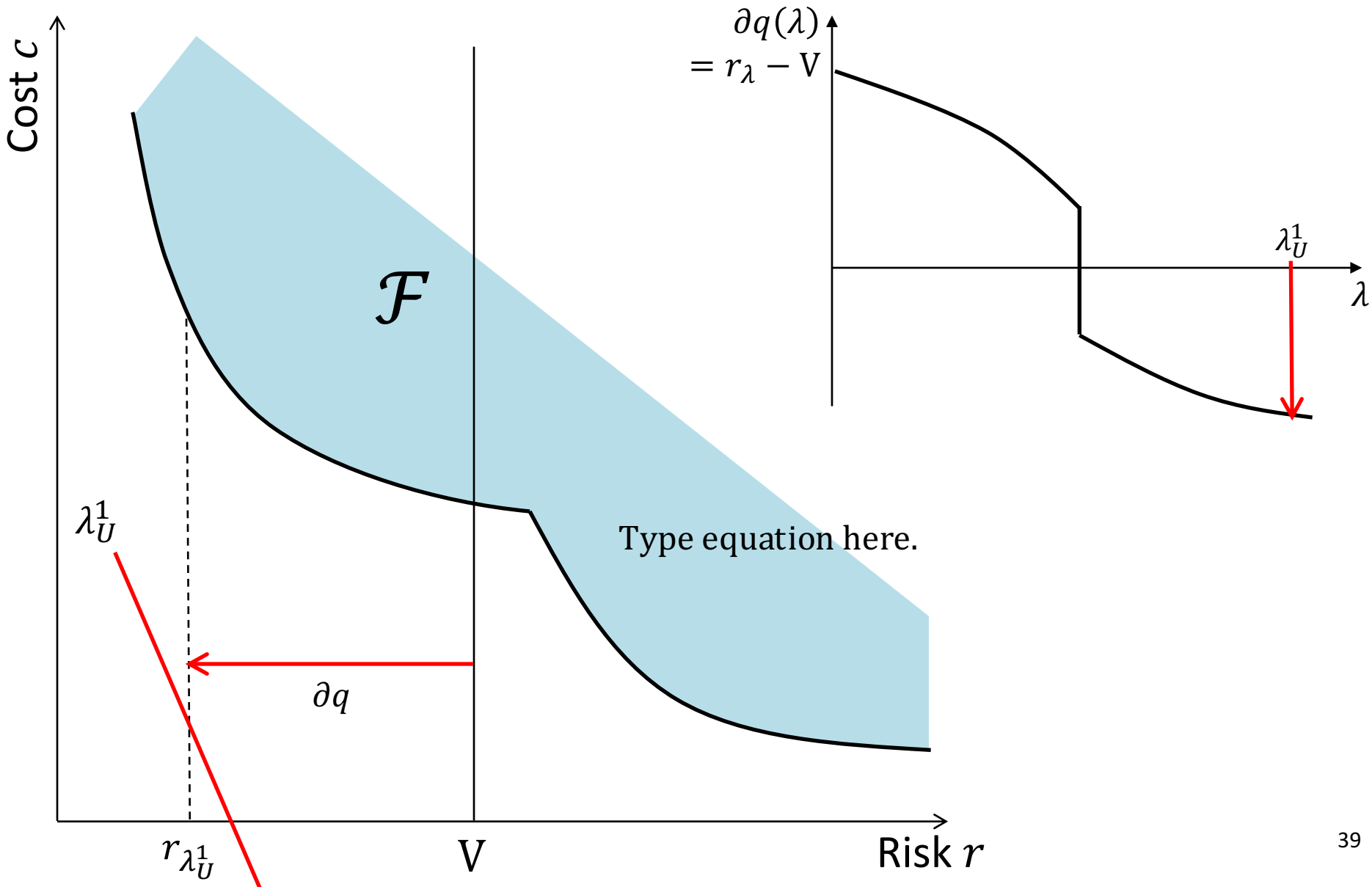
$$\max_{\lambda} \min_{u \in \mathcal{U}} \mathbb{E}[f_0(x,u)] + \lambda(\mathbb{E}[f_1(x,u)] - \mathrm{V})$$

$$\|$$

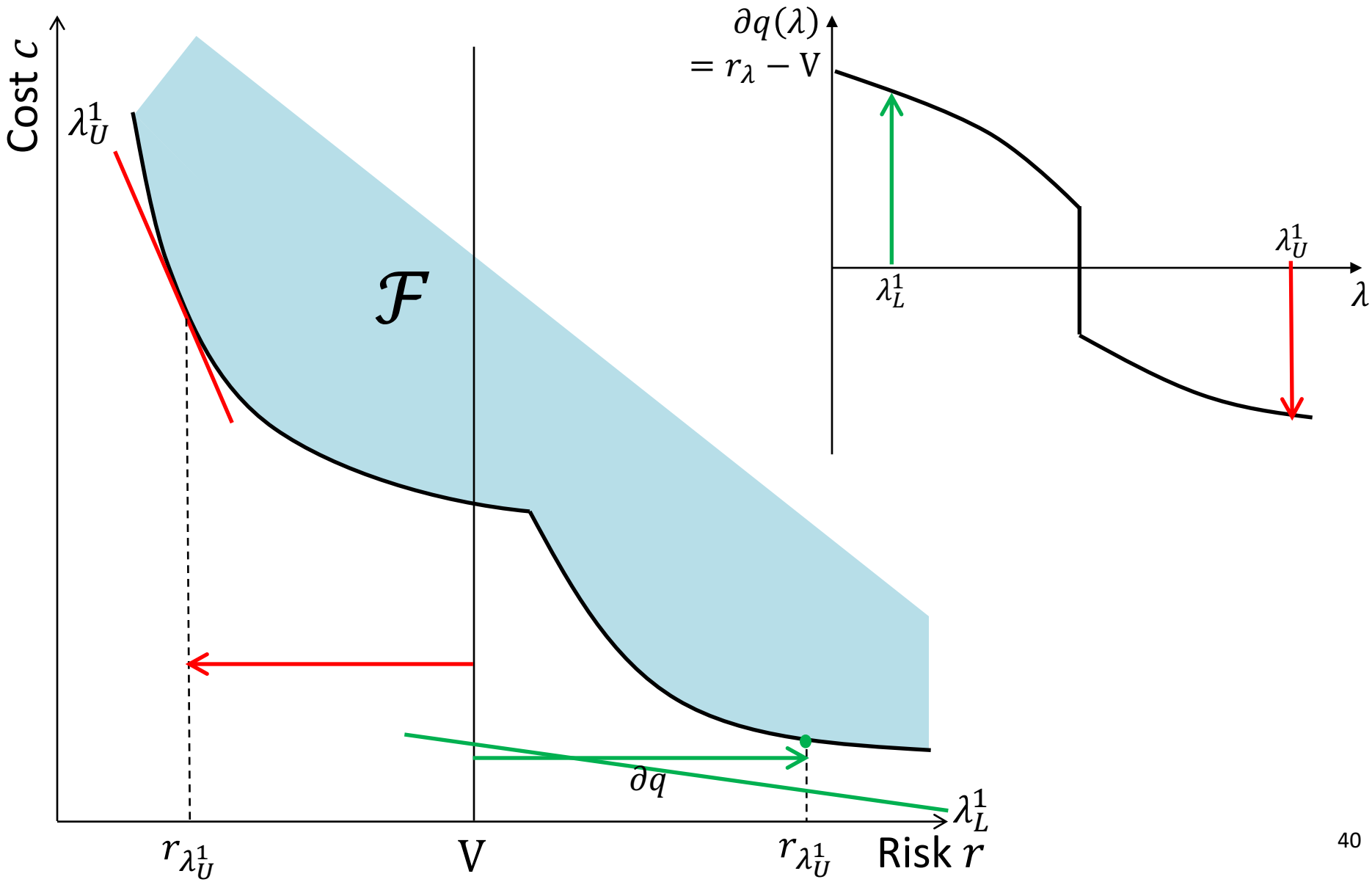
$$q(\lambda)$$

- $q(\lambda)$ is always concave
- Optimality condition: $0 \in \partial q(\lambda)$
  - $\partial q(\lambda)$: subgradient
- Therefore, dual optimization is reduce to a zero-finding problem over $\partial q(\lambda)$

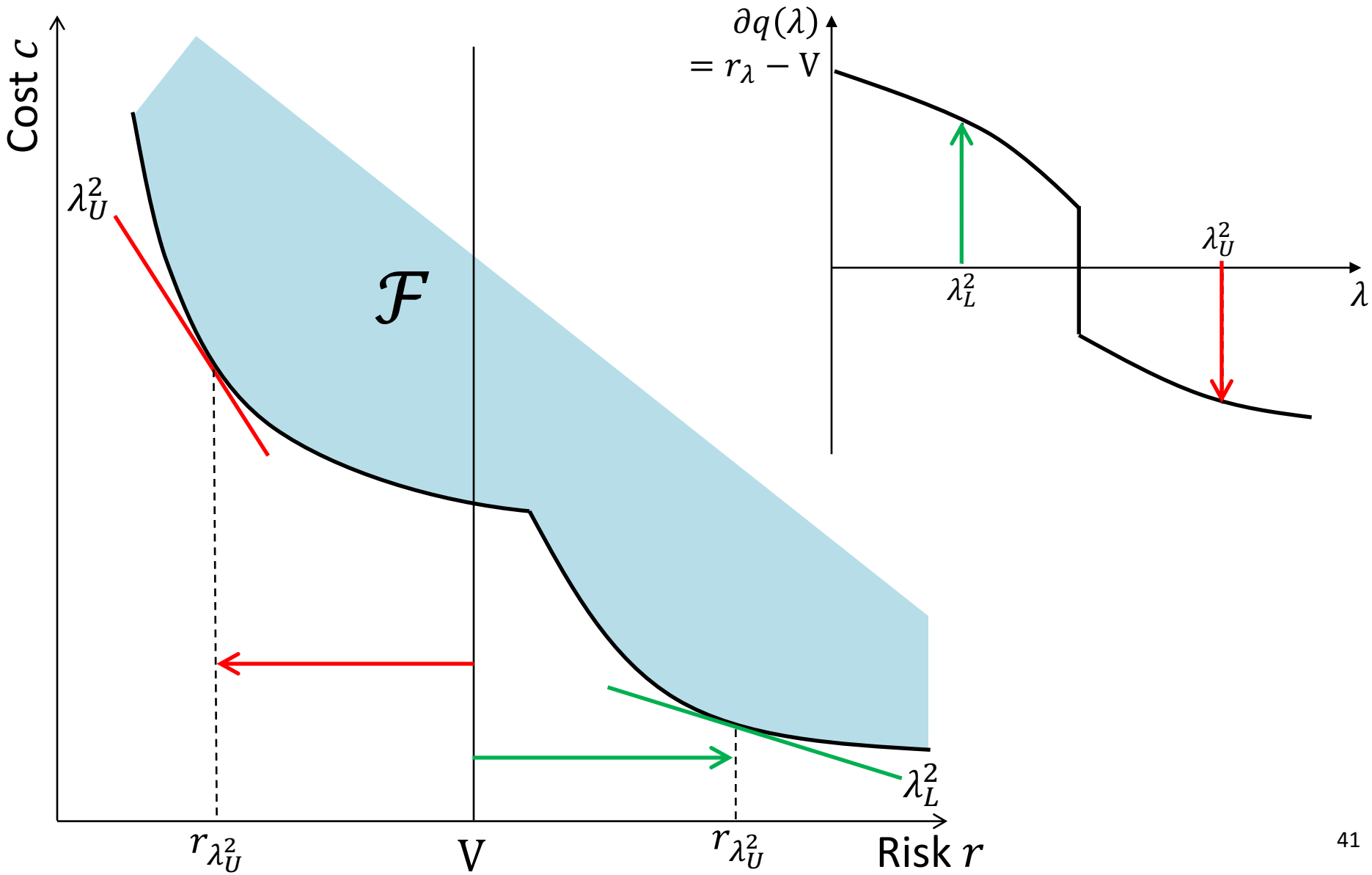
Dual Solution through Root Finding



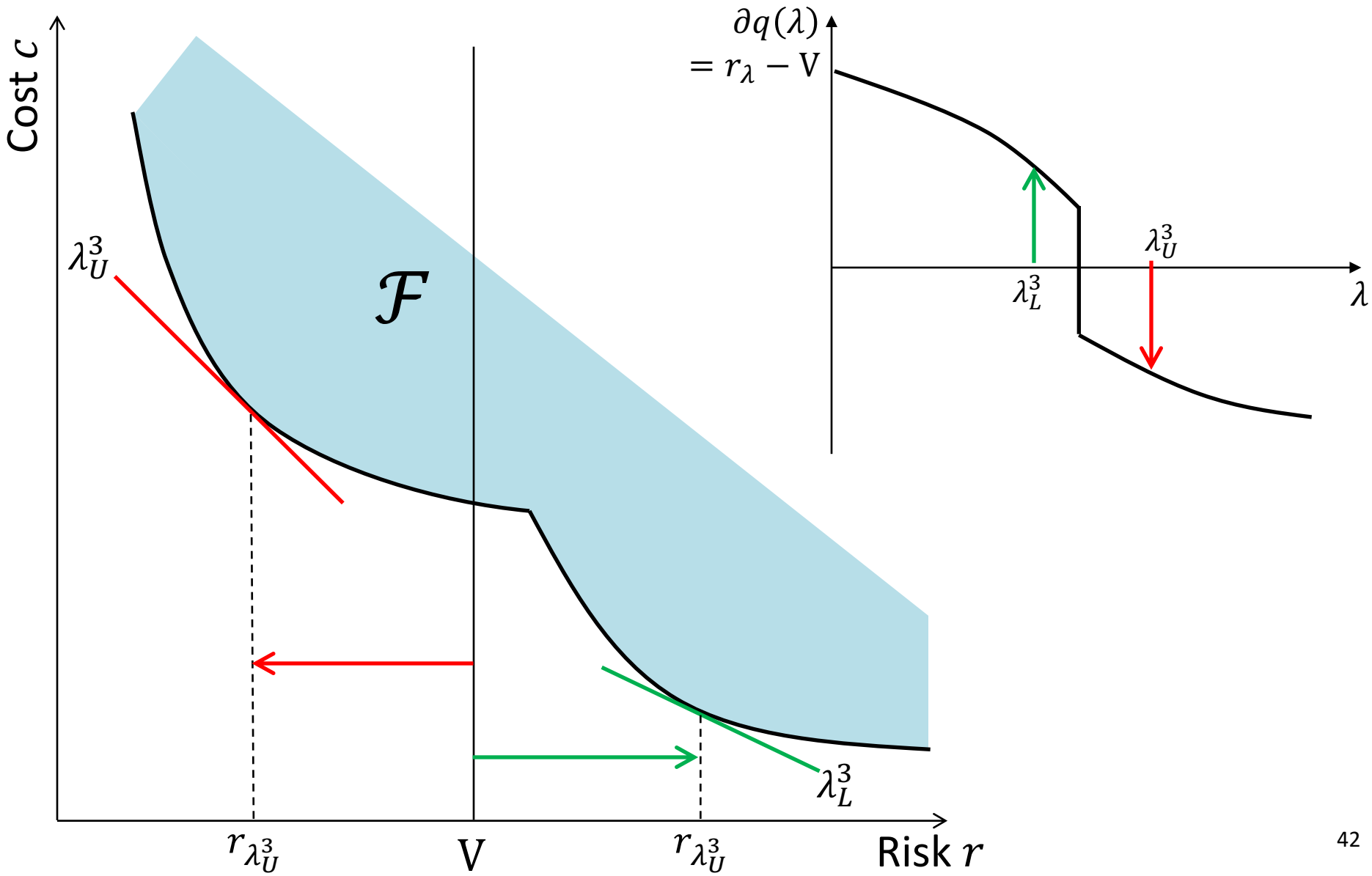
Dual Solution through Root Finding



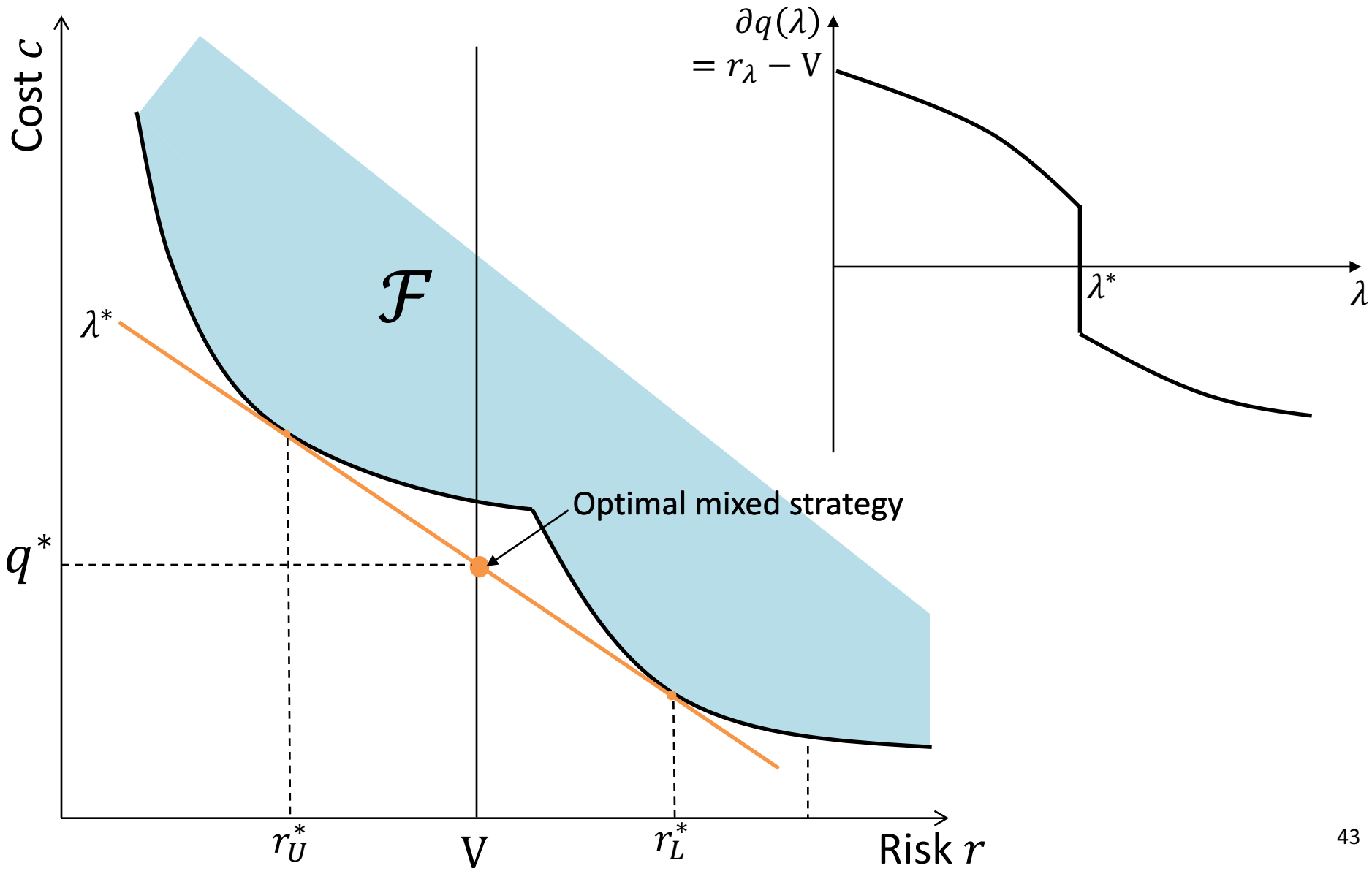
Dual Solution through Root Finding



Dual Solution through Root Finding



Dual Solution through Root Finding



Piecewise Linear Approximation of SMPC

MILP

$$\min \sum_{k=1}^{N}\sum_{i=1}^{m} v_{ki} + \lambda\left(\sum_{i=1}^{M}\sum_{k=1}^{N}\delta_{ik} - V\right)$$

$$\text{s.t. } v_{kd} \geq u_{kd},\ v_{kd} \geq -u_{kd}, Pu_k \preceq q$$

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + w_k$$

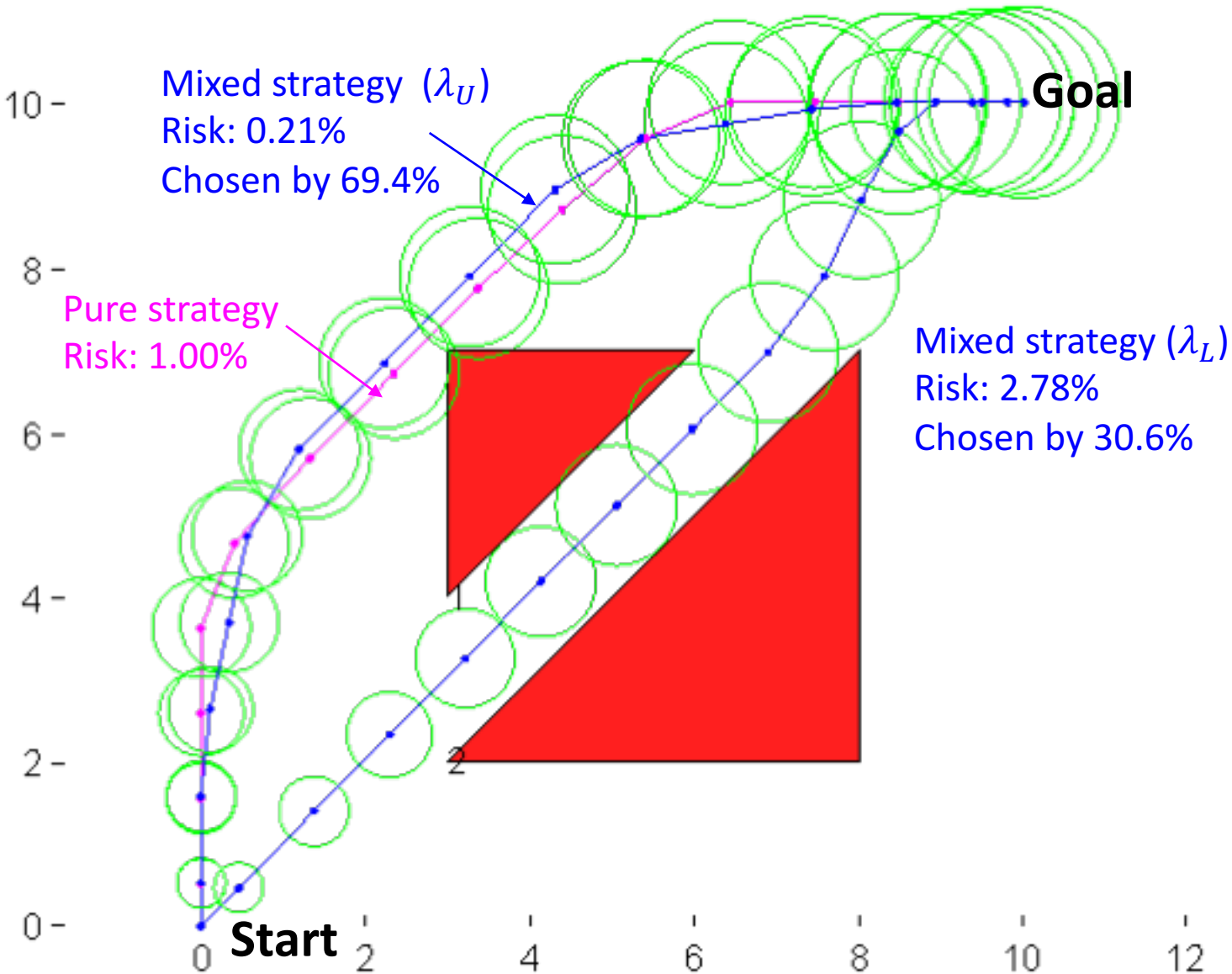
$$a_l\left(\frac{h_{ij}\bar{x}_k - g_{ij}}{h_{ij}\Sigma_{x_k}h_{ij}^T}\right) + b_l \leq \delta_{ik} + \mathcal{M}z_{ij}$$

$$\sum_{j=1}^{R_i} z_{ij} \leq R_i - 1, \quad z_{ij} \in \{0,1\}$$

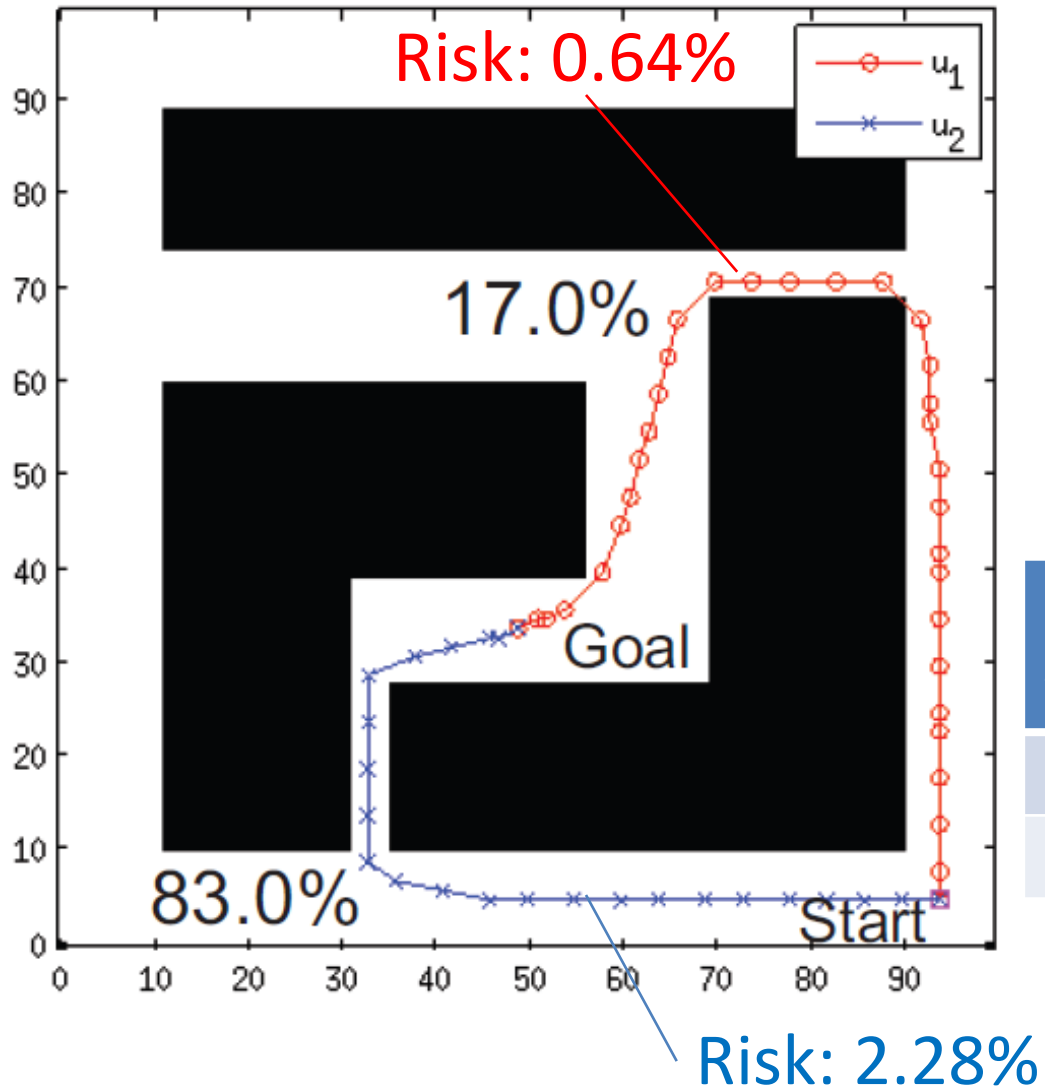
$$k = 1\cdots N, d = 1\cdots m,\ i = 1\cdots M,$$

$$j = 1\cdots R_i,\ l = 1\cdots L.$$

Path Planning with CCMPC



Application to Path Planning



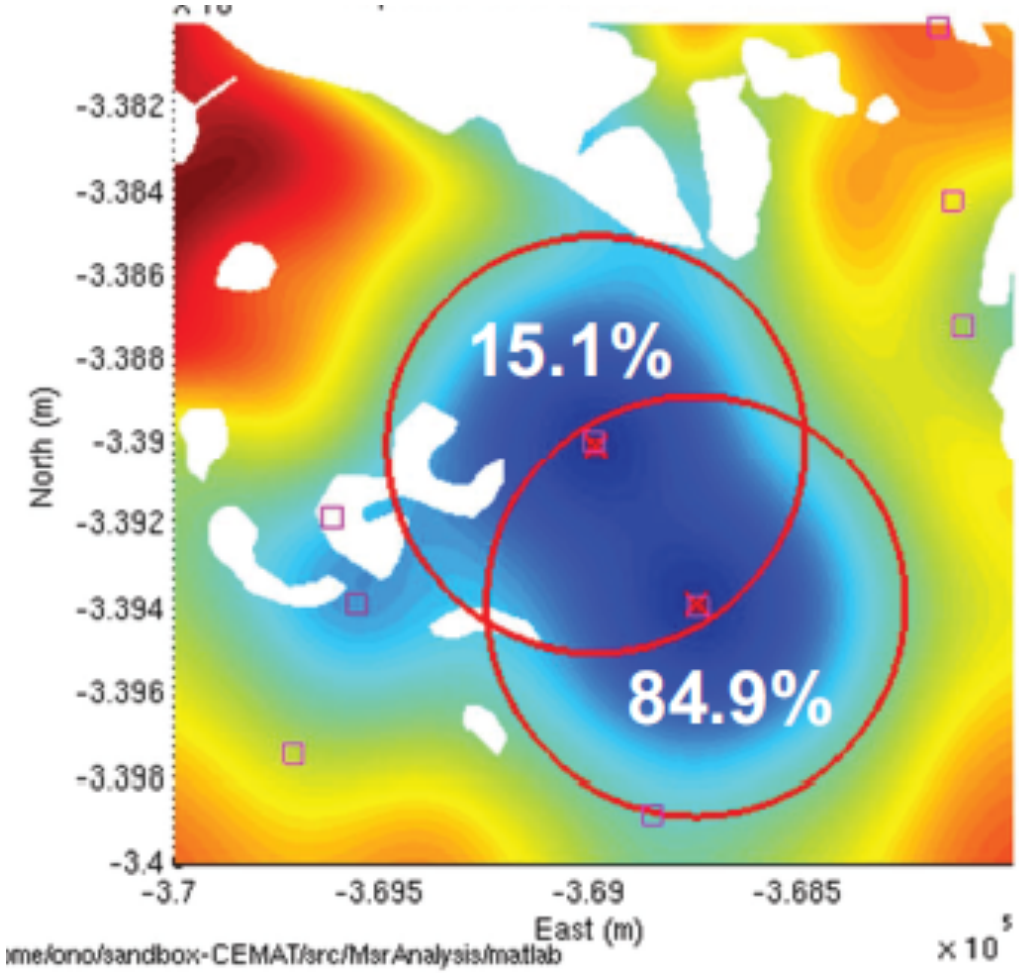
Problem:

- Minimize path length
- Risk $\leq$ 2%

- $x_{k+1} = x_k + u_k + w_k$
- 100x100 discrete state space

Control strategy	Expected path length	Risk
Pure	130.8	0.64%
Mixed	104.2	2.0%

Mars Landing Scenario



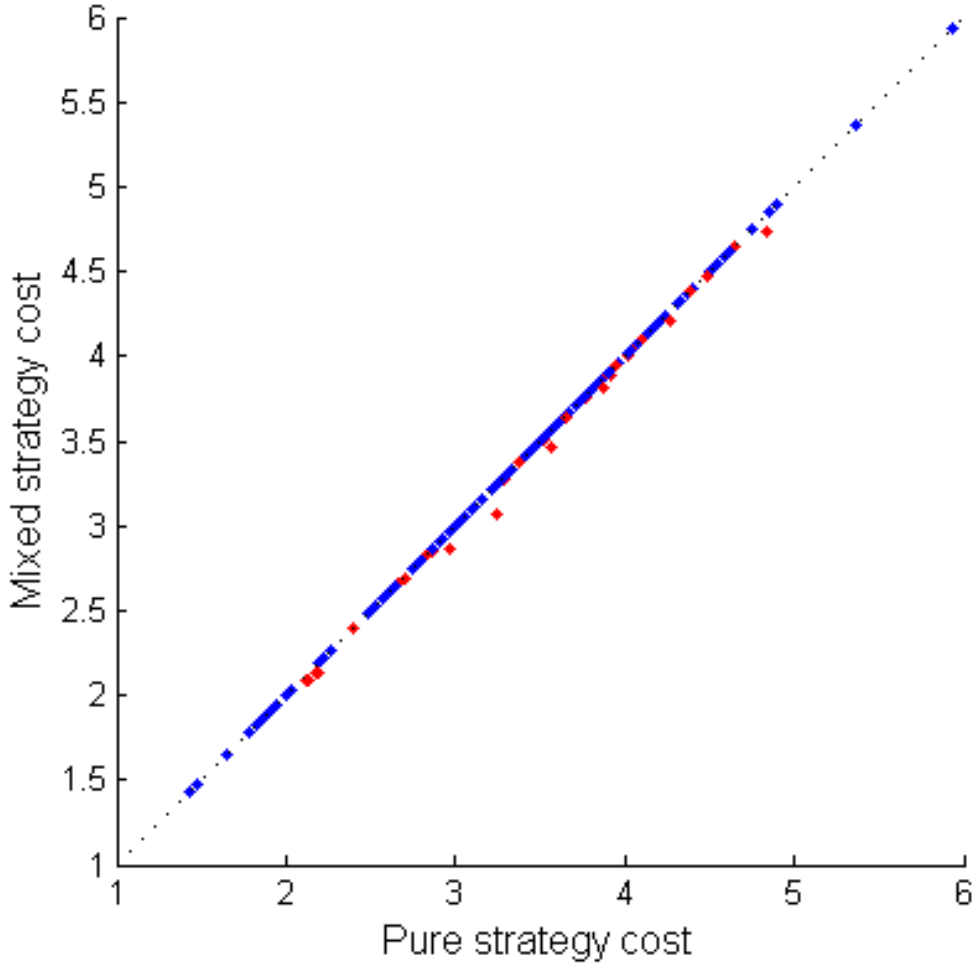
Problem:

- Minimize driving distance to visit two science targets after landing
- Risk $\leq$ 0.1%

- 2000x2000 discrete state space
- Used terrain data at E. Margaritifer on Mars

Control strategy	Expected cost	Risk
Pure	645.49	0.016%
Mixed	644.81	0.1%

Validation with Monte Carlo Simulation

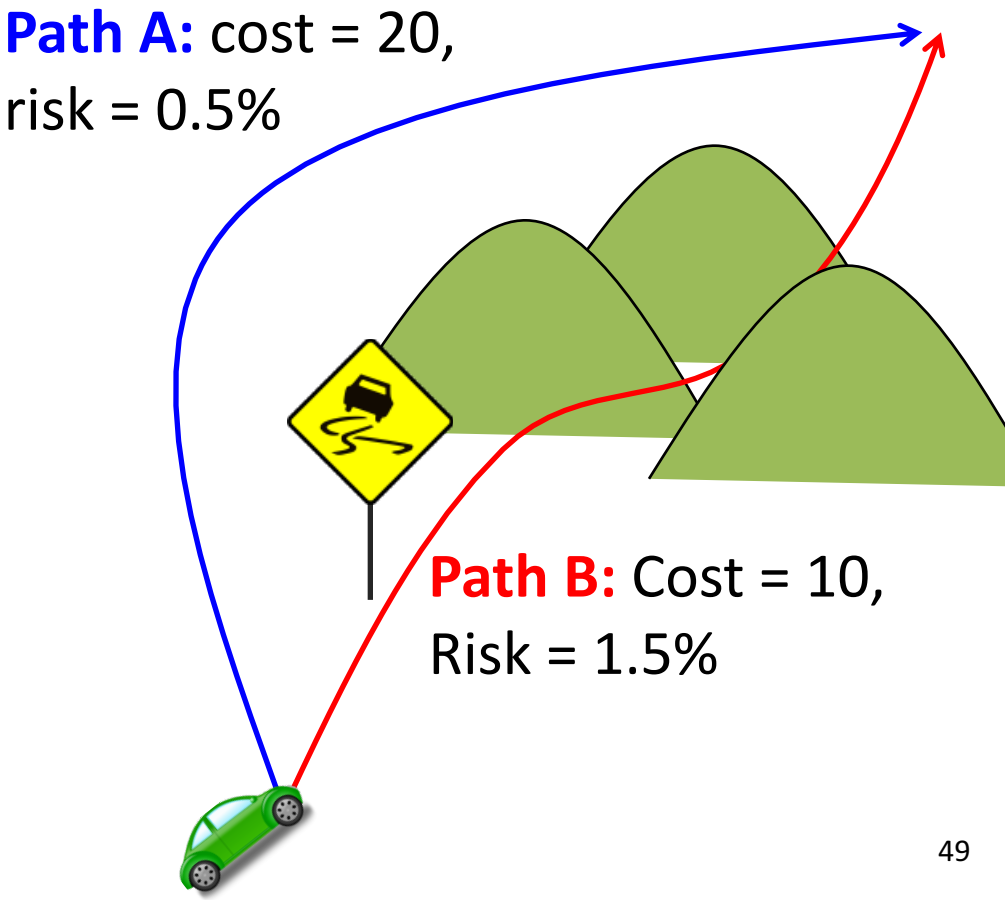


- 200 CCMPC problems with random obstacle locations
- Mix solutions outperforms in red dots
- No pure solutions outperforms the mixed ones

A Paradox

- Before the coin flip, risk is 1%
- But after seeing the result of the coin flip, the risk is NOT 1% any more!
- If you *know* that Path B is chosen, the chance constraint is violated!

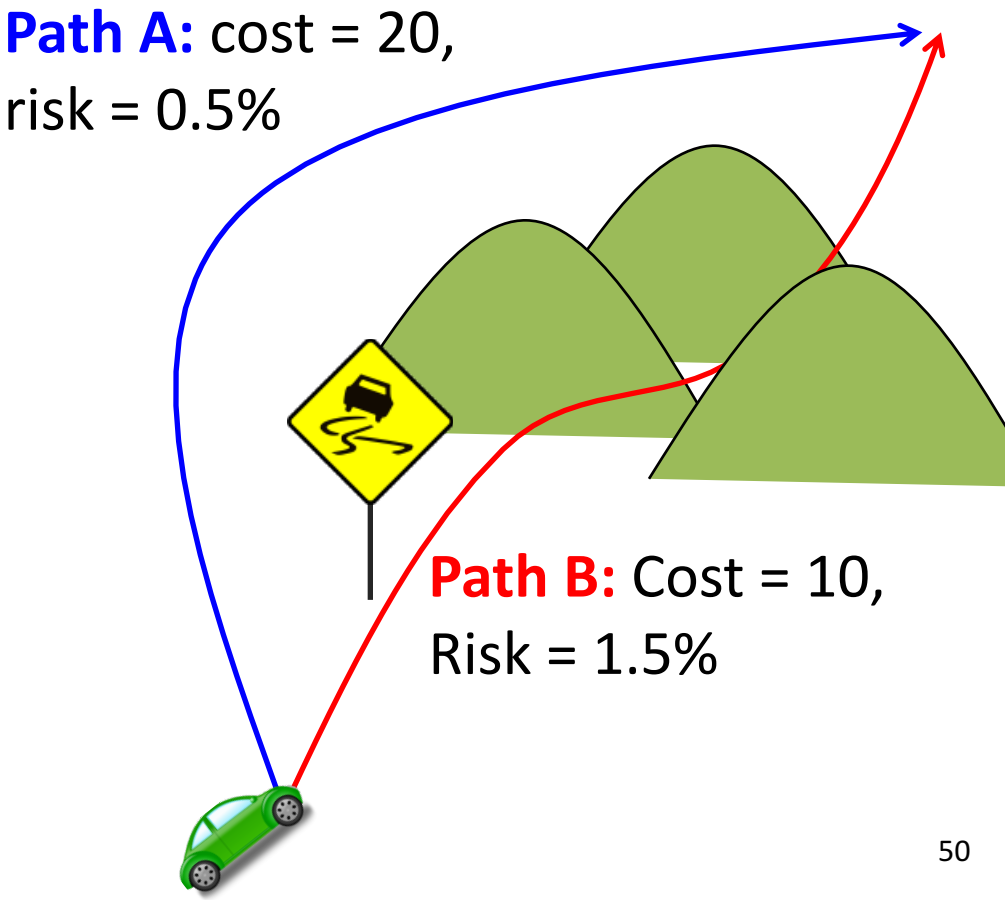
Optimal mixed strategy
Path A: 50%, Path B: 50%
Expected cost = 15, risk = 1%



Solution to the Paradox

- Don't let the operator know the result of the coin flip
  - Then the probability of failure stays at 1%
- Like Schrödinger's cat, the vehicle's state is a superimposition of Path A and Path B

Optimal mixed strategy
Path A: 50%, Path B: 50%
Expected cost = 15, risk = 1%



Conclusions

- In a nonconvex constrained stochastic optimal control problem, randomizing control may result in a reduced expected cost
- If there are *K* stochastic constraints, *K*-randomization is sufficient
- Developed a solution approach to the mixed strategy control through dual optimization
- Developed an efficient solution approach to a problem with K=1 using root finding

- Funded by the ONR Science of Autonomy Program
  - Office of Naval Research Grant N00014-15-IP-00052