Distributionally Robust Optimization and Its Applications in Communication and Networking

Zhu Han, IEEE/AAAS Fellow
University of Houston

Thanks to Drs. Lei Fan, Zhuang Ling, Kai-Chu Tsai, Hongliang Zhang, Yali Chen, Yang Yang, and Yuan Zi

http://wireless.egr.uh.edu/research.htm

UNIVERSITY of HOUSTON
COLLEGE of TECHNOLOGY

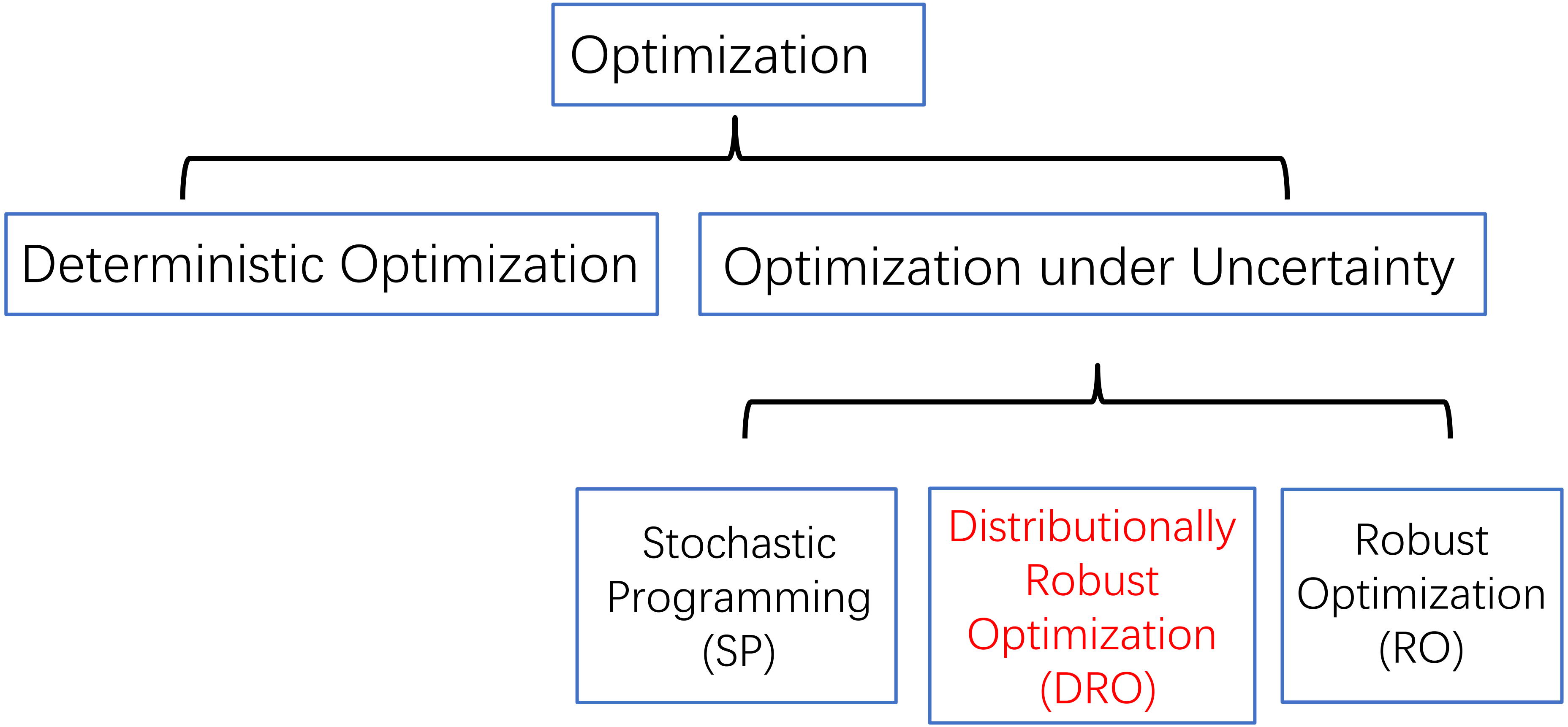
Outline

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

- [illegible]
- C[illegible]1: D[illegible]-[illegible]D[illegible]
  - ✓ [illegible]
  - ✓ [illegible]D[illegible]
  - ✓ A[illegible]:C[illegible]
- C[illegible]2: D[illegible], [illegible]A[illegible]C[illegible]
  - ✓ [illegible]C[illegible]
  - ✓ [illegible]-[illegible]-[illegible]([illegible]) [illegible]C[illegible]-[illegible]-[illegible](C[illegible])
  - ✓ A[illegible]:A[illegible]
- C[illegible]3: D[illegible]
  - ✓ D[illegible]
  - ✓ A[illegible]:[illegible]
- C[illegible]

Introduction

✓ Optimization Classification



✓ Stochastic Programming (SP)

The *probability distribution* of random parameters is *known* (inferred from the historical data). Objective is to find a decision x that minimizes a functional of the ***expected cost***.

$$\min_{x \in \chi} E_P[h(x, \xi)]$$

x -- Decision variables
χ -- Convex set of feasible solutions
ξ -- Uncertain parameter follows a certain distribution P
$h(x, \xi)$ -- Objective function in x that depends on parameters ξ

Classical assumptions in stochastic programming:

- The probability distribution of the random parameter vector is independent of decisions
- The "true" probability distribution of the random parameter vector is known relaxing it requires addressing distributional uncertainty.

Introduction: Robust Optimization

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

✓ Robust Optimization (RO)

The probability distribution of random parameter is unknown, but its fluctuation range is known. Objective is to find a decision x that minimizes the worst-case expected cost over an uncertainty set.

$$\min_{x \in \chi} \max_{\xi \in U} h\,(x, \xi)$$

x	-- Decision variables
ξ	-- Uncertain Parameter
χ	-- Convex set of feasible solutions
U	-- Uncertain set
$h(x, \xi)$	-- Objective function in x that depends on parameters ξ

- First find a decision x that minimizes the cost.
- Then a parameter ξ which leads to the maximum cost (worst case for given decision)

Introduction: DRO

✓ Distributionally Robust Optimization (DRO)

- In practice, the random parameters are *uncertain*.

- Although the exact distribution of the random variables may not be known, people usually know partial statistic information via certain observed samples.

Choose an intermediate approach to obtain a robust form of distributed optimization problem ***(DRO):***

$$\min_{x \in \chi} E_P[h(x,\xi)] \Rightarrow \min_{x \in \chi} \max_{P \in \mathcal{P}} E_P[h(x,\xi)]$$

$\mathcal{P}$ is an uncertain set of probability distributions constructed from the samples.

Objective is to find a decision x that minimizes the *worst-case expected cost* over an *ambiguity set*.

Introduction: Uncertainty Sets

✓ Key question: How to build ambiguity sets (uncertain sets)?

The probability distribution quantifying the model parameter uncertainty is known ***ambiguously***. When choosing ambiguity sets, we need to consider the following:

- Rich enough to contain the true data-generating distribution with high confidence.
- Small enough to exclude pathological distributions.
- Tractability
- Practical (Statistical) Meanings
- Performance (the potential loss comparing to the benchmark cases)

The form of ambiguity sets can be used to classify the distributionally robust optimization problems.

- ***Moment-based ambiguity sets :*** $\mathcal{P} = \{\xi: E[\xi] \leq \mu, E[\xi^T \xi] \leq \Sigma, \dots\}$
- ***Discrepancy-based ambiguity sets:*** $\mathcal{P} = \{P: d(\widehat{P_N}, P) \leq \rho\}$

Outline

- Introduction
- C□□□□1: D□□□□□□□□□-□□□□□D□□ □□□□□ □□□□□□□
  - ✓ □ □□□□□□□□□□ □□□□□□□□
  - ✓ □□□□□□□□□□□□□ □□□□□□□□□D□□
  - ✓ A□□□□□□□□□□:C□□ □□□□□□□□ □ □□□□□□□□□
- C□□□□2: D□□, □□□□A□□□□□□□, □□□□C□□□
- C□□□□3: D□□ B□□□□□ □□□□□□□□□□□ □□□□□□□□□□□□
- C□□□□□□□□□

Discrepancy-based DRO

- In many situations, we have an empirical estimate of the underlying probability distributions.
- A natural way to hedge against the distributional ambiguity is to consider a neighborhood of the empirical probability distribution

D

Ambiguity sets based on probability distance:

$$\mathcal{P} = \{P : d(\widehat{P_N}, P) \leq \varepsilon\}$$

$\widehat{P_N}$ -- Empirical probability
ε -- Radius
$d(\widehat{P_N}, P)$ -- ***Metric*** of the similarity of two distributions

By selecting a suitable *metric*, certain *infinite-dimensional* convex DRO problems can be transformed into *finite-dimensional* convex optimization problems

Is there a metric that is simple to calculate and suitable for discrete / continuous distributions? → **Wasserstein distance**

Discrepancy-based DRO

used to measure the distance between two distributions.

Definition:

$$d_W(P_1, P_2) = \inf_{\gamma \sim \prod(P_1, P_2)} \mathbb{E}_{(x,y)\sim\gamma}\,[||x - y||]$$

$\prod(P_1, P_2)$: the set of all possible joint distributions of P_1 and P_2.

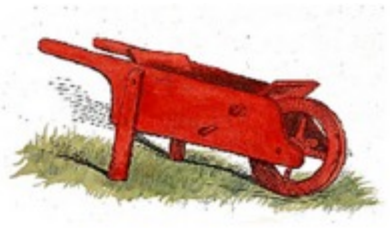
$(x, y)\sim\gamma$: samples under joint distribution γ

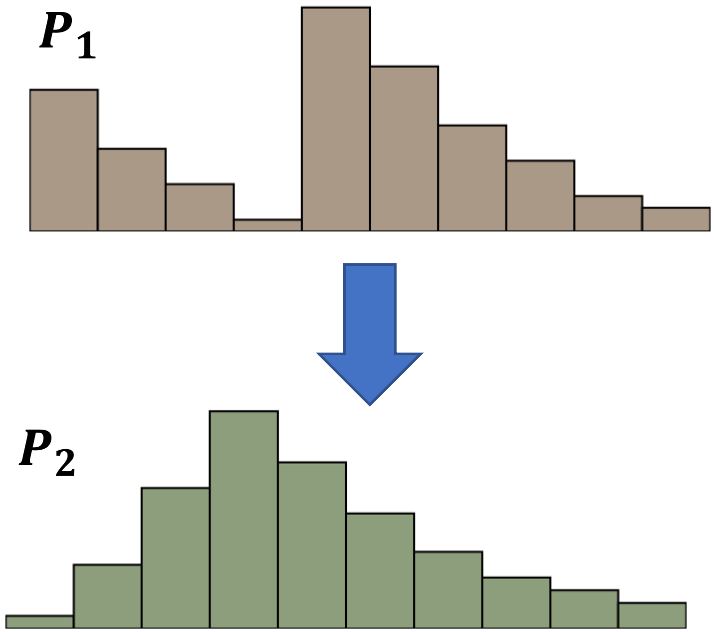
$||x - y||$: sample distance

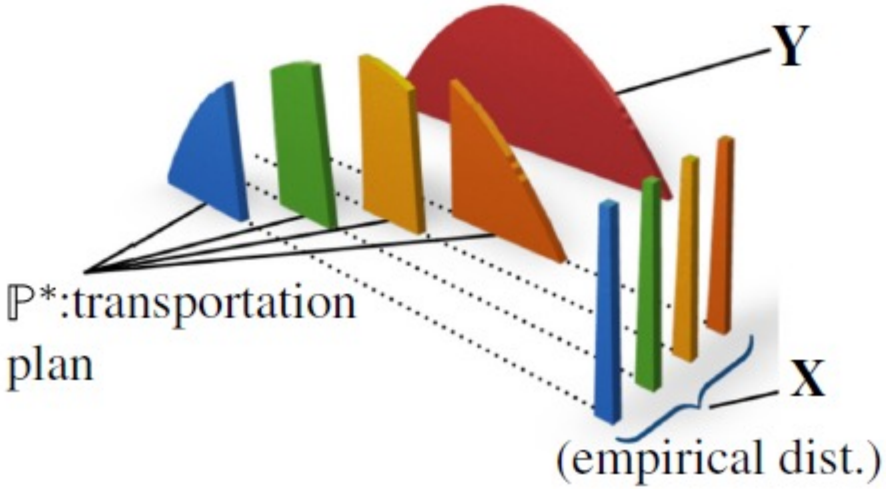
$\mathbb{E}_{(x,y)\sim\gamma}[||x - y||]$: expectation of distance for sample x and y under joint distribution γ

Wasserstein distance of P_1 and P_2: the lower bound of this expectation.

Discrepancy-based DRO







$\prod(P_1, P_2)$: transportation plan

$\|x - y\|$: distance the soil moves

$Y(x, y)$: amount of moving soil from x to y

$\mathbb{E}_{(x,y)\sim\gamma}[\|x - y\|]$: bulldozing cost

Move mass P_1 into the shape and position of P_2.

Bulldozing cost : amount of moving soil multiplied by the distance the soil moves.

Wasserstein distance: the smallest bulldozing cost from P_1 to P_2.

Discrepancy-based DRO

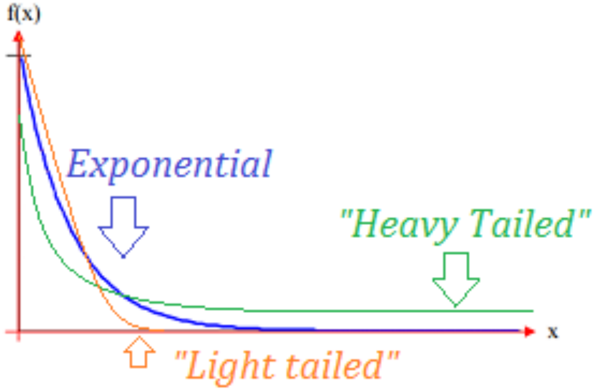
Wasserstein distance-based ambiguity set:

$$\mathbb{B}_{\varepsilon}\left(\widehat{P_N}\right) = \{Q: d_W\ (\widehat{P_N}, Q) \leq \varepsilon\}$$

- The ambiguity set Q can be viewed as a Wasserstein ball which contains all probability distributions whose Wasserstein distance to the empirical distribution $\widehat{P_N}$ is less than ε.
- Q will cover the true distribution with a higher probability with a larger value of ε.
  - There exists a trade-off between the accuracy and the complexity
  - It is important to well design the value of ε

Discrepancy-based DRO

How to calculate ε of ambiguity set



Light-tailed distribution assumption: Distribution $\mathbb{P}$ is call light-tailed i $a > 1$ such that

$$\mathbb{E}^{\mathbb{P}}\left[\exp(\|\xi\|^a)\right] = \int_{\Xi} \exp(\|\xi\|^a)\, \mathbb{P}(\mathrm{d}\xi) < \infty.$$

This assumption requires the tail of the distribution to decay at an exponential rate. The assumption guarantees that the ambiguity set can cover most of the possible distributions.

Radius selection: With this assumption, suppose that $\widehat{\mathbb{P}_N}$ is the empirical distribution, m is related to the dimension and cost parameter, for $m \neq 2$ and $c_1, c_2 > 0$, under a confidence level of $1-\beta$, we have

$$\varepsilon_N(\beta) := \begin{cases} \left(\frac{\log(c_1\beta^{-1})}{c_2 N}\right)^{1/\max\{m,2\}} & \text{if } N \geq \frac{\log(c_1\beta^{-1})}{c_2}, \\ \left(\frac{\log(c_1\beta^{-1})}{c_2 N}\right)^{1/a} & \text{if } N < \frac{\log(c_1\beta^{-1})}{c_2}. \end{cases}$$

Number of samples

Discrepancy-based DRO

With Wasserstein ball, the DRO problem can be rewritten as

$$\min_{x \in \chi} \max_{P \in \mathbb{B}_\varepsilon(\widehat{P_N})} E_P[h(x,\xi)]$$

Assumption:

- The uncertainty set is convex and closed
- $h(x,\xi)$ is convex with respect to ξ

Challenges to Compute Wasserstein Distances:

Computing the Wasserstein distance between two distributions $\widehat{P_N}$ and P_N is NP-hard if one of them is continuous since the dimension will be infinite.

Solution:

Rewrite the problem into a finite-dimensional convex program by leveraging tools from robust optimizations

Discrepancy-based DRO

How to transform an infinite-dimensional optimization problem into a finite-dimensional convex program:

$$\max_{P\in\mathbb{B}_\rho(\widehat{P_N})} E_P[h(x,\xi)] = \begin{cases} \sup_{\boldsymbol{\Pi},\boldsymbol{P}} \int_{\Xi} \boldsymbol{h}(x,\xi)\boldsymbol{P}(\boldsymbol{d\xi}) \\ \boldsymbol{s.t.} \quad \int_{\Xi^2} \|\boldsymbol{\xi}-\boldsymbol{\xi}'\| \boldsymbol{\Pi}(\boldsymbol{d\xi},\boldsymbol{d\xi}') \le \boldsymbol{\varepsilon} \end{cases}$$

$$\widehat{P_N} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\widehat{\xi}_i}$$

Convex reduction

$$= \begin{cases} \sup_{\boldsymbol{P_i}\in\mathcal{M}(\Xi)} \frac{\boldsymbol{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \int_{\Xi} \boldsymbol{h}(x,\xi)\boldsymbol{P_i}(\boldsymbol{d\xi}) \\ \boldsymbol{s.t.} \quad \frac{\boldsymbol{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \int_{\Xi} \|\boldsymbol{\xi}-\widehat{\boldsymbol{\xi}}_{\boldsymbol{i}}\| \boldsymbol{P_i}(\boldsymbol{d\xi}) \le \boldsymbol{\varepsilon} \end{cases}$$

$\boldsymbol{\Pi}$ is a joint distribution of $\boldsymbol{\xi}'$ and $\boldsymbol{\xi}$ with marginals distribution P and $\widehat{\boldsymbol{P_N}}$ of $\boldsymbol{\xi}'$ given $\boldsymbol{\xi}' = \widehat{\boldsymbol{\xi}}_{\boldsymbol{i}}$,and conditional $\boldsymbol{P_i}$ of $\boldsymbol{\xi}$. N is number of samples. Due to the law of total probability, we have

$$\Pi = \frac{1}{N}\sum_{i=1}^{N} \delta_{\widehat{\xi}_i} \otimes P_i$$

Discrepancy-based DRO

Using a standard duality argument, we obtain

$$\max_{P\in\mathbb{B}_\varepsilon(\widehat{P_N})} E_P[h(x,\xi)]$$

$$= \sup_{P_i\in\mathcal{M}(\Xi)} \inf_{\lambda\geq 0} \frac{1}{N}\sum_{i=1}^{N}\int_{\Xi} h(x,\xi)P_i\,(d\xi) + \lambda\left(\varepsilon - \frac{1}{N}\sum_{i=1}^{N}\int_{\Xi} \|\xi - \hat{\xi}_i\| P_i(d\xi)\right)$$

$$\leq \inf_{\lambda\geq 0} \sup_{P_i\in\mathcal{M}(\Xi)} \lambda\varepsilon + \frac{1}{N}\sum_{i=1}^{N}\int_{\Xi} \left(h(x,\xi) - \lambda\|\xi - \hat{\xi}_i\|\right)P_i\,(d\xi)$$

$$= \inf_{\lambda\geq 0} \lambda\varepsilon + \frac{1}{N}\sum_{i=1}^{N}\sup_{\xi\in\Xi}\left(h(x,\xi) - \lambda\|\xi - \hat{\xi}_i\|\right)$$

Maximum minima is always less than minimum maxima

The uncertainty set contains all distributions

Discrepancy-based DRO

Introducing epigraphical auxiliary variables s_i, $i \leq N$:

$$\max_{P \in \mathbb{B}_\varepsilon(\widehat{P_N})} E_P[h(x,\xi)] = \begin{cases} \inf_{\lambda, s_i} \boldsymbol{\lambda\varepsilon} + \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{s_i} \\ \boldsymbol{s.t.} \sup_{\xi \in \Xi} (\boldsymbol{h}(\boldsymbol{x},\xi) - \lambda\|\xi - \hat{\xi}_{\boldsymbol{i}}\|) \leq \boldsymbol{s_i}, \ \forall \boldsymbol{i} \leq \boldsymbol{N} \\ \boldsymbol{\lambda} \geq \mathbf{0} \end{cases}$$

As such, the problem is transformed into a finite convex program and can be solved by existing convex optimization techniques.

Discrepancy-based DRO

□□□□□□C□□□

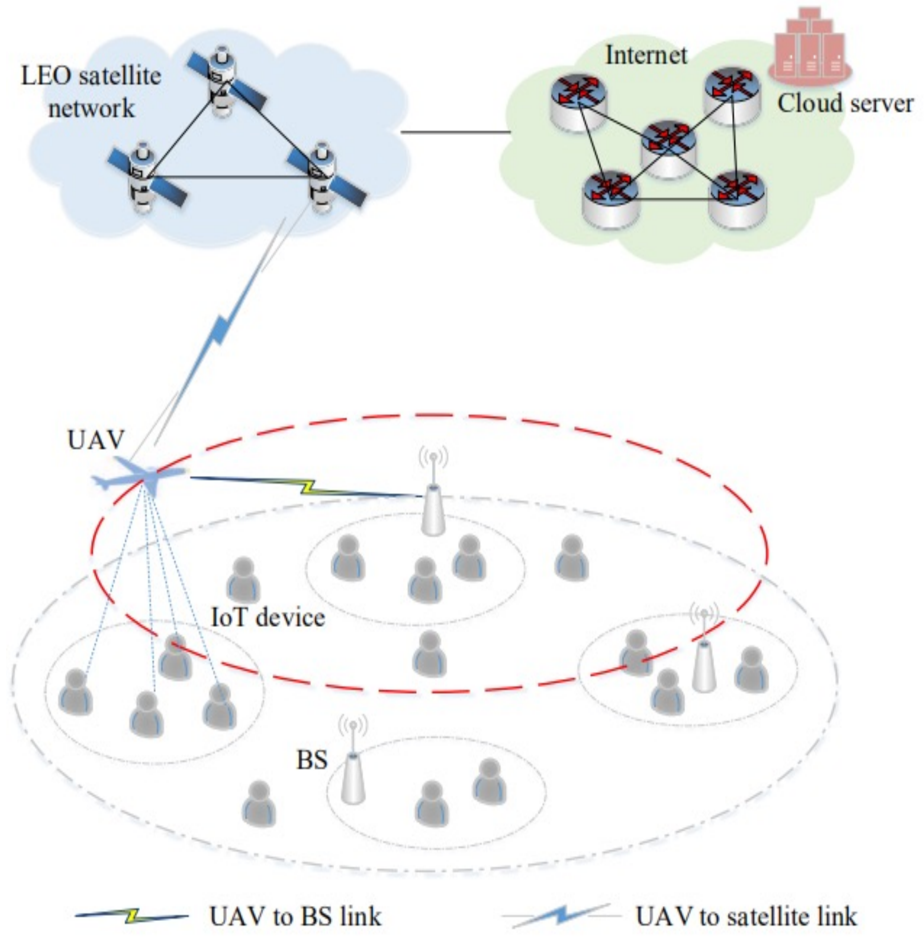
For the case of quadratic loss (possibly non-convex and non-concave), we still can derive its exact form.

Assume that $x(\xi) = \xi^T Q\xi + 2q^T\xi$ is a (possibly indefinite) quadratic loss function. Then the worst-case risk coincides with the optimal value of a tractable semidefinite program (SDP), that is,

$$\max_{P \in \mathbb{B}_\varepsilon(\widehat{P_N})} E_P[h(x,\xi)] = \begin{cases} \inf\limits_{\gamma, \boldsymbol{s_i}} \quad \gamma\boldsymbol{\varepsilon^2} + \frac{\boldsymbol{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{s_i} \\ \boldsymbol{s.t.} \begin{pmatrix} \gamma - Q & q + \gamma\widehat{\boldsymbol{\xi_i}} \\ q^T + \gamma\widehat{\boldsymbol{\xi_i}}^{\boldsymbol{T}} & \boldsymbol{s_i} + \|\widehat{\boldsymbol{\xi_i}}\|_2^2 \end{pmatrix} \geq 0 \end{cases}$$

Application 1: Computation Offloading

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING



Space-Air-Ground Integrated Networks:

- IoT devices: request services
- UAV: collect data from IoT devices
  - Determine to offload to a nearby BS or offload to a certain satellite and utilize the cloud server
  - Determine the proportion of tasks to offload, and the others will be done by the UAV
- The task request received by the UAV is uncertain

Yali Chen, Bo Ai, Yong Niu, Hongliang Zhang, and Zhu Han, "Energy-Constrained Computation Offloading in Space-Air-Ground Integrated Networks using Distributionally Robust Optimization," IEEE Transactions on Vehicular Technology (Volume: 70, Issue: 11, Nov. 2021)

Application 1: Problem Formulation

Objective: minimize the expected total latency of T time slots under the worst-case distribution realization in uncertainty set D

$$\min_{\mathbf{X},\mathbf{Y}} \max_{P \in D} \mathbb{E}_P[\psi(\mathbf{X}, \mathbf{Y}, \xi_k)]$$

$s.t.$

$$(a)\ x_{bt} \in \{0,1\},\ x_{st} \in \{0,1\},\ \forall b, s, t$$

$$(b)\ \sum_{b=1}^{N} x_{bt} + \sum_{s=1}^{M} x_{st} = 1,\ \forall t$$

$$(c)\ y_{ut}(\xi_k) + \sum_{b=1}^{N} y_{bt}(\xi_k) + \sum_{s=1}^{M} y_{st}(\xi_k) = \xi_k,\ \forall t, k$$

the sum of tasks processed by UAV, all BSs, and all LEO satellites equals to the amount of tasks arrived

$$(d)\ y_{ut}(\xi_k), y_{bt}(\xi_k), y_{st}(\xi_k) \geq 0,\ \forall b, s, t, k$$

$$(e)\ y_{ut}(\xi_k) \leq C_u, y_{bt}(\xi_k) \leq C_b x_{bt}, y_{st}(\xi_k) \leq C_s x_{st},\ \forall b, s, t, k$$

ensure the tasks assigned cannot exceed its capacity

$$(f)\ \psi(\mathbf{X}, \mathbf{Y}, \xi_k) = \sum_{t=1}^{T} \max\left(\frac{\delta y_{ut}(\xi_k)}{f_u}, \sum_{b=1}^{N} L_{ubt} + \sum_{s=1}^{M} L_{ust}\right),$$

total system latency

$$(g)\ E_u^{fly} + \sum_{b=1}^{N}\sum_{t=1}^{T} E_{ubt} + \sum_{s=1}^{M}\sum_{t=1}^{T} E_{ust} \leq E_{max},\ \forall k$$

energy constraint

Application 1: Algorithm and Results

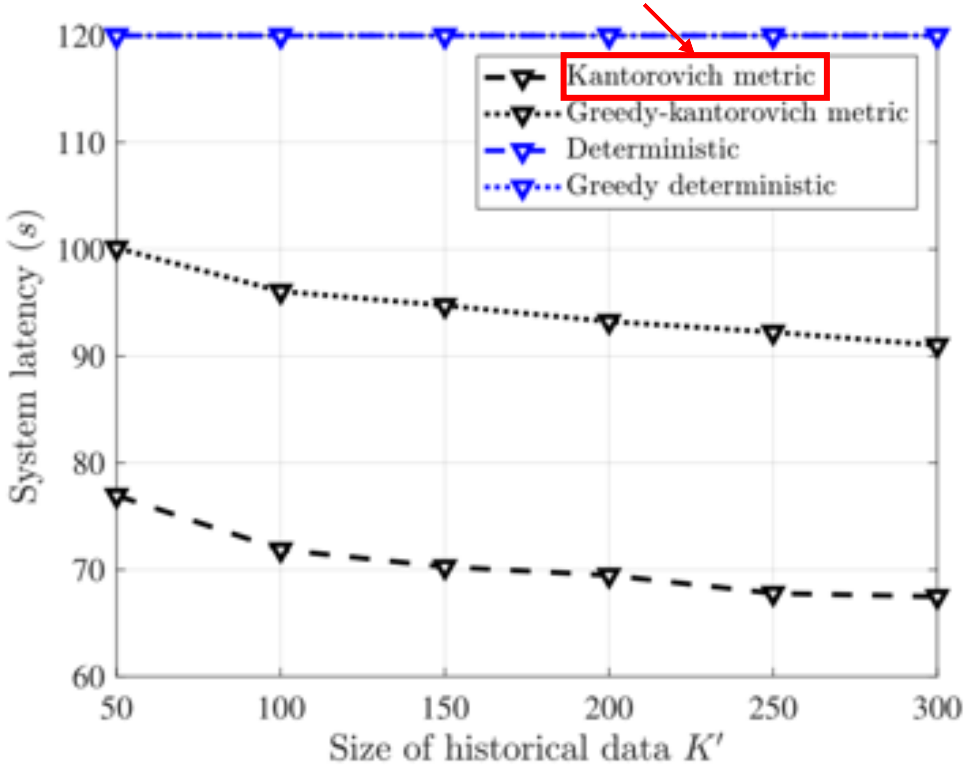
Algorithm 1 Distributionally Robust Latency Optimization Algorithm

Input: Sample space Ω, reference distribution P_0 based on a series of historical data

Output: $\mathbf{X}, \mathbf{Y}$

Solve the optimization problem (28) by the Gurobi Optimizer to obtain $P = \{p_1, p_2, ..., p_K\}$;
Solve the optimization problem (26) by the Gurobi Optimizer to get continuous $\mathbf{X}$ and $\mathbf{Y}$;
repeat
  Select the branch variable x^* according to the rule shown in (34);
  Add $x \leq 0$ as a constraint to the problem (26) and solve it to obtain the optimal result $Lat0$;
  Add $x \geq 1$ as a constraint to the problem (26) and solve it to obtain the optimal result $Lat1$;
  if $Lat0 < Lat1$ **then**
    Update problem (26) by adding a constraint $x \leq 0$;
  else
    Update problem (26) by adding a constraint $x \geq 1$;
  end if
until all elements in $\mathbf{X}$ are integers.

The same as Wasserstein metric



For deterministic scheme, many tasks cannot be processed and retransmission is required, leading to a higher latency

Outline

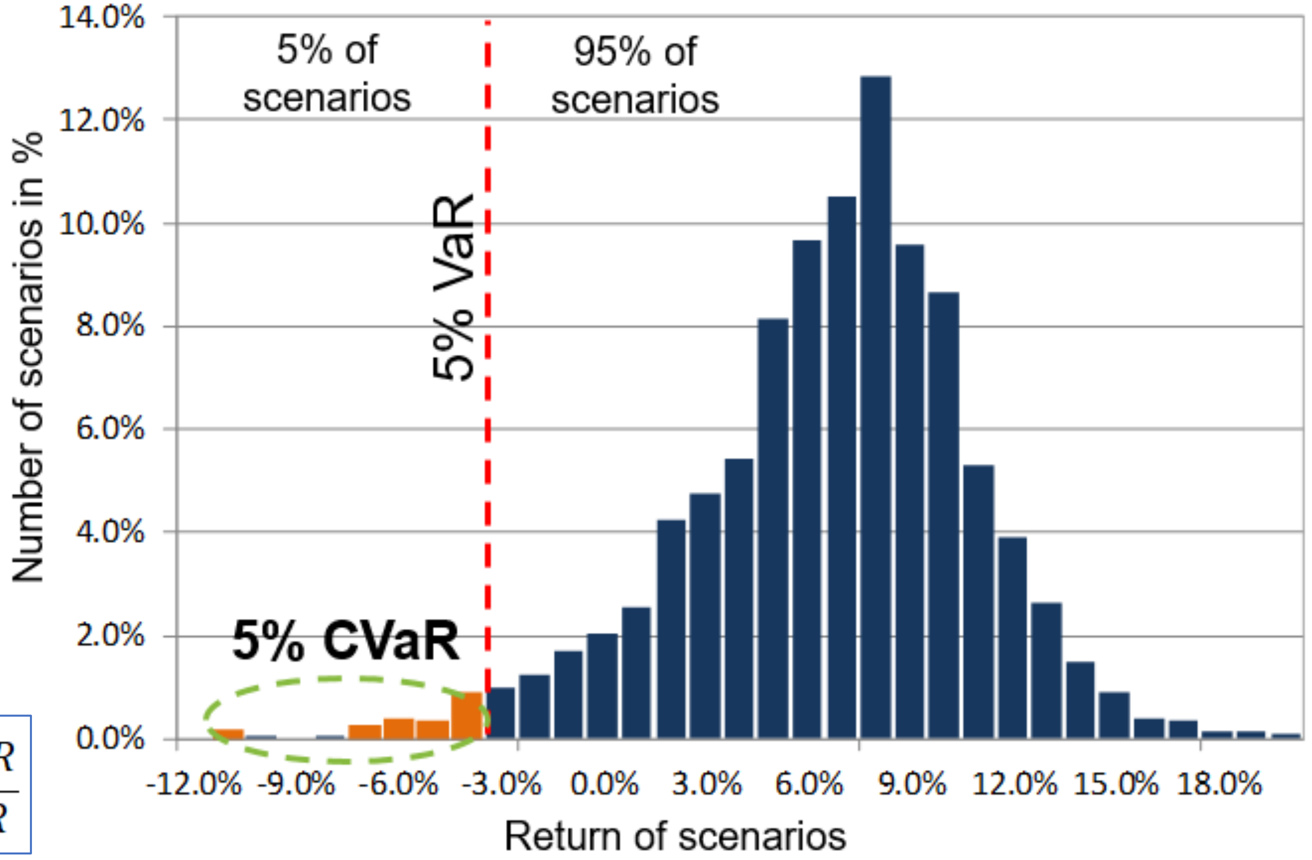
- Introduction
- C□1: D□-□D□ □ □
- C□2: D□, □A□ □C□
  - ✓ □-□-□(□) □C□-□-□(C□)
  - ✓ □ □C□ □ □ □
  - ✓ A□;A□ □
- C□3: D□ □ □ □
- C□

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

VaR& CVaR Representation

- [illegible]
- [illegible]

Risk functions: graphical representation of VaR, VaR Deviation, CVaR, CVaR Deviation, Max Loss, and Max Loss Deviation.



$$5\%\ CVaR = \frac{\sum return\ of\ scenarios < 5\%\ VaR}{number\ of\ scenarios < 5\%\ VaR}$$

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

VaR & CVaR

Let ***X*** be a random variable, with the *X* may have meaning of loss or gain.

$$F_X(z) = P\{X \le z\}$$

VaR Definition A lower α-percentile of the random variable X.

$$\mathrm{VaR}_\alpha(X) = \min\{z \mid F_X(z) \ge \alpha\}$$

$VaR_\alpha(X)$ is nonconvex and discontinuous function of the confidence level *α* for discrete distributions.

Difficult to control/optimize for non-normal distributions: VaR has many extremums for discrete distributions.

VaR& CVaR

CVaR Definition

The CVaR of X with confidence level $\alpha \in [0, 1]$ is the mean of the generalized α-tail distribution:

$$CVaR_\alpha(X) = \int_{-\infty}^{+\infty} z dF_X^\alpha(z)$$

where the distribution in question is the one with distribution function defined by

$$F_X^\alpha(z) = \begin{cases} 0 & \text{when } z < VaR_\alpha(X) \\ \dfrac{F_X(z) - \alpha}{1-\alpha} & \text{when } z \geq VaR_\alpha(X) \end{cases}$$

For random variables with continuous distribution functions, $CVaR_\alpha(X)$ equals the conditional expectation of X subject to $Z \geq VaR_\alpha(X)$.

CVaR is continuous with respect to α. CVaR is convex in X

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

VaR& CVaR Example

Main focus: **Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) [Rockafellar and Uryasev, 2000].**

if VAR(95) = 3%

5 = 100 - 95

5% chance to lose 3% or more on a given day

if CVAR(95) = 4.5%

5 = 100 - 95

In the worst 5% of returns

Your average loss will be 4.5%

CVaR gives us an average expected loss
VaR gives us a range of potential losses

- If more than lost 3% happens with 5%, but we still do not know average lost.
- CVAR describes the average lost conditioned on VaR happens.

Equivalence of Chance Constraints and VaR Constraints

In portfolio management, often it is required that portfolio loss at a certain future time is, with high reliability, at most equal to a certain value.

Let $f(x,y)$ be the loss associated with the decision vector x and the random vector y.

$$Prob\{f_i(x,y)\} \leq 0\} \geq p_i \quad , i = 1, \dots, m.$$

Let $VaR_\alpha\ (x)$ be the VaR_α of a loss function $f(x,y)$

$$VaR_\alpha(x) = \min\{\zeta : Prob\{f(x,y) \leq \zeta\} \geq \alpha\}$$

Then the following holds:

$$Prob\{f(x,y) \leq \zeta\} \geq \alpha \Leftrightarrow Prob\{f(x,y) > \zeta\} \leq 1 - \alpha \Leftrightarrow VaR_\alpha(\mathrm{x}) \leq \zeta$$

In general, $VaR_\alpha\ (x)$ is *nonconvex w.r.t.* x, (e.g., discrete distributions)

Minimization and CVaR Constraints

The underlying probability distribution of y will be assumed for convenience to have density $p(y)$. The α-CVaR of the loss associated with a decision x is the value

$$CVaR(x) = \frac{1}{1-\alpha}\int_{f(x,y)\geq VaR(x)} f(x,y)\,p(y)\,dy$$

The main idea is to define a function that can be used in connection with VaR and CVaR

$$F_\alpha(x,\zeta) = \zeta + \frac{1}{1-\alpha}\int_{y\in\mathbb{R}^m}[f(x,y)-\zeta]^+\,p(y)\,dy$$

$$F_\alpha(x,\zeta) = \zeta + \frac{1}{1-\alpha}E\{[f(x,y)-\zeta]^+\}, \quad \text{where } [t]^+ = \max\{0,t\}.$$

1. $F_\alpha(x,\zeta)$ is convex with respect to α;
2. $\mathrm{VaR}_\alpha(x)$ is a minimum point of function $F_\alpha(x,\zeta)$ w.r.t. ζ;
3. As a function of $\zeta \in \mathbb{R}$, $F_\alpha(x,\zeta)$ is finite and convex (hence continuous), :

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

Chance constrained programs

It is known that **distributionally robust chance constraints** can be conservatively approximated by Worst-Case Conditional Value-at-Risk (CVaR) constraints.

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & Pr\left(a_i(\tilde{\xi})^\top x \le b_i(\tilde{\xi}) \quad \forall i = 1, \ldots, m\right) \ge 1 - \epsilon \\ & x \in \mathcal{X}, \end{aligned}$$

x is the decision vector and c is a cost vector

Pr can be unknown arbitrary distribution

The *chance constraint* requires a set of m uncertainty-affected inequalities to be jointly satisfied with a probability of at least $1 - \epsilon$.

$$a_i(\tilde{\xi}) = a_i^0 + \sum_{j=1}^{k} a_i^j \tilde{\xi}_j \quad \text{and} \quad b_i(\tilde{\xi}) = b_i^0 + \sum_{j=1}^{k} b_i^j \tilde{\xi}_j.$$

$$y_i^j(x) = (a_i^j)^\top x - b_i^j, \quad i = 1, \ldots, n, \quad j = 0, \ldots, k.$$

$$Pr\left(y_i^0(x) + y_i(x)^\top \tilde{\xi} \le 0 \quad \forall i = 1, \ldots, m\right) \ge 1 - \epsilon$$

Distributionally robust approach

A natural way to *immunize* the chance constraint against uncertainty in the probability distribution is to adopt the **following ambiguous or distributionally robust chance constraint.**

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\left(y_i^0(x) + y_i(x)^\top \tilde{\xi} \leq 0 \quad \forall i = 1, \ldots, m \right) \geq 1 - \epsilon \\ & x \in \mathcal{X}, \end{aligned}$$

$\mathbb{P} \in \mathcal{P}$ denotes the distribution $\mathbb{P}$ belongs to an uncertainty set $\mathcal{P}$ with certain known structural properties.

Distributionally robust chance constraint means **the worst case** satisfies the probability at least $1 - \epsilon$ in the presence of channel uncertainties.

For *m* = 1, the feasible set is denoted by

$$\mathcal{X}^{\text{ICC}} = \left\{ x \in \mathbb{R}^n \; : \; \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\left(y^0(x) + y(x)^\top \tilde{\xi} \leq 0 \right) \geq 1 - \epsilon \right\}$$

CVaR chance constraints

The CVaR at ϵ level with respect to $\mathbb{P}$ is defined as

$$\mathbb{P}\text{-CVaR}_\epsilon(L(\tilde{\xi})) = \inf_{\beta \in \mathbb{R}} \left\{ \beta + \frac{1}{\epsilon} \mathbb{E}_{\mathbb{P}} \left((L(\tilde{\xi}) - \beta)^+ \right) \right\}$$

CVaR essentially evaluates the conditional expectation of loss above the $(1-\epsilon)$-quantile of the loss distribution. It can be shown that CVaR represents a convex functional of the random variable $L(\tilde{\xi})$

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\text{-CVaR}_\epsilon \left(y^0(x) + y(x)^\top \tilde{\xi} \right) \leq 0 \implies \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{P} \left(y^0(x) + y(x)^\top \tilde{\xi} \leq 0 \right) \geq 1 - \epsilon.$$

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & \boldsymbol{Pr}\left(a_i(\tilde{\xi})^\top x \leq b_i(\tilde{\xi}) \quad \forall i = 1, \ldots, m \right) \geq 1 - \epsilon \\ & x \in \mathcal{X}, \end{aligned}$$

$\Rightarrow$

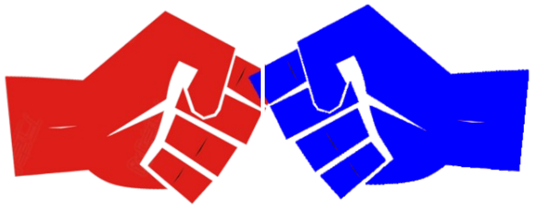
$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}\text{-CVaR}_\epsilon \left(y^0(x) + y(x)^\top \tilde{\xi} \right) \leq 0 \end{aligned}$$

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

Age of Information Minimization in Healthcare IoT Using Distributionally Robust Optimization [2]

How can I deal with channel state information (CSI) errors in Healthcare IoT system ?

- The perturbations in CSI are modeled to be statistically unbounded according to **Gaussian distribution** and **Rayleigh distribution.**
- These statistical channel assumptions **may not match** the healthcare IoT applications perfectly.



- The **retransmission scheme** is investigated against CSI errors to reduce the AoI.
- The retransmission will lead to the **energy consumption** of the IoT device **increase dramatically**, especially when the IoT device is used to retransmit the same updates for infinite times

A key open problem is to consider the case of imperfect CSI and investigate how CSI error effects may be mitigated through quantitative designs.

[2] Z. Ling, F. Hu, H. Zhang and Z. Han, "Age of Information Minimization in Healthcare IoT Using Distributionally Robust Optimization," *IEEE Internet of Things Journal*, to be appeared.

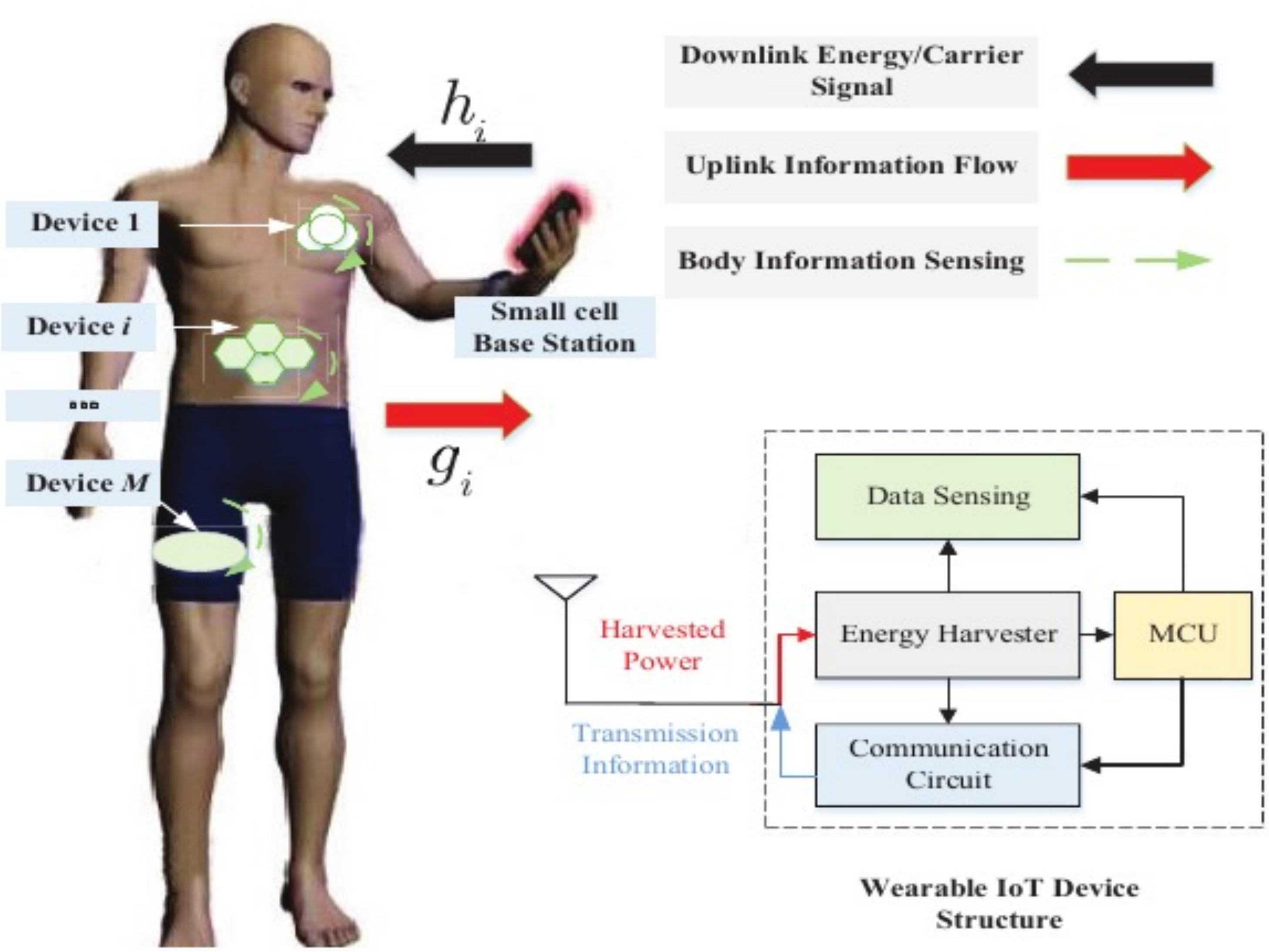
Application 2: Body Area Networks

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

System model

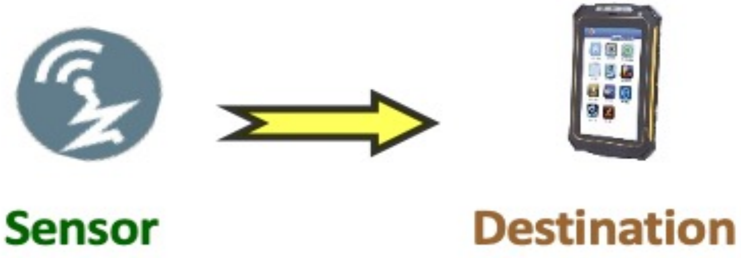
As shown in this picture, we consider **a single-hop cellular IoT** to support healthcare IoT applications:

1. the wearable IoT devices **harvest wireless RF energy** from a dedicated power cell BS
2. the wearable IoT devices send **their sensing physiological information signals** to a separate information receiving cell BS.



Application: Age of Information

Age of Information= Time – Timestamp



Measure the “age of information” that destination know about a source node.

Potential application: UAV, Uplink, Down link system in wireless communication IoT network, etc.

Evolutions of AoI function.

Application 2: Problem Formulation

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

$$\min_{\substack{\alpha_i(t),p_i(t),\gamma_i(t)_{\max},\tau_i(t)_{\max}\\ M_1,M_2,\beta_1,\beta_2}} \quad \sup_{\mathbb{P}\in\mathcal{P}} \bar{\Delta}_i(t)$$

$$s.t. \quad \begin{cases} \beta_1 + \frac{1}{\gamma_i(t)_{\max}}\mathrm{Tr}(\mathbf{\Omega}_1\mathbf{M}_1) \le 0, \\ \mathbf{M}_1 \in \mathbb{S}^2, \beta_1 \in \mathbb{R}, \\ \mathbf{M}_1 \succcurlyeq \mathbf{0}, \mathbf{M}_1 - \begin{bmatrix} q_1 & 0 \\ 0 & E_{C,i} - \beta_1 \end{bmatrix} \succcurlyeq \mathbf{0}, \end{cases}$$

CVaR constraint.

$$\begin{cases} \beta_2 + \frac{1}{\tau_i(t)_{\max}}\mathrm{Tr}(\mathbf{\Omega}_2\mathbf{M}_2) \le 0, \\ \mathbf{M}_2 \in \mathbb{S}^3, \beta_2 \in \mathbb{R}, \\ \mathbf{M}_2 \succcurlyeq \mathbf{0}, \mathbf{M}_2 - \begin{bmatrix} \boldsymbol{Q} & \mathbf{0} \\ \mathbf{0} & r_{U,\min}\sigma_A^2 - \beta_2 \end{bmatrix} \succcurlyeq \mathbf{0}, \end{cases}$$

CVaR constraint.

$$0 < \alpha_i(t) < \alpha_{max}$$

$$0 < p_i(t) \le P_{\max},$$

Power constraint

$$0 < \gamma_i(t) < 1,$$

Outage probability constraint for energy harvesting.

$$0 < \tau_i(t) < 1,$$

Outage probability constraint for transmission

$$\forall(\Delta_{h_i}, \Delta_{g_i}) \sim \mathbb{P} \in \mathcal{P},$$

$$\forall i \in \{1,2,\ldots,M\}, \forall t \in \{1,2,\ldots,K\},$$

Application 2: CVAR Algorithm

To get the optimal average AoI, we propose a **low-complexity upper bound of AoI minimization (UBAM)** algorithm in an iterative manner to address the **distributionally robust optimization problem**.

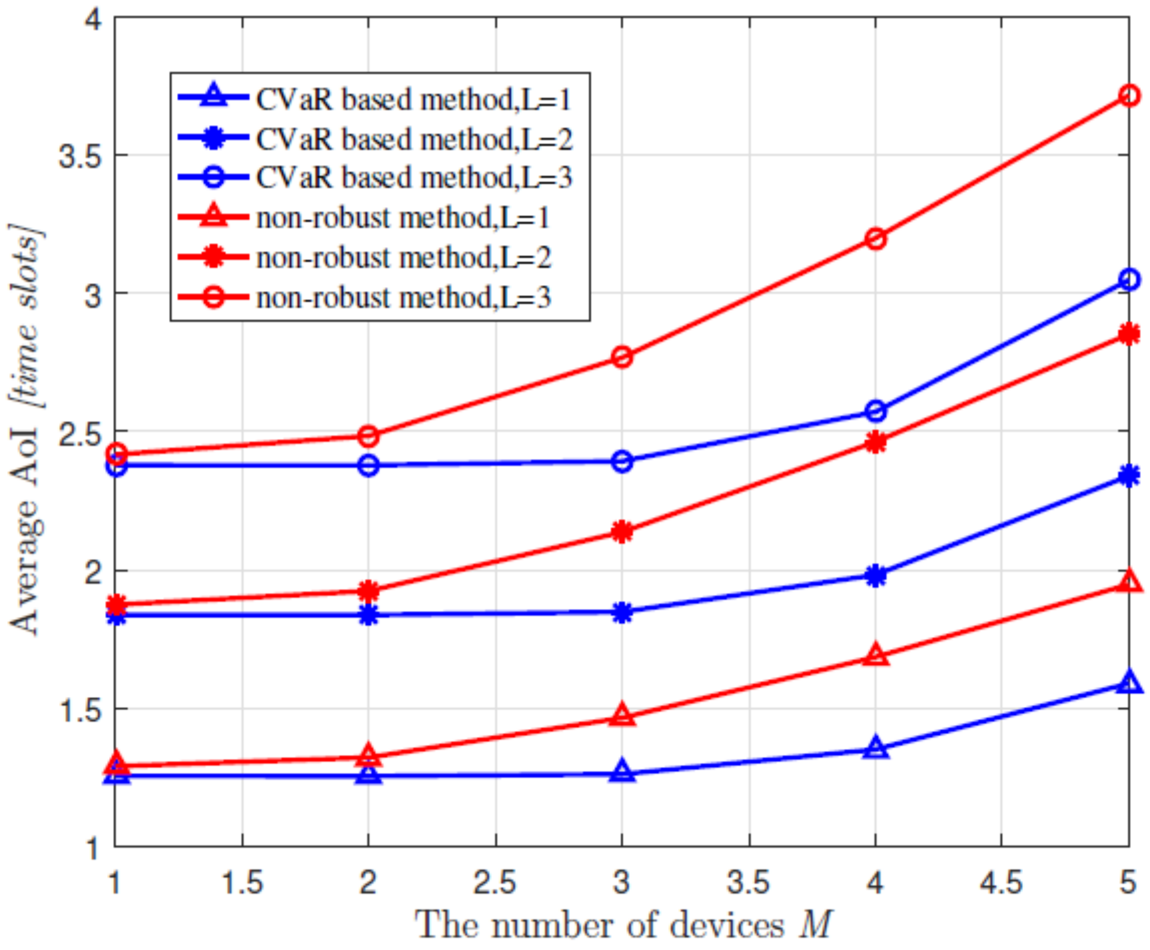
Algorithm 1 UBAM Algorithm

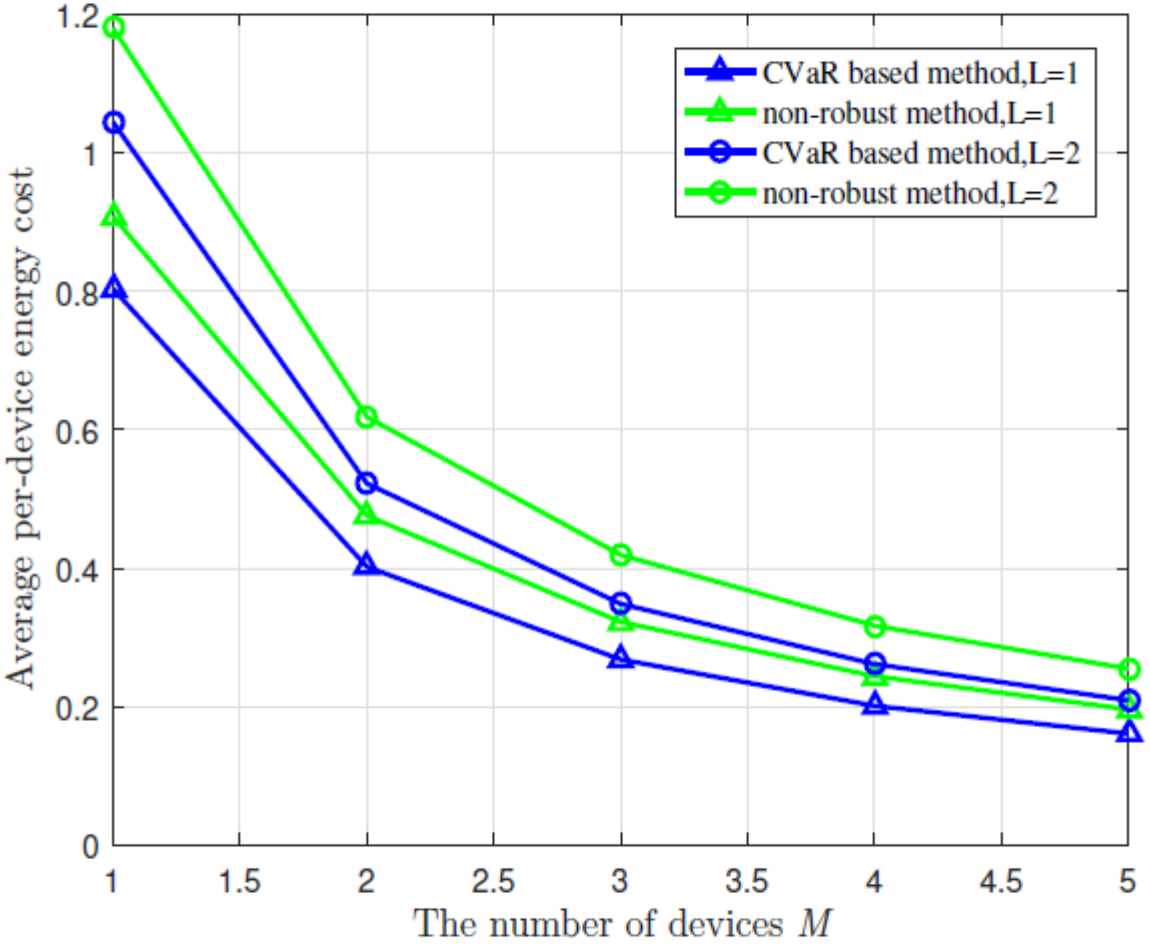
Input: Initialize $\boldsymbol{\alpha}^{\{0\}}$, $\boldsymbol{p}^{\{0\}}$, $\boldsymbol{\tau}^{\{0\}}$, $\gamma^{\{0\}}$, the terminated threshold $\Theta = 10^{-3}$, and $j = 0$.

repeat
    $j = j + 1$.
    Solve the power allocation subproblem for given $\gamma^{\{j\}}$ and $\boldsymbol{\tau}^{\{j\}}$, to obtian the feasible solution as $\boldsymbol{\alpha}^{\{j+1\}}$ and $\boldsymbol{p}^{\{j+1\}}$.
    Solve the outage probability subproblem for given $\boldsymbol{\alpha}^{\{j+1\}}$ and $\boldsymbol{p}^{\{j+1\}}$, to obtian the optimal solution as $\gamma^{\{j+1\}}$ and $\boldsymbol{\tau}^{\{j+1\}}$.
until $\left| \sup_{\mathbb{P}\in\mathcal{P}} \bar{\Delta}_i(t)^{\{j+1\}} - \sup_{\mathbb{P}\in\mathcal{P}} \bar{\Delta}_i(t)^{\{j\}} \right|$ satisfies termination criterion Θ.
return $\boldsymbol{p}^{\{j-1\}}$, $\boldsymbol{\alpha}^{\{j-1\}}$, $\gamma^{\{j-1\}}$, and $\boldsymbol{\tau}^{\{j-1\}}$.

Output: The optimal solution $\boldsymbol{\alpha}^*$, $\boldsymbol{p}^*$, γ^*, and $\boldsymbol{\tau}^*$.

Application 2: Results





The energy harvesting and information transmission successful opportunities for each link become lower with more wearable IoT devices. In this sense, we find an **AoI-energy tradeoff** in the healthcare IoT system.

Outline

- Introduction
- C[illegible]1: D[illegible]-[illegible]D[illegible]
- C[illegible]2: D[illegible], [illegible]A[illegible], [illegible]C[illegible]
- C[illegible]3: D[illegible] B[illegible]
  - ✓ D[illegible]
  - ✓ A[illegible];[illegible]
- C[illegible]

Motivation

- Previous works considering robustness in Reinforcement Learning
  - Mainly focus on **uncertain environment** for searching strategies to achieve best performance
    - ✓ Robust reinforcement learning
      - Guard against common systematic perturbation situations
    - ✓ Robust Adversarial Reinforcement Learning
      - Conquer rare, catastrophic events
- DRO RL
  - Not only consider uncertainty in environment
  - Limited collected samples when training
  - → Affect how to transit to new state based on current state and action
    - ✓ **Estimation error in Policy Iteration**

$$\pi_{t+1} \in \mathcal{G}(\tilde{V}_t)$$

$$\tilde{V}_{t+1} = \mathcal{T}^{\pi_{t+1}} \tilde{V}_t + \delta_t$$

Estimation error

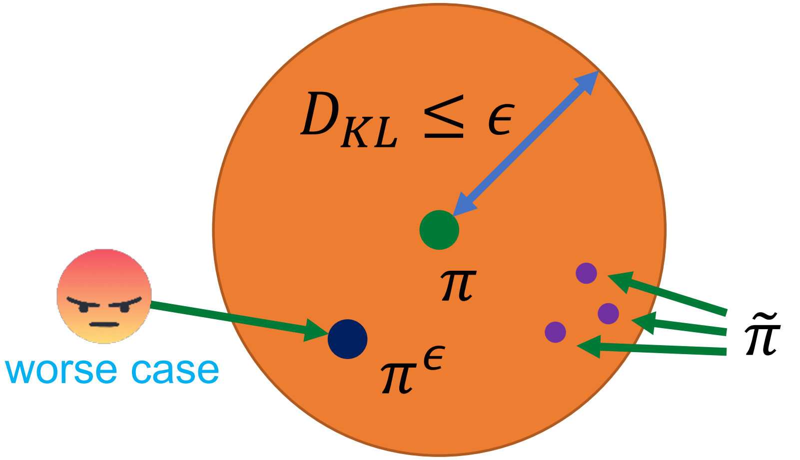
  - Concentrate on learning process for agent itself
    - ✓ Conservative policy in unknown environment
    - ✓ Optimistic policy in familiar environment

Distributionally robust policy evaluation

- Uncertainty set:

$$\mathcal{U}_\epsilon(\pi) := \{\tilde{\pi} \in \Delta_{\mathcal{A}}^{\mathcal{S}} \mid D_{\mathrm{KL}}(\tilde{\pi}(\cdot|s)\|\pi(\cdot|s)) \le \epsilon(s),\ \forall s \in \mathcal{S}\}$$

Considered possible policies ($\tilde{\pi}$)   Best policy ($\pi(\cdot|s)$)

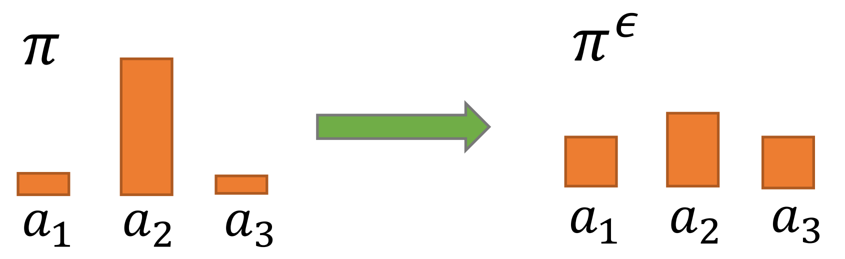


- Adversarial Bellman operator:
  - Evaluate with an adversarial policy $\mathcal{T}^{\pi^\epsilon} V := \min_{\tilde{\pi} \in \mathcal{U}_\epsilon(\pi)} \mathcal{T}^{\tilde{\pi}} V.$
- Rewrite the maximum problem by

$$[\mathcal{T}^{\pi^\epsilon} V](s) = \max_{\lambda(s)>0} \min_{\tilde{\pi} \in \Delta(\mathcal{A})} [\mathcal{T}^{\tilde{\pi}} V](s)$$
$$+ \lambda(s) D_{\mathrm{KL}}(\tilde{\pi}(\cdot|s)\|\pi(\cdot|s)) - \lambda(s)\epsilon(s)$$
$$= \min_{\lambda(s)>0} \max_{\pi \in \Delta(\mathcal{A})} [-\mathcal{T}^{\pi} V](s)$$
$$- \lambda(s) D_{\mathrm{KL}}(\tilde{\pi}(\cdot|s)\|\pi(\cdot|s)) + \lambda(s)\epsilon(s)$$

$$\pi^\epsilon(a|s;\lambda) \propto \exp(-Q_V(s,a)/\lambda(s))\pi(a|s)$$ adversarial policy

$$\lambda^\star(s) := \arg\min_{\lambda(s)>0} \Omega^\star(-Q_V/\lambda(s)) + \lambda(s)\epsilon(s)$$ Adversarial temperature



Re-weight policy action probabilities opposite to Q-values

- Distributionally robust modified policy iteration (MPI)
  - Policy improvement: $\pi_{t+1} \leftarrow \mathcal{G}(\tilde{V}_t)$
  - Policy evaluation: $\tilde{V}_{t+1} \leftarrow (\mathcal{T}^{\pi_{t+1}^{\epsilon_t}})^m \tilde{V}_t$

Regularized Bellman Optimality Operator

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

- *Rewritten of Bellman operator for value function*
  - $[T^{\pi}V](s) \coloneqq E_{a\sim\pi(\cdot|s)}\left[\sum_{t\geq 0}\gamma^t r_t \mid s_0 = s, \pi\right]$

$$= E_{a\sim\pi(\cdot|s)}\left[r(s,a) + \gamma\, E_{s'\sim p(s'|s,a)}[V(s')]\right] \quad Q_V(s,a)$$

$= < \pi(\cdot\,|s), Q_V(s,a) >$ … inner product of the policy and Q-function

- *Regularized Bellman optimality operator*
  - $[T^{\pi,\Omega}V](s) \coloneqq T^{\pi}V - \Omega(\pi)$ is a Fenchel dual function and can be seen as the normal value function minus the baseline.
  - *Set of optimal policies*
    - $\mathcal{G}^{\Omega}(V) := \{\pi : \pi \in \arg\max_{\pi\in\Pi} \mathcal{T}^{\pi,\Omega}V = \nabla\Omega^{\star}(Q_V)\}$

- *Lengendre-Fenchel duality (Conjugate function)*
  - Fenchel dual : $[\Omega^{\star}V](Q_V) = \max_{\pi\in\Delta A}(< \pi, Q_V > -\Omega(\pi))$
  - Gradient form: $[\nabla\Omega^{\star}V](Q_V) = \operatorname{argmax}_{\pi\in\Delta A}(< \pi, Q_V > -\Omega(\pi))$

- Soft distributionally robust modified policy
  $\Omega(\pi(\cdot|s)) = \alpha\mathcal{H}(\pi(\cdot|s))\ \forall s \in \mathcal{S}$
- Distributionally robust soft actor-critic
  $\Omega_{\lambda}(\pi(\cdot|s)) = \lambda\times(\ -D_{\mathrm{KL}}(\pi(\cdot|s)||\mu(\cdot|s)))$

Extension to Entropy-Regularized Policy Maximum Entropy

- Soft distributionally robust modified policy
  - Policy improvement: $\pi_{t+1} \leftarrow \mathcal{G}^{\Omega}(\tilde{V}_t)$
  - Policy evaluation: $\tilde{V}_{t+1} \leftarrow (\mathcal{T}^{\pi^{\epsilon_t}_{t+1},\Omega})^m \tilde{V}_t$
- Soft adversarial Bellman operator

$$\mathcal{T}^{\pi^{\epsilon},\Omega} V := \min_{\tilde{\pi} \in \mathcal{U}_{\epsilon}(\pi)} \mathcal{T}^{\tilde{\pi},\Omega} V,$$

$$\Omega(\pi(\cdot|s)) = \alpha \mathcal{H}(\pi(\cdot|s)) \; \forall s \in \mathcal{S}$$

entropy form

$\mathrm{E}_{a\sim\pi(\cdot|s)} \log[\pi(a|s)]$

  - Here, we take entropy form into consideration to obtain robustness guarantees on exploration process.
- Adversarial entropy-regularized policy

$$\pi^{\epsilon}(a|s;\alpha,\lambda) \propto \exp((1/\alpha - 1/\lambda(s)) Q_V(s,a))$$

  - Recall the adversarial policy $\pi^{\epsilon}(a|s;\lambda) \propto \exp(-Q_V(s,a)/\lambda(s))\pi(a|s)$

$\propto \exp(Q_V(s,a)/\alpha)$

Regularized policy

Extended to Continuous Control

- Apply this distributional robust MPI to the continuous control problem.
- The formulation is intractable and we get the **approximate** formula as follows:
  - Regularized Bellman operator can be written in terms of Fenchel conjugate

$$[\mathcal{T}^{\Omega} V](s) = \Omega^{\star}(Q_V(s,\cdot))$$

  - Regularization parameter

$$\Omega_{\lambda}(\pi(\cdot|s)) := \lambda \Omega(\pi(\cdot|s))$$
$$= \lambda \times (-D_{\mathrm{KL}}(\pi(\cdot|s)||\mu(\cdot|s)))$$

Prior policy

  - Fenchel conjugate

$$\Omega^{\star}_{\lambda}(Q_V(s,\cdot)) = \lambda \log \mathbb{E}_{a\sim\mu(\cdot|s)} \exp(Q_V(s,a)/\lambda)$$

Logarithm of moment-generating function

Taylor expansion

$$= \mathbb{E}_{a\sim\mu}[Q_V(s,a)] + \frac{1}{2\lambda}\mathrm{Var}_{a\sim\mu}(Q_V(s,a)) + \mathcal{O}\left(\frac{1}{\lambda^2}\right)$$

$$\simeq Q_V(s,a) + \frac{1}{2\lambda}\mathrm{Var}_{a\sim\mu}(Q_V(s,a))$$

Potential-based reward shaping

$$= \mathbb{E}_{a\sim\mu, s'\sim p(\cdot|s,a)}[r^{\Omega}(s,a,s') + \gamma V(s')]$$

reward shaping

$$r^{\Omega}(s,a,s') := r(s,a) + \gamma\Phi(s') - \Phi(s)$$
$$\Phi(s) := \frac{1}{2\lambda}\mathrm{Var}_{a\sim\mu}(Q_V(s,a))$$

- When it come to the implementation, it's quite simple. ➡ Just change the reward
  - encourage the agent to visit states **with smaller variance**

Distributionally Robust Soft Actor-Critic (DRSAC)

- In this **continuous control case**: the distribution of the policy is **Normal distribution**.
  - parametrized Gaussian policies

$$\pi_\theta(a|s) = \mathcal{N}(\mu_\theta(s), \Sigma_\theta^2(s))$$

  - Approximation the variance of Q-values by using the 1st order **Taylor approximation of Q-values** around the mean action

$$\mathrm{Var}_{a\sim\pi_\theta}(Q(s,a)) \simeq g_0(s)^T \Sigma_\theta(s) g_0(s)$$

1st order Taylor approximation of Q-values around the mean action

$$g_0(s) = \nabla_a Q(s,a)|_{a=\mu_\theta(s)}$$

Independent actions

$$\Sigma_\theta(s) = diag(\sigma_{1,\theta}(s), \ldots, \sigma_{K,\theta}(s))$$

$$\simeq ||g_0(s)||^2_{diag(\sigma_{1,\theta}(s),\ldots,\sigma_{K,\theta}(s))}$$

  - Combined with the soft-actor critic algorithm

$$\pi^* = \underset{\pi}{\mathrm{argmax}} \sum_{t=0}^{T} E_{(s_t,a_t)\sim\tau_\pi}[\gamma^t(r(s_t,a_t) + \alpha\mathcal{H}(\pi(.|s_t))] \qquad \mathcal{H}(\pi(.|s_t)) = -\log\pi(.|s_t).$$

Importance: entropy V.S. reward

Haarnoja, T., Zhou, A., Abbeel, P., and Levine, S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. arXiv preprint arXiv:1801.01290, 2018a.

Application 3: Signal Localization

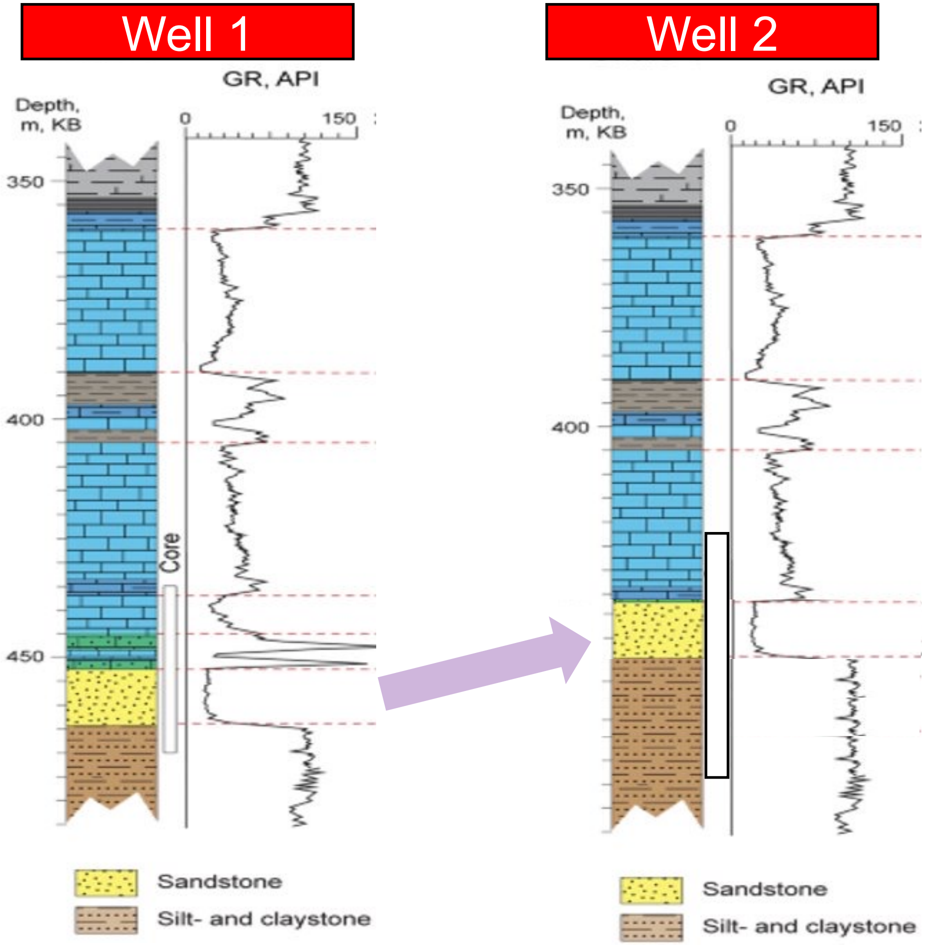
UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

Gamma-Ray Signal Localization Task

Many 1D signals containing the same signal pattern are collected by the well-logging process and sent to human-expert to do hand-picking for oil&gas localization.

Traditional methods

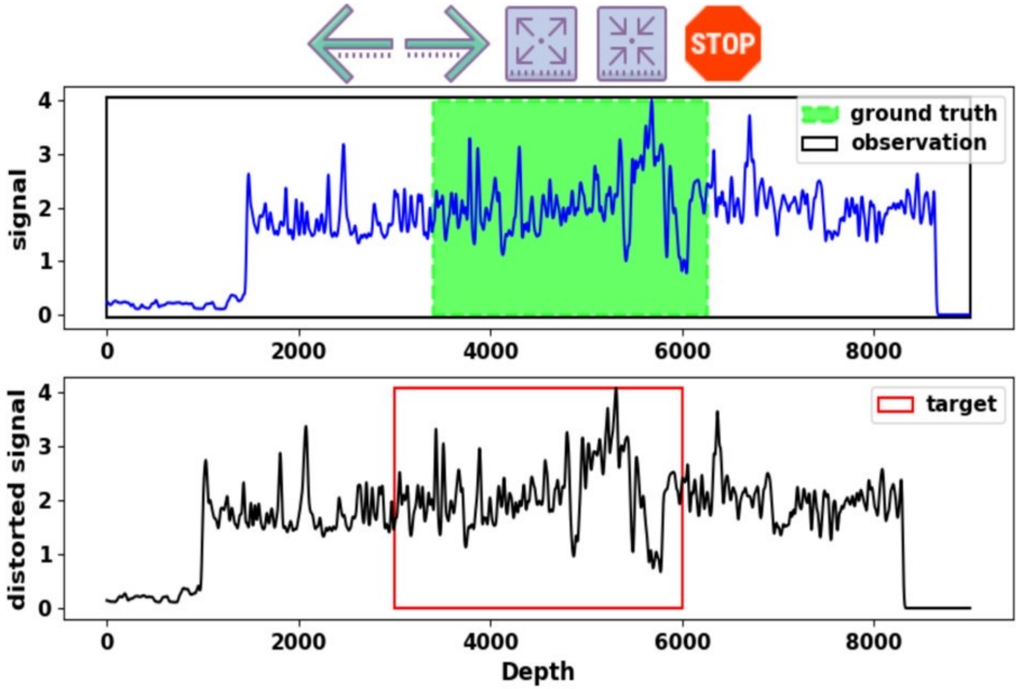
- The human experts pick the patterns associated with interested rock based on one reference signal:
  - Recognize the pattern with domain prior knowledges.
  - Roughly matching with correlation matching methods.



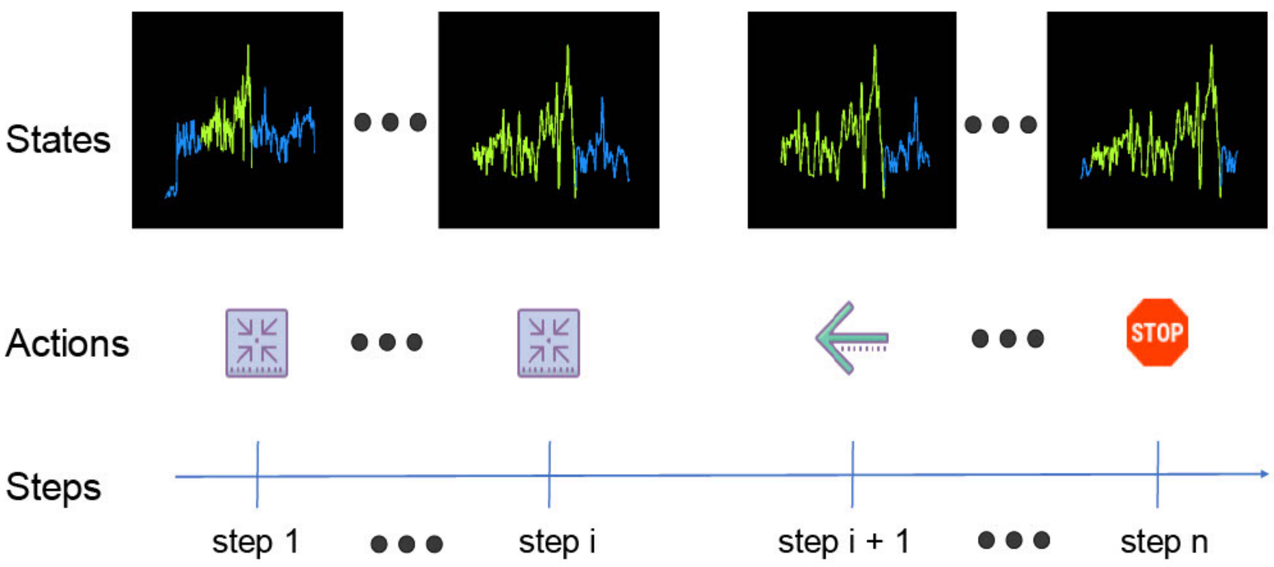
Yuan Zi, Lei Fan, Jiefu Chen, and Zhu Han, “Active Gamma-ray Log Pattern Localization with Distributional Robust Reinforcement Learning,” submitted to IEEE Transactions on Neural Networks and Learning Systems.

Application 3: Signal Localization

Reinforcement Learning Signal Localization Scheme



Sequence of attended regions to localize the object States



- Thoroughly studied log with interested signal fragment as reference/target.
- In the new log, there is a signal fragment that has the same pattern as the reference

- Initial the whole new log trace as the agent's observation.
- Let the agent move (left, right, expand, shrink) to search the reference pattern.

Conclusion

- Motivation
  - DRO accounts for the fact that one is never able to exactly specify a probability distribution in practice.
  - Due to the finite experience sample, the distribution of the real experience is hard to get.
- DRO Optimization
  - The Wasserstein distributionally robust problems can often be reformulated as (or tightly approximated by) finite convex programs within a certain Wasserstein distance from the empirical distribution constructed from training samples.
- CVAR
  - Distributionally robust chance constraints can be conservatively approximated by worst-case CVaR constraints.
- DRO RL
  - The risk-averse exploration in approximate RL setting is required and we can use the distributionally robust modified policy iteration scheme that implements safety in policy evaluation step w.r.t. estimation errors to avoid.
  - Implement DRO to the Reinforcement Learning is just add one safe regularizer to encourage agent visit the state with smaller variance.

Main References

UNIVERSITY of HOUSTON
CULLEN COLLEGE of ENGINEERING

- Rahimian H, and Mehrotra S.
  Distributionally robust optimization: A review. 2019
- Zymler S, Kuhn D, and Rustem B.
  Distributionally robust joint chance constraints with second-order moment information. 2013
- Esfahani P M, and Kuhn D.
  Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. 2018
- T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine.
  Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. 2018
- B. Scherrer, M. Ghavamzadeh, V. Gabillon, B. Lesner, and M. Geist.
  Approximate modified policy iteration and its application to the game of tetris. Journal of Machine Learning Research, 16:1629–1676, 2015
- E. Smirnova, E. Dohmatob, J. Mary.
  Distributionally robust reinforcement learning. 2019

Thank You Very Much for Your Attention !

Conditional value-at-risk in optimization

In problems of optimization under uncertainty, CVaR can be used in the objective or the constraints, or both.

Optimization shortcut

$$\min_{x} CVaR_{\alpha}(x) = \min_{(x,\zeta)} F_{\alpha}(x,\zeta)$$

$$(x^*, \zeta^*) \in \underset{(x,\zeta)\in X\times\mathbb{R}}{\operatorname{argmin}} F_{\alpha}(x,\zeta) \iff x^* \in \arg\min_{x\in X} CVaR(x) \quad \zeta^* \in \underset{\zeta\in\mathbb{R}}{\operatorname{argmin}} F_{\alpha}(x^*,\zeta)$$

CVaR accounts for losses exceeding VaR. So when CVaR is considering, VaR can be ignored