Measuring the Value of Randomized Solutions in Distributionally Robust Optimization

Erick Delage
HEC Montréal
CRC in Decision making under uncertainty

joint work with Ahmed Saif (Dalhousie U.)

BIRS Workshop on Distributionally Robust Optimization
March 9th, 2018

HEC MONTRÉAL

Canada Research Chairs
Chaires de recherche du Canada

Canada

Summary of Wolfram's Talk
Axiomatic Motivation for DRO

Definition: A risk measure ρ_0 is called ambiguity averse if it satisfies for all $X, Y \in \mathcal{L}_\infty(\Omega_0, \mathcal{F}_0, \mathcal{P}_0)$:

- ☑ **Ambiguity aversion:** If $\{F_X^{\mathbb{P}} : \mathbb{P} \in \mathcal{P}_0\} \subseteq \{F_Y^{\mathbb{P}} : \mathbb{P} \in \mathcal{P}_0\}$, then $\rho_0(X) \leq \rho_0(Y)$.
- ☑ **Ambiguity monotonicity:** If $\varrho_0(F_X^{\mathbb{P}}) \leq \varrho_0(F_Y^{\mathbb{P}})$ for all $\mathbb{P} \in \mathcal{P}_0$, then $\rho_0(X) \leq \rho_0(Y)$.

Proposition: Assume that $(\Omega_0, \mathcal{F}_0, \mathcal{P}_0)$ is non-atomic and that ρ_0 is ambiguity averse and translation invariant.

Then the risk measure satisfies

$$\rho_0(X) = \sup_{\mathbb{P} \in \mathcal{P}_0} \varrho_0(F_X^{\mathbb{P}}) \quad \forall X \in \mathcal{L}_\infty(\Omega_0, \mathcal{F}_0, \mathcal{P}_0).$$

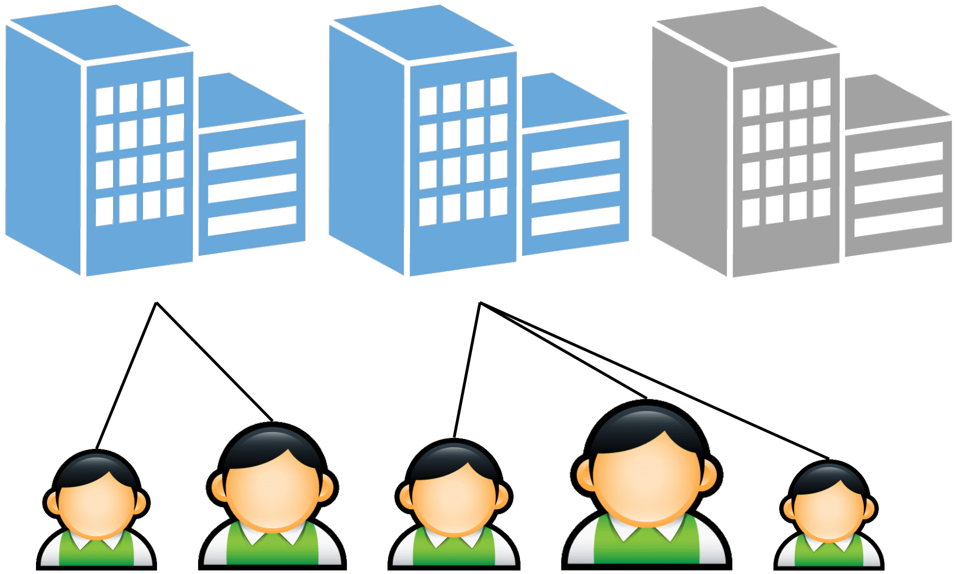
Summary of Wolfram's Talk

Randomisation proof/receptiveness

	Randomization Receptive	Randomization Proof
Distributional Ambiguity	every RM with Lebesgue property	convex RM's with convex $\mathcal{X}_0$
Stochastic Uncertainty	M… me… / M… mean semi-deviati…	convex RM's … $\mathcal{X}_0$ / …asi- …RM's

In DRO problems, if the set of reachable random cost is non-convex, then a <u>randomized strategy</u> might outperform <u>all deterministic ones</u>.

Distributionally Robust Uncapacitated Facility Location Problem



$\mathcal{J}$: set of potential facility locations
$\mathcal{I}$: set of customers

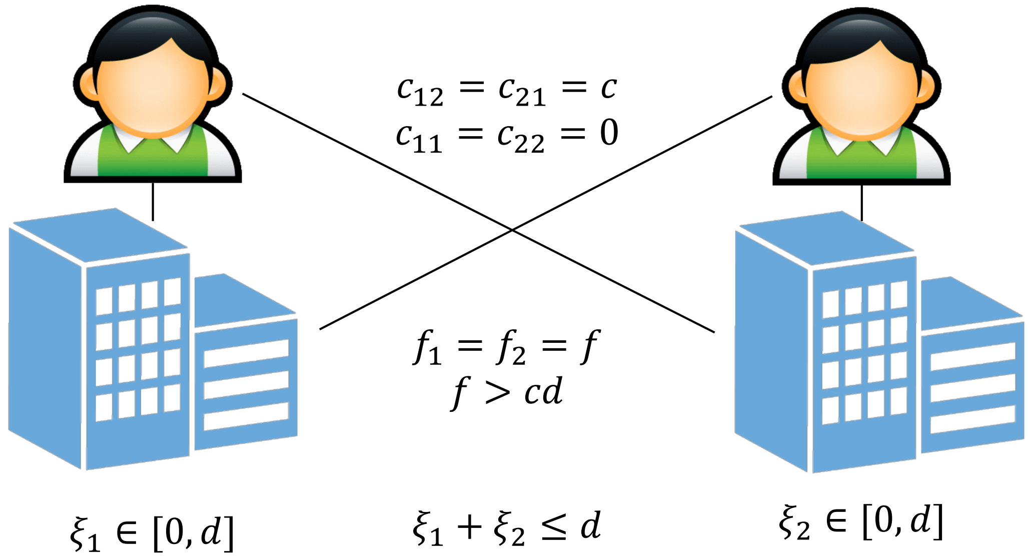
f_j: setup cost of facility j
ξ_i: demand of customer i
c_{ij}: unit shipping cost from j to i

$x_j = 1$ if facility j is opened, 0 otherwise
$y_{ij} = 1$ if customer i's demand is served by facility j

The UFLP is formulated as:

$$\begin{aligned}
\underset{x,y}{\text{minimize}} \quad & \max_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi} \left[\sum_{j \in \mathcal{J}} f_j x_j + \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{I}} \xi_i c_{ij} y_{ij} \right] & \\
\text{subject to} \quad & \sum_{j \in \mathcal{J}} y_{ij} = 1 \qquad , \ \forall i \in \mathcal{I} & (1) \\
& y_{ij} \le x_j \qquad , \ \forall i \in \mathcal{I}, \, j \in \mathcal{J} & (2) \\
& x \in \{0, 1\}^{|\mathcal{J}|} \, , \; y \in \{0, 1\}^{|\mathcal{I}| \times |\mathcal{J}|} & (3)
\end{aligned}$$

Illustrative Example



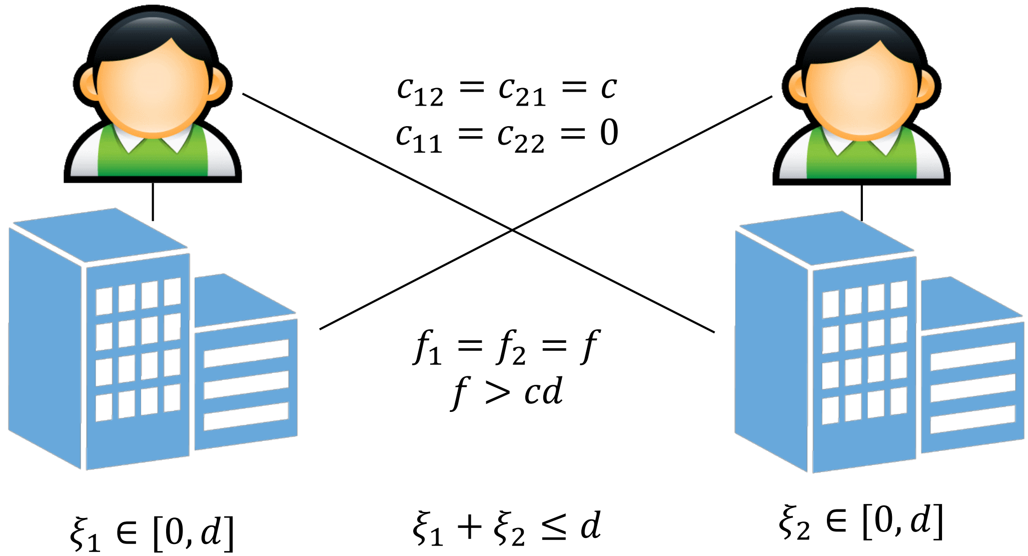
Without randomization

- $v_d^* = \min_{x,y} \sup_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi}[f(x_1 + x_2) + c(\xi_1 y_{12} + \xi_2 y_{21})]$
- Solution 1:
  $x_1^{*1} = 1,\ y_{21}^{*1} = 1,\ v^{*1} = f + cd$
- Solution 2:
  $x_2^{*2} = 1,\ y_{12}^{*2} = 1,\ v^{*2} = f + cd$
- Optimal value : $v_d^* = f + cd$, the worst-case distribution puts all the mass on the wrong node.

With randomization

- $v_r^* = \min_{p_1, p_2} \sup_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi}[\sum_k p_k(f(x_1^{*k} + x_2^{*k}) + c(\xi_1 y_{12}^{*k} + \xi_2 y_{21}^{*k}))]$
- Randomized solution
  $p^* : p_1^* = p_2^* = 0.5$
- Randomized optimal value :
  $v_r^* = f + cd/2$
- Randomization ensures that each unit of demand has 50% chance of staying at same node.

Illustrative Example



Without randomization

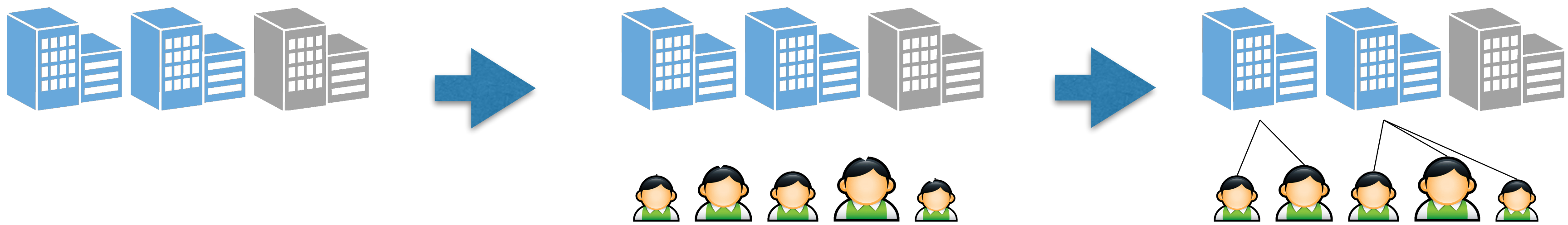
- $v_d^* = \min_{x,y} \sup_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi}[f(x_1 + x_2) + c(\xi_1 y_{12} + \xi_2 y_{21})]$
- Solution 1: $x_1^{*1} = 1,\ y_{21}$ ...
- Solution 2: $x_2^{*2} = 1,\ y_{12}$ ...
- Optimal value : ... worst-case distribution p... the mass on the wrong node.

With randomization

- $v_r^* = \min_{p_1, p_2} \sup_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi}[\sum_k p_k(f(x_1^{*k} + x_2^{*k}) + c(\xi_1 y_{12}^{*k} + \xi_2 y_{21}^{*k}))]$
- ...dom... solution
- ...al value :
- ...nsures that each unit ... mand has 50% chance of staying at same node.

Randomization reduces worst-case expected cost by «cd/2» (i.e. up to 25%)

Two-Stage Distributionally Robust Capacitated Facility Location Problem



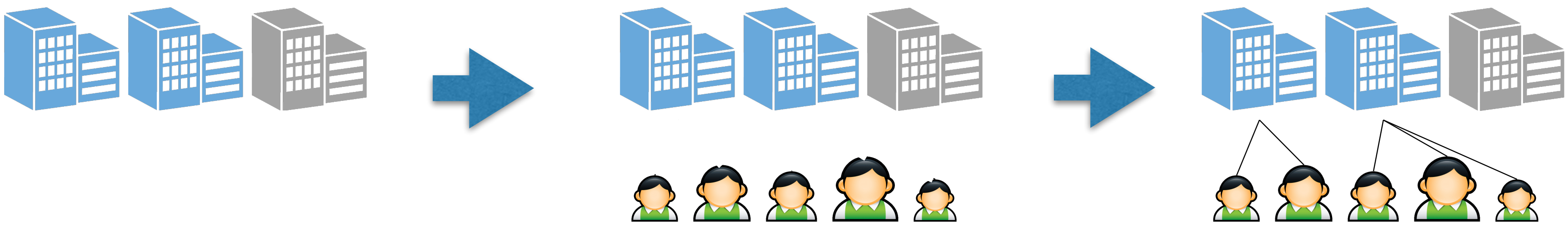
The Two-Stage DR CFLP is formulated as:

$$\underset{x \in \{0,1\}^{|\mathcal{J}|}}{\text{minimize}} \quad \max_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi} \left[\sum_{j \in \mathcal{J}} f_j x_j + G(x, \xi) \right]$$

where

$$\begin{aligned} G(x,\xi) \; := \quad \min_{z \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{J}|}} \quad & \sum_{i,j} c_{i,j} z_{i,j} \\ \text{subject to} \quad & \sum_{j} z_{i,j} \geq \xi_i \qquad \forall\, i \in \mathcal{I} \\ & \sum_{i} z_{i,j} \leq V_j x_j \qquad \forall\, j \in \mathcal{J} \end{aligned}$$

Two-Stage Distributionally Robust Capacitated Facility Location Problem



The Two-Stage DR CFLP is formulated as:

$$\underset{x\in\{0,1\}^{|\mathcal{J}|}}{\text{minimize}} \quad \max_{F_\xi\in\mathcal{D}} \mathbb{E}_{F_\xi}\left[\sum_{j\in\mathcal{J}} f_j x_j + G(x,\xi)\right]$$

where

$$G(x,\xi) \;:=\; \min_{z\in\mathbb{R}^{|\mathcal{I}|\times|\mathcal{J}|}} \quad \sum_{i,j} c_{i,j} z_{i,j}$$

How much improvement can be achieved using randomization here ???

$$\forall i \in \mathcal{I}$$

$$\sum_i z_{i,j} \le V_j x_j \qquad \forall j \in \mathcal{J}$$

The Value of Randomized Solutions

We define the pure strategy problem

$$v_d := \min_{x \in \mathcal{X}_0} \sup_{F_\xi \in \mathcal{D}} \rho_{\xi \sim F_\xi} [h(x, \xi)] \qquad \text{(PSP)}$$

We define the randomized strategy problem

$$v_r := \min_{F_x \in \Delta(\mathcal{X}_0)} \sup_{F_\xi \in \mathcal{D}} \rho_{(X,\xi) \sim F_x \times F_\xi} [h(X, \xi)] \qquad \text{(RSP)}$$

We define the value of randomized solutions

$$\text{VRS} := v_d - v_r$$

Agenda

1. ~~Introduction~~
2. A Tractable Upper Bound for VRS
   - ☑ Bound obtained from the Convex Relaxations
   - ☑ Implications for Totally Unimodular MILP
3. Column-Generation Approach for TS-DRO Problems
4. Numerical experiments
   - ☑ DR Assignment Problems
   - ☑ DR Uncapacitated Facility Location Problems
   - ☑ Two-Stage DR Capacitated Facility Location Problems

Bounding VRS using Convex Hull of $\mathcal{X}_0$

Theorem 1. *Given that* ρ *is a convex risk measure and* $h(x,\xi)$ *a convex function with respect to* x *for all* $\xi \in \mathbb{R}^m$. *Let* $\mathcal{X}$ *be any set known to contain the convex hull of* $\mathcal{X}_0$, *then*

$$VRS \leq v_d - \min_{x\in\mathcal{X}} \sup_{F_\xi \in \mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi)) \, .$$

Bounding VRS using Convex Hull of $\mathcal{X}_0$

Theorem 1. *Given that ρ is a convex risk measure and $h(x,\xi)$ a convex function with respect to x for all $\xi \in \mathbb{R}^m$. Let $\mathcal{X}$ be any set known to contain the convex hull of $\mathcal{X}_0$, then*

$$VRS \leq v_d - \min_{x\in\mathcal{X}} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi))\,.$$

Proof:

$$\begin{aligned}
\min_{x\in\mathcal{X}} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi)) &\leq \min_{x\in\mathcal{C}(\mathcal{X}_0)} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi)) \\
&= \min_{g(\cdot)\in\mathcal{G}} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(g(\xi)) \\
&= \min_{F_g\in\Delta(\mathcal{G})} \sup_{F_\xi\in\mathcal{D}} \rho_{(G,\xi)\sim F_g\times F_\xi}(G(\xi)) \\
&= \min_{F_g\in\Delta(\mathcal{G}_0)} \sup_{F_\xi\in\mathcal{D}} \rho_{(G,\xi)\sim F_g\times F_\xi}(G(\xi)) \\
&= \inf_{F_x\in\Delta(\mathcal{C}(\mathcal{X}_0))} \sup_{F_\xi\in\mathcal{D}} \rho_{(X,\xi)\sim F_x\times F_\xi}(h(X,\xi)) \\
&\leq \inf_{F_x\in\Delta(\mathcal{X}_0)} \sup_{F_\xi\in\mathcal{D}} \rho_{(X,\xi)\sim F_x\times F_\xi}[h(X,\xi)] = v_r\,,
\end{aligned}$$

where

$$\mathcal{G} := \{g : \mathbb{R}^m \to \mathbb{R} \,|\, \exists x \in \mathcal{C}(\mathcal{X}_0),\ g(\xi) \geq h(x,\xi) \forall\, \xi\}$$

$$\mathcal{G}_0 := \{g : \mathbb{R}^m \to \mathbb{R} \,|\, \exists x \in \mathcal{C}(\mathcal{X}_0),\ g(\xi) = h(x,\xi) \forall\, \xi\}$$

Bounding VRS using Convex Hull of $\mathcal{X}_0$

Theorem 1. *Given that ρ is a convex risk measure and $h(x,\xi)$ a convex function with respect to x for all $\xi \in \mathbb{R}^m$. Let $\mathcal{X}$ be any set known to contain the convex hull of $\mathcal{X}_0$, then*

$$VRS \leq v_d - \min_{x\in\mathcal{X}} \sup_{F_\xi \in \mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi))\,.$$

Proof:

$$\min_{x\in\mathcal{X}} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi)) \leq \min_{x\in\mathcal{C}(\mathcal{X}_0)} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(h(x,\xi))$$

$$= \min_{g(\cdot)\in\mathcal{G}} \sup_{F_\xi\in\mathcal{D}} \rho_{\xi\sim F_\xi}(g(\xi))$$

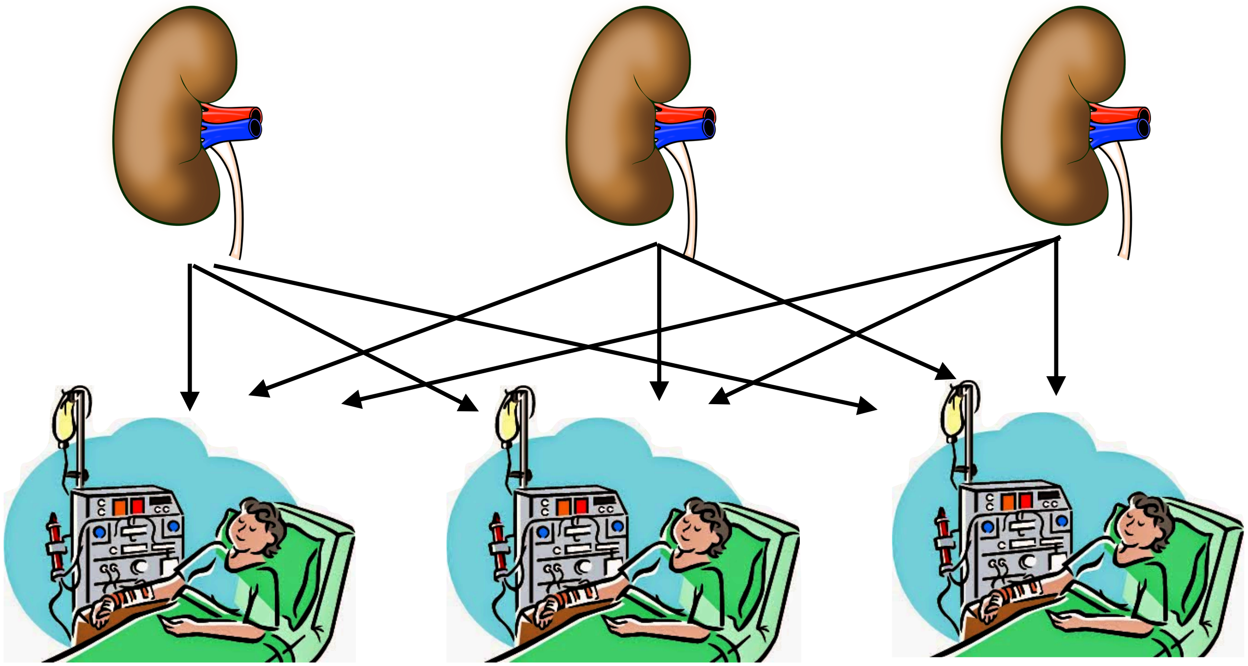
$$= \min_{F_g\in\Delta(\mathcal{G})} \sup_{F_\xi\in\mathcal{D}} \rho_{(G,\xi)\sim F_g\times F_\xi}(G(\xi))$$

$$= \min_{F_g\in\Delta(\mathcal{G}_0)} \sup_{F_\xi\in\mathcal{D}} \rho_{(G,\xi)\sim F_g\times F_\xi}(G(\xi))$$

$$= \inf \; \sup \; \rho_{(X,\xi)\sim F_x\times F_\xi}(h(X,\xi))$$

If $\mathcal{X} = \mathcal{C}(\mathcal{X}_0)$, the risk measure is the expected value, and the cost function is linear in x, then this bound is tight and is achieved by any $F_x^* \in \Delta(\mathcal{X}_0) : E_{X\sim F_x^*}[X] = x^*$.

Distributionally Robust Assignment Problem



$\mathcal{J}$: set of potential recipient

$\mathcal{I}$: set of donated organs

$\xi_{\mathbf{i,j}}$: **incompatibility score between organ i and recipient j**

$\mathbf{x_{i,j}} = \mathbf{1}$ **if we assign organ i to recipient j**

The DR Assignment Problem is formulated as:

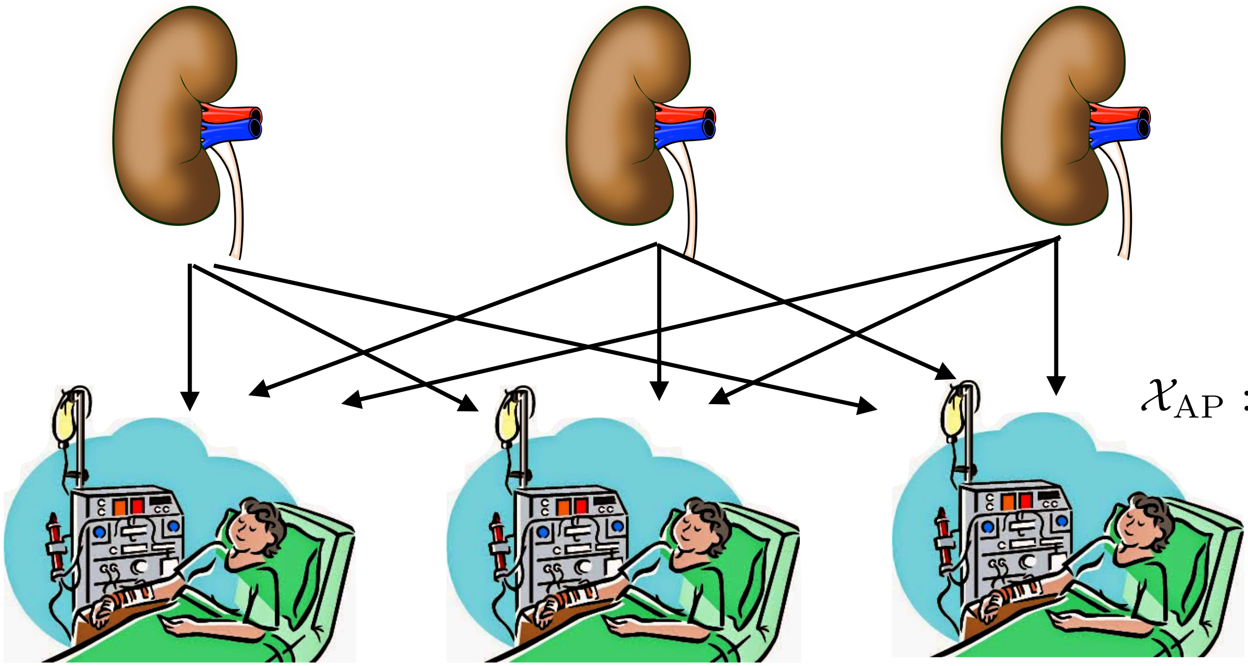
$$\min_{x \in \mathcal{X}_{\text{AP}}} \max_{F_\xi \in \mathcal{D}} \mathbb{E}_{\xi \sim F_\xi} \left[\sum_{i \in I} \sum_{j \in J} \xi_{ij} x_{ij} \right]$$

where

$$\mathcal{X}_{\text{AP}} := \left\{ x \in \{0,\, 1\}^{n_i \times n_j} \;\middle|\; \begin{array}{l} \sum_{j \in J} x_{ij} = 1\,,\ \forall\, i \\ \sum_{i \in I} x_{ij} = 1\,,\ \forall\, j \end{array} \right\}$$

HEC MONTREAL

VRS for DR Assignment Problem



$$\min_{x \in \mathcal{X}_{\text{AP}}} \max_{F_\xi \in \mathcal{D}} \mathbb{E}_{\xi \sim F_\xi} \left[\sum_{i \in I} \sum_{j \in J} \xi_{ij} x_{ij} \right]$$

$$\mathcal{X}_{\text{AP}} := \left\{ x \in \{0,\, 1\}^{n_i \times n_j} \;\middle|\; \begin{array}{l} \sum_{j \in J} x_{ij} = 1\,,\ \forall\, i \\ \sum_{i \in I} x_{ij} = 1\,,\ \forall\, j \end{array} \right\}$$

$$\text{VRS} = \min_{x \in \mathcal{X}_{\text{AP}}} \delta^*(x|\mathcal{U}) - \min_{x' \in \mathcal{X}'_{\text{AP}}} \delta^*(x'|\mathcal{U})$$

where

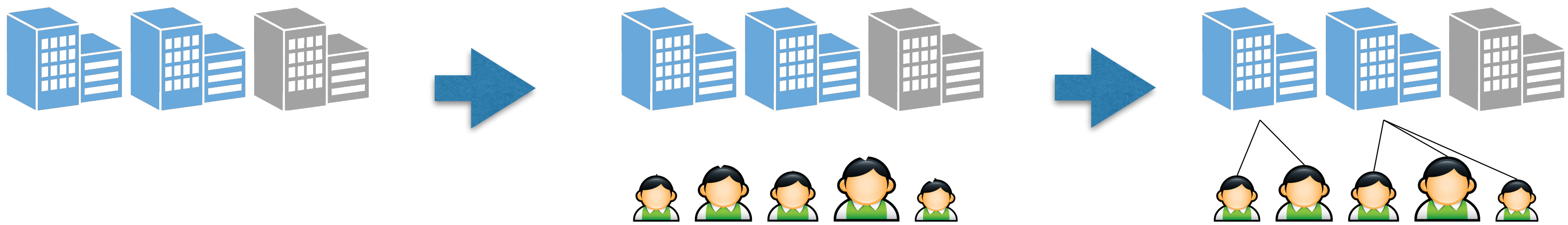
$$\left\{ \qquad \middle| \; \sum x_{ij} = 1\,,\ \forall\, i \right\}$$

An optimal randomized strategy can be obtained by solving: $\underset{p \in \Delta}{\text{minimize}} \, \|x'^* - \sum_{k \in \mathcal{K}} p_k x^k\|$.

Agenda

1. ~~Introduction~~
2. ~~A Tractable Upper Bound for VRS~~
   - ☑ ~~Bound obtained from the Convex Relaxations~~
   - ☑ ~~Implications for Totally Unimodular MILP~~
3. Column-Generation Approach for TS-DRO Problems
4. Numerical experiments
   - ☑ DR Assignment Problems
   - ☑ DR Uncapacitated Facility Location Problems
   - ☑ Two-Stage DR Capacitated Facility Location Problems

Two-Stage Distributionally Robust Capacitated Facility Location Problem



The Two-Stage DR CFLP is formulated as:

$$\underset{x \in \{0,1\}^{|\mathcal{J}|}}{\text{minimize}} \quad \max_{F_\xi \in \mathcal{D}} \mathbb{E}_{F_\xi} \left[\sum_{j \in \mathcal{J}} f_j x_j + G(x, \xi) \right]$$

where

$$\begin{aligned} G(x,\xi) \;:=\; & \min_{z \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{J}|}} && \sum_{i,j} c_{i,j} z_{i,j} \\ & \text{subject to} && \sum_{j} z_{i,j} \geq \xi_i && \forall\, i \in \mathcal{I} \\ & && \sum_{i} z_{i,j} \leq V_j x_j && \forall\, j \in \mathcal{J} \end{aligned}$$

Two-Stage Distributionally Robust Linear Program with RHS Uncertainty

A Two-Stage DR LP with RHS uncertainty is formulated as:

$$\underset{x\in\mathcal{X}\subseteq\{0,1\}^n}{\text{minimize}} \ \sup_{F_\xi\in\mathcal{D}} \ \mathbb{E}_{\xi\sim F_\xi}[c_1^T x + G(x,\xi)]$$

where

$$G(x,\xi) := \min_{y\in\mathbb{R}^m} \ c_2^T y$$

$$\text{s.t.} \ \ Ax + By \leq W\xi + b$$

We wish to identify the optimal randomized strategy

$$\underset{F_x\in\Delta(\mathcal{X})}{\text{minimize}} \ \sup_{F_\xi\in\mathcal{D}} \ \mathbb{E}_{(X,\xi)\sim F_x\times F_\xi}[c_1^T X + G(X,\xi)]$$

which can be simplified since feasible set $\mathcal{X}$ is discrete

$$\underset{p\in\Delta\subset\mathbb{R}_+^{|\mathcal{X}|}}{\text{minimize}} \ \sup_{F_\xi\in\mathcal{D}} \ \sum_{k\in\mathcal{K}} p_k c_1^T x^k + \mathbb{E}_{\xi\sim F_\xi}[\underbrace{\sum_{k\in\mathcal{K}} p_k G(x^k,\xi)}_{g(p,\xi)}]$$

Under the Wasserstein Ambiguity Set

Given a sample set $\{\hat{\xi}_1, \ldots, \hat{\xi}_N\}$, we assume the ambiguity set is a Wasserstein ball:

$$\mathcal{D}(\epsilon) := \left\{ F_\xi \;\middle|\; \begin{array}{l} \mathbb{P}_{F_\xi}(\xi \in \Xi) = 1 \\ d_W(F_\xi, \hat{F}_\xi) \leq \epsilon \end{array} \right\}$$

Based on Mohajerin Esfahani & Kuhn (2017), the DRO can be reformulated as:

$$\begin{array}{ll} \underset{p\in\Delta,\lambda\in\mathbb{R},s}{\text{minimize}} & \displaystyle\sum_{k\in\mathcal{K}} c_1^T x^k p_k + \lambda\epsilon + (1/N)\sum_n s_n \\ \text{subject to} & \displaystyle\sup_{\xi\in\Xi} g(p,\xi) - \lambda\|\xi - \xi_n\| \leq s_n \;,\; \forall n \\ & \displaystyle\sum_{k\in\mathcal{K}} p_k = 1 \end{array}$$

Under the Wasserstein Ambiguity Set

Given a sample set $\{\hat{\xi}_1, \ldots, \hat{\xi}_N\}$, we assume the ambiguity set is a Wasserstein ball:

$$\mathcal{D}(\epsilon) := \left\{ F_\xi \;\middle|\; \begin{array}{l} \mathbb{P}_{F_\xi}(\xi \in \Xi) = 1 \\ d_W(F_\xi, \hat{F}_\xi) \leq \epsilon \end{array} \right\}$$

Based on Mohajerin Esfahani & Kuhn (2017), the DRO can be reformulated as:

Equivalently as:

$$\begin{aligned} \underset{p \in \Delta, \lambda \in \mathbb{R}, s}{\text{minimize}} \quad & \sum_{k \in \mathcal{K}} c_1^T x^k p_k + \lambda \epsilon + (1/N) \sum_n s_n \\ \text{subject to} \quad & \sup_{(\xi, \gamma) \in \Xi'_n} g(p, \xi) - \lambda \gamma \leq s_n \,, \; \forall n \\ & \sum_{k \in \mathcal{K}} p_k = 1 \end{aligned}$$

where

$$\Xi'_n := \{(\xi, \gamma) | \xi \in \Xi, \; \|\xi - \hat{\xi}_n\| \leq \gamma\}$$

Formulating a Large-scale LP

When 1-norm is used and Ξ is polyhedral, each Ξ'_n is polyhedral, since $g(p,\cdot)$ is convex, one can instead enumerate the vertices

$$\begin{aligned}
\underset{p\in\Delta,\lambda\in\mathbb{R},s}{\text{minimize}} \quad & \sum_{k\in\mathcal{K}} c_1^T x^k p_k + \lambda\epsilon + (1/N)\sum_n s_n \\
\text{subject to} \quad & \sum_{k\in\mathcal{K}} p_k G(x^k, \xi^{h_n}) - \lambda\gamma^{h_n} \leq s_n \ , \ \forall n \ , \ \forall\, h_n \in \mathcal{H}_n \\
& \sum_{k\in\mathcal{K}} p_k = 1
\end{aligned}$$

One can solve this large scale linear program using a two-layer column and constraint generation

Formulating a Large-scale LP

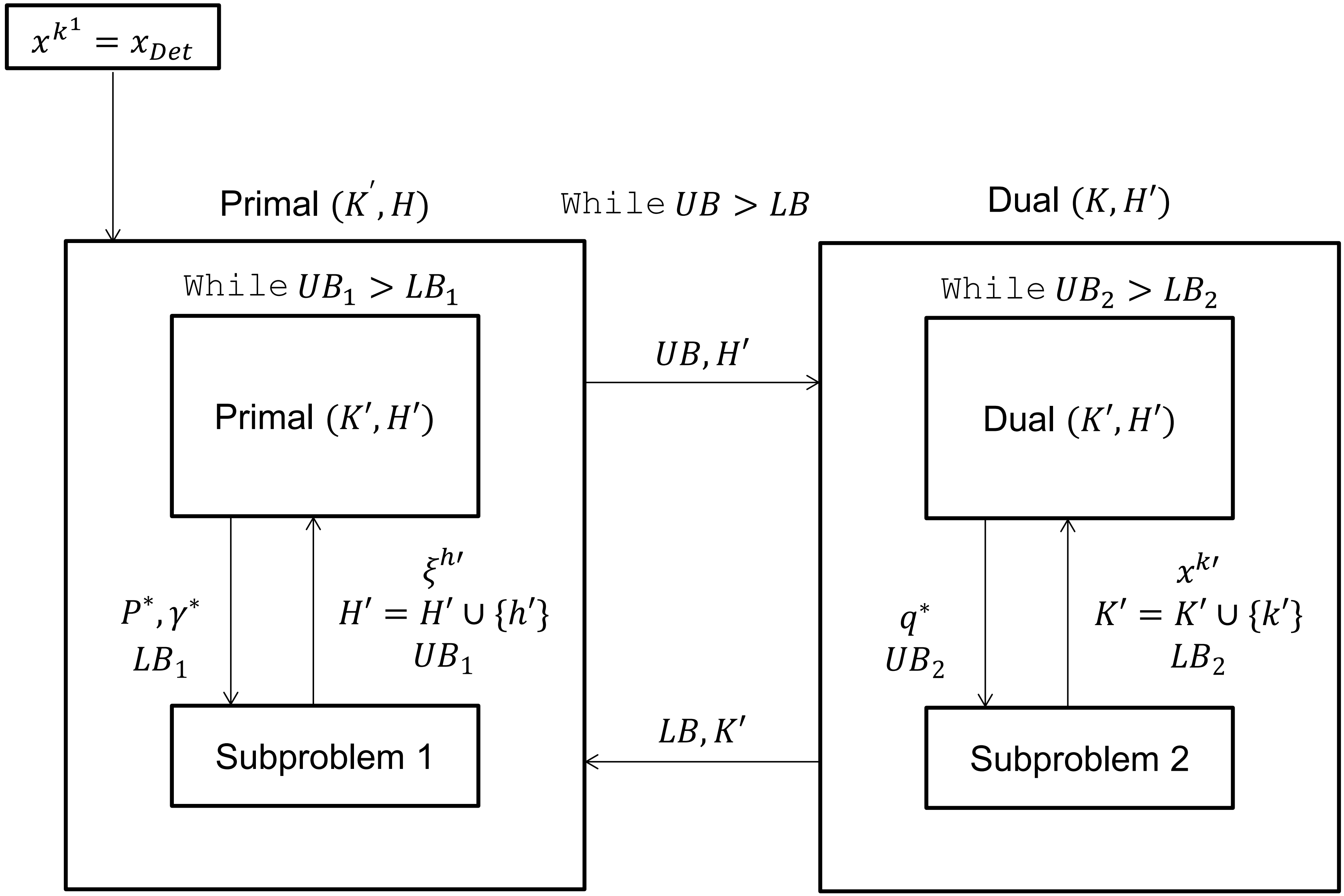
When 1-norm is used, each Ξ'_n is polyhedral, since $g(p,\cdot)$ is convex, one can instead enumerate the vertices

$$\begin{aligned}
\underset{p\in\Delta,\lambda\in\mathbb{R},s}{\text{minimize}} \quad & \sum_{k\in\mathcal{K}'} c_1^T x^k p_k + \lambda\epsilon + (1/N)\sum_n s_n \\
\text{subject to} \quad & \sum_{k\in\mathcal{K}'} p_k G(x^k, \xi^{h_n}) - \lambda\gamma^{h_n} \leq s_n \;,\; \forall n \;,\; \forall\, h_n \in \mathcal{H}'_n \\
& \sum_{k\in\mathcal{K}'} p_k = 1
\end{aligned}$$

One can solve this large scale linear program using a two-layer column and constraint generation

$$\mathcal{K}' \rightarrow \mathcal{K} \qquad \mathcal{H}' := \{\mathcal{H}'_n\}_{n=1}^N \rightarrow \mathcal{H} := \{\mathcal{H}_{n=1}^N\}_n \;,\; \forall\, n$$

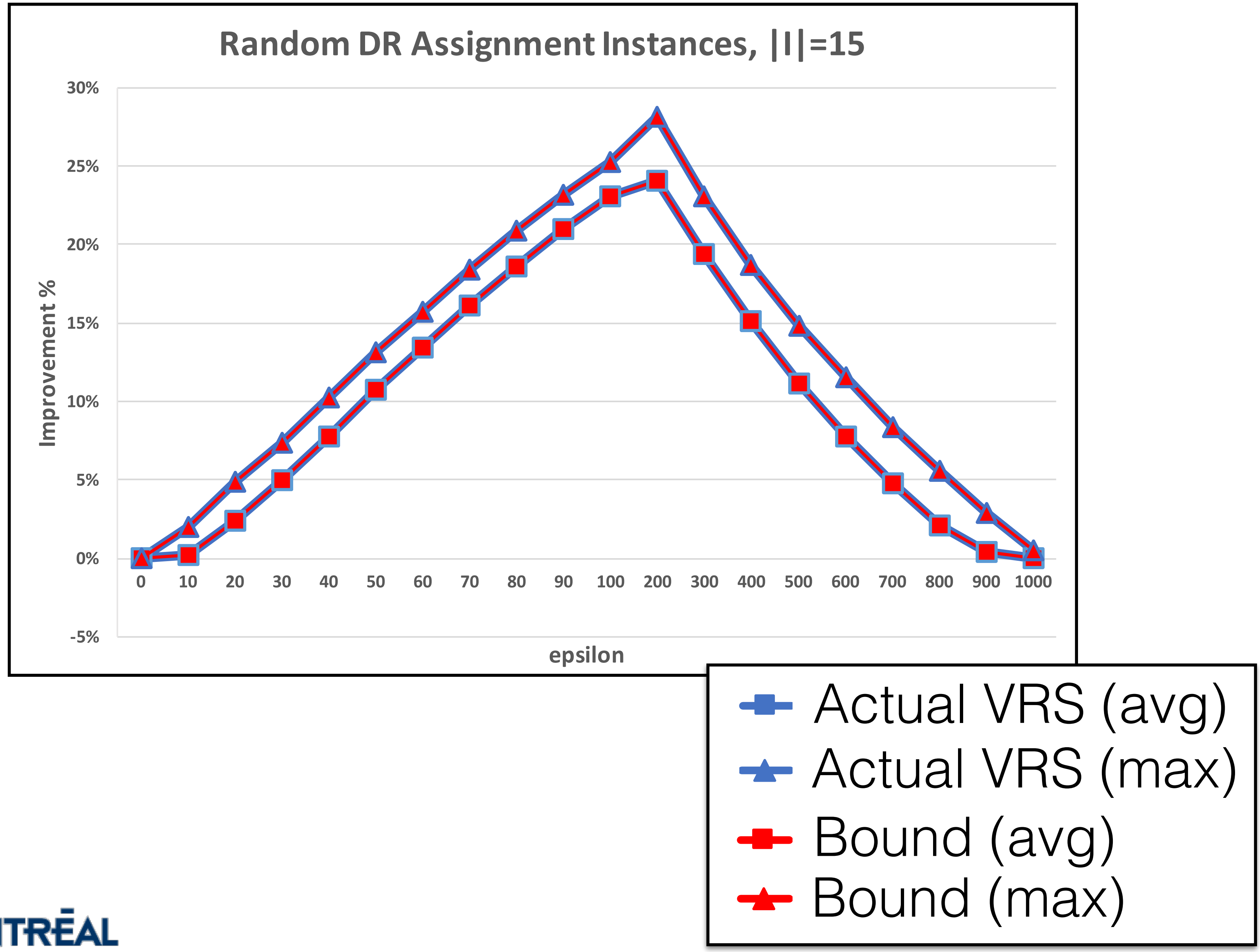
Two-layer Column Generation Algorithm



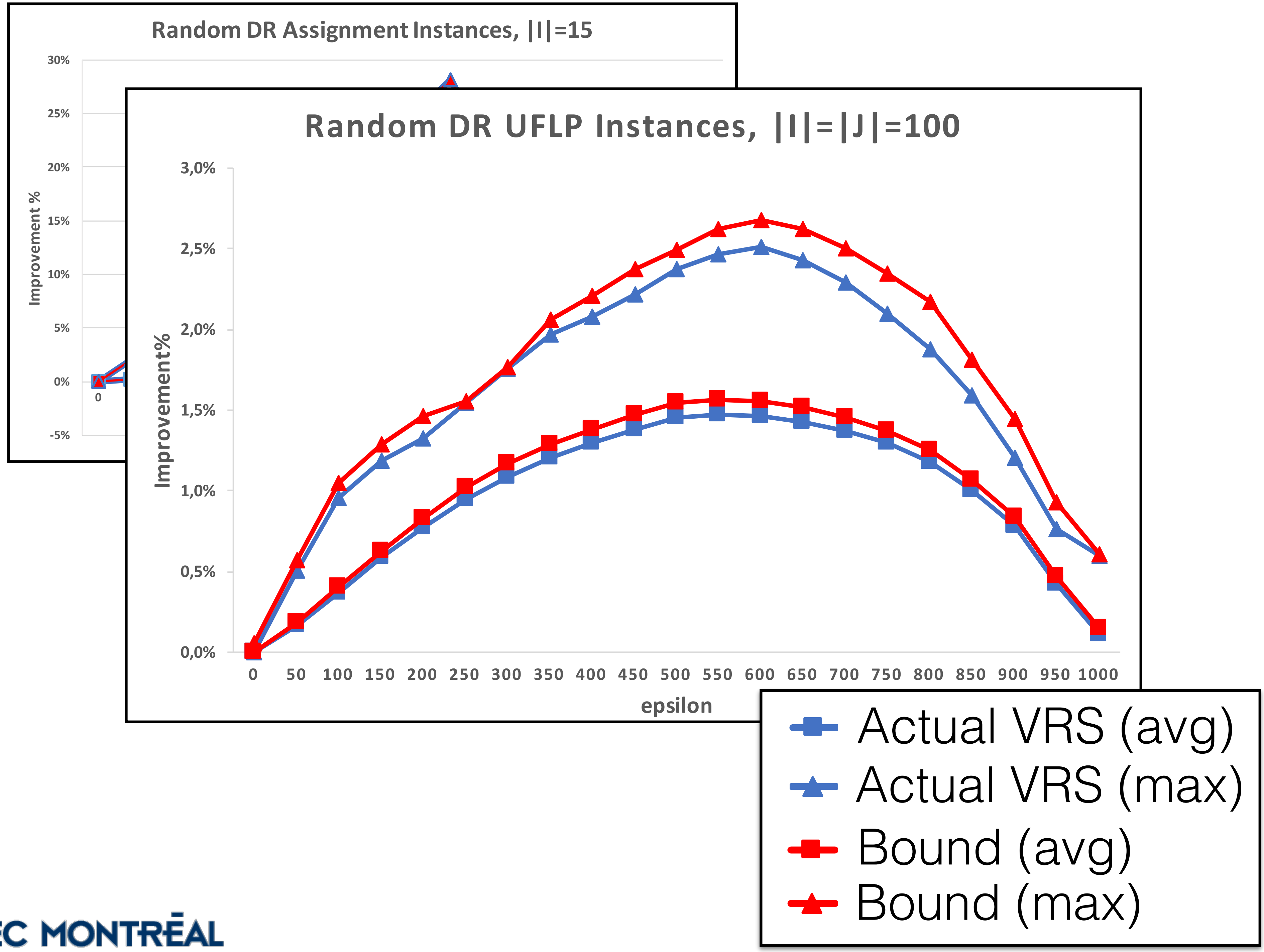
Agenda

1. ~~Introduction~~
2. ~~A Tractable Upper Bound for VRS~~
   - ☑ ~~Bound obtained from the Convex Relaxations~~
   - ☑ ~~Implications for Totally Unimodular MILP~~
3. ~~Column Generation Approach for TS-DRO Problems~~
4. Numerical experiments
   - ☑ DR Assignment Problems
   - ☑ DR Uncapacitated Facility Location Problems
   - ☑ Two-Stage DR Capacitated Facility Location Problems

Quality of Bound & True VRS

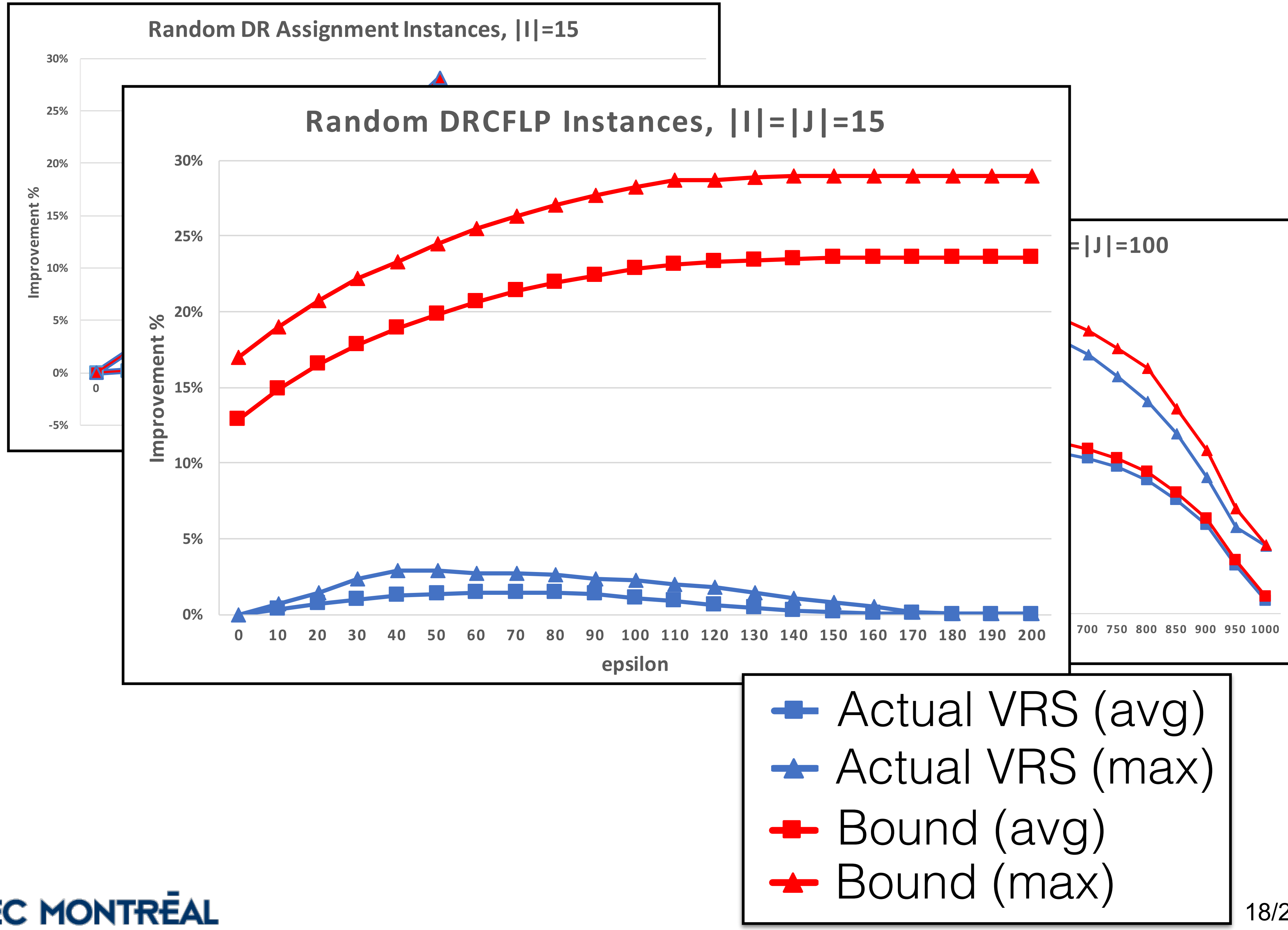


Quality of Bound & True VRS

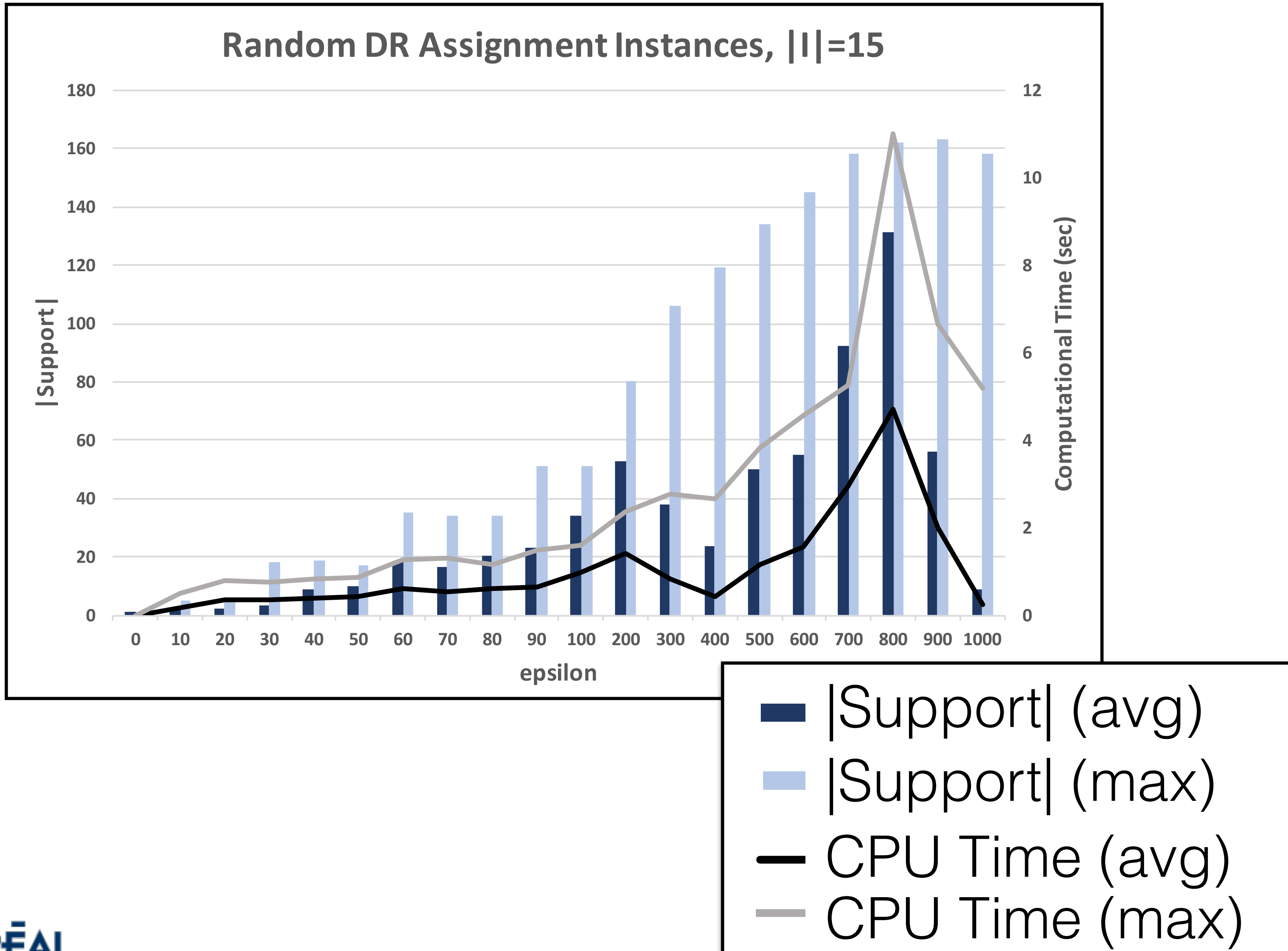


HEC MONTRĒAL

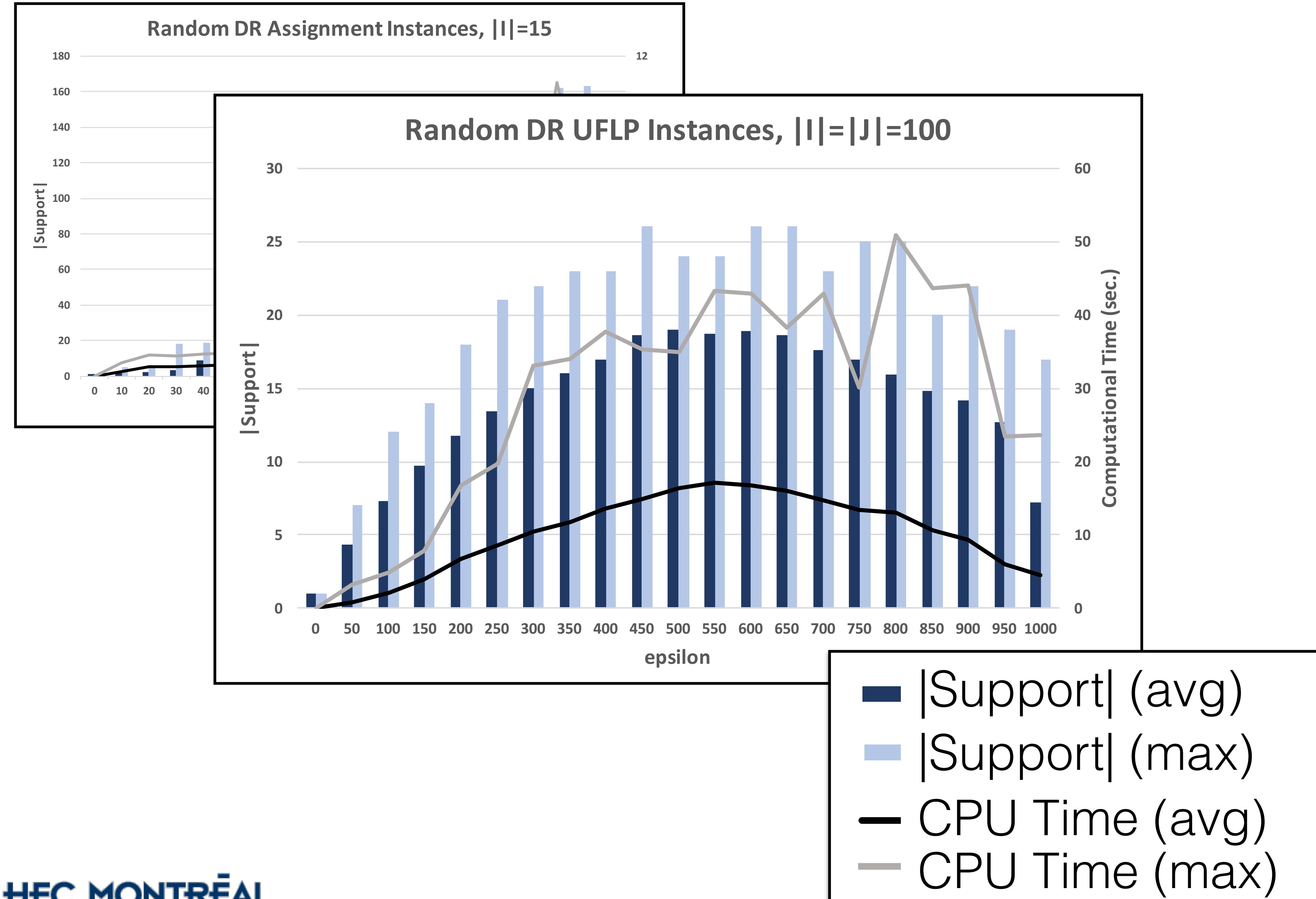
Quality of Bound & True VRS



Optimal Support Size & Computation Time

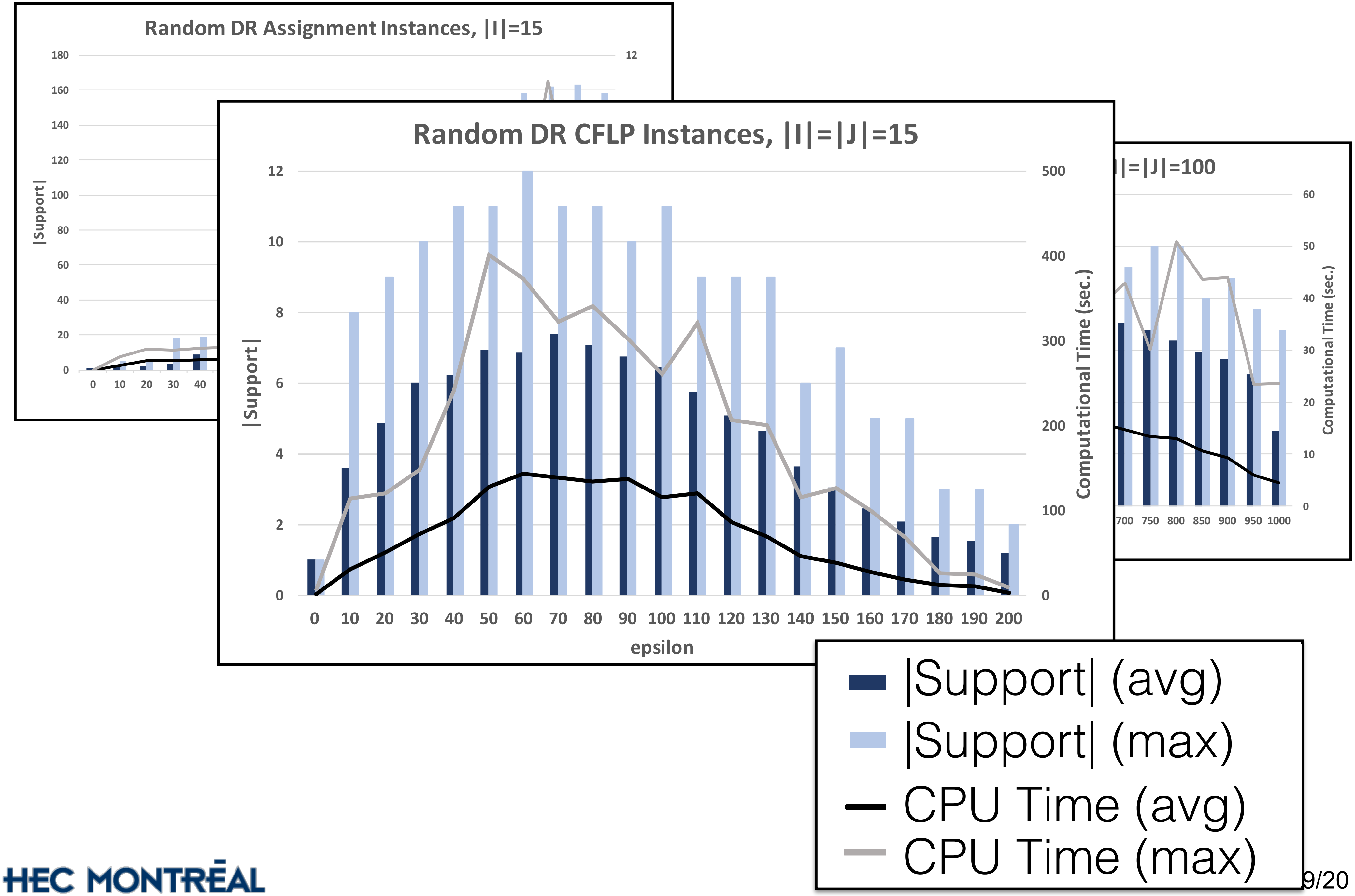


Optimal Support Size & Computation Time



HEC MONTRÉAL

Optimal Support Size & Computation Time



HEC MONTRÉAL

Bibliography

Randomized decisions in economics:

[1] M. Agranov and P. Ortoleva. *Stochastic choice and preferences for randomization.* Journal of Political Economy, 125(1):40-68, 2017.

Randomized decisions in algorithm design:

[2] R. Motwani and P. Raghavan. *Randomized algorithms.* Cambridge University Press, 1995.

Randomized decisions in Markov decision processes:

[3] Y. Le Tallec. *Robust, risk-sensitive, and data-driven control of Markov decision processes.* PhD thesis, MIT, 2007.

[4] W. Wiesemann, D. Kuhn and B. Rustem. *Robust Markov decision processes.* Math. OR 62(6):1358-1316, 2014.

Randomized decisions in SP and DRO:

[5] G. Pflug. Version-independence and nested distributions in multistage stochastic optimization. SIAM J. Optim. 20(3):1406-1420, 2010.

[6] E. Delage, D. Kuhn and W. Wiesemann. *"Dice"-sion making under uncertainty: When can a random decision reduce risk?* Available on Opt. Online, 2016.

[7] E. Delage and A. Saif, *The Value of Randomization under Distributional Ambiguity*, working draft, 2018.

Accounting for time inconsistencies with a commitment mechanism:

[8] R. H. Strotz. *Myopia and Inconsistency in Dynamic Utility Maximization.* The Review of Economic Studies, 23(3):165-180, 1955.