Numerical method for approximately optimal solutions of two-stage distributionally robust optimization with marginal constraints

Qikun Xiang

Nanyang Technological University, Singapore

June 30, 2022

Joint work with:

Ariel Neufeld,
Nanyang Technological University, Singapore

Two-stage distributionally robust optimization (DRO)

- Two-stage distributionally robust optimization (DRO) problem:

$$\begin{aligned} \phi_{\mathrm{DRO}} := \quad & \underset{\boldsymbol{a}}{\text{minimize}} \quad \langle \boldsymbol{c}_1, \boldsymbol{a} \rangle + \sup_{\mu \in \mathcal{P}_{\boldsymbol{\mathcal{X}}}} \int_{\boldsymbol{\mathcal{X}}} Q(\boldsymbol{a}, \boldsymbol{x})\, \mu(\mathrm{d}\boldsymbol{x}) \\ & \text{subject to} \quad \mathbf{L}_{\mathrm{in}} \boldsymbol{a} \le \boldsymbol{q}_{\mathrm{in}}, \ \mathbf{L}_{\mathrm{eq}} \boldsymbol{a} = \boldsymbol{q}_{\mathrm{eq}}, \ \boldsymbol{a} \in \mathbb{R}^{K_1}. \end{aligned}$$

- First-stage decision variable $\boldsymbol{a} \in \mathbb{R}^{K_1}$.
- Uncertain quantity $\boldsymbol{x} \in \boldsymbol{\mathcal{X}} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_N \subset \mathbb{R}^N$, where $\mathcal{X}_1, \ldots, \mathcal{X}_N$ are compact subsets of $\mathbb{R}$.
- $\mathcal{P}_{\boldsymbol{\mathcal{X}}} \subseteq \mathcal{P}(\boldsymbol{\mathcal{X}})$ is the ambiguity set of probability measures.
- $\langle \boldsymbol{c}_1, \boldsymbol{a} \rangle$ is the first-stage cost, $Q(\boldsymbol{a}, \boldsymbol{x})$ is the second-stage cost:

$$\begin{aligned} Q(\boldsymbol{a}, \boldsymbol{x}) := \quad & \underset{\boldsymbol{z}}{\text{minimize}} \quad \langle \boldsymbol{c}_2, \boldsymbol{z} \rangle \\ & \text{subject to} \quad \mathbf{A}_{\mathrm{in}} \boldsymbol{z} \le \mathbf{V}_{\mathrm{in}} \boldsymbol{a} + \mathbf{W}_{\mathrm{in}} \boldsymbol{x} + \boldsymbol{b}_{\mathrm{in}}, \\ & \qquad\qquad\quad \mathbf{A}_{\mathrm{eq}} \boldsymbol{z} = \mathbf{V}_{\mathrm{eq}} \boldsymbol{a} + \mathbf{W}_{\mathrm{eq}} \boldsymbol{x} + \boldsymbol{b}_{\mathrm{eq}}, \\ & \qquad\qquad\quad \boldsymbol{z} \in \mathbb{R}^{K_2}. \end{aligned}$$

Two-stage distributionally robust optimization (DRO)

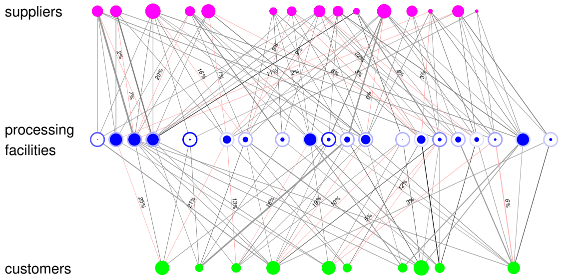
- We consider the case where $\mathcal{P}_{\boldsymbol{\mathcal{X}}}$ is the set of couplings of fixed marginals $\mu_1 \in \mathcal{P}(\mathcal{X}_1), \ldots, \mu_N \in \mathcal{P}(\mathcal{X}_N)$:

$$\mathcal{P}_{\boldsymbol{\mathcal{X}}} = \Gamma(\mu_1, \ldots, \mu_N) := \Big\{\mu \in \mathcal{P}(\boldsymbol{\mathcal{X}}) : \text{the marginal of } \mu \text{ on } \mathcal{X}_i \text{ is } \mu_i \ \forall 1 \leq i \leq N\Big\}.$$

- **Motivation:** there is much less ambiguity about the marginal distributions than the dependence structure (Eckstein, Kupper, and Pohl 2020).

Example: supply chain network design

- $\boldsymbol{a}$: investment for the processing facilities
- $\boldsymbol{x}$: demands of product & failure of edges
- $\langle \boldsymbol{c}_1, \boldsymbol{a} \rangle$: investment cost
- $Q(\boldsymbol{a}, \boldsymbol{x})$: transportation & processing costs



- Additional examples include: task scheduling, assemble-to-order system, ...

Contributions

- We develop a relaxation scheme for two-stage DRO with marginal constraints such that the relaxation error can be controlled to be arbitrarily close to 0.
- We develop a numerical algorithm which computes:
  - an **approximately optimal solution** $\hat{\boldsymbol{a}}$ of ϕ_{DRO};
  - **upper bound** $\phi^{\mathsf{UB}}_{\mathrm{DRO}}$ and **lower bound** $\phi^{\mathsf{LB}}_{\mathrm{DRO}}$ such that $\phi^{\mathsf{LB}}_{\mathrm{DRO}} \le \phi_{\mathrm{DRO}} \le \phi^{\mathsf{UB}}_{\mathrm{DRO}}$;
  - $\hat{\epsilon} := \phi^{\mathsf{UB}}_{\mathrm{DRO}} - \phi^{\mathsf{LB}}_{\mathrm{DRO}}$ **measuring the sub-optimality** of $\hat{\boldsymbol{a}}$, where $\hat{\epsilon}$ can be controlled to be arbitrarily close to 0 by the inputs of the algorithm.
- We perform numerical experiments to demonstrate the proposed algorithm in prominent decision problems including scheduling, assemble-to-order system, and supply chain network design.

Related studies

- Gao and Kleywegt (2017) consider DRO with marginal constraints & Wasserstein distance based constraint.
  - Computation is only tractable when the marginals are discrete and when the second-stage cost is the maximum of finitely many affine functions.
- Chen, Ma, Natarajan, Simchi-Levi, and Yan (2021) deal with a particular class of DRO problem with marginal constraints, most notably appointment scheduling.
  - Focuses on analysing the theoretical computational complexity.
  - No concrete numerical algorithm is provided.
- **Connection with multi-marginal optimal transport (MMOT)**: Carlier, Oberman, Oudet (2015), Pass (2015), Benamou, Carlier, and Nenna (2019), Peyré and Cuturi (2019), Benamou (2021), Neufeld and Xiang (2022), . . .

Step 1: Augmentation

- Recall that the two-stage DRO problem

$$\begin{aligned} \phi_{\mathrm{DRO}} := \quad & \underset{\boldsymbol{a}}{\text{minimize}} \quad \langle \boldsymbol{c}_1, \boldsymbol{a} \rangle + \sup_{\mu \in \mathcal{P}_{\boldsymbol{\mathcal{X}}}} \int_{\boldsymbol{\mathcal{X}}} Q(\boldsymbol{a}, \boldsymbol{x}) \, \mu(\mathrm{d}\boldsymbol{x}) \\ & \text{subject to} \quad \mathbf{L}_{\mathrm{in}} \boldsymbol{a} \le \boldsymbol{q}_{\mathrm{in}}, \ \mathbf{L}_{\mathrm{eq}} \boldsymbol{a} = \boldsymbol{q}_{\mathrm{eq}}, \ \boldsymbol{a} \in \mathbb{R}^{K_1} \end{aligned}$$

  has a min-max-min structure.

- To begin, take the dual of the second-stage problem and represent $Q(\boldsymbol{a}, \boldsymbol{x}) = \max_{\boldsymbol{\lambda} \in S_2^*} \{ \langle \mathbf{V}\boldsymbol{a} + \mathbf{W}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{\lambda} \rangle \}$ for some polytope $S_2^* \subset \mathbb{R}^{K_2^*}$. This transforms the problem into a min-max-max problem.

Step 1: Augmentation

- Next, to combine the two maximization steps, we augment $Q(\cdot,\ \cdot)$ and $\Gamma(\mu_1, \ldots, \mu_N)$:

$$Q_{\mathrm{aug}}(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{\lambda}) := \langle \mathbf{V}\boldsymbol{a} + \mathbf{W}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{\lambda} \rangle \qquad \forall \boldsymbol{a},\ \forall \boldsymbol{x},\ \forall \boldsymbol{\lambda},$$

$$\Gamma_{\mathrm{aug}}(\mu_1, \ldots, \mu_N) := \Big\{ \mu_{\mathrm{aug}} \in \mathcal{P}(\boldsymbol{\mathcal{X}} \times S_2^*) : \text{the marginal of } \mu_{\mathrm{aug}} \text{ on } \mathcal{X}_i \text{ is } \mu_i\ \forall 1 \le i \le N \Big\}.$$

Lemma (Augmentation)

The following equality holds:

$$\begin{aligned} \phi_{\mathrm{DRO}} = \underset{\boldsymbol{a}}{\text{minimize}} \quad & \langle \boldsymbol{c}_1, \boldsymbol{a} \rangle + \sup_{\mu_{\mathrm{aug}} \in \Gamma_{\mathrm{aug}}(\mu_1, \ldots, \mu_N)} \int_{\boldsymbol{\mathcal{X}} \times S_2^*} Q_{\mathrm{aug}}(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{\lambda})\, \mu_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda}) \\ \text{subject to} \quad & \mathbf{L}_{\mathrm{in}}\boldsymbol{a} \le \boldsymbol{q}_{\mathrm{in}},\ \mathbf{L}_{\mathrm{eq}}\boldsymbol{a} = \boldsymbol{q}_{\mathrm{eq}},\ \boldsymbol{a} \in \mathbb{R}^{K_1}. \end{aligned}$$

Step 2: Relaxation

- Relax the marginal constraints into finitely many linear constraints.
  - Fixing a marginal can be seen as having infinitely many linear constraints.
- Augmented couplings are replaced by a *moment set*:

Definition (Moment set (Neufeld and X. 2022))

For $i = 1, \ldots, N$, $[\mu_i]_{\mathcal{G}_i}$ is called a moment set centered at μ_i characterized by functions $\mathcal{G}_i \subset \mathcal{L}^1(\mathcal{X}_i, \mu_i)$: $\quad \nu_i \in [\mu_i]_{\mathcal{G}_i} \quad \Leftrightarrow \quad \int_{\mathcal{X}_i} g_i \, \mathrm{d}\mu_i = \int_{\mathcal{X}_i} g_i \, \mathrm{d}\nu_i \ \forall g_i \in \mathcal{G}_i$.
Moreover,

$$\Gamma_{\mathrm{aug}}([\mu_1]_{\mathcal{G}_1}, \ldots, [\mu_N]_{\mathcal{G}_N}) := \Big\{ \mu_{\mathrm{aug}} \in \Gamma_{\mathrm{aug}}(\nu_1, \ldots, \nu_N) : \nu_i \in [\mu_i]_{\mathcal{G}_i} \ \forall 1 \le i \le N \Big\}.$$

- **Relaxed two-stage DRO problem**:

$$\begin{aligned} \underset{\boldsymbol{a}}{\text{minimize}} \quad & \langle \boldsymbol{c}_1, \boldsymbol{a} \rangle + \sup_{\mu_{\mathrm{aug}} \in \Gamma_{\mathrm{aug}}([\mu_1]_{\mathcal{G}_1}, \ldots, [\mu_N]_{\mathcal{G}_N})} \int_{\boldsymbol{\mathcal{X}} \times S_2^*} Q_{\mathrm{aug}}(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{\lambda}) \, \mu_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda}) \\ \text{subject to} \quad & \mathbf{L}_{\mathrm{in}} \boldsymbol{a} \le \boldsymbol{q}_{\mathrm{in}}, \ \mathbf{L}_{\mathrm{eq}} \boldsymbol{a} = \boldsymbol{q}_{\mathrm{eq}}, \ \boldsymbol{a} \in \mathbb{R}^{K_1}. \end{aligned}$$

Step 3: Dualization

- Let $\mathcal{G}_i = \{g_{i,1}, \ldots, g_{i,m_i}\}$ for $i = 1, \ldots, N$ and let $m := \sum_{i=1}^N m_i$.
  - $\boldsymbol{g}(x_1, \ldots, x_N) := \big(g_{1,1}(x_1), \ldots, g_{N,m_N}(x_N)\big)^\mathsf{T} \quad \forall (x_1, \ldots, x_N) \in \boldsymbol{\mathcal{X}}$.
  - $\boldsymbol{v} := \Big(\int_{\mathcal{X}_1} g_{1,1} \,\mathrm{d}\mu_1, \ldots, \int_{\mathcal{X}_N} g_{N,m_N} \,\mathrm{d}\mu_N \Big)^\mathsf{T}$.
- Replacing the relaxed inner maximization problem by its dual yields the following linear semi-infinite programming (LSIP) problem:

$$\begin{aligned}
\underset{\boldsymbol{a}, y_0, \boldsymbol{y}}{\text{minimize}} \quad & \langle \boldsymbol{c}_1, \boldsymbol{a} \rangle + y_0 + \langle \boldsymbol{v}, \boldsymbol{y} \rangle \\
\text{subject to} \quad & y_0 + \langle \boldsymbol{g}(\boldsymbol{x}), \boldsymbol{y} \rangle - \langle \mathbf{V}^\mathsf{T} \boldsymbol{\lambda}, \boldsymbol{a} \rangle \geq \langle \mathbf{W}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{\lambda} \rangle \\
& \qquad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}},\ \forall \boldsymbol{\lambda} \in S_2^*, \\
& \mathbf{L}_{\mathrm{in}} \boldsymbol{a} \leq \boldsymbol{q}_{\mathrm{in}},\ \mathbf{L}_{\mathrm{eq}} \boldsymbol{a} = \boldsymbol{q}_{\mathrm{eq}}, \\
& \boldsymbol{a} \in \mathbb{R}^{K_1},\ y_0 \in \mathbb{R},\ \boldsymbol{y} \in \mathbb{R}^m.
\end{aligned} \tag{LSIP}$$

- Each approximately optimal solution $(\hat{\boldsymbol{a}}, \hat{y}_0, \hat{\boldsymbol{y}})$ of (LSIP) provides:
  - an **approximately optimal solution** $\hat{\boldsymbol{a}}$ of the DRO problem,
  - an **upper bound** $\langle \boldsymbol{c}_1, \hat{\boldsymbol{a}} \rangle + \hat{y}_0 + \langle \boldsymbol{v}, \hat{\boldsymbol{y}} \rangle$ for ϕ_{DRO} (with controlled quality).

Step 4: Bounding from below

- The problem (LSIP) admits the following dual:

$$
\begin{aligned}
\underset{\boldsymbol{\xi}_{\mathrm{in}},\boldsymbol{\xi}_{\mathrm{eq}},\mu_{\mathrm{aug}}}{\text{maximize}} \quad & \langle \boldsymbol{q}_{\mathrm{in}}, \boldsymbol{\xi}_{\mathrm{in}} \rangle + \langle \boldsymbol{q}_{\mathrm{eq}}, \boldsymbol{\xi}_{\mathrm{eq}} \rangle + \int_{\boldsymbol{\mathcal{X}} \times S_2^*} \langle \mathbf{W}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{\lambda} \rangle \, \mu_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda}) \\
\text{subject to} \quad & \mathbf{L}_{\mathrm{in}}^{\mathsf{T}} \boldsymbol{\xi}_{\mathrm{in}} + \mathbf{L}_{\mathrm{eq}}^{\mathsf{T}} \boldsymbol{\xi}_{\mathrm{eq}} - \mathbf{V}^{\mathsf{T}} \Big(\int_{\boldsymbol{\mathcal{X}} \times S_2^*} \boldsymbol{\lambda} \, \mu_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda}) \Big) = \boldsymbol{c}_1, \\
& \boldsymbol{\xi}_{\mathrm{in}} \leq \mathbf{0}, \; \mu_{\mathrm{aug}} \in \Gamma_{\mathrm{aug}}\big([\mu_1]_{\mathcal{G}_1}, \ldots, [\mu_N]_{\mathcal{G}_N}\big), \\
& \boldsymbol{\xi}_{\mathrm{in}} \in \mathbb{R}^{n_{\mathrm{in}}}, \; \boldsymbol{\xi}_{\mathrm{eq}} \in \mathbb{R}^{n_{\mathrm{eq}}}, \; \mu_{\mathrm{aug}} \in \mathcal{P}(\boldsymbol{\mathcal{X}} \times S_2^*).
\end{aligned} \tag{LSIP*}
$$

- For each approximately optimal solution $(\hat{\boldsymbol{\xi}}_{\mathrm{in}}, \hat{\boldsymbol{\xi}}_{\mathrm{eq}}, \hat{\mu}_{\mathrm{aug}})$ of (LSIP*):
  - if $\tilde{\mu}_{\mathrm{aug}} \in \Gamma_{\mathrm{aug}}(\mu_1, \ldots, \mu_N)$ satisfies

$$\mathbf{L}_{\mathrm{in}}^{\mathsf{T}} \hat{\boldsymbol{\xi}}_{\mathrm{in}} + \mathbf{L}_{\mathrm{eq}}^{\mathsf{T}} \hat{\boldsymbol{\xi}}_{\mathrm{eq}} - \mathbf{V}^{\mathsf{T}} \Big(\int_{\boldsymbol{\mathcal{X}} \times S_2^*} \boldsymbol{\lambda} \, \tilde{\mu}_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda}) \Big) = \boldsymbol{c}_1,$$

    then $\langle \boldsymbol{q}_{\mathrm{in}}, \hat{\boldsymbol{\xi}}_{\mathrm{in}} \rangle + \langle \boldsymbol{q}_{\mathrm{eq}}, \hat{\boldsymbol{\xi}}_{\mathrm{eq}} \rangle + \int_{\boldsymbol{\mathcal{X}} \times S_2^*} \langle \mathbf{W}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{\lambda} \rangle \, \tilde{\mu}_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda})$ is a **lower bound** for ϕ_{DRO} (with controlled quality).

Step 4: Bounding from below

Definition (Partial reassembly)

Let $\bar{\mathcal{X}}_i := \mathcal{X}_i$ for $i = 1, \ldots, N$. $\tilde{\mu}_{\mathrm{aug}}$ is called a partial reassembly of $\hat{\mu}_{\mathrm{aug}} \in \mathcal{P}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times S_2^*)$ with the marginals $\mu_1, \ldots, \mu_N$ if there exists a probability measure $\gamma \in \mathcal{P}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times \bar{\mathcal{X}}_1 \times \cdots \times \bar{\mathcal{X}}_N \times S_2^*)$ such that:

1. the marginal of γ on $\mathcal{X}_1 \times \cdots \times \mathcal{X}_N \times S_2^*$ is $\hat{\mu}_{\mathrm{aug}}$;
2. for $i = 1, \ldots, N$, the marginal $\gamma_i \in \Gamma(\hat{\mu}_i, \mu_i)$ of γ on $\mathcal{X}_i \times \bar{\mathcal{X}}_i$ satisfies $\int_{\mathcal{X}_i \times \bar{\mathcal{X}}_i} |x - y| \, \gamma_i(\mathrm{d}x, \mathrm{d}y) = W_1(\hat{\mu}_i, \mu_i)$;
3. the marginal of γ on $\bar{\mathcal{X}}_1 \times \cdots \times \bar{\mathcal{X}}_N \times S_2^*$ is $\tilde{\mu}_{\mathrm{aug}}$.

The set of *partial reassemblies* is denoted by

$$R_{\mathrm{part}}(\hat{\mu}_{\mathrm{aug}}; \mu_1, \ldots, \mu_N) \subset \Gamma_{\mathrm{aug}}(\mu_1, \ldots, \mu_N).$$

- **Idea:** morphing $\hat{\mu}_{\mathrm{aug}}$ in an “optimal” way to turn its marginals on $\mathcal{X}_1, \ldots, \mathcal{X}_N$ into $\mu_1, \ldots, \mu_N$ while leaving its marginal on S_2^* unchanged.
- One can construct a partial reassembly using Sklar’s theorem from copula theory.

Controlling the approximation error

Theorem (Approximation of two-stage DRO with marginal constraints)

Suppose that:

1. *for* $i = 1, \ldots, N$, $\mathcal{G}_i$ *contains only continuous functions;*
2. $(\hat{\boldsymbol{a}}, \hat{y}_0, \hat{\boldsymbol{y}})$ *is an* ϵ*-optimal solution of* (LSIP) *for* $\epsilon > 0$;
3. $(\hat{\boldsymbol{\xi}}_{\mathrm{in}}, \hat{\boldsymbol{\xi}}_{\mathrm{eq}}, \hat{\mu}_{\mathrm{aug}})$ *is an* ϵ^**-optimal solution of* (LSIP*) *for* $\epsilon^* > 0$;
4. $\tilde{\mu}_{\mathrm{aug}} \in R_{\mathrm{part}}(\hat{\mu}_{\mathrm{aug}}; \mu_1, \ldots, \mu_N)$;
5. $\phi^{\mathsf{UB}}_{\mathrm{DRO}} := \langle \boldsymbol{c}_1, \hat{\boldsymbol{a}} \rangle + \hat{y}_0 + \langle \boldsymbol{v}, \hat{\boldsymbol{y}} \rangle$;
6. $\phi^{\mathsf{LB}}_{\mathrm{DRO}} := \langle \boldsymbol{q}_{\mathrm{in}}, \hat{\boldsymbol{\xi}}_{\mathrm{in}} \rangle + \langle \boldsymbol{q}_{\mathrm{eq}}, \hat{\boldsymbol{\xi}}_{\mathrm{eq}} \rangle + \int_{\boldsymbol{\mathcal{X}} \times S_2^*} \langle \mathbf{W}\boldsymbol{x} + \boldsymbol{b}, \boldsymbol{\lambda} \rangle \, \tilde{\mu}_{\mathrm{aug}}(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{\lambda})$;
7. $\tilde{\epsilon} := \epsilon + \epsilon^* + \left(\sum_{i=1}^{N} \sup_{\nu_i \in [\mu_i]_{\mathcal{G}_i}} \left\{ W_1(\mu_i, \nu_i) \right\} \right) \sup_{\boldsymbol{\lambda} \in S_2^*} \left\{ \|\mathbf{W}^{\mathsf{T}} \boldsymbol{\lambda}\|_\infty \right\}$.

Then,

- $\phi^{\mathsf{LB}}_{\mathrm{DRO}} \le \phi_{\mathrm{DRO}} \le \phi^{\mathsf{UB}}_{\mathrm{DRO}}$ *with* $\phi^{\mathsf{UB}}_{\mathrm{DRO}} - \phi^{\mathsf{LB}}_{\mathrm{DRO}} \le \tilde{\epsilon}$;
- $\hat{\boldsymbol{a}}$ *is an* $\hat{\epsilon}$***-optimal solution*** *of the two-stage DRO problem, where* $\hat{\epsilon} := \phi^{\mathsf{UB}}_{\mathrm{DRO}} - \phi^{\mathsf{LB}}_{\mathrm{DRO}} \le \tilde{\epsilon}$.

Practical questions

- **Question 1**: for any $\epsilon > 0$, can we obtain ϵ-optimal solutions of (LSIP) and (LSIP*)?
  - **Answer 1**: yes, we develop a suitable cutting-plane algorithm (inspired by Conceptual Algorithm 11.4.1 of Goberna and López (1998)) to obtain ϵ-optimal solutions of (LSIP) and (LSIP*).
- **Question 2**: can we numerically evaluate an integral with respect to a $\tilde{\mu}_{\mathrm{aug}} \in R_{\mathrm{part}}(\hat{\mu}_{\mathrm{aug}}; \mu_1, \dots, \mu_N)$?
  - **Answer 2**: yes, using the copula theory, we develop an algorithm to explicitly construct a partial reassembly and efficiently generate independent random samples from it.
- **Question 3**: can we control $\sup_{\nu_i \in [\mu_i]_{\mathcal{G}_i}} \{ W_1(\mu_i, \nu_i) \}$ to be arbitrarily close to 0 for $i = 1, \dots, N$?
  - **Answer 3**: yes, for any $\epsilon_i > 0$ we can explicitly construct a finite collection $\mathcal{G}_i$ of continuous piece-wise affine functions with $\sup_{\nu_i \in [\mu_i]_{\mathcal{G}_i}} \{ W_1(\mu_i, \nu_i) \} \leq \epsilon_i$.

Numerical algorithm

- Putting these pieces together, we develop a numerical algorithm, whose properties are summarized as follows.

Theorem (Properties of the proposed algorithm)

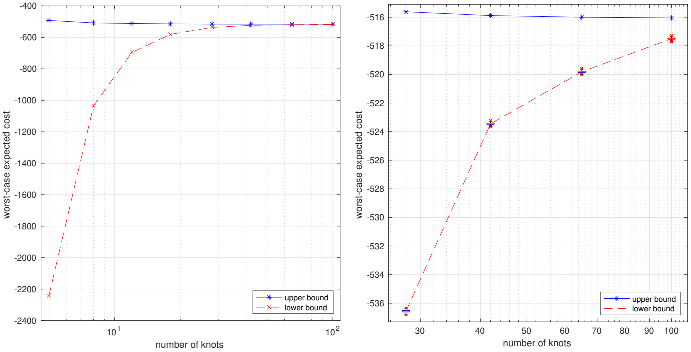
Under suitable conditions, for any $\tilde{\epsilon} > 0$, there exists inputs to the proposed algorithm such that it produces the outputs: $\hat{\boldsymbol{a}}$, $\phi^{\mathsf{LB}}_{\mathrm{DRO}}$, $\phi^{\mathsf{UB}}_{\mathrm{DRO}}$ such that

1. $\phi^{\mathsf{LB}}_{\mathrm{DRO}} \leq \phi_{\mathrm{DRO}} \leq \phi^{\mathsf{UB}}_{\mathrm{DRO}}$;
2. $\hat{\boldsymbol{a}}$ *is an* $\hat{\epsilon}$-***optimizer*** *of the two-stage DRO problem, where* $\hat{\epsilon} := \phi^{\mathsf{UB}}_{\mathrm{DRO}} - \phi^{\mathsf{LB}}_{\mathrm{DRO}} \leq \tilde{\epsilon}$.

- **Remark:** the sub-optimality measure $\hat{\epsilon}$ can be computed, and it is often much less conservative than its theoretical upper bound $\tilde{\epsilon}$.

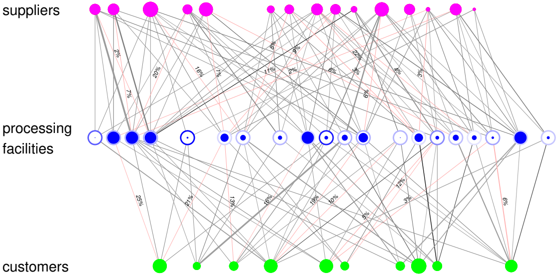
Convergence of the bounds

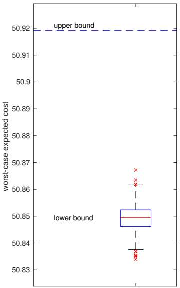
- When appropriately chosen continuous piece-wise affine functions are incrementally added to $(\mathcal{G}_i)_{i=1:N}$, the difference between the upper bound $\phi^{\mathrm{UB}}_{\mathrm{DRO}}$ and the lower bound $\phi^{\mathrm{LB}}_{\mathrm{DRO}}$ goes to 0.



Numerical example: supply chain network design

- **Settings:**
  - We consider 15 suppliers, 20 processing facilities, 10 customers, 150 edges with 25 susceptible to failure. ($N = 10 + 25$, $K_1 = 170$, $K_2 = 150$);
  - $\mathcal{X}_1 = \mathcal{X}_2 = \cdots = \mathcal{X}_{10} = [0, 2]$, $\mathcal{X}_{11} = \mathcal{X}_{12} = \cdots = \mathcal{X}_{35} = \{0, 1\}$.
  - $\mu_1, \ldots, \mu_{10}$ are mixture of truncated normal distributions.
  - Parameters in the model are randomly generated.
- **Result:**





The difference between the upper bound and the lower bound is ~ 0.07.

References

1. A. Neufeld and Q. Xiang. Numerical method for approximately optimal solutions of two-stage distributionally robust optimization with marginal constraints. Preprint, arXiv:2205.05315, 2022.
URL: https://arxiv.org/abs/2205.05315

2. R. Gao, A. J. Kleywegt. Data-driven robust optimization with known marginal distributions. Working paper, 2017.

3. L. Chen, W. Ma, K. Natarajan, D. Simchi-Levi, Z. Yan. Distributionally robust linear and discrete optimization with marginals. *Operations Research*, 2021.

4. M. A. Goberna and M. A. López. *Linear semi-infinite optimization.* John Wiley & Sons, 1998.