

Yutian He¹, Beste Basciftci², Qihang Lin²

¹Department of Mathematics, University of Iowa; ²Department of Business Analytics, University of Iowa

1

Introduction

Two-stage stochastic program

$$\text{cost}^* = \min_{x \in X, w(\xi)} g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] \quad \text{s.t.} \quad A_\xi x + B_\xi w(\xi) \leq h_\xi, \quad \forall \xi$$

- x : First-stage decision chosen from a convex closed set X .
- ξ : Random variable.
- $w(\xi)$: Second-stage decision.
- g : Continuous function for the first-stage cost.
- Q : Continuous function (linear) for the second-stage cost.
- $\mathbb{E}$: The expectation over ξ .
- cost^* : Minimum cost without considering fairness.

Wasserstein p-Distance

Definition

Wasserstein p -Distance $W_p(\cdot, \cdot)$ is define as:

$$W_p(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{Q}} \left\{ \left(\int_{\Xi \times \Xi} \|\zeta_1 - \zeta_2\|^p Q(d\zeta_1, d\zeta_2) \right)^{1/p} \right\}$$

where $\mathbb{Q}$ is the joint distribution of ζ_1 and ζ_2 with marginals $\mathbb{P}_1$ and $\mathbb{P}_2$, respectively. $\|\cdot\|$ is a norm and $p \in [1, \infty)$.

Distributionally fair two-stage single level and two-stage bilevel problem

DF-TSSL

DF-TSBL

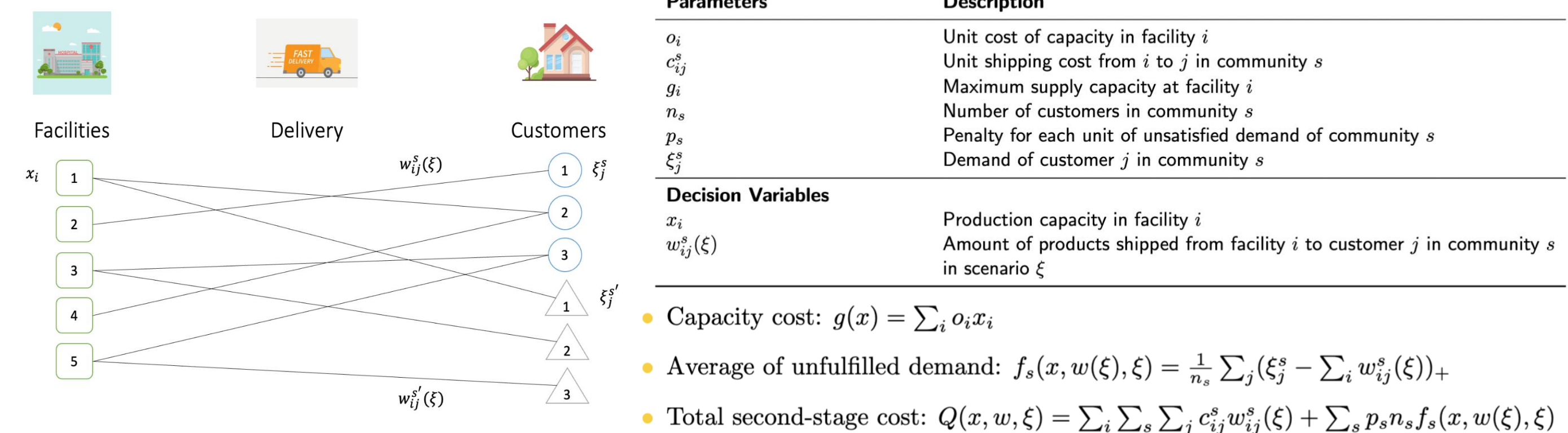
$$\begin{aligned} \min_{x \in X, w(\xi)} \max_{s, s'} W_p^p(f_s(x, w(\xi), \xi), f_{s'}(x, w(\xi), \xi)) & \quad \min_{x \in X, w(\xi)} \max_{s, s'} W_p^p(f_s(x, w(\xi), \xi), f_{s'}(x, w(\xi), \xi)) \\ \text{s.t.} \quad g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] & \leq \text{cost}^* \cdot \kappa & \text{s.t.} \quad g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] & \leq \text{cost}^* \cdot \kappa \\ A_\xi x + B_\xi w(\xi) & \leq h_\xi, \quad \forall \xi & w(\xi) \in \arg \min_w \{Q(x, w, \xi) : A_\xi x + B_\xi w & \leq h_\xi\}, \quad \forall \xi \end{aligned}$$

- $f_s(x, w(\xi), \xi)$: Utility functions for community s , $s \in S$ under decisions x and $w(\xi)$ in scenario ξ .
- κ : Tolerance on cost, ≥ 1 .

2

Motivation

Capacity planning problem

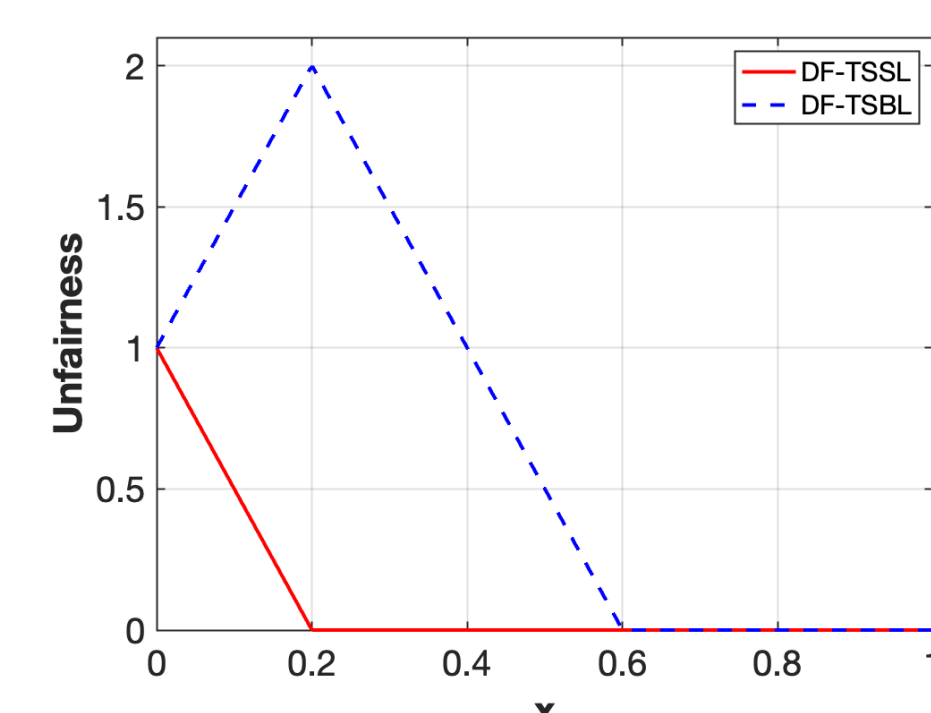


$$\begin{aligned} \min \quad & \sum_i \alpha_i x_i + \frac{1}{m} \sum_{\xi \in \Xi} \left(\sum_i \sum_s \sum_j c_{ij} w_{ij}^s(\xi) + \sum_s p_s n_s f_s(x, w(\xi), \xi) \right) \\ \text{s.t.} \quad & \sum_s \sum_j w_{ij}^s(\xi) - g_i x_i \leq 0, \quad \forall i, \xi \in \Xi, \quad \sum_s \sum_j w_{ij}^s(\xi) - \xi_j^s \leq 0, \quad \forall j, \xi \in \Xi, s \\ & x_i \in [0, 1], \quad \forall i, \quad w_{ij}^s(\xi) \geq 0, \quad \forall i, j, s, \xi \in \Xi \end{aligned}$$

- m : Number of scenarios.
- ξ : Scenario uniformly sampled from a discrete support set $\Xi = \{\xi_1, \xi_2, \dots, \xi_m\}$.

Assume $S = 2$, $m = 1$, $n_s = 1$, $n = 1$, $\xi_1^1 = 2$, $\xi_1^2 = 1$, $c_s = 1$, $g_1 = 5$, $p_1 = 1$, $p_2 = 2$ and $\kappa = \infty$.

- In **DF-TSSL**, both first- and second-stage decisions aim to minimize the fairness while incorporating total cost into consideration.
- In **DF-TSBL**, first-stage decisions balance the trade-off between cost and fairness, while second-stage decisions focus only on minimizing the second-stage costs.



3

Reformulations

Lemma (Ye et al, 2024)

For any $s \in S$ and a fixed decision $(x, w(\xi))$, the $W_p^p(f_s(x, w(\xi), \xi), f_{s'}(x, w(\xi), \xi))$ (we use $W_p^p(f_s, f_{s'})$ for simplicity) can be expressed as:

$$W_p^p(f_s, f_{s'}) = \int_0^1 |F_s^{-1}(y|x, w(\xi)) - F_{s'}^{-1}(y|x, w(\xi))|^p dy$$

where F_s^{-1} is the inverse distribution function of the random function $f_s(x, w(\xi), \xi)$ for each $s \in S$.

Under a discrete support set $\Xi = \{\xi_1, \xi_2, \dots, \xi_m\}$, then the result in the lemma can be reformulated as:

$$W_p^p(f_s, f_{s'}) = \sum_{k \in [m]-1} \frac{1}{m} |t_{ks} - t_{ks'}|^p$$

where t_{ks} is the $\frac{k}{m}$ quantile of $f_s(x, w(\xi), \xi)$, having a mixed-integer convex representation.

Quantile formulation by mixed-integer convex programs

$$\begin{aligned} \min_{x \in X, w(\xi), t} \quad & \max_{s, s'} \sum_{k \in [m]-1} \frac{1}{m} |t_{ks} - t_{ks'}|^p \\ \text{s.t.} \quad & g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] \leq \text{cost}^* \cdot \kappa \\ & A_\xi x + B_\xi w(\xi) \leq h_\xi, \quad \forall \xi \\ & t_{ks} = \frac{k}{m} \text{ quantile of } f_s(x, w(\xi), \xi), \quad \forall k \in [m]-1, s \in S. \end{aligned}$$

DF-TSSL

DF-TSBL

Version 1: Bilevel constraint based on complementary slackness

$$\lambda_\xi \geq 0, \quad \lambda_\xi^T (A_\xi x + B_\xi w(\xi) - h_\xi) = 0, \quad \nabla_w Q(x, w(\xi), \xi) + B_\xi^T \lambda_\xi = 0, \quad \forall \xi.$$

Big-M reformulation for Version 1

$$0 \leq \lambda_\xi \leq M y_\xi, \quad -M(1 - y_\xi) \leq A_\xi x + B_\xi w(\xi) - h_\xi \leq 0, \quad y_\xi \in \{0, 1\}, \quad \forall \xi.$$

Version 2: Bilevel constraint based on strong duality

$$\begin{aligned} \lambda_\xi \geq 0, \quad Q(x, w(\xi), \xi) & \leq Q(x, w'(\xi), \xi) + \lambda_\xi^T (A_\xi x + B_\xi w'(\xi) - h_\xi), \\ \nabla_w Q(x, w'(\xi), \xi) + B_\xi^T \lambda_\xi & = 0, \quad \lambda_\xi \geq 0, \quad \forall \xi. \end{aligned}$$

Difference-of-Convex (DC) program

DC constrained problem

Proposition

$$W_p^p(f_s(x, w(\xi), \xi), f_{s'}(x, w(\xi), \xi)) := W_p^+(x, w(\xi)) - W_p^-(x, w(\xi))$$

where W_p^+ and W_p^- are convex functions in x and $w(\xi)$.

Proof sketch: W_p^p can be written as the minimum of finitely many continuous functions of x and $w(\xi)$.

$$\begin{aligned} \min_{x \in X, w(\xi)} \quad & W_p^+(x, w(\xi)) - W_p^-(x, w(\xi)) \\ \text{s.t.} \quad & g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] \leq \text{cost}^* \cdot \kappa, \quad A_\xi x + B_\xi w(\xi) \leq h_\xi, \quad \forall \xi \\ & Q(x, w(\xi), \xi) - Q^*(x, \xi) \leq 0, \quad \forall \xi \\ & \text{where } Q^*(x, \xi) := \min_w \{Q(x, w(\xi), \xi) : A_\xi x + B_\xi w(\xi) \leq h_\xi\}, \quad \forall \xi \end{aligned}$$

DF-TSBL

DC constrained problem general form

$$\begin{aligned} \min_{z \in Z} \quad & h_0(z) := h_0^+(z) - h_0^-(z) \\ \text{s.t.} \quad & h_i(z) := h_i^+(z) - h_i^-(z) \leq 0, \quad i \in [n] \\ & \Rightarrow \min_{v \in Z} H_0(v) := h_0^+(v) - \hat{h}_0^-(v; z^{(k)}) \\ & \text{s.t. } H_i(v) := h_i^+(v) - \hat{h}_i^-(v; z^{(k)}) \leq 0, \quad i \in [n], \end{aligned}$$

- h_i^+ and h_i^- for $i = 0, \dots, n$ are real-valued, convex functions.
- $\hat{h}_i^-(v; z) := h_i^-(z) + (h_i^-)^T(v - z)$.
- h_i^- : Arbitrary subgradient in $\partial h_i^-(z)$.

Algorithm Difference-of-Convex Algorithm (DCA), (Sriperumbudur and Lanckriet, 2012)

- Input:** A feasible $z^{(0)}$, the number of iterations K .
- for** $k = 0, 1, \dots, K-1$ **do**
- Find $z^{(k+1)}$ as an optimal solution of the subproblem.
- end for**

Computationally efficient lower bound

Convex lower bound of the objective function based on Jensen inequality

Lemma

For and $p \geq 1$, $W_p^p(\cdot, \cdot)$ is bounded by:

$$\begin{aligned} \max_{s, s'} W_p^p(f_s(x, w_\xi, \xi), f_{s'}(x, w_\xi, \xi)) & \geq \\ \max_{s, s'} |\mathbb{E}[f_s(x, w_\xi, \xi)] - \mathbb{E}[f_{s'}(x, w_\xi, \xi)]|^p & \end{aligned}$$

where $\mathbb{E}$ is the expectation over ξ .

Valid inequality for bilevel constraint based on strong duality (Kleinert and Schmidt, 2023)

$$A_{i\xi}^+ := \max_{x \in X, w(\xi)} A_{i\xi} x \quad \text{s.t.} \quad A_\xi x + B_\xi w(\xi) \leq h_\xi, \quad g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] \leq \text{cost}^* \cdot \kappa, \quad \forall \xi.$$

Formulation

$$\begin{aligned} \min_{x \in X, w(\xi)} \quad & \max_{s, s'} |\mathbb{E}[f_s(x, w_\xi, \xi)] - \mathbb{E}[f_{s'}(x, w_\xi, \xi)]|^p \\ \text{s.t.} \quad & g(x) + \mathbb{E}[Q(x, w(\xi), \xi)] \leq \text{cost}^* \cdot \kappa \\ & A_\xi x + B_\xi w(\xi) \leq h_\xi, \quad \forall \xi \\ & Q(x, w(\xi), \xi) \leq Q(x, w'(\xi), \xi) + \lambda_\xi^T (A_\xi x + B_\xi w'(\xi) - h_\xi), \quad \forall \xi \\ & \nabla_w Q(x, w'(\xi), \xi) + B_\xi^T \lambda_\xi = 0, \quad \lambda_\xi \geq 0, \quad \forall \xi \end{aligned}$$

where A_ξ^+ denotes the vector of upper bounds $A_{i\xi}^+$.

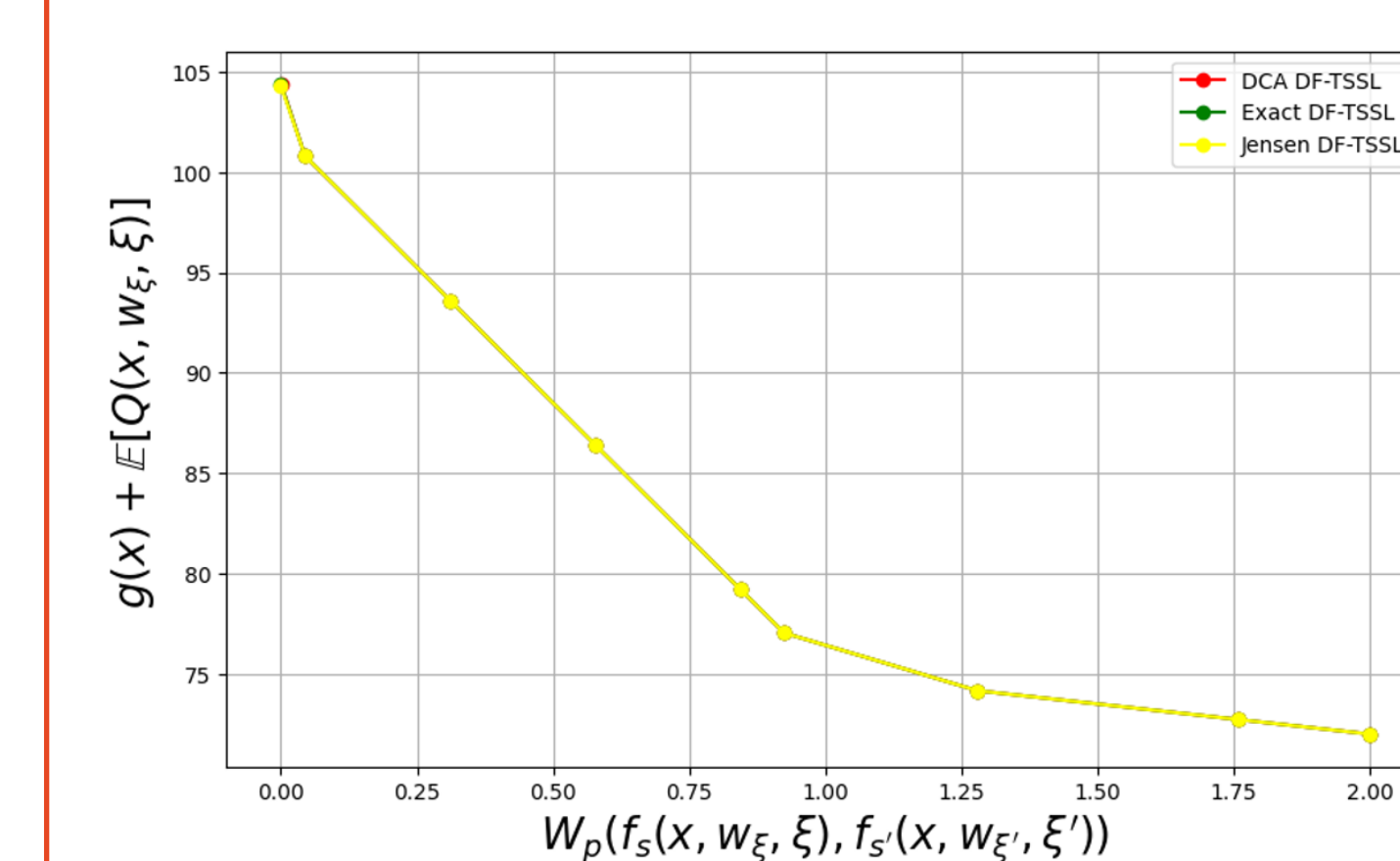
4

Computational Study

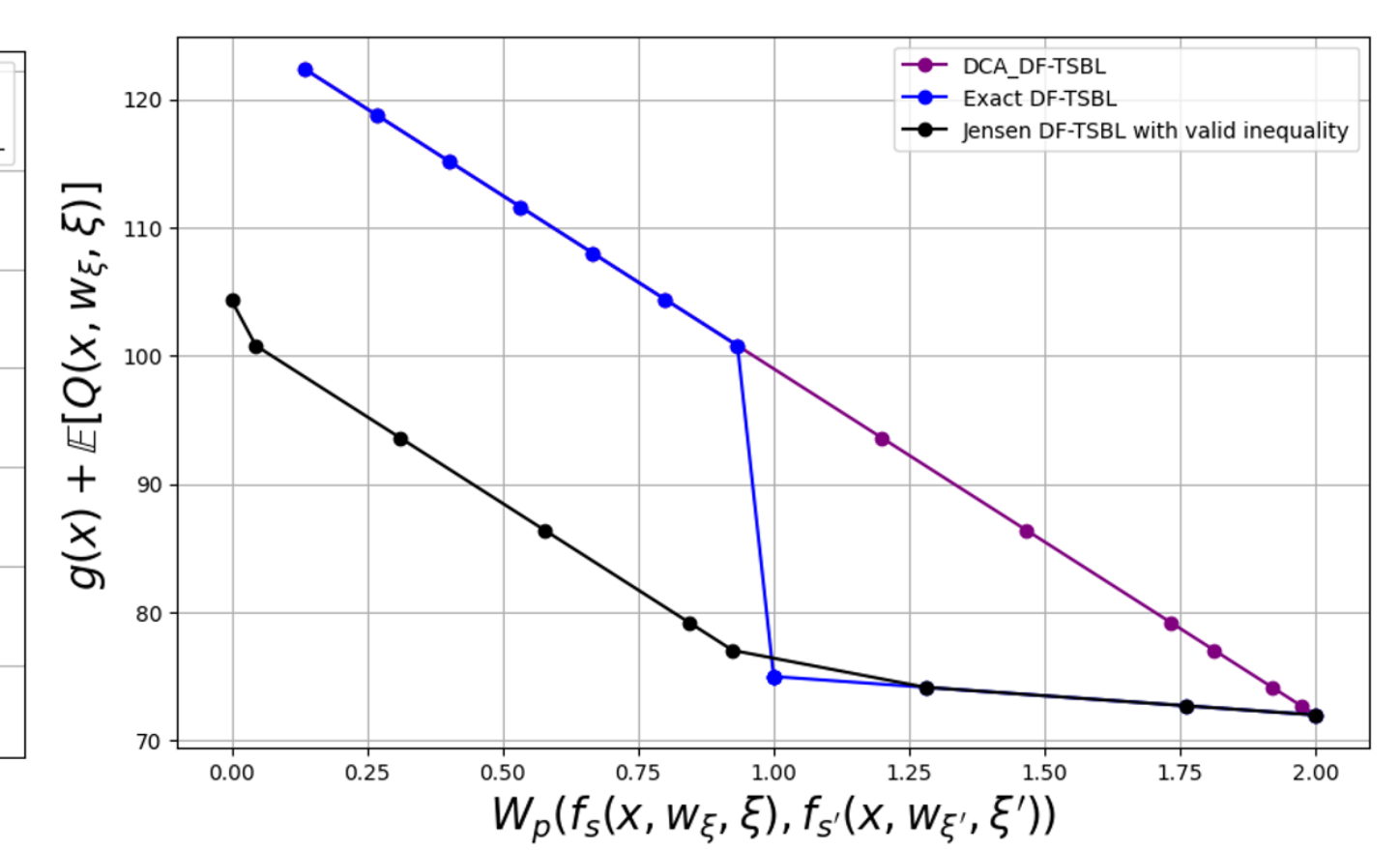
Problem setting

- $S = 2$, $m = 10$, $n_s = 3$, $n = 5$, $c_s = 5$, $p_1 = 5$, $p_2 = 15$, $\xi_j^1 = 2$, $\xi_j^2 = 1$.
- κ varies from 1 to 1.7.

Numerical results for solving DF-TSSL



Numerical results for solving DF-TSBL



Running time for solving DF-TSBL

κ	Exact DF-TSBL	Exact DF-TSBL with big-M method	DCA	Jensen relaxation with valid inequality
1.00	4.19	4.20	4.11	1.23
1.01	7.16	15.67	3.99	1.05
1.03	7.08	31.54	3.99	1.07
1.07	TL	16.51	5.95	1.08
1.10	TL	36.65	5.96	1.07
1.20	TL	94.81	11.99	1.07
1.30	TL	2704.95	16.02	1.06
1.40	TL	1941.68	22.07	1.47
1.45	TL	1070.18	25.91	1.06
1.50	TL	44.20	29.97	1.15
1.55	TL	106.52	36.04	1.15
1.60	TL	35.06	44.04	1.07
1.65	TL	744.82	53.86	1.07
1.70	TL	86.79	71.94	1.07

TL: Time limit is 3600 seconds

- In **DF-TSSL**: all three methods yield identical results.
- In **DF-TSBL**: DCA obtains the optimal solution when the unfairness is small; the lower bound is tight when unfairness is high.

5

References

- Thomas Kleinert and Martin Schmidt. Why there is no need to use a big-M in linear bilevel optimization: A computational study of two ready-to-use approaches. *Computational Management Science*, 20(1):3-7, 2023.
- Bharath K. Sriperumbudur and Gert R. G. Lanckriet. A proof of convergence of the concave-convex procedure using zangwill's theory. *Neural Computation*, 24(6):1391-1407, 2012.
- Qing Ye, Grani A. Hanasusanto, and Weijun Xie. Distributionally Fair Stochastic Optimization using Wasserstein Distance. *arXiv preprint 2402.01872*, 2024.