The complexity of pricing for the two-stage vehicle routing problem

Ricardo Fukasawa

Department of Combinatorics & Optimization
University of Waterloo

Column Generation Workshop 2023
May 18, 2023
Joint work with J. Gunter and M. Ota

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS
Department of Combinatorics
and Optimization

Why I couldn't do pricing for the two-stage vehicle routing problem

Ricardo Fukasawa

Department of Combinatorics & Optimization
University of Waterloo

Column Generation Workshop 2023
May 18, 2023
Joint work with J. Gunter and M. Ota

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS
Department of Combinatorics
and Optimization

Outline

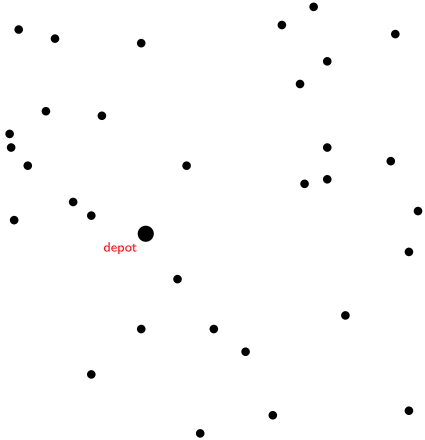
Takeaway 1:

- Spliet (2023):
  "Complexity result for pricing not enough for a CG talk"
  (paraphrased)

Takeaway 1:

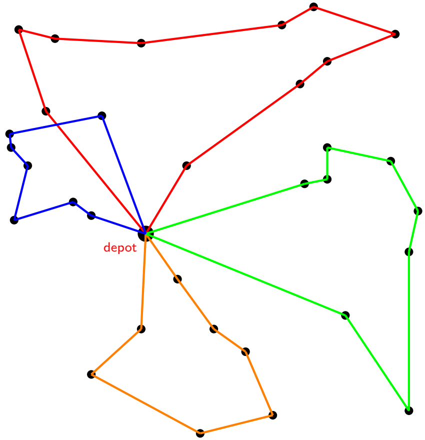
- Spliet (2023):
  "Complexity result for pricing not enough for a CG talk"
  (paraphrased)
- Conclusion:
  I shouldn't be giving this talk!

The two-stage vehicle routing problem



- $G = (V, E)$
- $V = \{0\} \cup V_+$
- Edge lengths $\ell_e,\ e \in E$
- K vehicles, capacity C
- Client demands $D_i, \forall i \in V_+$ are random variables.
- Let S_j be the set of clients served by route j.
  Then
  $\mathbb{E}[D(S_j)] := \sum_{u \in S_j} \mathbb{E}[D_u] \leq C$

The two-stage vehicle routing problem



- $G = (V, E)$
- $V = \{0\} \cup V_+$
- Edge lengths $\ell_e,\ e \in E$
- K vehicles, capacity C
- Client demands $D_i, \forall i \in V_+$ are random variables.
- Let S_j be the set of clients served by route j.
  Then
  $\mathbb{E}[D(S_j)] := \sum_{u \in S_j} \mathbb{E}[D_u] \leq C$

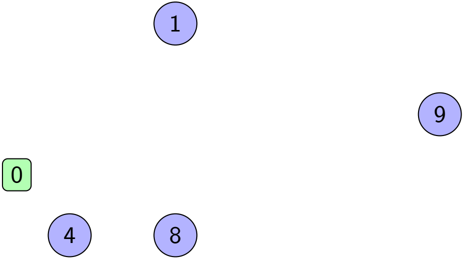
Two-stage stochastic VRP: Costs

- Routes are decided a-priori (first stage): Incur a First stage cost
- After demand realization, truck will follow route and pay second stage cost if capacity is exceeded
- Goal: Minimize expected cost (=first + expected second stage)

Recourse actions

Simple recourse: $C = 10$

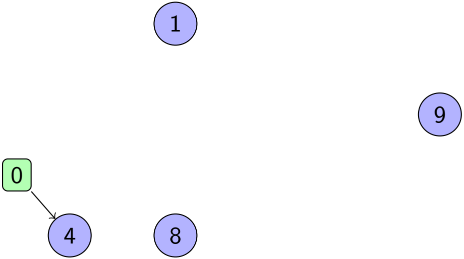
Cap used:



Recourse actions

Simple recourse: $C = 10$

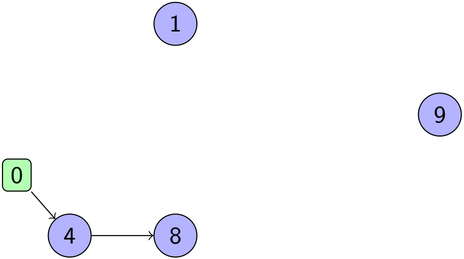
Cap used: 4



Recourse actions

Simple recourse: $C = 10$

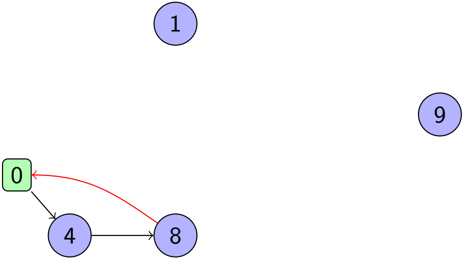
Cap used: 12



Recourse actions

Simple recourse: $C = 10$

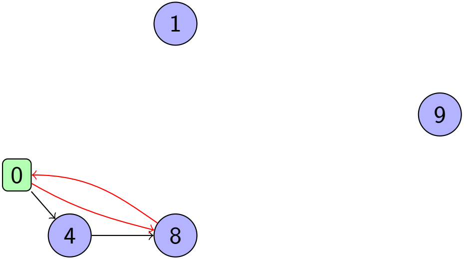
Cap used: 2



Recourse actions

Simple recourse: $C = 10$

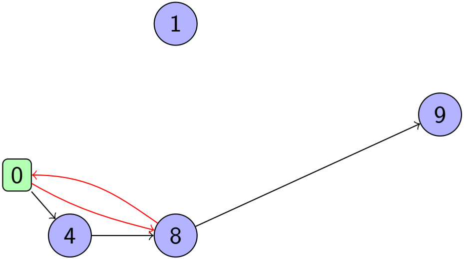
Cap used: 2



Recourse actions

Simple recourse: $C = 10$

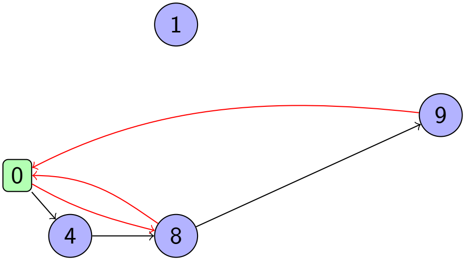
Cap used: 11



Recourse actions

Simple recourse: $C = 10$

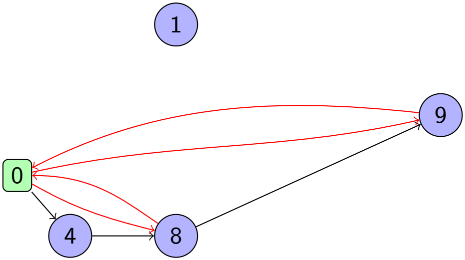
Cap used: 1



Recourse actions

Simple recourse: $C = 10$

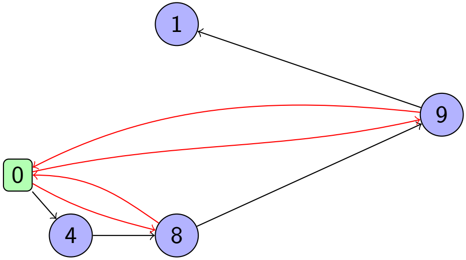
Cap used: 1



Recourse actions

Simple recourse: $C = 10$

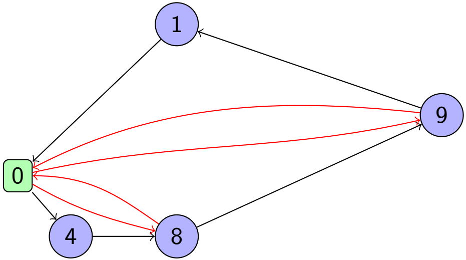
Cap used: 2



Recourse actions

Simple recourse: $C = 10$

Cap used: 2



The 2-stage VRPSD: Problem definition

Given:

- $G = (V, E)$, $V = \{0\} \cup V_+$ (assume complete graph)
- Edge lengths ℓ_e, $e \in E$
- K vehicles, capacity C
- Random demands $D_j, \forall j \in V_+$

Given a route $r = (0, v_1, \ldots, v_k, 0)$, with $v_0 = v_{k+1} = 0$.

- It is feasible if $\sum_{j=1}^{k} \mathbb{E}[D_{v_j}] \leq C$
- Its first-stage cost c'_r is the sum of the edge lengths, i.e. $c'_r := \sum_{j=0}^{k} \ell_{v_j, v_{j+1}}$
- Its second-stage cost c''_r is the expected cost due to failures
  - Let $D(r, i) := \sum_{j=1}^{i} D_{v_j}$
  - The expected failure cost at the i-th vertex v_i is $EFC(r, i) := \sum_{u=1}^{\infty} 2\ell_{0,v_i}(\mathbb{P}[D(r, i-1) \leq uC \text{ and } D(r, i) > uC])$
  - $c''_r = \sum_{j=1}^{k} EFC(r, j)$
- Total expected cost: $\tilde{c}_r := c'_r + c''_r$

Literature review

VRPSD (2-stage)

- Heuristics: Stewart & Golden (1983), Dror & Trudeau (1986), Savelsbergh & Goetschalckx (1995), Novoa et al. (2006), Secomandi and Margot (2009), ...
- Integer L-Shaped: Gendreau et al. (1994), Laporte et al. (2002), ...
- Branch-and-cut: Laporte et al. (1989), ...
- Branch-and-price: Christiansen et al. (2007)
- Branch-and-cut-and-price: Gauvin et al. (2014)
- Complex recourse policies: Florio et al. (2020, 2021, 2022), Salavati-Khoshghalb et al. (2019), Louveaux and Salazar-González (2018).

Note: All approaches rely on strong assumptions on demands (independent random variables and/or particular distribution, like normal)

Literature review

VRPSD (2-stage)

- Heuristics: Stewart & Golden (1983), Dror & Trudeau (1986), Savelsbergh & Goetschalckx (1995), Novoa et al. (2006), Secomandi and Margot (2009), . . .
- Integer L-Shaped: Gendreau et al. (1994), Laporte et al. (2002), . . .
- Branch-and-cut: Laporte et al. (1989), . . .
- Branch-and-price: Christiansen et al. (2007)
- Branch-and-cut-and-price: Gauvin et al. (2014)
- Complex recourse policies: Florio et al. (2020, 2021, 2022), Salavati-Khoshghalb et al. (2019), Louveaux and Salazar-González (2018).

Note: All approaches rely on strong assumptions on demands (independent random variables and/or particular distribution, like normal)

Stochastic IP:

- Sample average approximation approach widely used (Shultz 1996, Ahmed and Shapiro 2002, Wang and Ahmed 2008, ...)
- Idea: Sample true distribution and use these samples as "proxy" for it
- Reduces problem to finite discrete distribution

Literature review

VRPSD (2-stage)

- Heuristics: Stewart & Golden (1983), Dror & Trudeau (1986), Savelsbergh & Goetschalckx (1995), Novoa et al. (2006), Secomandi and Margot (2009), ...
- Integer L-Shaped: Gendreau et al. (1994), Laporte et al. (2002), ...
- Branch-and-cut: Laporte et al. (1989), ...
- Branch-and-price: Christiansen et al. (2007)
- Branch-and-cut-and-price: Gauvin et al. (2014)
- Complex recourse policies: Florio et al. (2020, 2021, 2022), Salavati-Khoshghalb et al. (2019), Louveaux and Salazar-González (2018).

Note: All approaches rely on strong assumptions on demands (independent random variables and/or particular distribution, like normal)

Stochastic IP:

- Sample average approximation approach widely used (Shultz 1996, Ahmed and Shapiro 2002, Wang and Ahmed 2008, ...)
- Idea: Sample true distribution and use these samples as "proxy" for it
- Reduces problem to finite discrete distribution

Goal:

Develop models assuming that we are given a finite discrete distribution (scenarios).

Set partitioning with q-routes IP formulation (Balinsky and Quandt(1964), Christofides, Mingozzi and Toth, 1981)

Definition: A q-route is a walk that starts at the depot, traverses a sequence of customers with total demand at most C, and returns to the depot.
One binary variable (z_r) per possible q-route:

$$\begin{array}{lrcll} \min & \sum_{r \in \mathcal{R}} \tilde{c}_r z_r & & & \\ \text{s.t.} & \sum_{r \in \mathcal{R}} z_r & = & k & \\ & \sum_{r \in \mathcal{R}} a_{ir} z_r & = & 1 & , \forall i \in V \\ & z_r & \in & \{0,1\} & \end{array}$$

where:

- $\mathcal{R}$: set of all possible vehicle q-routes
- a_{ir} number of times q-route r goes through customer i.

Pricing problem

After incorporating dual variables in costs, pricing problem amounts to solving a problem of the form:

$$\min_{r \in \mathcal{R}} \tilde{c}_r \qquad \text{(2SQ)}$$

Pricing problem

After incorporating dual variables in costs, pricing problem amounts to solving a problem of the form:

$$\min_{r \in \mathcal{R}} \tilde{c}_r \qquad \text{(2SQ)}$$

Main question

Can (2SQ) be solved in pseudo-polynomial time?

Pricing problem

After incorporating dual variables in costs, pricing problem amounts to solving a problem of the form:

$$\min_{r \in \mathcal{R}} \tilde{c}_r \tag{2SQ}$$

Main question

Can (2SQ) be solved in pseudo-polynomial time? NO

Pricing problem

After incorporating dual variables in costs, pricing problem amounts to solving a problem of the form:

$$\min_{r \in \mathcal{R}} \tilde{c}_r \qquad \text{(2SQ)}$$

Main question

Can (2SQ) be solved in pseudo-polynomial time? NO

Observations:

Pricing problem

After incorporating dual variables in costs, pricing problem amounts to solving a problem of the form:

$$\min_{r \in \mathcal{R}} \tilde{c}_r \tag{2SQ}$$

Main question

Can (2SQ) be solved in pseudo-polynomial time? NO

Observations:

- Hardness of pricing q-routes implies hardness of pricing ng-routes

Pricing problem

After incorporating dual variables in costs, pricing problem amounts to solving a problem of the form:

$$\min_{r \in \mathcal{R}} \tilde{c}_r \tag{2SQ}$$

Main question

Can (2SQ) be solved in pseudo-polynomial time? NO

Observations:

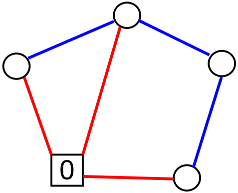
- Hardness of pricing q-routes implies hardness of pricing ng-routes
- Hardness of pricing with "simple recourse" indicates hardness of pricing with more complex recourse (though not necessarily implies it)

Outline

1. Introduction
2. Pricing complexity results
   - Scenarios
   - Independent normal
   - Revisiting scenarios
3. Conclusion

The scenarios case

Consider a graph $G = (V_+ \cup \{0\}, E')$.



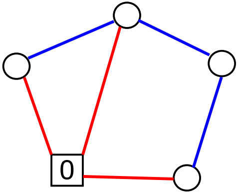
and scenarios $s \in \{1, \ldots, n = |V_+|\}$ with $p_s = \frac{1}{n}$:

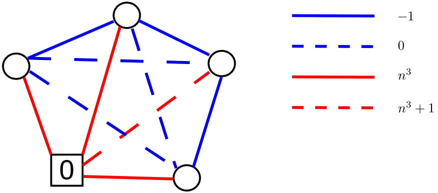
$$D_j^s = \begin{cases} n, & \text{if } j = s \\ 1, & \text{otherwise.} \end{cases}$$

with $C = 2n - 1$.

The scenarios case

Consider a graph $G = (V_+ \cup \{0\}, E')$. Construct G' as follows:





and scenarios $s \in \{1, \ldots, n = |V_+|\}$ with $p_s = \frac{1}{n}$:

$$D_j^s = \begin{cases} n, & \text{if } j = s \\ 1, & \text{otherwise.} \end{cases}$$

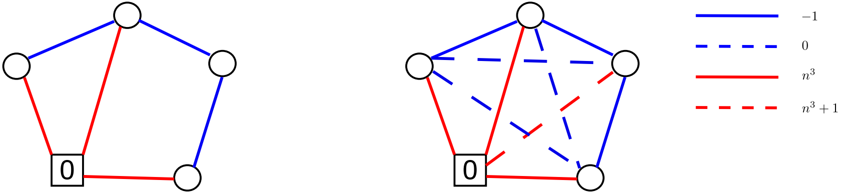
with $C = 2n - 1$.

Lemma (F. and Gunter '22)

Elementary routes r have $c''_r = 0$. Nonelementary routes r have $c''_r \geq n^2$.

The scenarios case

Consider a graph $G = (V_+ \cup \{0\}, E')$. Construct G' as follows:



and scenarios $s \in \{1, \dots, n = |V_+|\}$ with $p_s = \frac{1}{n}$:

$$D_j^s = \begin{cases} n, & \text{if } j = s \\ 1, & \text{otherwise.} \end{cases}$$

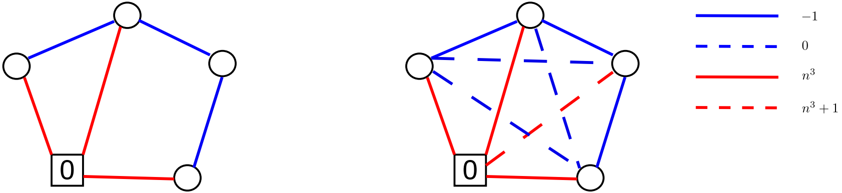
with $C = 2n - 1$.

Lemma (F. and Gunter '22)

A q-route is feasible if and only if it has length $\leq n$.

The scenarios case

Consider a graph $G = (V_+ \cup \{0\}, E')$. Construct G' as follows:



and scenarios $s \in \{1, \ldots, n = |V_+|\}$ with $p_s = \frac{1}{n}$:

$$D_j^s = \begin{cases} n, & \text{if } j = s \\ 1, & \text{otherwise.} \end{cases}$$

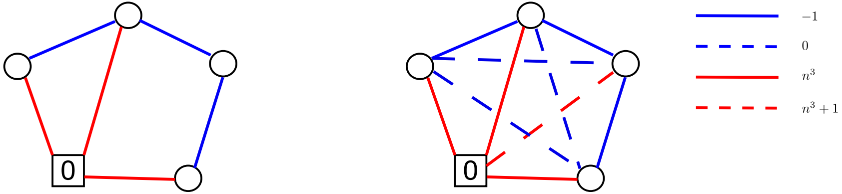
with $C = 2n - 1$.

Lemma (F. and Gunter '22)

A minimum 2-stage cost q-route must be elementary.

The scenarios case

Consider a graph $G = (V_+ \cup \{0\}, E')$. Construct G' as follows:



and scenarios $s \in \{1, \ldots, n = |V_+|\}$ with $p_s = \frac{1}{n}$:

$$D_j^s = \begin{cases} n, & \text{if } j = s \\ 1, & \text{otherwise.} \end{cases}$$

with $C = 2n - 1$.

Theorem (F. and Gunter '22)

G has a Hamiltonian-cycle if and only if the minimum 2-stage cost q-route is a Hamiltonian cycle.

Outline

Independent normal

Theorem (F. and Gunter '22)

Suppose one can solve (2SQ) under the following assumptions:

1. Demands are independent and identically distributed normal $N(\mu, \sigma^2)$
2. μ and σ^2 are constant integers which do not grow in n.
3. C polynomially bounded in n.
4. All elementary operations can be performed in $O(1)$ time.
5. $RF(\mu, \sigma^2)$ is computable in polynomial time.

Then there exists an algorithm using polynomially many operations that solves the Hamiltonian cycle problem with polynomially many calls to this algorithm.

Independent normal

Theorem (F. and Gunter '22)

Suppose one can solve (2SQ) under the following assumptions:

1. Demands are independent and identically distributed normal $N(\mu, \sigma^2)$
2. μ and σ^2 are constant integers which do not grow in n.
3. C polynomially bounded in n.
4. All elementary operations can be performed in $O(1)$ time.
5. $RF(\mu, \sigma^2)$ is computable in polynomial time.

Then there exists an algorithm using polynomially many operations that solves the Hamiltonian cycle problem with polynomially many calls to this algorithm.

Comments:

- Indicates strong NP-hardness but does not prove it
- Knapsack is polytime solvable with either small weights or small costs

P=NP?

Christiansen and Lysgaard (2007), Gauvin et al (2014):
Propose a pseudo-polynomial time algorithm for pricing for independent normal demands $N(\mu_i, \sigma_i^2)$, with μ_i, σ_i^2 integers.

P=NP?

Christiansen and Lysgaard (2007), Gauvin et al (2014):
Propose a pseudo-polynomial time algorithm for pricing for independent normal demands $N(\mu_i, \sigma_i^2)$, with μ_i, σ_i^2 integers.

Issue:

What is the probability distribution of the total demand of r? (Assume $\mu_i = \sigma_i^2 = 1$)

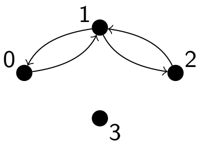


Figure: Route $r = (0, 1, 2, 1, 0)$.

P=NP?

Christiansen and Lysgaard (2007), Gauvin et al (2014):
Propose a pseudo-polynomial time algorithm for pricing for independent normal demands $N(\mu_i, \sigma_i^2)$, with μ_i, σ_i^2 integers.

Issue:

What is the probability distribution of the total demand of r? (Assume $\mu_i = \sigma_i^2 = 1$)

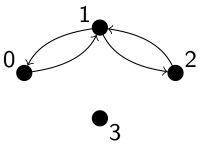


Figure: Route $r = (0, 1, 2, 1, 0)$.

- Our work (Customer independence):
  Calculates second stage cost $\tilde{c}$ based on $2D_1 + D_2 \sim N(3, 5)$

P=NP?

Christiansen and Lysgaard (2007), Gauvin et al (2014):
Propose a pseudo-polynomial time algorithm for pricing for independent normal demands $N(\mu_i, \sigma_i^2)$, with μ_i, σ_i^2 integers.

Issue:

What is the probability distribution of the total demand of r? (Assume $\mu_i = \sigma_i^2 = 1$)

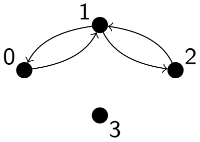


Figure: Route $r = (0, 1, 2, 1, 0)$.

- Our work (Customer independence):
  Calculates second stage cost $\tilde{c}$ based on $2D_1 + D_2 \sim N(3, 5)$
- Christiansen and Lysgaard (2007) (Route independence):
  Calculates a different cost $\hat{c}$, based on $N(3, 3)$ - Sum of three independent $N(1, 1)$ random variables:

P=NP?

Christiansen and Lysgaard (2007), Gauvin et al (2014):
Propose a pseudo-polynomial time algorithm for pricing for independent normal demands $N(\mu_i, \sigma_i^2)$, with μ_i, σ_i^2 integers.

Issue:

What is the probability distribution of the total demand of r? (Assume $\mu_i = \sigma_i^2 = 1$)

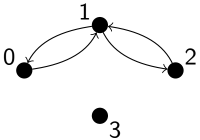


Figure: Route $r = (0, 1, 2, 1, 0)$.

- Our work (Customer independence):
  Calculates second stage cost $\tilde{c}$ based on $2D_1 + D_2 \sim N(3, 5)$
- Christiansen and Lysgaard (2007) (Route independence):
  Calculates a different cost $\hat{c}$, based on $N(3, 3)$ - Sum of three independent $N(1, 1)$ random variables:
- On elementary routes, both approaches calculate the same expected cost.

Outline

Can this idea help with scenarios?

What if we solve

$$\min_{r \in \mathcal{R}} \hat{c}_r \tag{2SQ'}$$

where $\hat{c}_r = \tilde{c}_r$ for all elementary r.

Can this idea help with scenarios?

What if we solve

$$\min_{r \in \mathcal{R}} \hat{c}_r \qquad \text{(2SQ')}$$

where $\hat{c}_r = \tilde{c}_r$ for all elementary r.

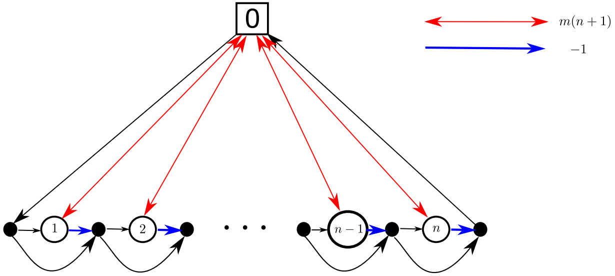
Theorem (F. and Ota '23)

(2SQ') is strongly NP-hard when demands are given as scenarios

Proof idea:

Given $G = (V, E)$ with n vertices, m edges, we wish to find maximum cardinality independent set.

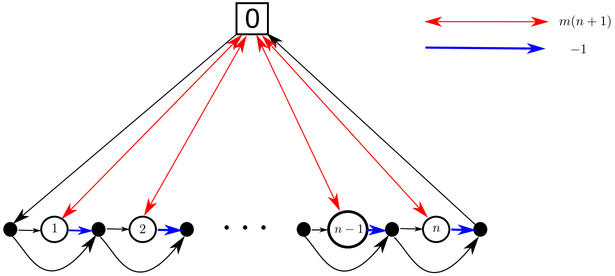
Solve (2SQ') on the following graph:



Observations:

Proof idea:

Given $G = (V, E)$ with n vertices, m edges, we wish to find maximum cardinality independent set.
Solve (2SQ') on the following graph:

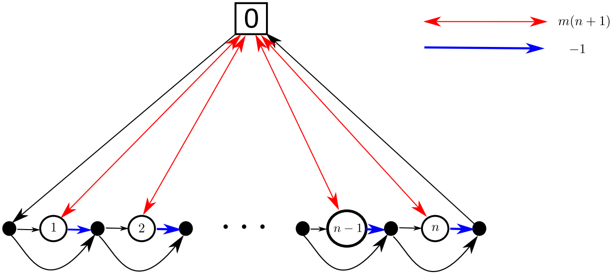


Observations:

- Graph above is independent of G

Proof idea:

Given $G = (V, E)$ with n vertices, m edges, we wish to find maximum cardinality independent set.
Solve (2SQ') on the following graph:



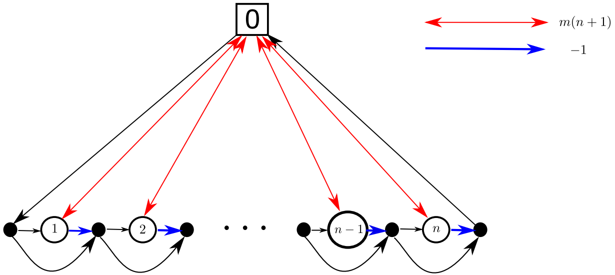
Observations:

- Graph above is independent of G
- All routes are elementary, thus (2SQ') and (2SQ) are the same

Proof idea:

$C = 2m - 1$, and create $m = |E|$ scenarios:

- Scenario for edge $e = ij$ is:
$$D^e(w) = \begin{cases} m, & \text{if } w = i \text{ or } j \\ 0, & \text{otherwise.} \end{cases}$$



- If a route goes through both i and j, then in scenario D^{ij}, capacity is exceeded, so total cost will be positive
- There exists a route of cost 0
- Negative cost route goes through a subset of vertices S that form an independent set in G

Conclusion

Main takeaway

Pricing with correlations is hard
(also pointed out in Gendreau, Jabali and Rey 2016)

Conclusion

Main takeaway

Pricing with correlations is hard
(also pointed out in Gendreau, Jabali and Rey 2016)

Future work:

- Can specific reduced cost structure help?
- Do we need to solve (2SQ)?
- Branch-and-cut and branch-and-price for the problem (and other variants)

THANK YOU!