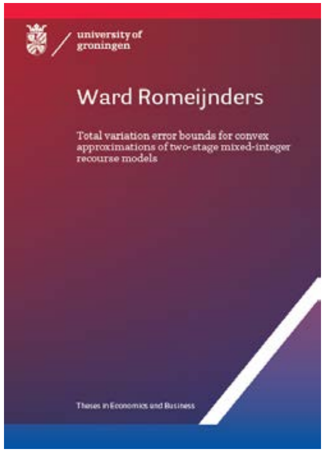
Gijs de Leve Prize 2015-2017

Total variation error bounds for convex approximations of two-stage mixed-integer recourse models

Ward Romeijnders



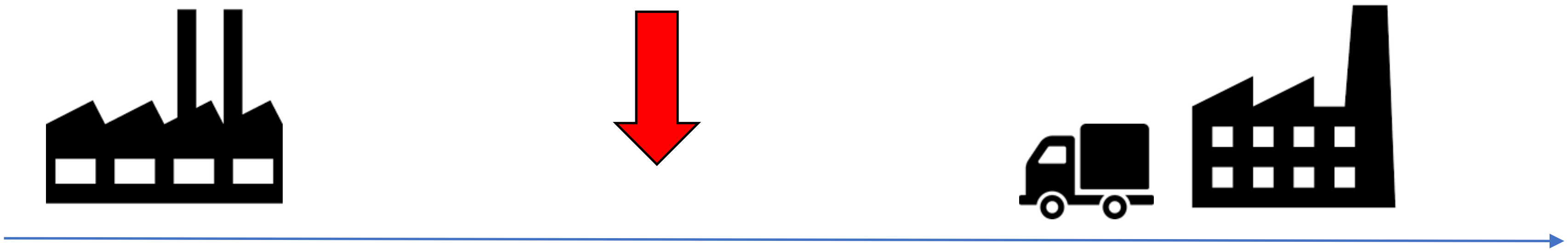
Supervisors:

Maarten H. van der Vlerk†
Willem K. Klein Haneveld

Production planning example

- Timeline:

Observe demand ω



Produce $x \in \mathbb{R}_+$

If $\omega > x$, then buy $y \in \mathbb{Z}_+$ from competitor in batches of size 1

- Minimize total expected costs:

$$\min_{x \geq 0} \; cx + q\,\mathbb{E}_\omega\left[\left\lceil \omega - x \right\rceil^+\right]$$

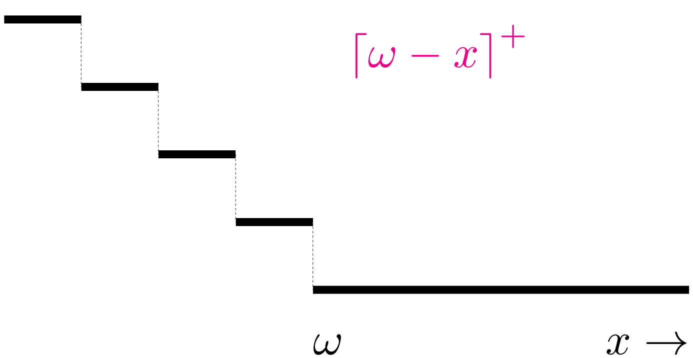
$y = \max\{\lceil \omega - x \rceil, 0\}$

production costs

expected future purchasing costs

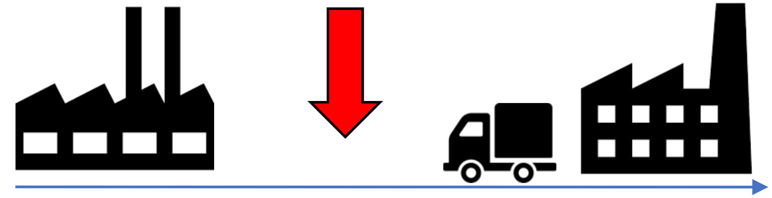
- For fixed demand ω:

$\lceil \omega - x \rceil^+$



Non-convex!

Non-convex objective function?

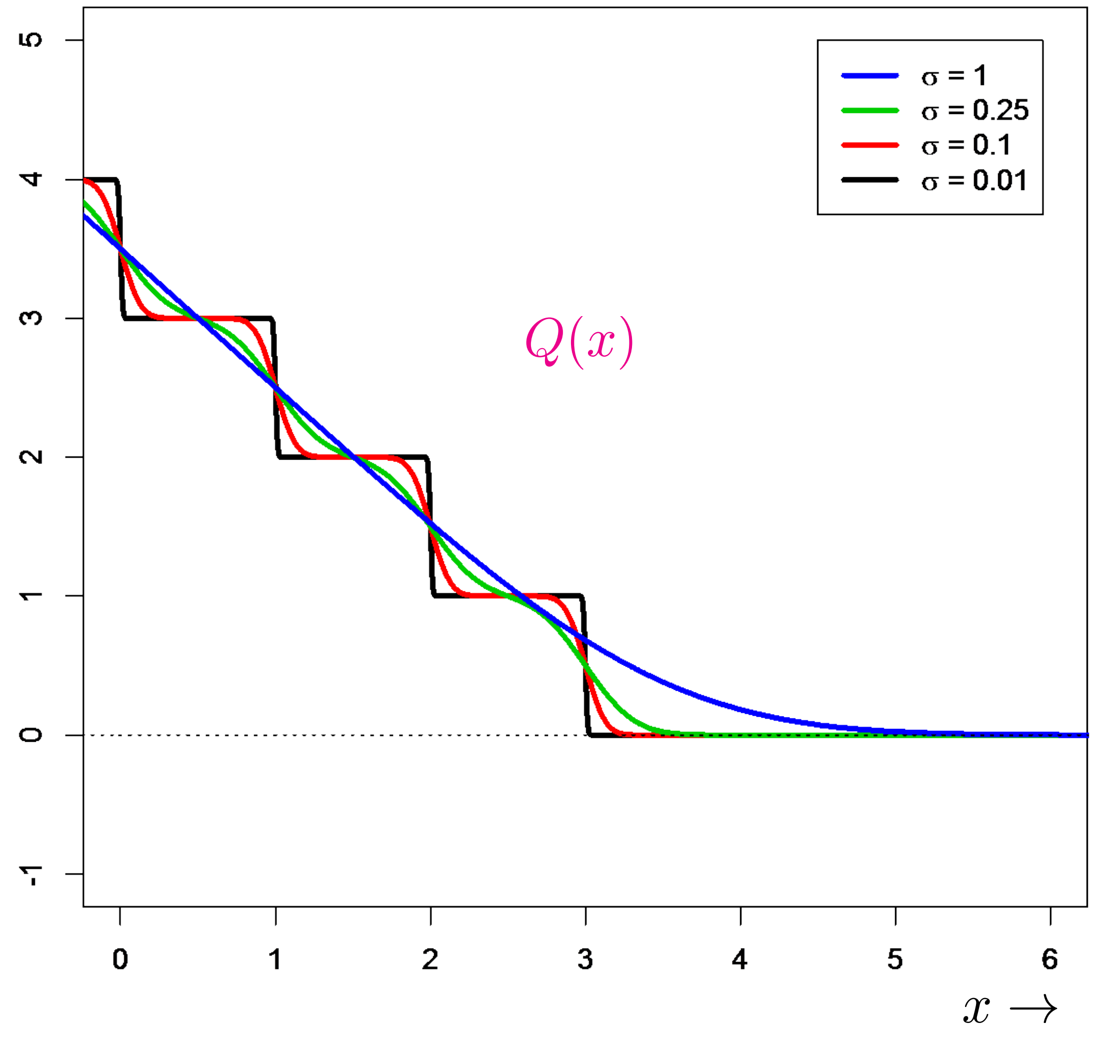


- Minimize total expected costs:

$$\min_{x \geq 0} \; cx + \underbrace{q\,\mathbb{E}_\omega\left[\lceil \omega - x \rceil^+\right]}_{Q(x)}$$

- Numerical example:

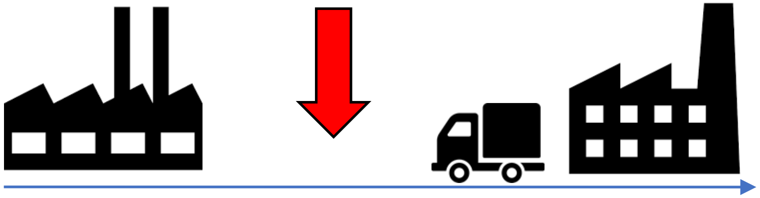
$$\omega \sim N(\mu, \sigma^2)$$
$$\mu = 3$$
$$q = 1$$



> Observation: Q "convexer"
> if σ increases

Who cares?!

Mixed-integer recourse models

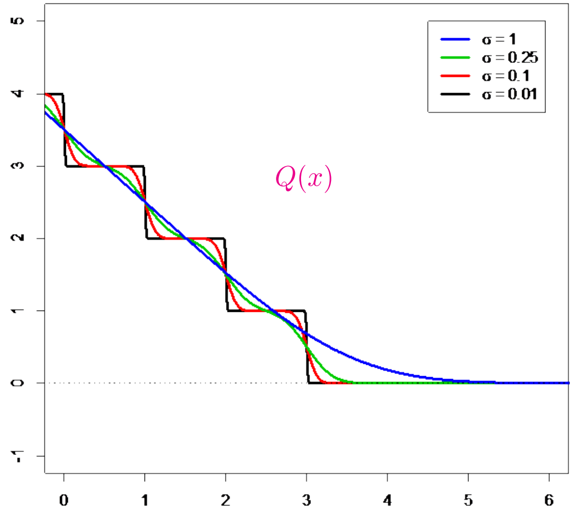


- Models from Stochastic Programming

  Characteristics:

  - random parameters (demand)
  - integer decision variables (batches)

⟶ production planning problem is special case



- Research questions:
  - Are MIR models approximately convex?
  - Can we use this to solve MIR models?

- Applications in:

Finance

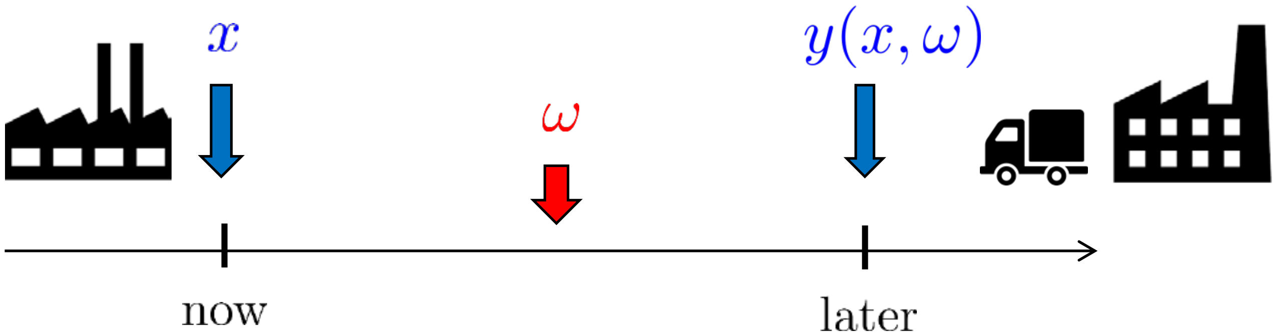
Energy

Logistics

Healthcare

Mathematical formulation

- Two-stage mixed-integer recourse models



- Optimization problem:

$$\min_{x \in X} \left\{ c^\top x + \underbrace{Q(x)}_{:= \mathbb{E}_\omega \left[\underbrace{v(\omega, x)}_{:= \min\limits_{y \in Y} \left\{ q^\top y : \; Wy \geq h(\omega) - T(\omega)x \right\}} \right]} : \; Ax = b \right\}$$

Expected Value Function (EVF)

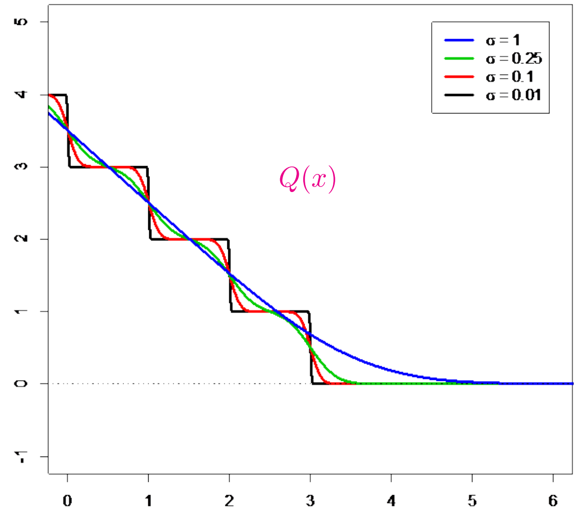
2^nd^-stage value function

↑ Y contains integrality restrictions

Computational challenges

- Difficulty of solving MIR:

  - Second-stage value function v is MIP $\longrightarrow$ NP-hard
  - Expected value function Q is non-convex



- Traditional approach:

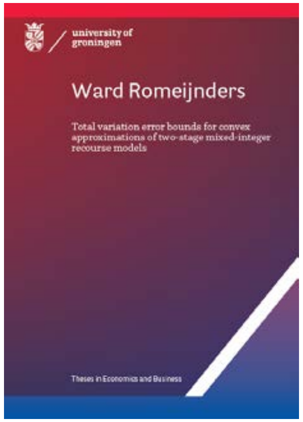
  - combine methods for
    - deterministic MIP
    - continuous SP

$\longrightarrow$ cannot solve large problem instances in general

- Alternative approach:

  - Construct convex approximation $\hat{Q}$ of Q $\longrightarrow$ can be solved efficiently
  - Derive error bounds $\longrightarrow$ performance guarantee

Main contributions PhD thesis



1. Fundamentally different approach for solving MIR

2. Convex approximations for MIR
   - simple integer recourse models
   - totally unimodular integer recourse models
   - general two-stage mixed-integer recourse models

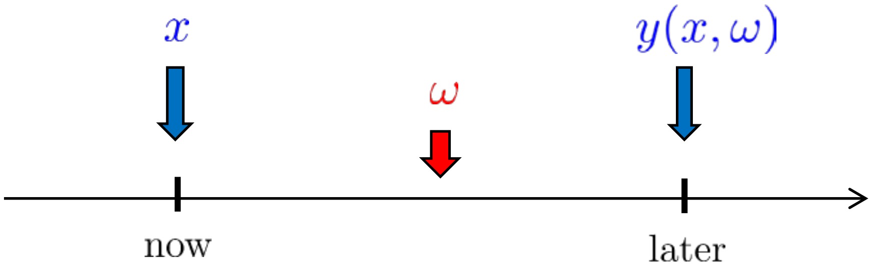
3. Error bounds for convex approximations
   - depend on total variations of pdfs of random parameters in the model
   - more variability $\longrightarrow$ tighter bounds
   - total variation bounds on the expectation of periodic functions

4. Numerical assessment of the quality of convex approximations
   - fleet allocation and routing problem
   - stochastic activity network investment

Simple Integer Recourse models (Van der Vlerk, 1995)

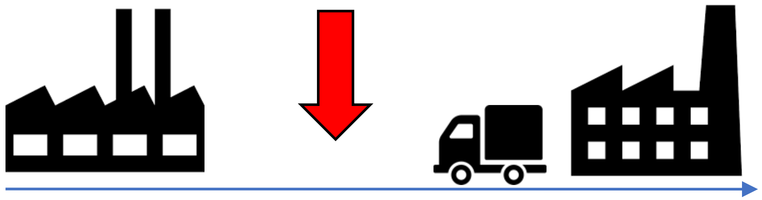
- Special case of general model

$$v(x,\omega) := \min_{y \in Y} \left\{ qy : \; Wy \geq h(\omega) - T(\omega)x \right\}$$



- Consider one-dimensional variant here $(x, \omega \in \mathbb{R})$
  - $W = I, T = I, h(\omega) = \omega$

- Production planning is SIR:



- Mathematical formulation:

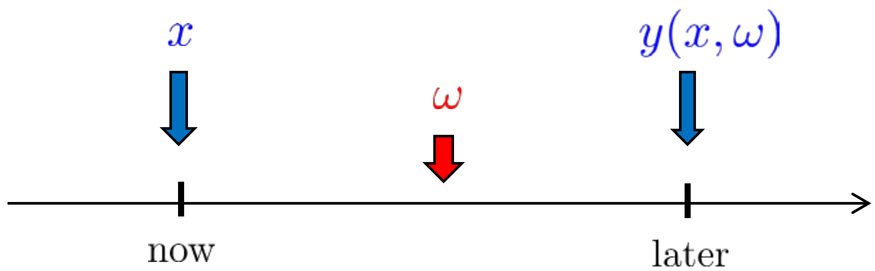
$$v(x,\omega) = \min_{y} \; qy : \quad y \geq \omega - x$$
$$y \in \mathbb{Z}_+$$

$$= q \lceil \omega - x \rceil^+$$

Simple Integer Recourse models (Van der Vlerk, 1995)

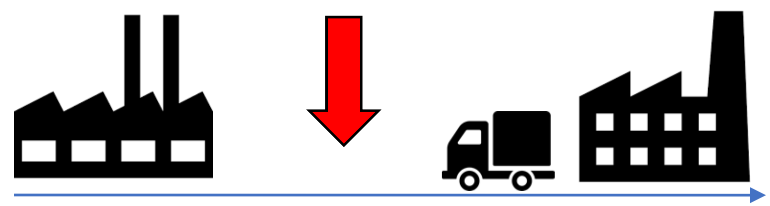
- Special case of general model

$$v(x,\omega) := \min_{y \in Y} \left\{ qy : \ Wy \geq h(\omega) - T(\omega)x \right\}$$



- Consider one-dimensional variant here ($x, \omega \in \mathbb{R}$)
  - $W = I, T = I, h(\omega) = \omega$

- Production planning is SIR:



- Mathematical formulation:

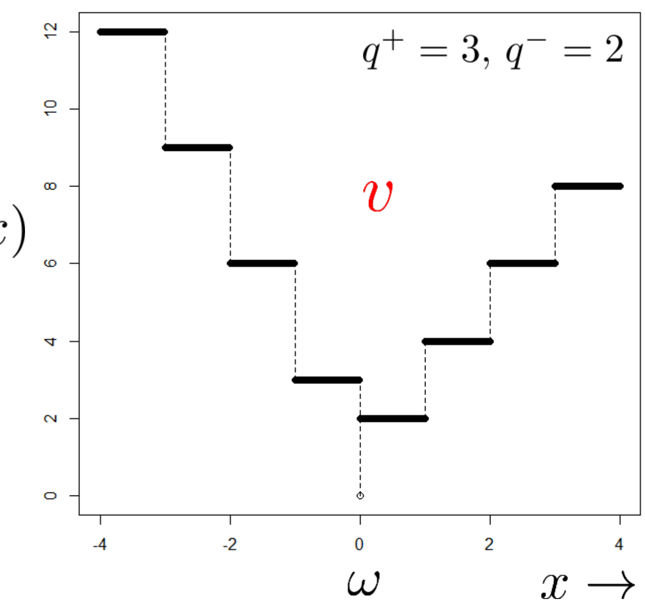
$$v(x,\omega) = \min_{y^+, y^-} \ q^+ y^+ + q^- y^- : \quad y^+ \geq \omega - x$$

$$y^- \geq -(\omega - x)$$

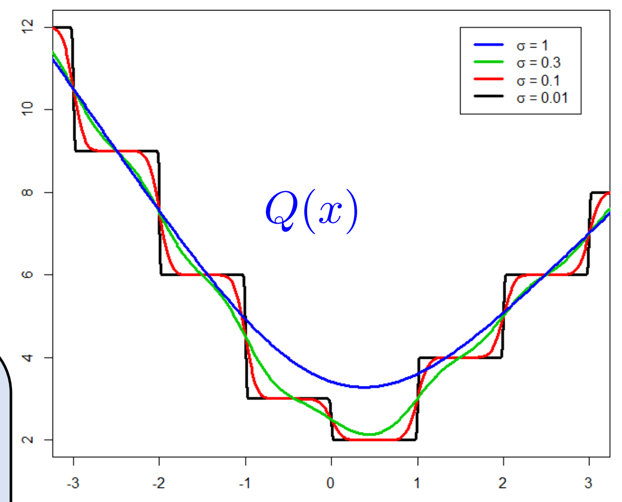
$$y^+, y^- \in \mathbb{Z}_+$$

$$= q^+ \lceil \omega - x \rceil^+ + q^- \lfloor \omega - x \rfloor^-$$

$$Q(x) = q^+ \mathbb{E}_\omega \Big[\lceil \omega - x \rceil^+ \Big] + q^- \mathbb{E}_\omega \Big[\lfloor \omega - x \rfloor^- \Big]$$



Convexity of SIR models



- $Q(x) = q^+ \mathbb{E}_\omega \left[\lceil \omega - x \rceil^+ \right] + q^- \mathbb{E}_\omega \left[\lfloor \omega - x \rfloor^- \right]$
- Notation: f is pdf of ω

Theorem Klein Haneveld et al. (2006):

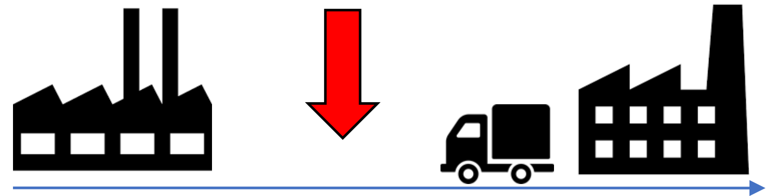
Q is convex $\Leftrightarrow$ $f(s) = G(s+1) - G(s)$ $(\star)$

for some cdf G with finite mean

- If f does not satisfy $(\star)$ $\longrightarrow$ convex approximations
  1. α-approximations: $\hat{Q}_\alpha$ $\qquad \hat{G}_\alpha(s) := F(\lfloor s \rfloor_\alpha)$
  2. Forward difference approximation: $\hat{Q}$ $\qquad \hat{G}(s) := F(s - 1/2)$
- Focus on FWD:
  - Intuition
  - Derivation error bound

Forward difference approximation

- Production planning example:
  - $v(x,\omega) = [\omega - x]^+$ (unit costs 1)
  - $Q(x) = \mathbb{E}_\omega\left[v(x,\omega)\right]$

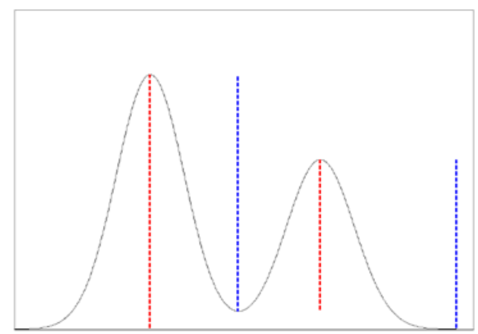


FWD approximation: $\hat{Q}$

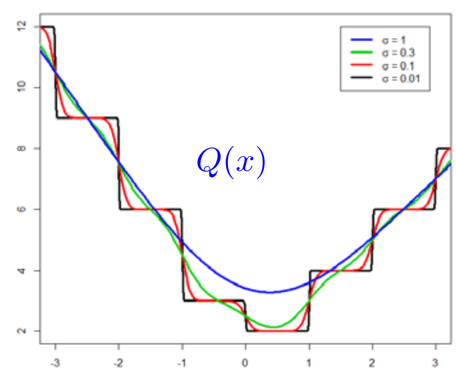
- $\hat{v}(x,\omega) := (\omega + 1/2 - x)^+$
- $\hat{Q}(x) = \mathbb{E}_\omega\left[\hat{v}(x,\omega)\right]$
- Good "on average"
- Error bound:

$$\|Q - \hat{Q}\|_\infty \leq \frac{1}{16}|\Delta|f$$

- Total variation $|\Delta|f$ of f:



$|\Delta|f =$ "Total increase" + "Total decrease"

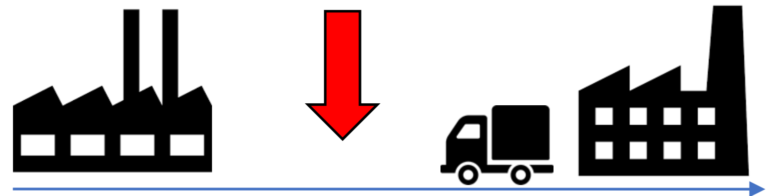


- Intuition error bound:
  - For normally distributed ω:

    Larger variance $\sigma^2 \longrightarrow$ better bound

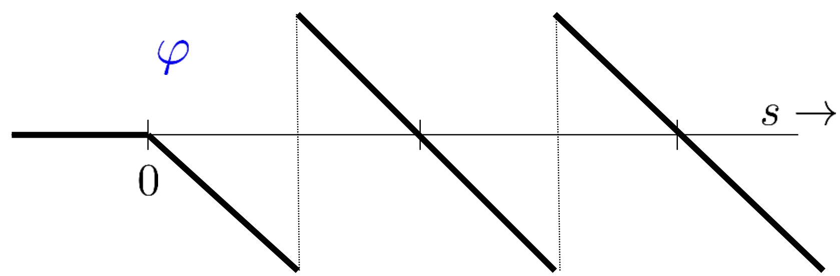
Derivation error bound

- Production planning example:



- Value functions:
  - $v(x,\omega) = \lceil \omega - x \rceil^+$
  - $\hat{v}(x,\omega) = (\omega + 1/2 - x)^+$

- Underlying difference function φ
  - $Q(x) - \hat{Q}(x) = \mathbb{E}_\omega\Big[\lceil \omega - x \rceil^+ - (\omega + 1/2 - x)^+\Big] = \mathbb{E}_\omega\Big[\varphi(\omega + 1/2 - x)\Big]$
  - $\varphi(s) = \lceil s - 1/2 \rceil^+ - (s)^+$
  - φ is half-periodic
    - with period $p = 1$
    - with mean $\nu = 0$



- Expected difference

Expectation of periodic function

  - $Q(x) - \hat{Q}(x) = \int_0^\infty \varphi(s) f(s) ds$

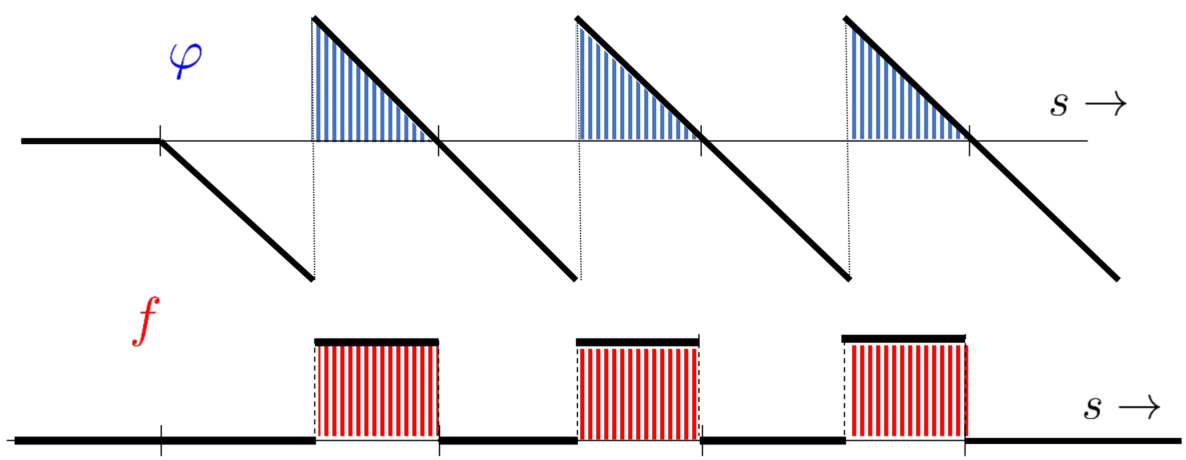
pdf of $\omega + 1/2 - x$ for fixed x

Intuition total variation error bound

- Expected difference for fixed x
  - $Q(x) - \hat{Q}(x) = \int_0^\infty \varphi(s) f(s) ds$

pdf of $\omega + 1/2 - x$ for fixed x

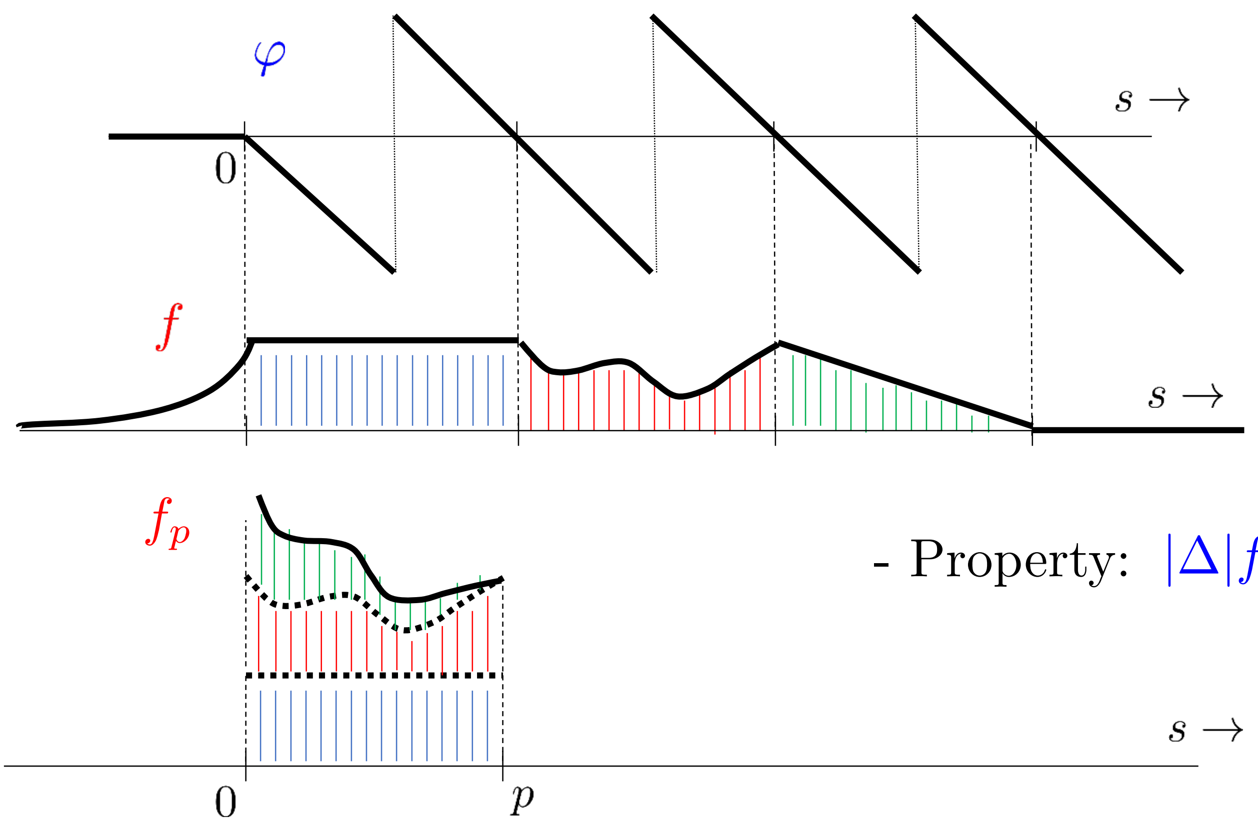
- Worst-case density with $|\Delta| f \leq B$:
  - $\sup_f \left\{ \int_0^\infty \varphi(s) f(s) ds : \; |\Delta| f \leq B \right\}$



Packed densities (for general periodic functions with period p)

- Definition: $f_p(x) := \sum_{k=0}^{\infty} f(x+pk), \qquad x \in [0,p]$

- Property: $Q(x) - \hat{Q}(x) = \int_0^{\infty} \varphi(s) f(s) ds = \int_0^{p} \varphi(s) f_p(s) ds$



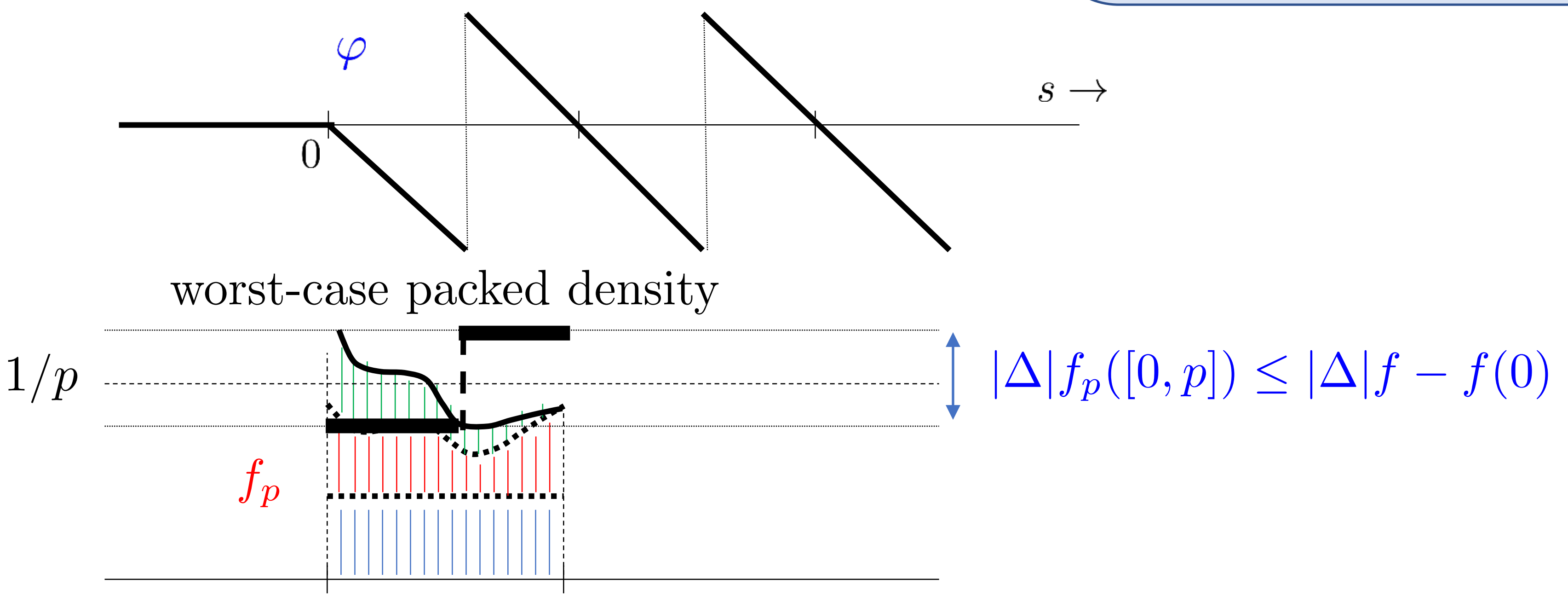
- Property: $|\Delta| f_p([0,p]) \leq |\Delta| f - f(0)$

Total variation error bound

- Expected difference:

$$Q(x) - \hat{Q}(x) = \int_0^p \varphi(s) f_p(s) ds$$

- Bound on packed density:



- Total variation error bound:

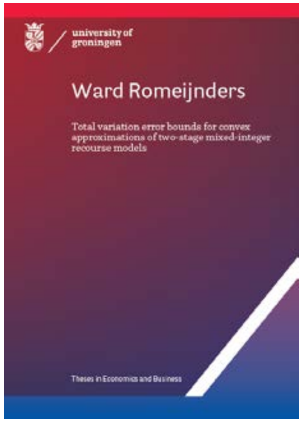
$$\|Q - \hat{Q}\|_\infty \leq \frac{1}{2} \int_0^p |\varphi(s)| ds |\Delta| f = \frac{1}{16} |\Delta| f$$

FWD approximation: $\hat{Q}$

- $\hat{v}(x, \omega) := (\omega + 1/2 - x)^+$
- $\hat{Q}(x) = \mathbb{E}_\omega \left[\hat{v}(x, \omega) \right]$
- Good "on average"
- Error bound:

$$\|Q - \hat{Q}\|_\infty \leq \frac{1}{16} |\Delta| f$$

Main contributions PhD thesis



1. Fundamentally different approach for solving MIR

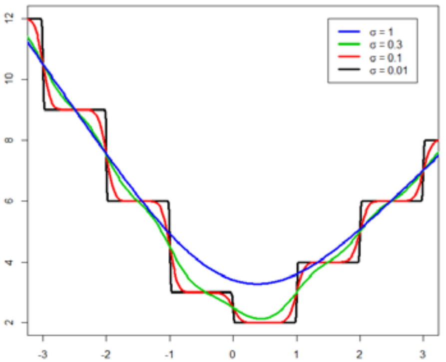
2. Convex approximations for MIR
   - simple integer recourse models
   - totally unimodular integer recourse models
   - general two-stage mixed-integer recourse models

3. Error bounds for convex approximations
   - depend on total variations of pdfs of random parameters in the model
   - more variability $\longrightarrow$ tighter bounds
   - total variation bounds on the expectation of periodic functions

4. Numerical assessment of the quality of convex approximations
   - fleet allocation and routing problem
   - stochastic activity network investment

General two-stage MIR

- Goal
  - Construct convex approximation $\hat{Q}$
  - Derive error bound: $\|Q - \hat{Q}\|_\infty \leq G(|\Delta| f_i, \dots)$
  - $G(|\Delta| f_i, \dots) \to 0$ as $|\Delta| f_i \to 0$



- Approach
  - Exploit periodicity in mixed-integer value function v
  - Construct convex approximation $\hat{v}$ such that $v - \hat{v}$ is 'periodic'
  - Use total variation bounds on expectation of periodic functions

Periodicity in mixed-integer linear programming problems

- Alternative formulation of $v(s)$

$$
\begin{aligned}
v(s) &= \min_{y} \left\{ q^\top y : \, Wy = s, \, y \in \mathbb{Z}_+^{n_2} \times \mathbb{R}_+^{n_3} \right\} \\
&= \min_{y_B, y_N} \left\{ q_B^\top y_B + q_N^\top y_N : \, By_B + Ny_N = s, \quad y_B \in \mathbb{Z}_+^{n_B} \times \mathbb{R}_+^{m-n_B}, \right. \\
&\qquad \left. y_N \in \mathbb{Z}_+^{n_N} \times \mathbb{R}_+^{n-n_N} \right\}
\end{aligned}
$$

- Substituting $y_B = B^{-1}s - B^{-1}Ny_N$:

$$
\begin{aligned}
v(s) = q_B^\top B^{-1} s + \min_{y_N} \quad & \bar{q}_N^\top y_N \\
\text{s.t.} \quad & B^{-1}s - B^{-1}Ny_N \in \mathbb{Z}_+^{n_B} \times \mathbb{R}_+^{m-n_B} \\
& y_N \in \mathbb{Z}_+^{n_N} \times \mathbb{R}_+^{n-n_N}
\end{aligned}
$$

- with reduced costs $\bar{q}_N^\top := q_B^\top - q_N^\top B^{-1} N \geq 0$

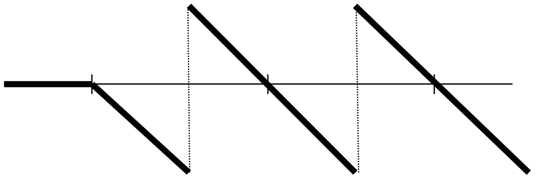
Gomory relaxation v_B

- Relax non-negativity constraints on y_B:

$$v_B(s) = q_B^\top B^{-1} s + \min_{y_N} \; \bar{q}_N^\top y_N$$
$$\text{s.t.} \quad B^{-1}s - B^{-1}N y_N \in \mathbb{Z}^{n_B} \times \mathbb{R}^{m-n_B}$$
$$y_N \in \mathbb{Z}_+^{n_N} \times \mathbb{R}_+^{n-n_N}$$

- Properties of optimal solutions $y_N^*(s)$
  - B-periodicity
  - Boundedness
  - Optimal for $v(s)$ if $B^{-1}s - B^{-1}N y_N^*(s) \geq 0$

B-periodicity of $y_N^*(s)$



- Gomory relaxation:

$$
\begin{aligned}
v_B(s) = q_B^\top B^{-1} s + \min_{y_N} \quad & \bar{q}_N^\top y_N \\
\text{s.t.} \quad & B^{-1} s - B^{-1} N y_N && \in \mathbb{Z}^{n_B} \times \mathbb{R}^{m-n_B} \\
& y_N && \in \mathbb{Z}_+^{n_N} \times \mathbb{R}_+^{n-n_N}
\end{aligned}
$$

- Definition
  - A function $\psi : \mathbb{R}^m \mapsto \mathbb{R}^m$ is called B-periodic if and only if $\psi(s + Bl) = \psi(s)$ for all $s \in \mathbb{R}^m$ and $l \in \mathbb{Z}^m$.

- B-periodicity of $y_N^*(\cdot)$
  - Let $s' = s + Bl$ with $l \in \mathbb{Z}^m$ (so that $B^{-1}s' = B^{-1}s + l$)
  - Then, fractional values of $B^{-1}s'$ and $B^{-1}s$ are the same
  - Thus, $y_N^*(s + Bl) = y_N^*(s)$

Boundedness of $y_N^*(s)$

- Gomory relaxation:

$$v_B(s) = q_B^\top B^{-1} s + \min_{y_N} \ \bar{q}_N^\top y_N$$
$$\text{s.t.} \quad B^{-1}s - B^{-1}N y_N \in \mathbb{Z}^{n_B} \times \mathbb{R}^{m-n_B}$$
$$y_N \in \mathbb{Z}_+^{n_N} \times \mathbb{R}_+^{n-n_N}$$

- Observations
  - $\bar{q}_N \geq 0$
  - If W is integer, then for $z = \det(B) l$ with $l \in \mathbb{Z}^m$,

$$B^{-1}Nz = \left[\det(B)^{-1}\text{adj}(B)\right] N \left[\det(B) l\right] = \text{adj}(B)Nl \in \mathbb{Z}^m$$

- Consequence
  - $y_N^*(s) \leq |\det(B)| e_n$,
  - where e_n is the all-one vector

Relation between v and v_B

- Optimality condition

  - If $B^{-1}s - B^{-1}Ny_N^*(s) \geq 0$, then $y_N^*(s)$ is optimal for $v(s)$

  - Equivalent to $s - Ny_N^*(s) \in \Lambda$ with $\Lambda := \{t \in \mathbb{R}^m : B^{-1}t \geq 0\}$

- Deriving a sufficient condition

  - use boundedness of $y_N^*(s)$

  - implies $\|Ny_N^*(s)\|_2 \leq |\det(B)| \sum_{j=1}^{n} \|N_j\|_2 =: d$

- Sufficient condition

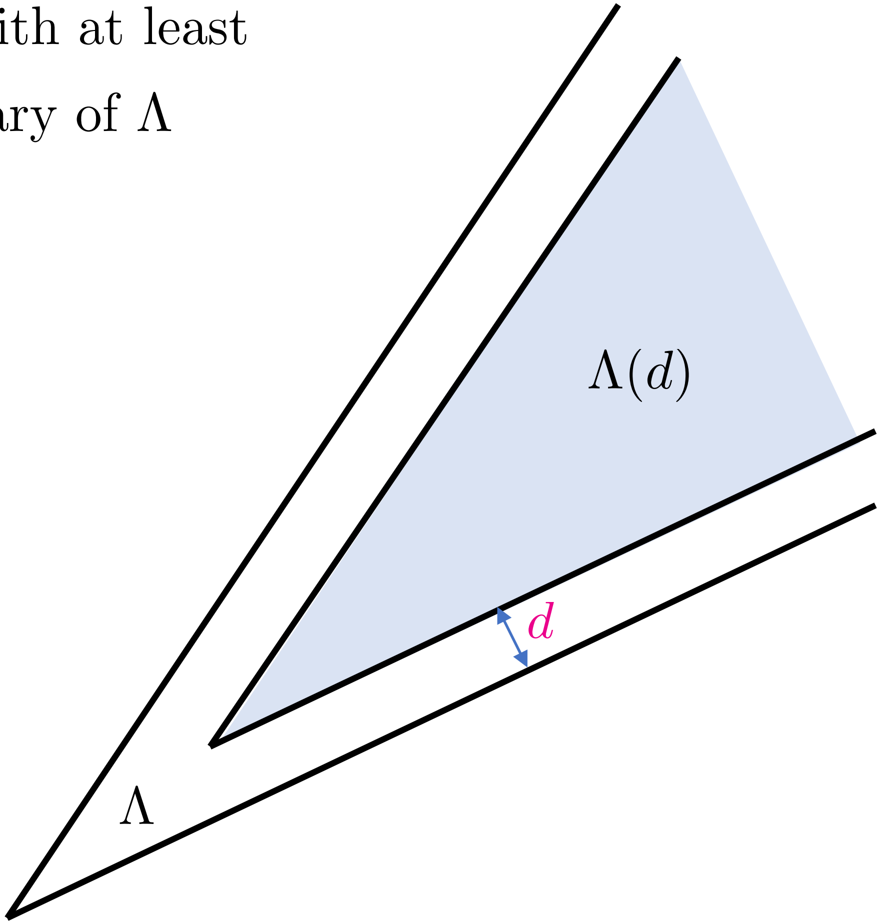
  - If s has at least distance d from the boundary of Λ, then $y_N^*(s)$ is optimal for $v(s)$

Illustration

- Definition
  - Let $\Lambda(d)$ denote all points $s \in \Lambda$ with at least Euclidean distance d to the boundary of Λ

- Lemma
  - For every $s \in \Lambda(d)$,

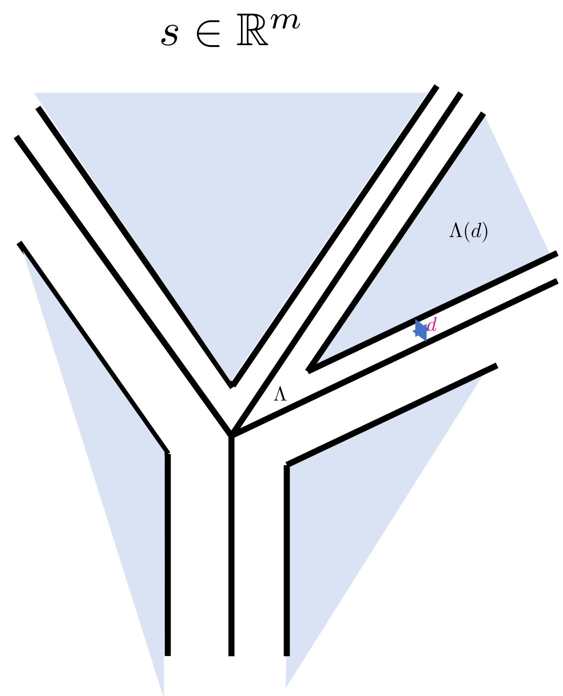
$$v(s) = \underbrace{q_B^\top B^{-1} s}_{\text{linear}} + \underbrace{\bar{q}_N^\top y_N^*(s)}_{\text{periodic}}$$



Periodicity in mixed-integer linear programming

- Mixed-integer value function

$$v(s) = \min\{q^\top y : \; Wy = s, \; y \in \mathbb{Z}_+^{n_2} \times \mathbb{R}_+^{n_3}\}$$



- Theorem
  - There exist
    - · dual feasible basis matrices B^k
    - · simplicial cones Λ^k
    - · distances d_k
    - · B^k-periodic functions ψ^k
  - such that for every $s \in \Lambda^k(d_k)$,

$$v(s) = \underbrace{q_{B^k}^\top (B^k)^{-1} s}_{\text{linear}} + \underbrace{\psi^k(s)}_{\text{periodic}}$$

Convex approximation

- Definition Γ_k
  - $\Gamma_k := (p_k)^{-m} \int_0^{p_k} \cdots \int_0^{p_k} \psi^k(s) ds_1 \dots ds_m$
  - $p_k = |\det(B^k)|$

- Idea of convex approximation

  - $v(s) = \underbrace{q_{B^k}^\top (B^k)^{-1} s}_{\text{linear}} + \underbrace{\psi^k(s)}_{\text{periodic}}$

  Average of ψ^k

  - Replace periodic function ψ^k by a constant Γ_k

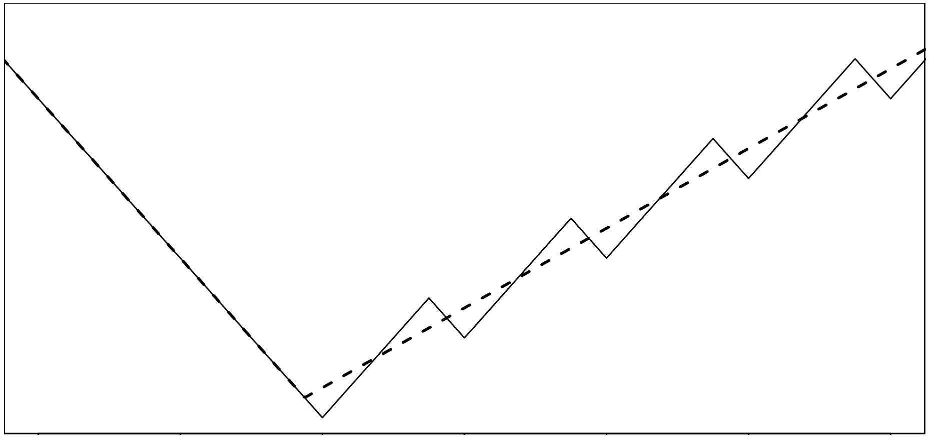
- Definition convex approximation $\hat{v}$

  - $\hat{v}(s) = \max_{k=1,\dots,K} \left\{ q_{B^k}^\top (B^k)^{-1} s + \Gamma_k \right\}$

  - SIR: $\hat{v}(s) \longrightarrow$ FWD

- Example:

$$v(s) = \min \ y_1 + 2y_2 + 2y_3$$
$$\text{s.t.} \ \ y_1 + y_2 - y_3 = s$$
$$y_1 \in \mathbb{Z}_+, \ y_2, y_3 \in \mathbb{R}_+$$

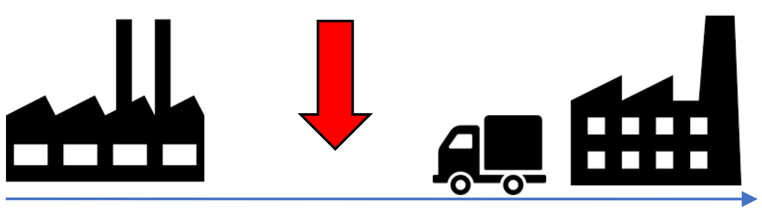


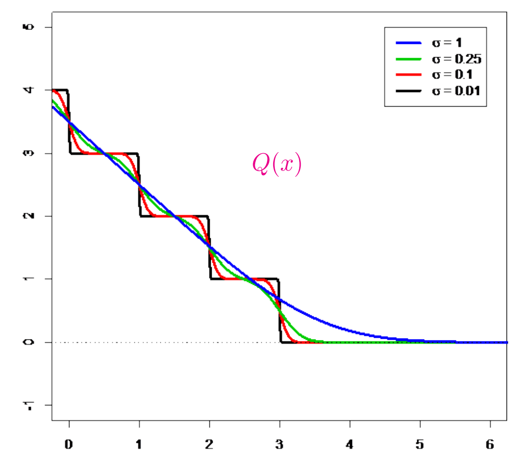
Asymptotic error bound

- Theorem for ω with independently distributed components
  - There exists a constant $C > 0$ such that for every continuous random vector ω:

$$\|Q - \hat{Q}\|_\infty \leq C \sum_{i=1}^{m} |\Delta| f_i$$

- Interpretation
  - Error converges to zero if total variations converge to zero
  - Holds for all two-stage mixed-integer recourse models





Future research

- Convex approximations for
  - risk-averse MIR
  - distributionally robust MIR
  - Multistage MIR
- Improved error bounds
  - using dual norm framework
  - using moment conditions
  - for special cases or particular applications
- Computational improvements
  - generalizing α-approximations
- ...