Peeling for L0-Regularized Least-Squares

Théo Guyard[⋆], Gilles Monnoyer[†], Cédric Herzet[⋆] and Clément Elvira[‡]
[⋆]SIMSMART, Inria, France
[†]ICTEAM, UCLouvain, Belgium
[‡]IETR, CentraleSupelec, France

GRETSI - Colloque Francophone de Traitement du Signal et des Images
Grenoble, France
2023

L0-Regularized Least-Squares

Sparse linear model

Linear approximation $\longleftrightarrow$ **Sparse solution**

Sparse linear model

Machine Learning
Network design
Signal processing
...

Sparse linear model

Linear approximation $\longleftrightarrow$ **Sparse linear model** $\longleftrightarrow$ **Sparse solution**

Machine Learning
Network design
Signal processing
...

ℓ_0-regularized least-squares

$$\min_x \quad \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$$

Sparse linear model

Linear approximation $\longleftrightarrow$ **Sparse solution**

Sparse linear model

Machine Learning
Network design
Signal processing
...

ℓ_0-regularized least-squares

$$\min_x \quad \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$$
$$\text{s.t.} \quad L \leq x \leq U$$

Bounds needed for numerical tractability

Sparse linear model

Linear approximation ⟷ **Sparse solution**

Sparse linear model

Machine Learning
Network design
Signal processing
...

ℓ_0-regularized least-squares

$$\min_x \quad \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_0$$
$$\text{s.t.} \quad L \leq x \leq U$$

Bounds needed for numerical tractability

NP hard

Specialized Branch-and-Bound solvers

Branch-and-Bound

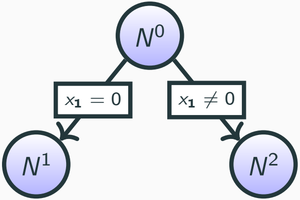
Tree search

$\min_{\mathbf{x} \in [\mathbf{L}, \mathbf{U}]} \ \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0$

N^0

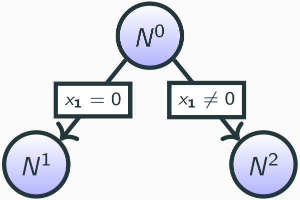
Tree search

$\min_{\mathbf{x} \in [\mathbf{L}, \mathbf{U}]} \quad \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_0$



Tree search

$\min_{\mathbf{x}\in[\mathbf{L},\mathbf{U}]} \quad \frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$

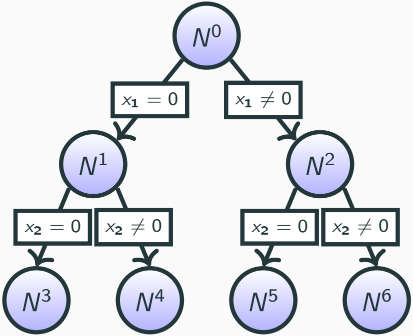


Same problem + constraint $x_1 = 0$

Same problem + constraint $x_1 \neq 0$

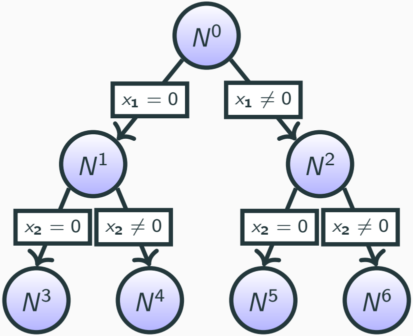
Tree search

$\min_{\mathbf{x}\in[\mathbf{L},\mathbf{U}]} \quad \frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$



Tree search

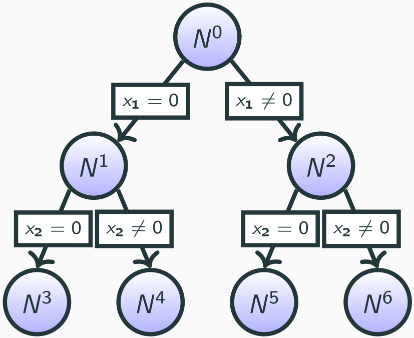
$\min_{\mathbf{x} \in [\mathbf{L}, \mathbf{U}]} \; \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$

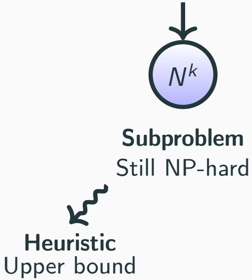


Subproblem
Still NP-hard

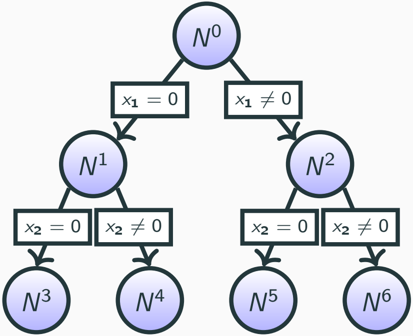
Tree search

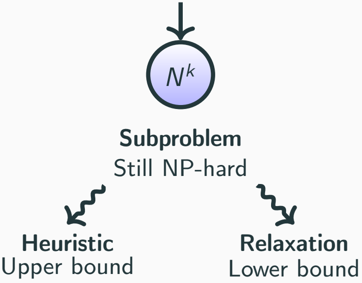
$\min_{\mathbf{x}\in[\mathbf{L},\mathbf{U}]} \ \frac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$



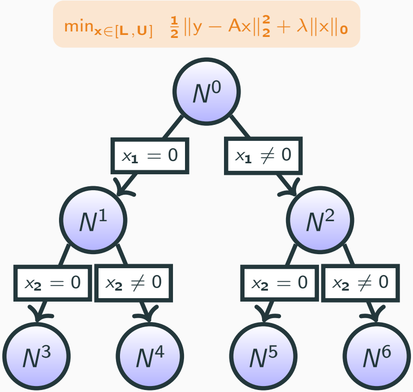


$$\min_{\mathbf{x}\in[\mathbf{L},\mathbf{U}]} \ \tfrac{1}{2}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$$

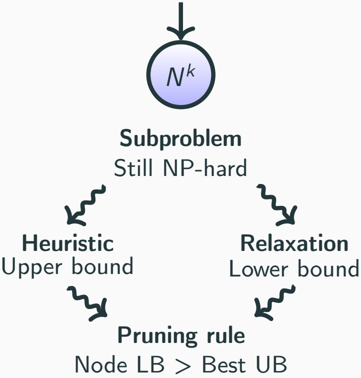




Tree search

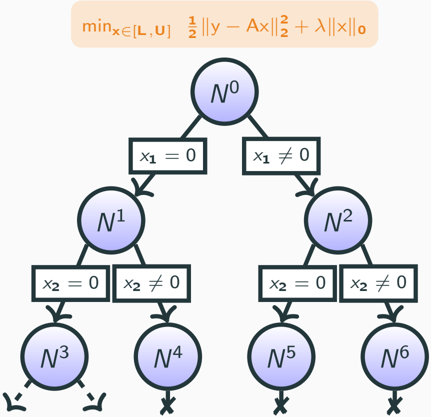


$\min_{\mathbf{x} \in [\mathbf{L}, \mathbf{U}]} \ \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$
N^0
$x_1 = 0$
$x_1 \neq 0$
N^1
N^2
$x_2 = 0$
$x_2 \neq 0$
$x_2 = 0$
$x_2 \neq 0$
N^3
N^4
N^5
N^6

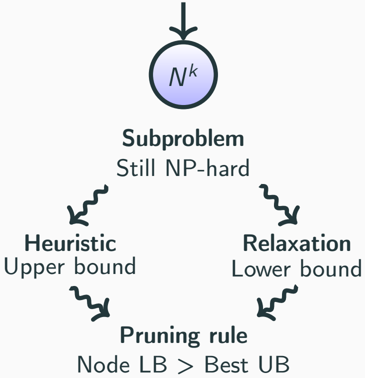


N^k
Subproblem
Still NP-hard
Heuristic
Upper bound
Relaxation
Lower bound
Pruning rule
Node LB > Best UB

Tree search



$\min_{\mathbf{x} \in [\mathbf{L},\mathbf{U}]} \ \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_0$
N^0
$x_1 = 0$
$x_1 \neq 0$
N^1
N^2
$x_2 = 0$
$x_2 \neq 0$
$x_2 = 0$
$x_2 \neq 0$
N^3
N^4
N^5
N^6



N^k
Subproblem
Still NP-hard
Heuristic
Upper bound
Relaxation
Lower bound
Pruning rule
Node LB > Best UB

Relaxation construction

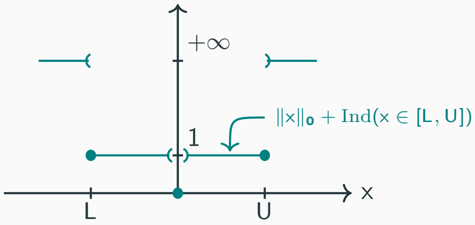
Node problem

$\min_{\mathbf{x} \in [\mathbf{L}, \mathbf{U}]} \quad \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$

Relaxation construction

Node problem

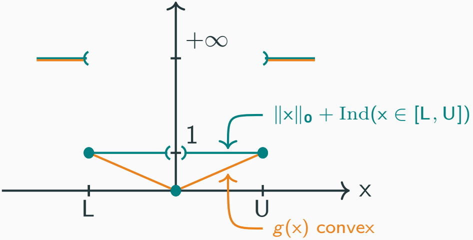
$\min_{\mathbf{x}\in[\mathbf{L},\mathbf{U}]} \; \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$



Relaxation construction

Node problem

$\min_{x \in [L,U]} \ \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$



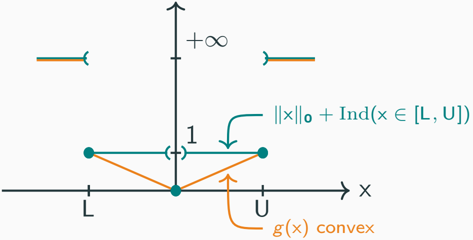
Relaxation construction

Node problem

$\min_{x \in [L,U]} \;\; \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$

Relaxation

$\min_x \frac{1}{2}\|y - Ax\|_2^2 + \lambda g(x)$



Relaxation construction

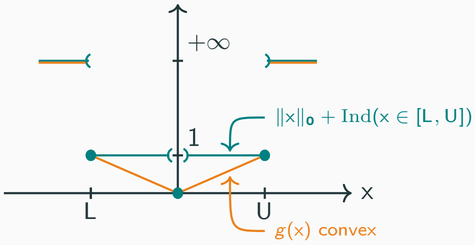
Node problem

$\min_{x \in [L,U]} \ \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$

↓

Relaxation

$\min_x \frac{1}{2}\|y - Ax\|_2^2 + \lambda g(x)$



Dilemma in the choice of $[L, U]$

- Wide interval $[L, U]$ to recover the right solution
- Tight interval $[L, U]$ to build strong relaxations

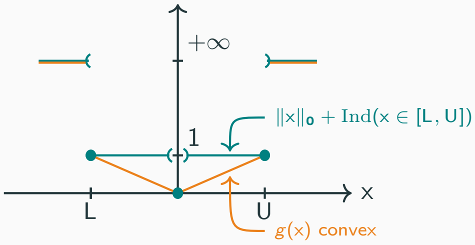
Relaxation construction

Node problem

$\min_{x \in [L, U]} \ \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_0$

Relaxation

$\min_x \frac{1}{2}\|y - Ax\|_2^2 + \lambda g(x)$



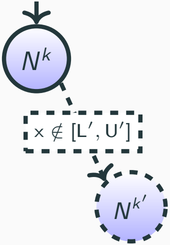
Dilemma in the choice of $[L, U]$

- Wide interval $[L, U]$ to recover the right solution
- Tight interval $[L, U]$ to build strong relaxations

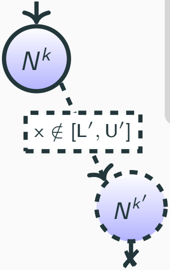
Our solution : Peeling

Peeling

Bound tightening

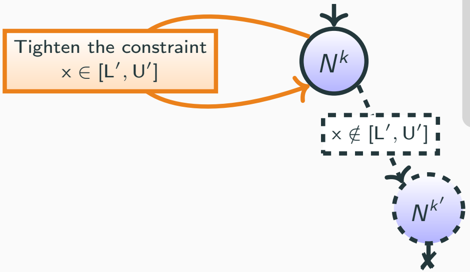


Bound tightening



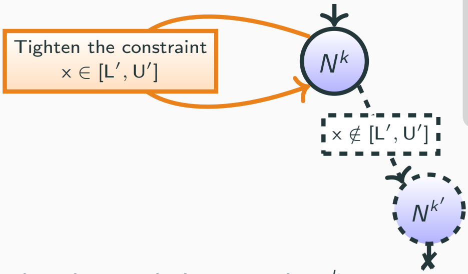
Find some $[L', U'] \subseteq [L, U]$ such that $N^{k'}$ is pruned

Bound tightening



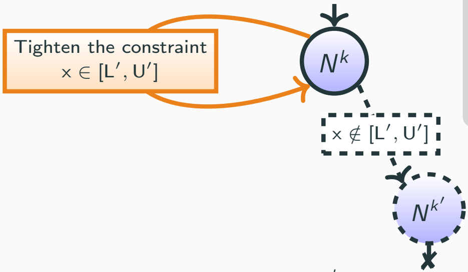
Find some
$[L', U'] \subseteq [L, U]$
such that $N^{k'}$ is pruned

Bound tightening



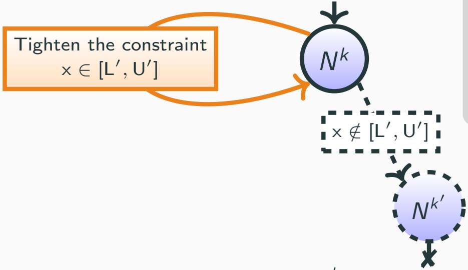
Find some
$[L', U'] \subseteq [L, U]$
such that $N^{k'}$ is pruned

Bound tightening

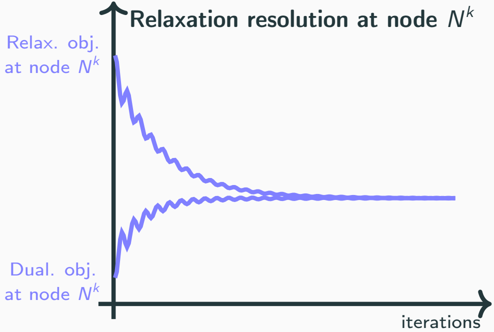


Find some
$[L', U'] \subseteq [L, U]$
such that $N^{k'}$ is pruned

Bound tightening

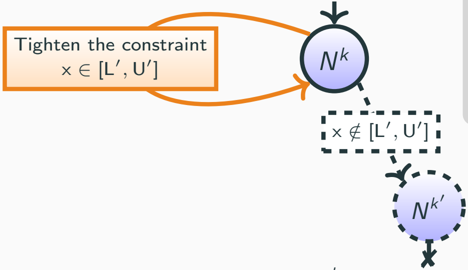


Find some
$[L', U'] \subseteq [L, U]$
such that $N^{k'}$ is pruned



Ingredients

- Already-computed quantities
- Cost-free dual evaluation

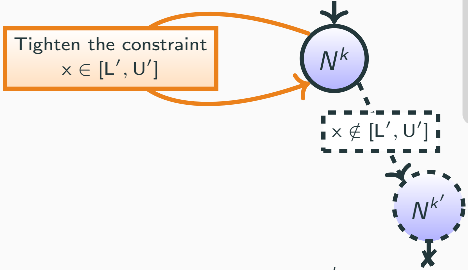


Find some
[L′, U′] ⊆ [L, U]
such that $N^{k'}$ is pruned

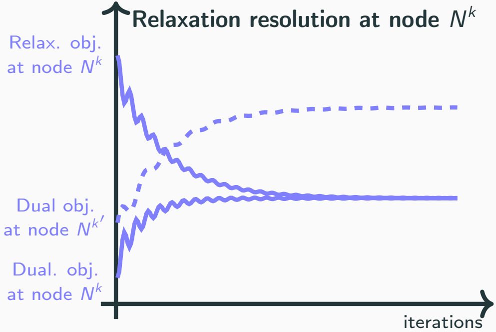
Ingredients

- Already-computed quantities
- Cost-free dual evaluation
- Dual link between nodes

Bound tightening



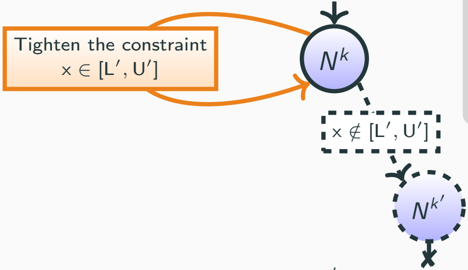
Find some
$[L', U'] \subseteq [L, U]$
such that $N^{k'}$ is pruned



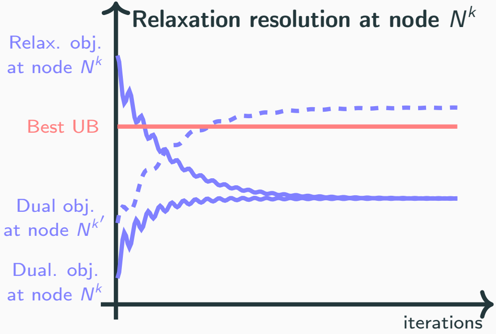
Ingredients

- Already-computed quantities
- Cost-free dual evaluation
- Dual link between nodes

Bound tightening



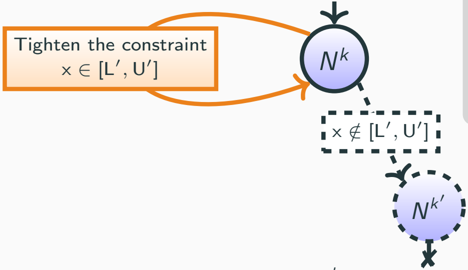
Find some $[L', U'] \subseteq [L, U]$ such that $N^{k'}$ is pruned



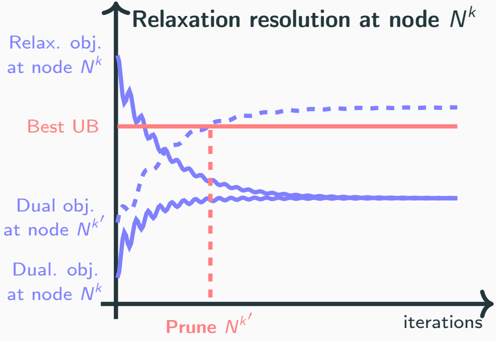
Ingredients

- Already-computed quantities
- Cost-free dual evaluation
- Dual link between nodes

Bound tightening



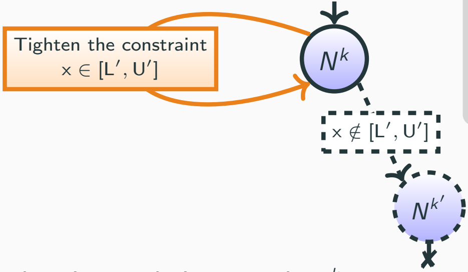
Find some $[L', U'] \subseteq [L, U]$ such that $N^{k'}$ is pruned



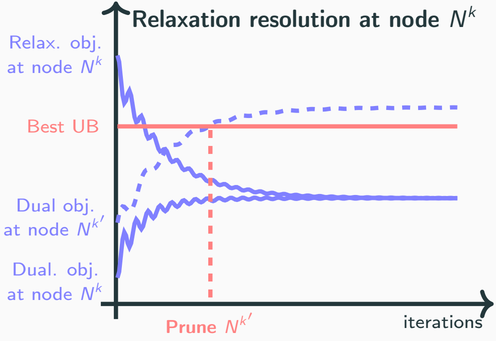
Ingredients

- Already-computed quantities
- Cost-free dual evaluation
- Dual link between nodes

Bound tightening



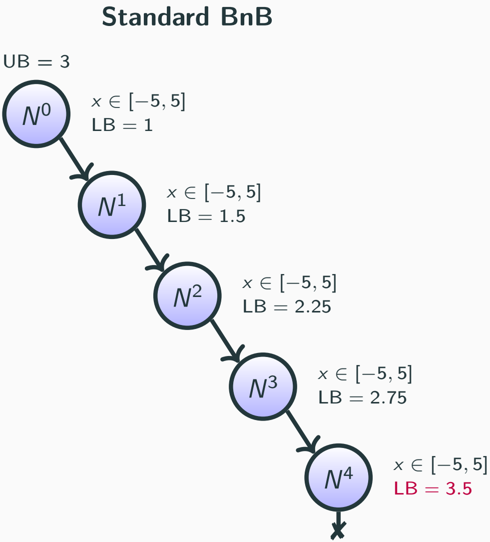
Find some $[L', U'] \subseteq [L, U]$ such that $N^{k'}$ is pruned



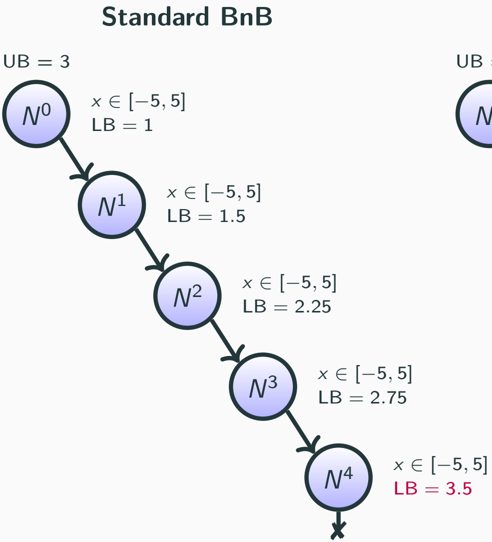
Ingredients

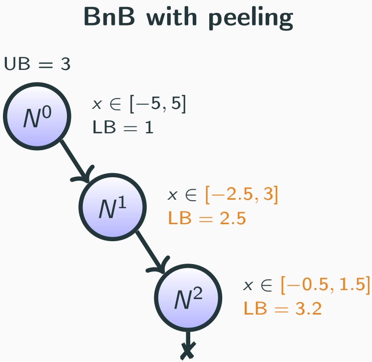
- Already-computed quantities
- Cost-free dual evaluation
- Dual link between nodes

Propagation along branches



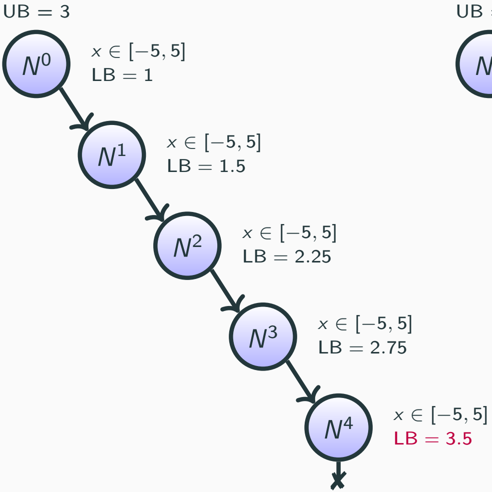
Propagation along branches



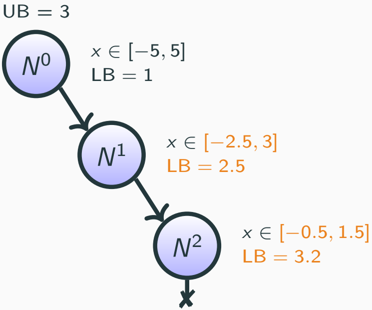


Propagation along branches

Standard BnB



BnB with peeling



- ✔ Branches pruned early-on
- ✔ Less nodes explored
- ✔ Reduce solve time

Numerical results

Numerical results

$$\min_{x \in [L,U]} \ \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$$

Data A **and** y : Sparse regression

Bounds L **and** U : $-M \leq x \leq M$

Parameter λ : Set statistically

Bounds spead : $M = \gamma\|x^\star\|_\infty$

Numerical results

$$\min_{x\in[L,U]} \ \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0$$

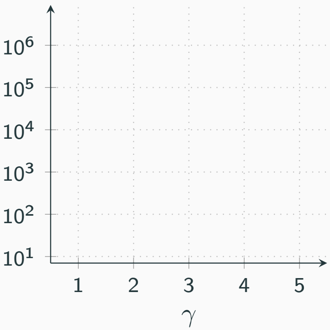
Data A **and** y : Sparse regression

Bounds L **and** U : $-M \leq x \leq M$

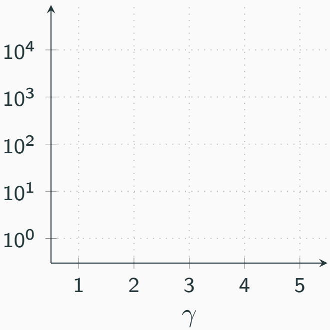
Parameter λ : Set statistically

Bounds spead : $M = \gamma\|x^\star\|_\infty$

Node count



Solve time (sec.)



Numerical results

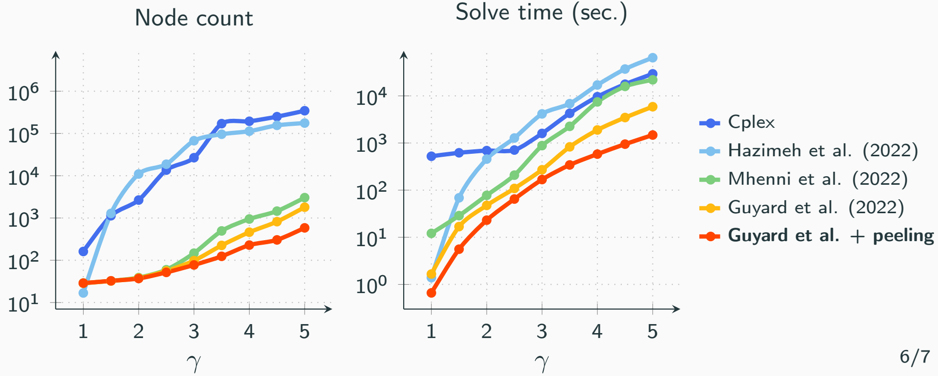
$$\min_{x \in [L,U]} \ \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_0$$

Data A **and** y : Sparse regression

Bounds L **and** U : $-M \leq x \leq M$

Parameter λ : Set statistically

Bounds spead : $M = \gamma \|x^\star\|_\infty$



Take-home message

ℓ_0-regularized least-squares

$$\begin{aligned} \min_x \quad & \tfrac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_0 \\ \text{s.t.} \quad & L \leq x \leq U \end{aligned}$$

Wide interval $[L, U]$ to recover the solution

Tight interval $[L, U]$ to build strong relaxations

Peeling

Start with wide interval $[L, U]$ and refine it along the BnB branches

Question time

🤓