Multi-Objective Parametric Query Optimization

Immanuel Trummer, Christoph Koch

September 2, 2015

Paper Goal

Generalize Query Optimization

Traditional Query Optimization

- Input: query
- One cost metric
- Plan cost known
- Goal: find optimal plan

Traditional Query Optimization

- Input: query
- One cost metric
- Plan cost known
- Goal: find optimal plan

$$\text{Cost}(\text{Plan}) \in \mathbb{R}$$

Parametric Query Optimization

- Input: query **template**
- One cost metric
- Plan cost **unknown**
- Goal: find optimal plan **for each parameter region**

Parametric Query Optimization

- Input: query **template**
- One cost metric
- Plan cost **unknown**
- Goal: find optimal plan **for each parameter region**

$$\text{Cost(Plan)} \in \mathbb{R}^N \rightarrow \mathbb{R}$$

Multi-Objective Query Optimization

- Input: query
- **Multiple** cost metrics
- Plan cost known
- Goal: find **Pareto-optimal** plans

Multi-Objective Query Optimization

- Input: query
- **Multiple** cost metrics
- Plan cost known
- Goal: find **Pareto-optimal** plans

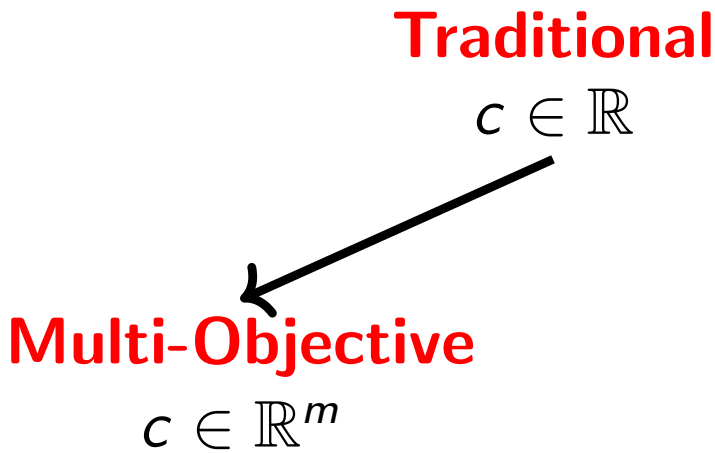
$$\text{Cost(Plan)} \in \mathbb{R}^N$$

Cost Models of Query Optimization Variants

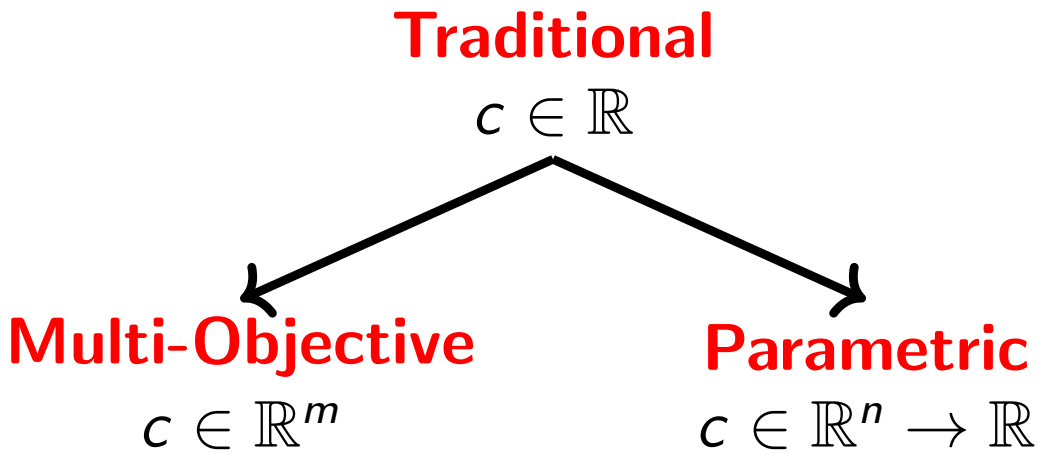
Traditional

$c \in \mathbb{R}$

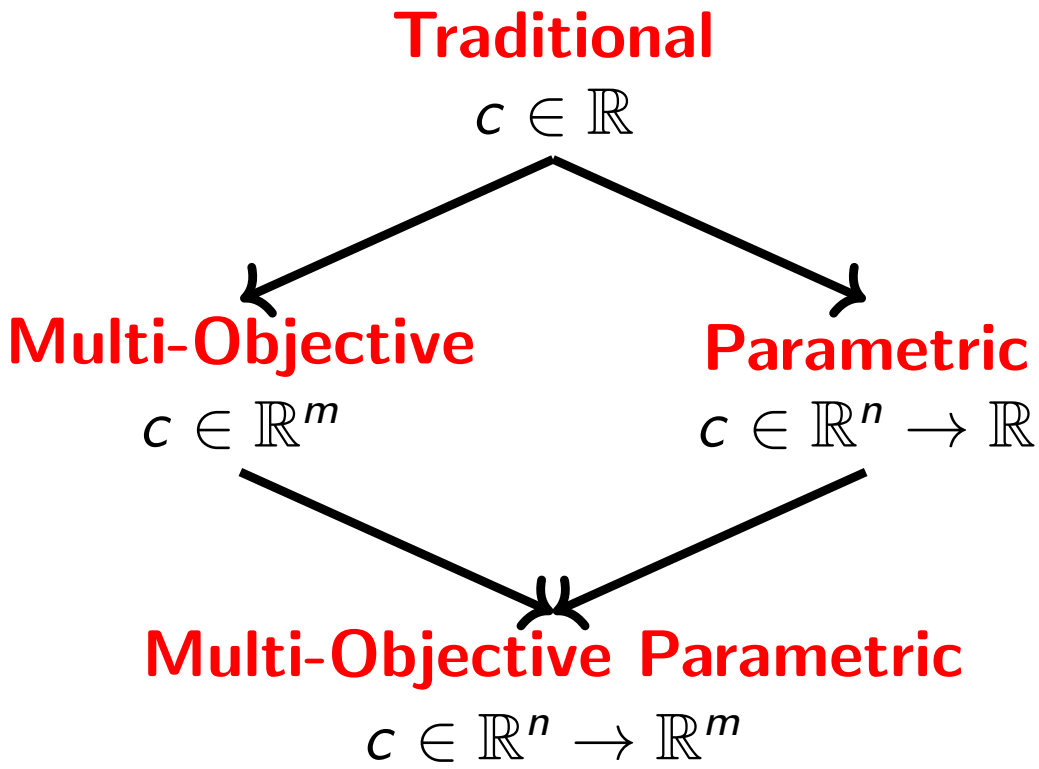
Cost Models of Query Optimization Variants



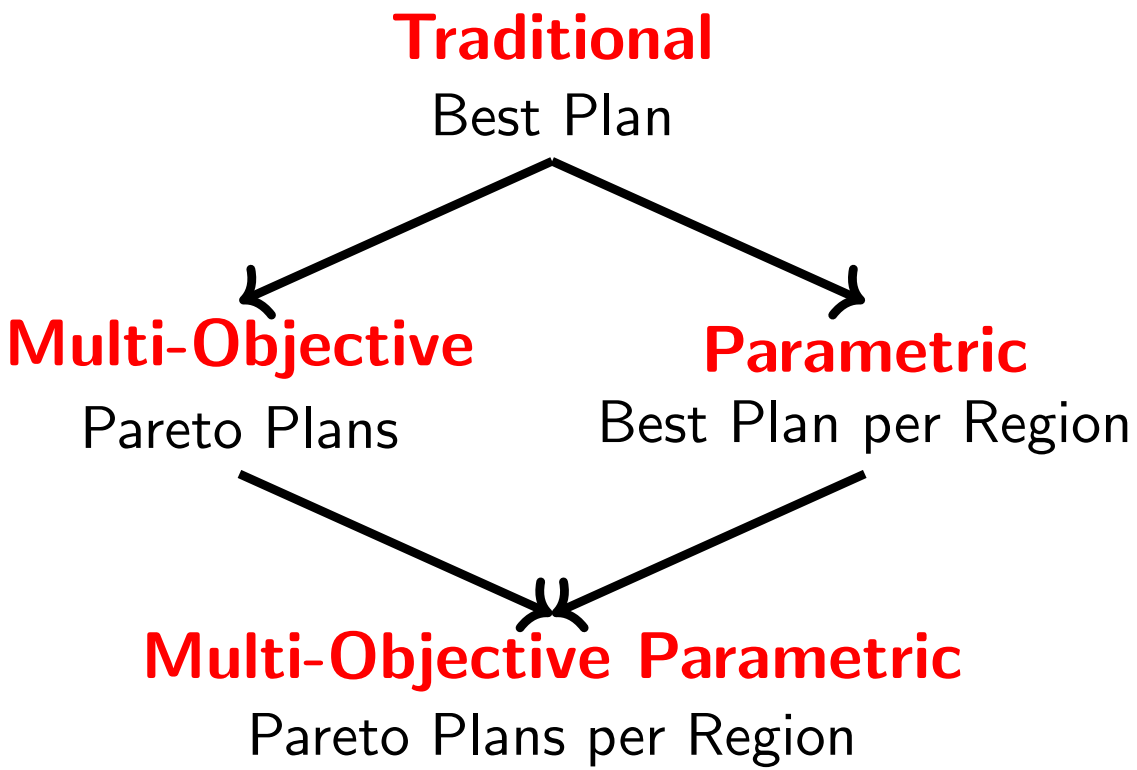
Cost Models of Query Optimization Variants



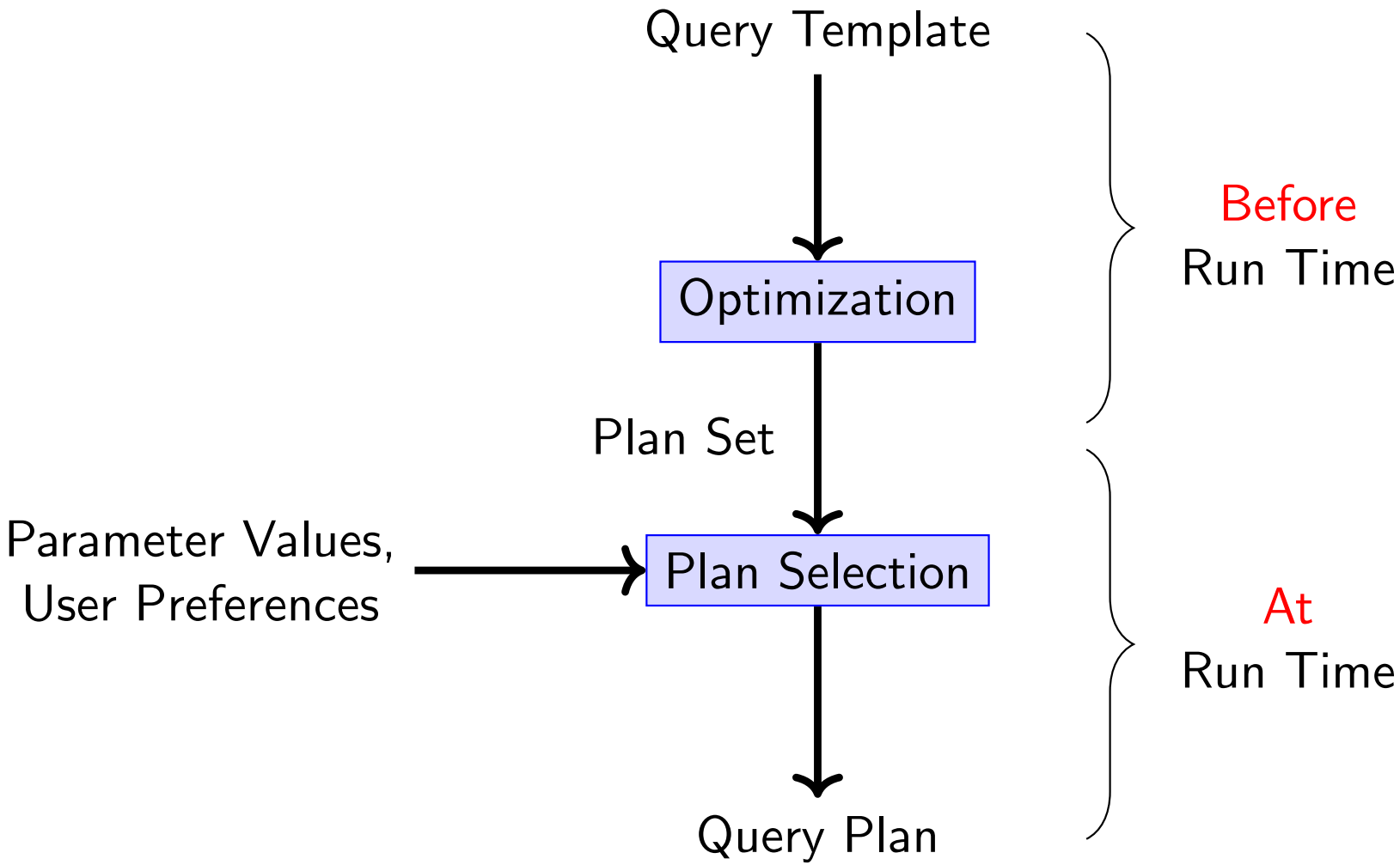
Cost Models of Query Optimization Variants



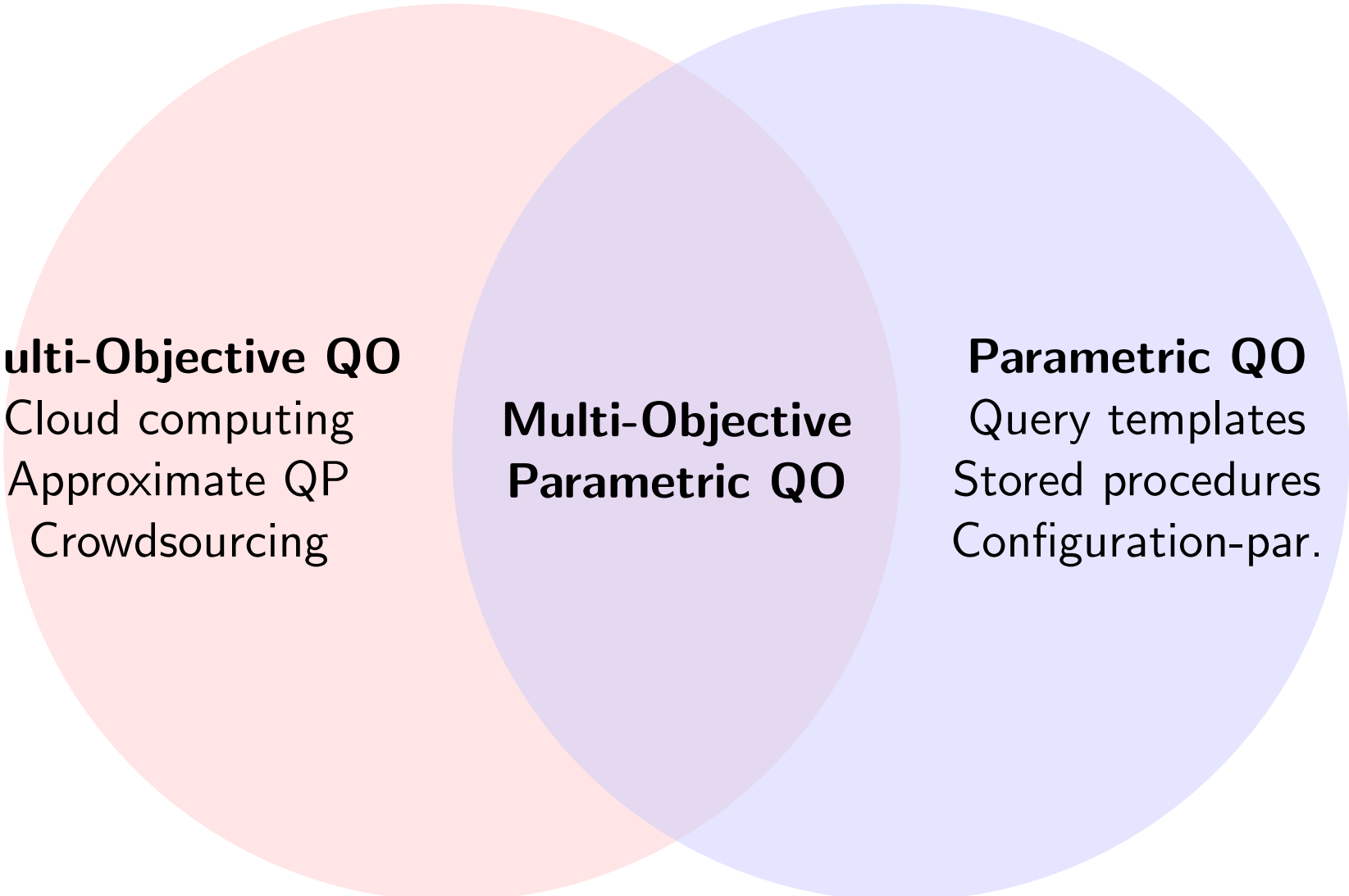
Optimization Goals



Context of MPQ



Application Scenarios



Paper Contribution Overview

- Formalize & analyze MPQ
- Propose first algorithm
- Theoretical analysis of algorithm
- Experimental evaluation

Dynamic Programming Scheme

Dynamic Programming Scheme

QUERY: $R \bowtie S \bowtie T$

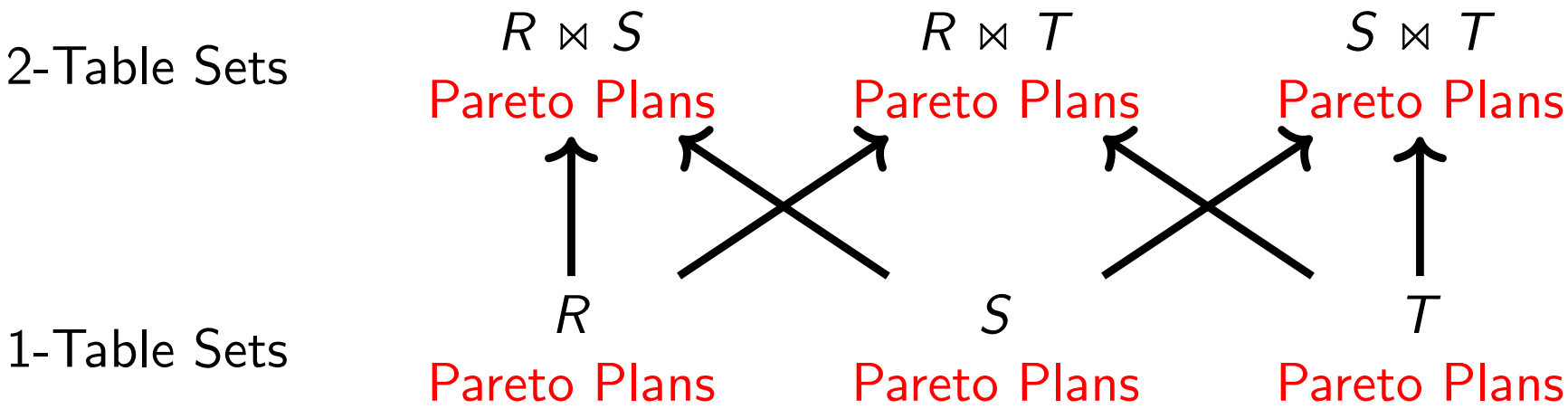
Dynamic Programming Scheme

QUERY: $R \bowtie S \bowtie T$

1-Table Sets	R	S	T
	Pareto Plans	Pareto Plans	Pareto Plans

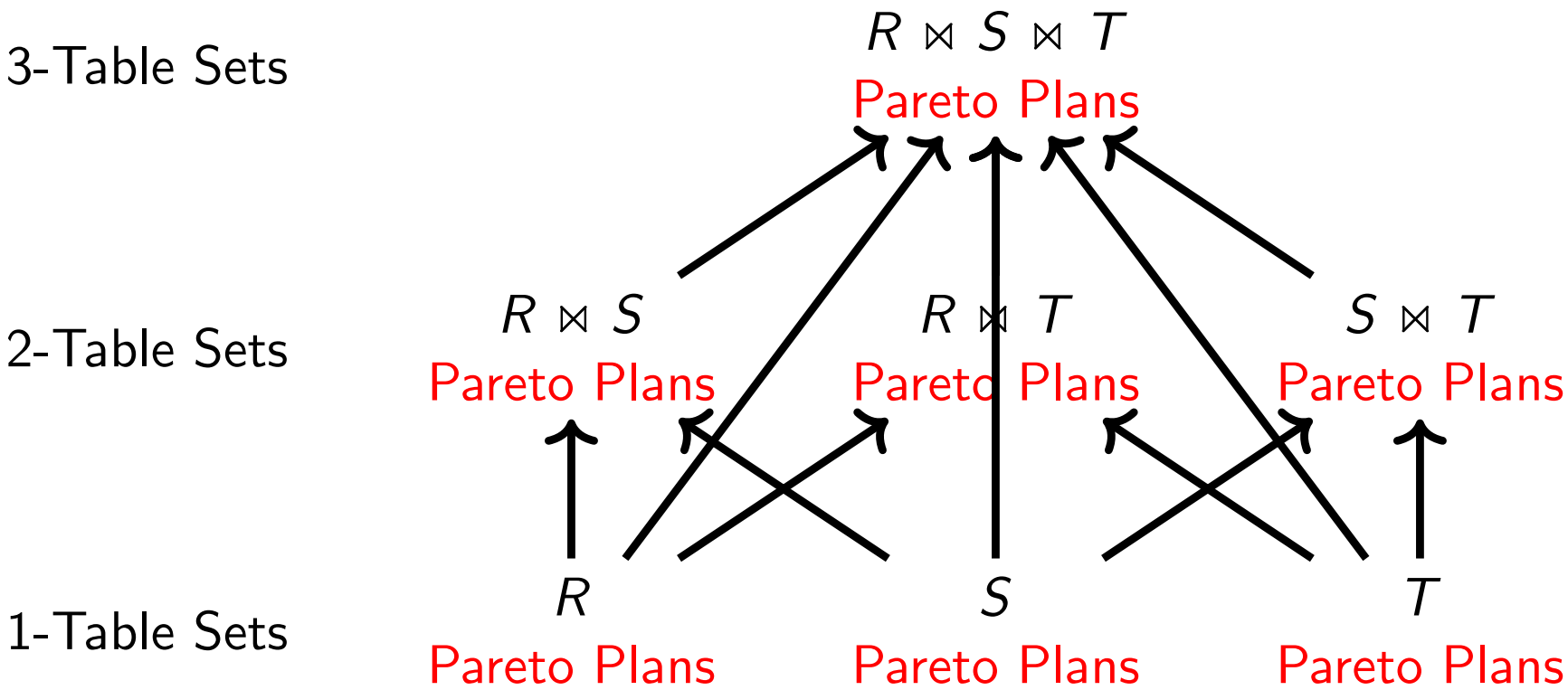
Dynamic Programming Scheme

QUERY: $R \bowtie S \bowtie T$



Dynamic Programming Scheme

QUERY: $R \bowtie S \bowtie T$



Pruning Overview

- Each query plan p maps to a Pareto region $R(p)$
- Compare plans pairwise during pruning:
  If p_1 dominates p_2 in X then $R(p_2) \leftarrow R(p_2) \setminus X$
- Plans with empty Pareto region are discarded

Questions

- Shape of Pareto regions?
- How to represent them?
- How to implement comparison, reduction, emptiness check?

Questions

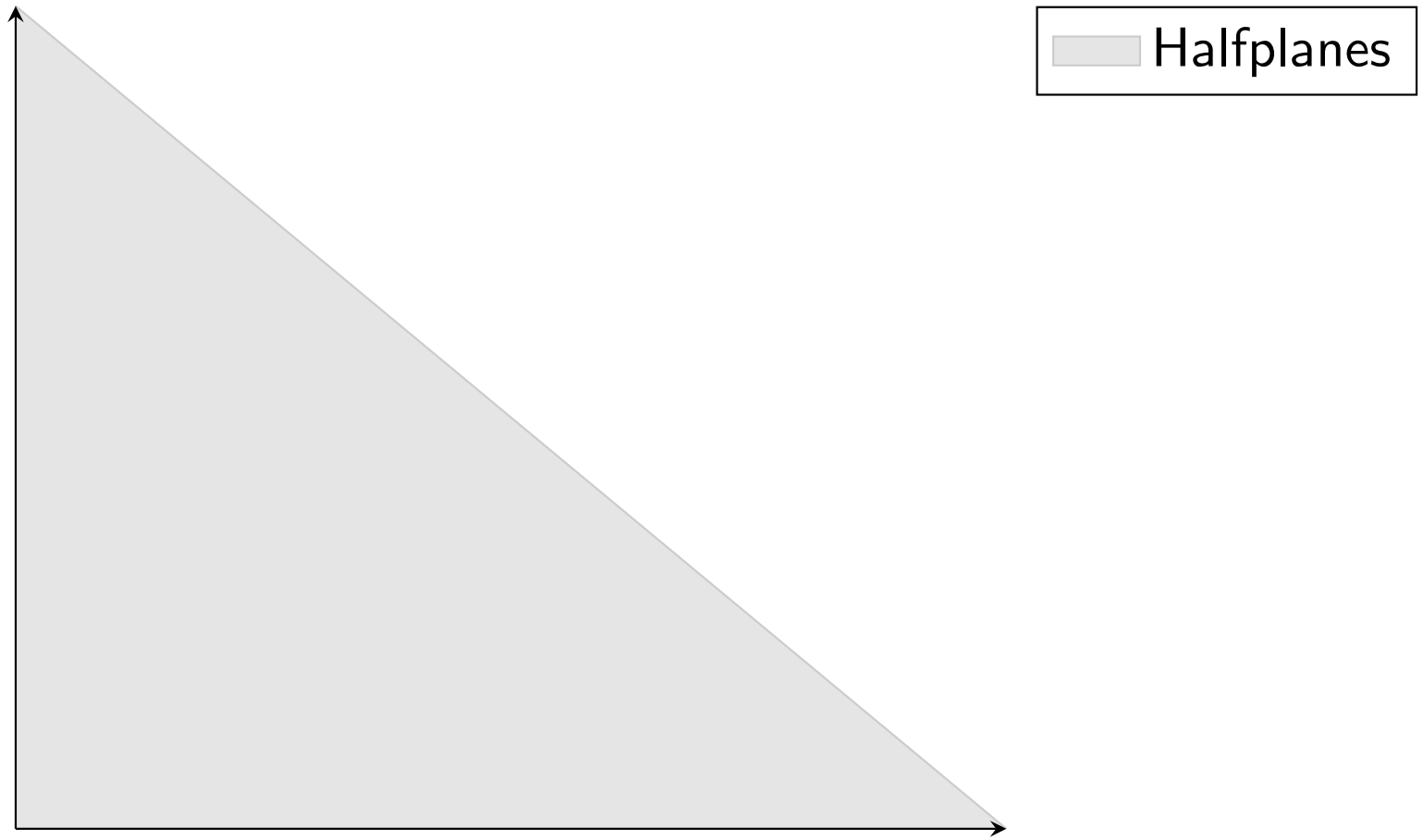
- Shape of Pareto regions?
- How to represent them?
- How to implement comparison, reduction, emptiness check?

Choose: Piece-wise Linear Plan Cost Functions

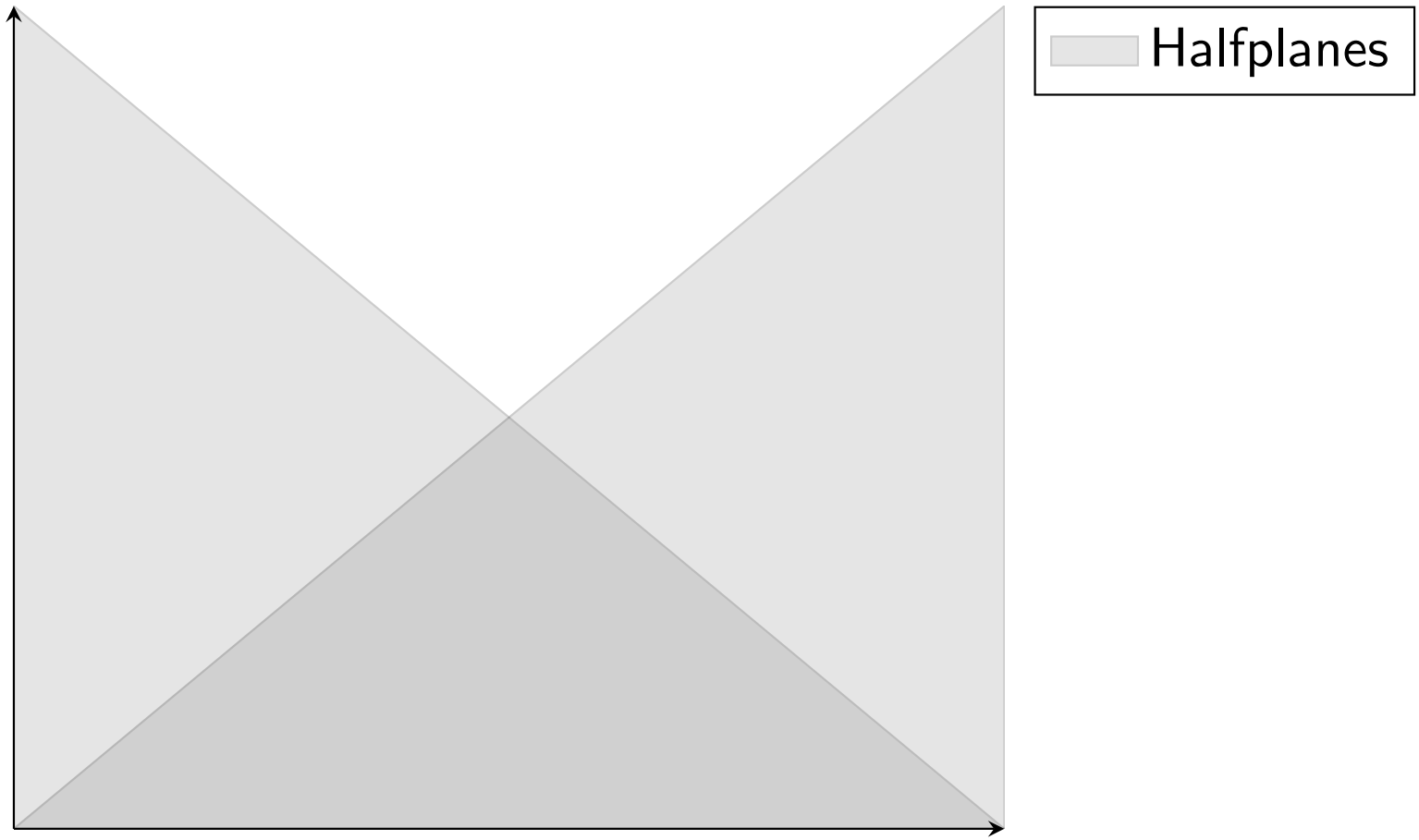
Answer

- Region representation based on convex polytopes

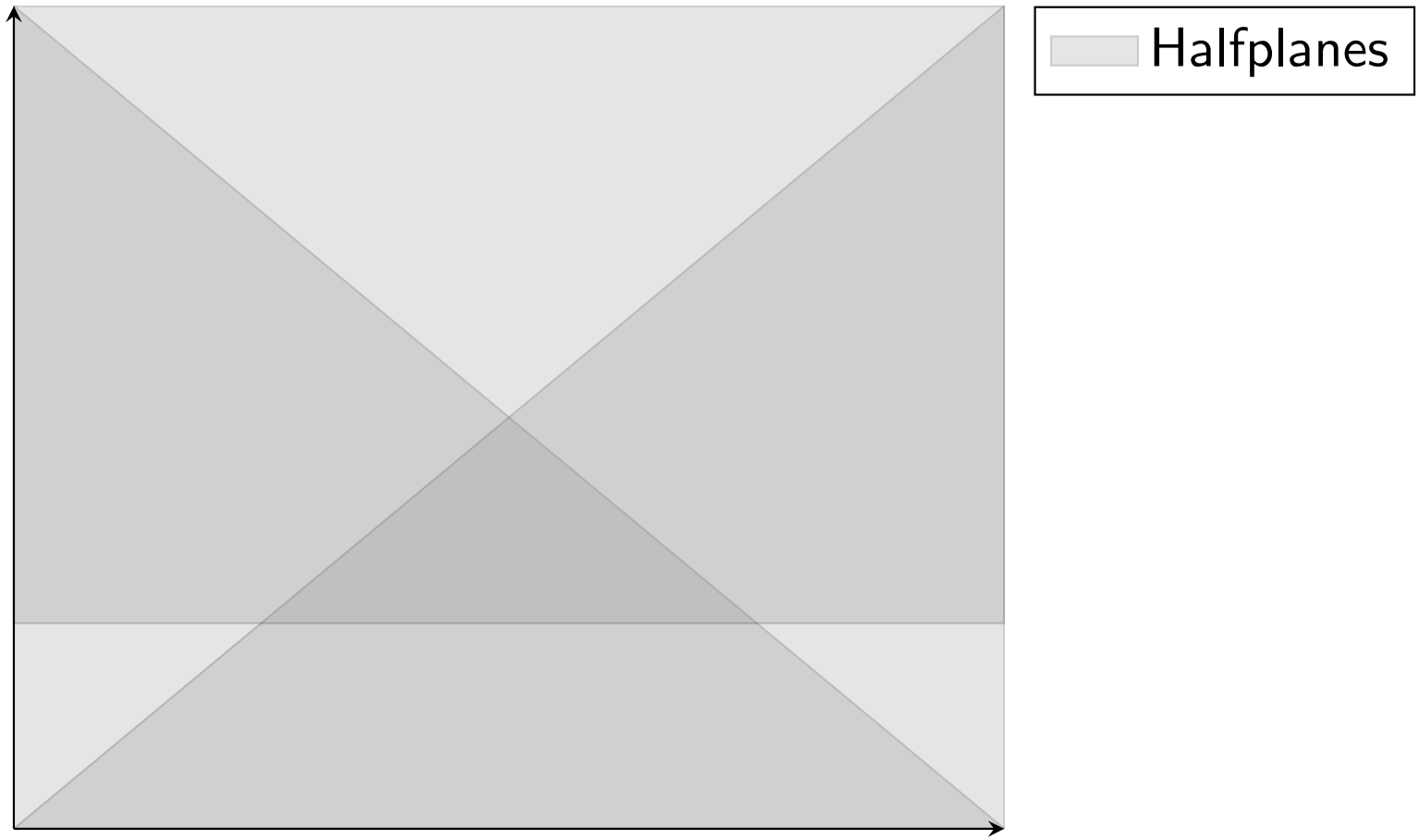
Convex Polytopes: Intersections of Half-Planes



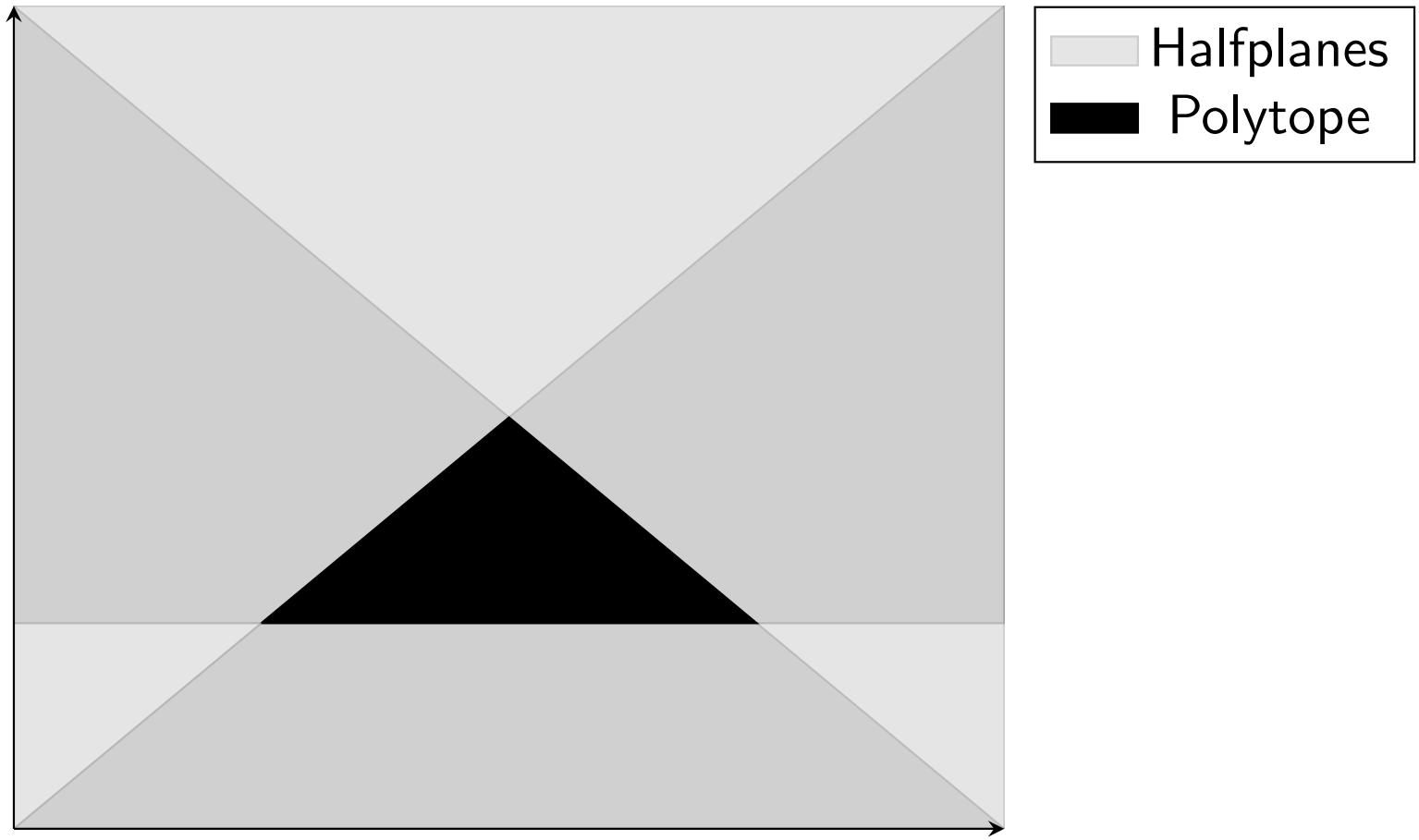
Convex Polytopes: Intersections of Half-Planes



Convex Polytopes: Intersections of Half-Planes



Convex Polytopes: Intersections of Half-Planes



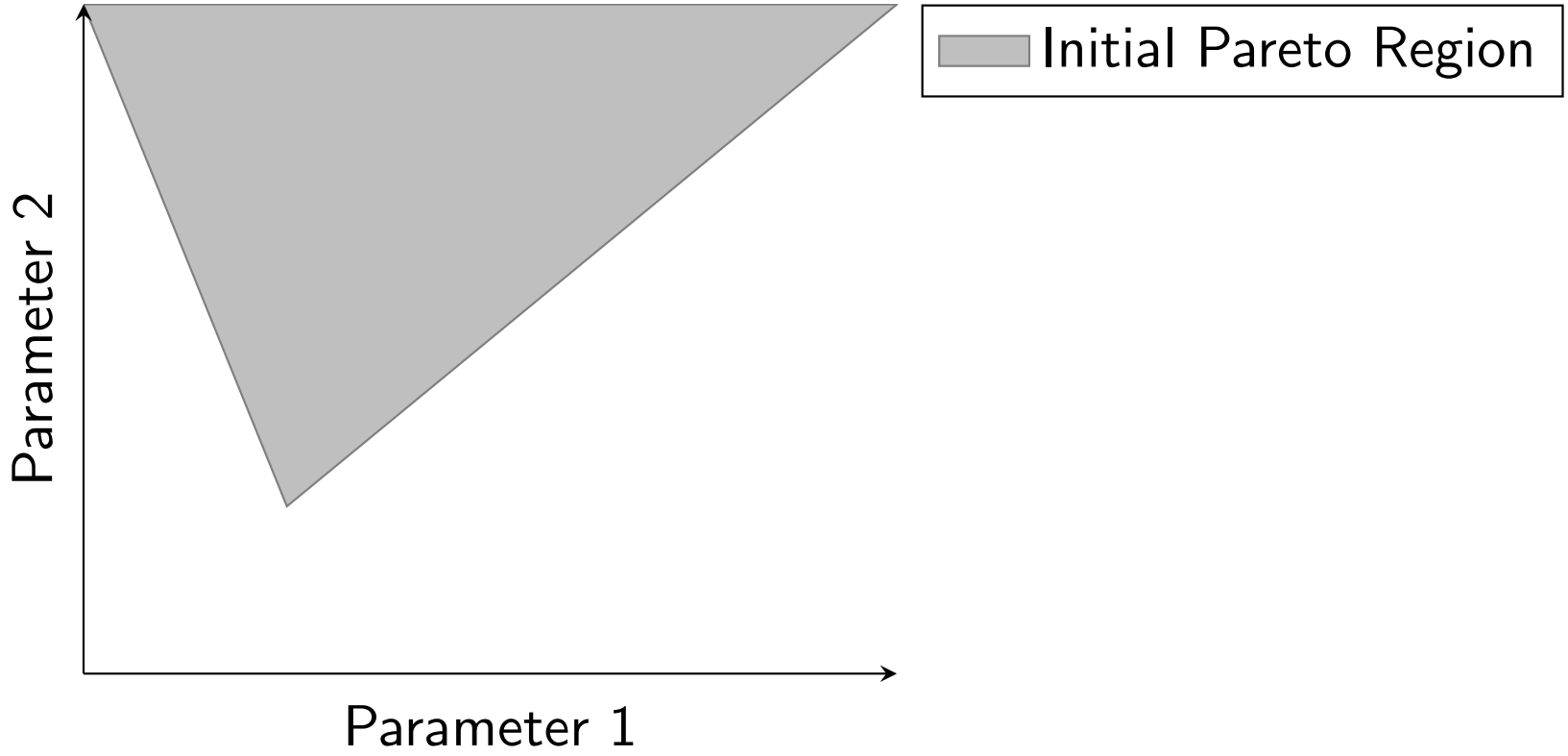
Answer

- Region representation based on convex polytopes

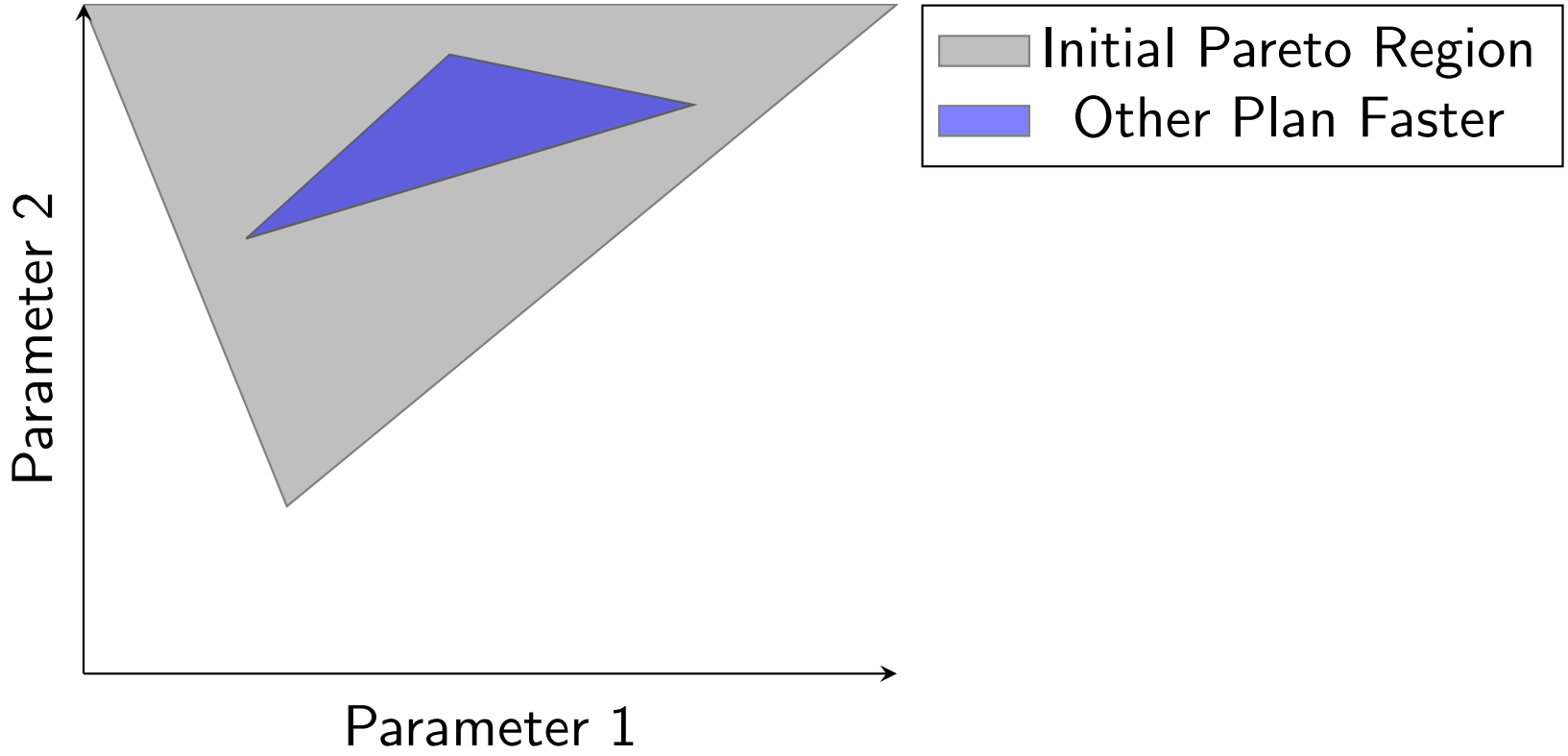
Answer

- Region representation based on convex polytopes
- Pareto region: union of sub-regions
- Sub-region: convex polytope minus union of convex polytopes
- Implement basic operations using linear programming solver

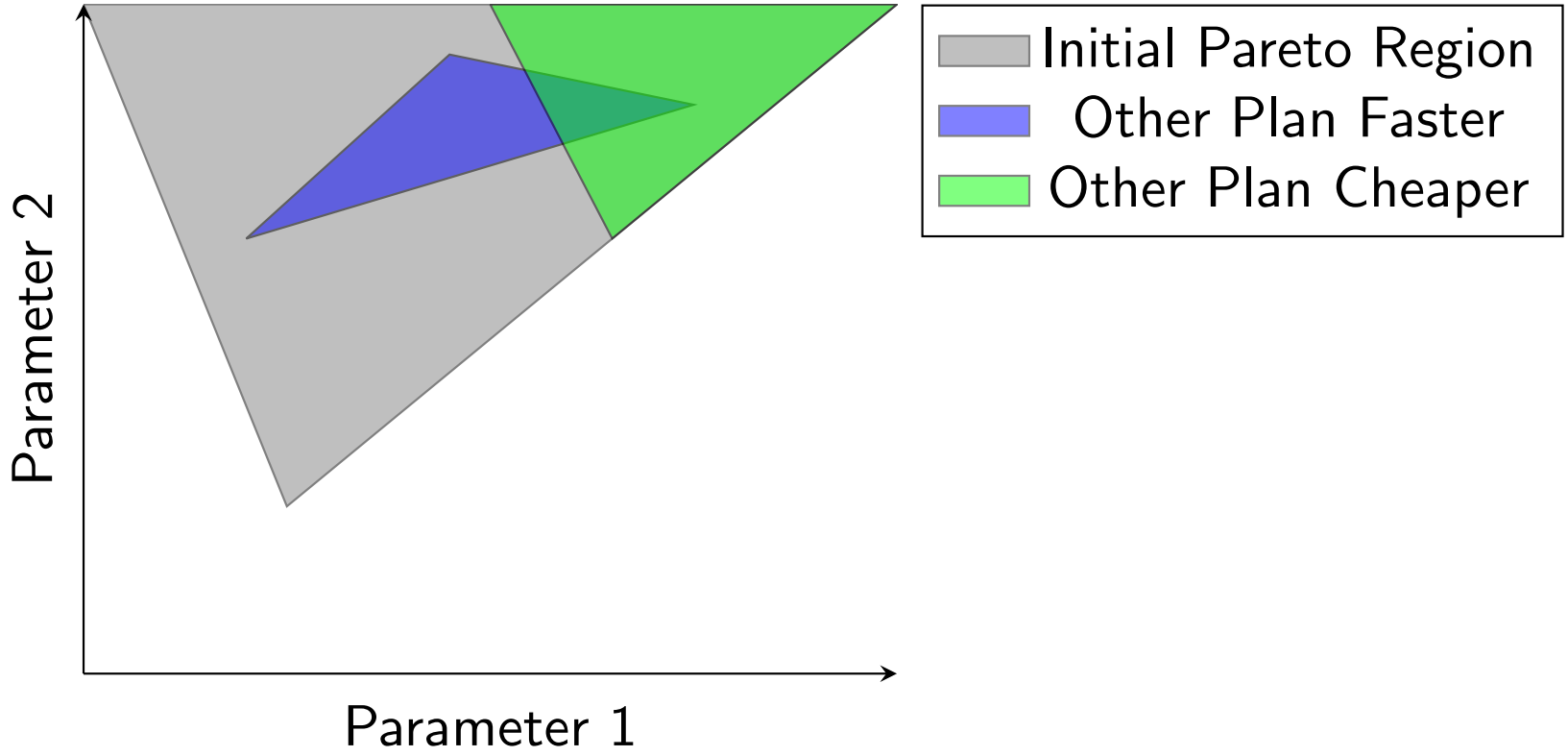
Pruning Illustration



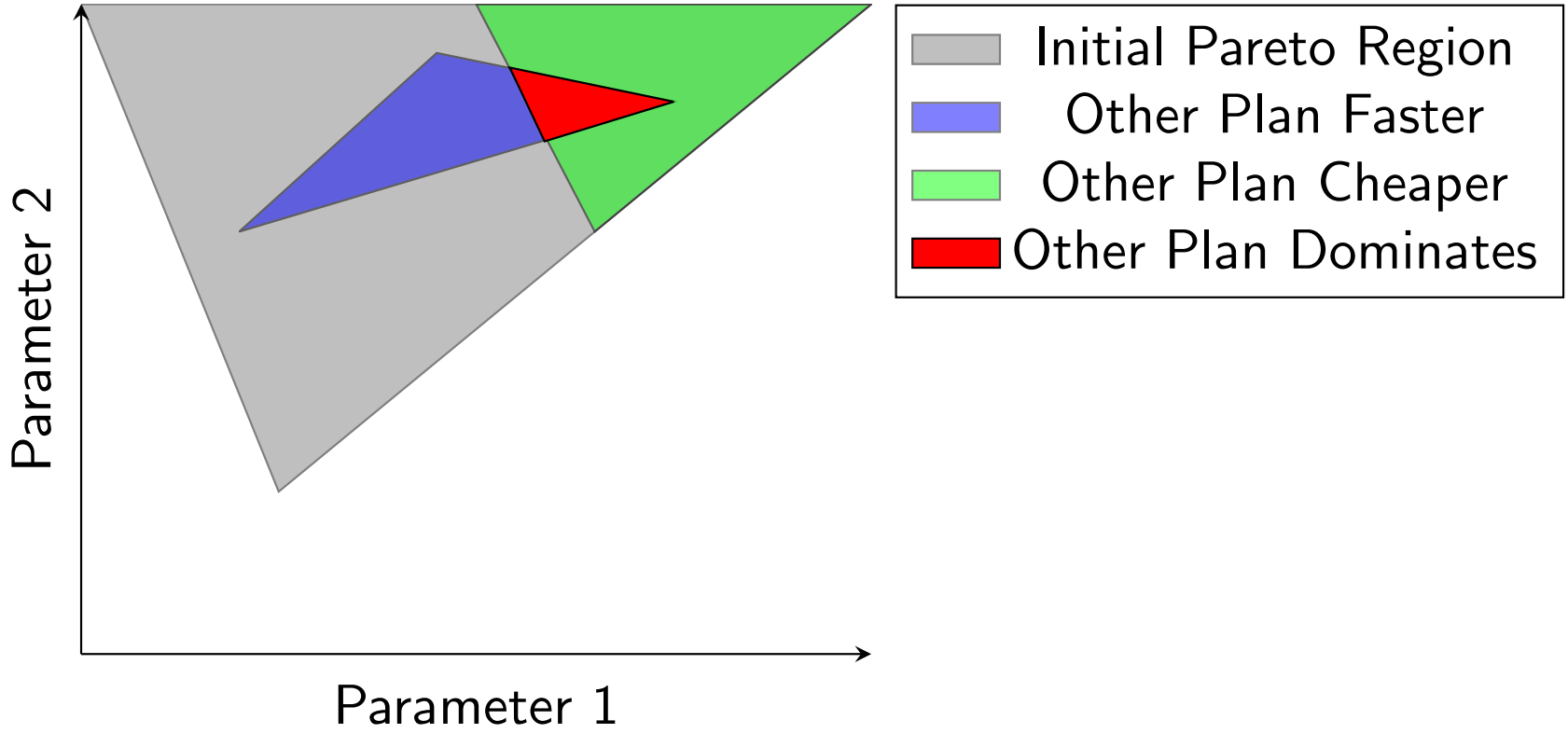
Pruning Illustration



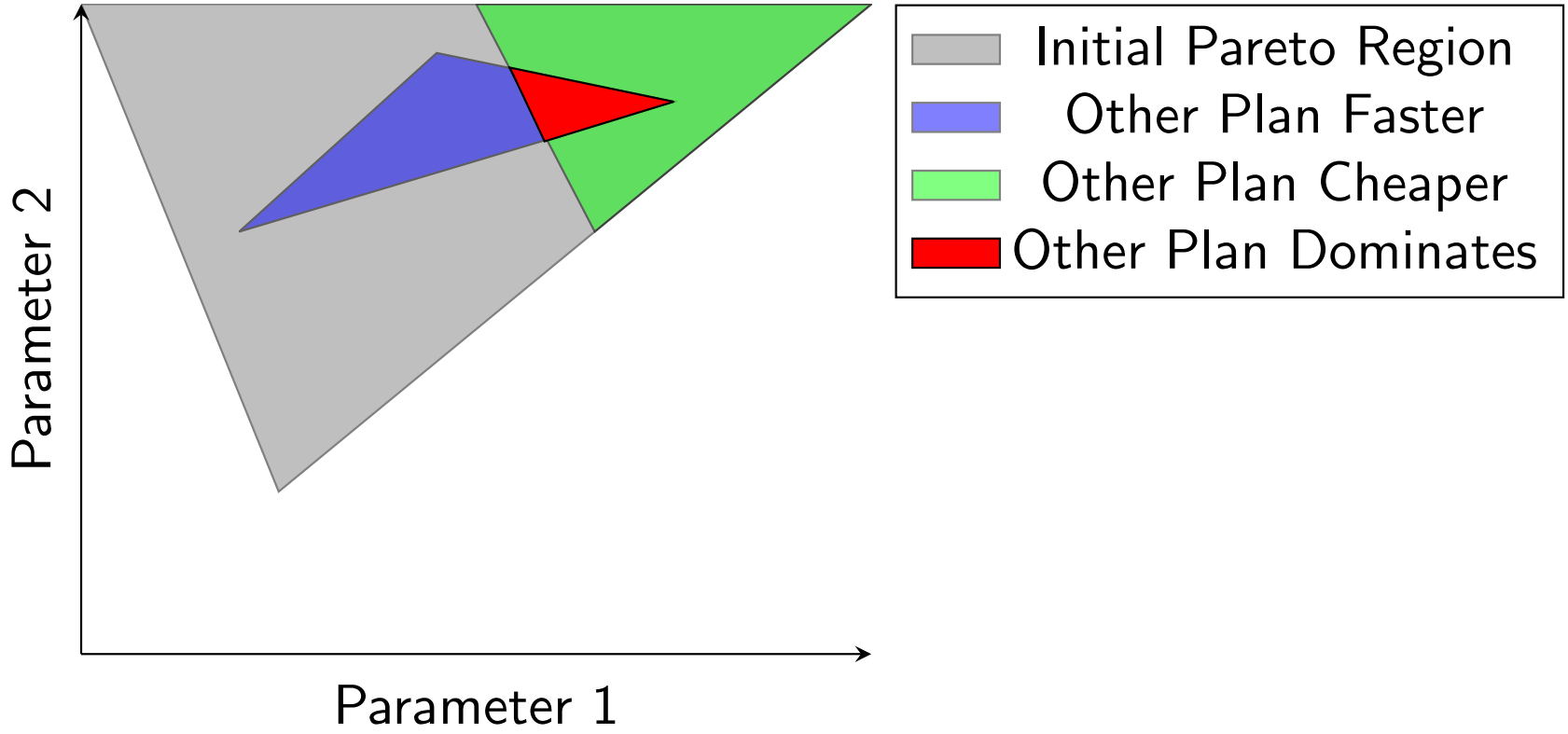
Pruning Illustration



Pruning Illustration

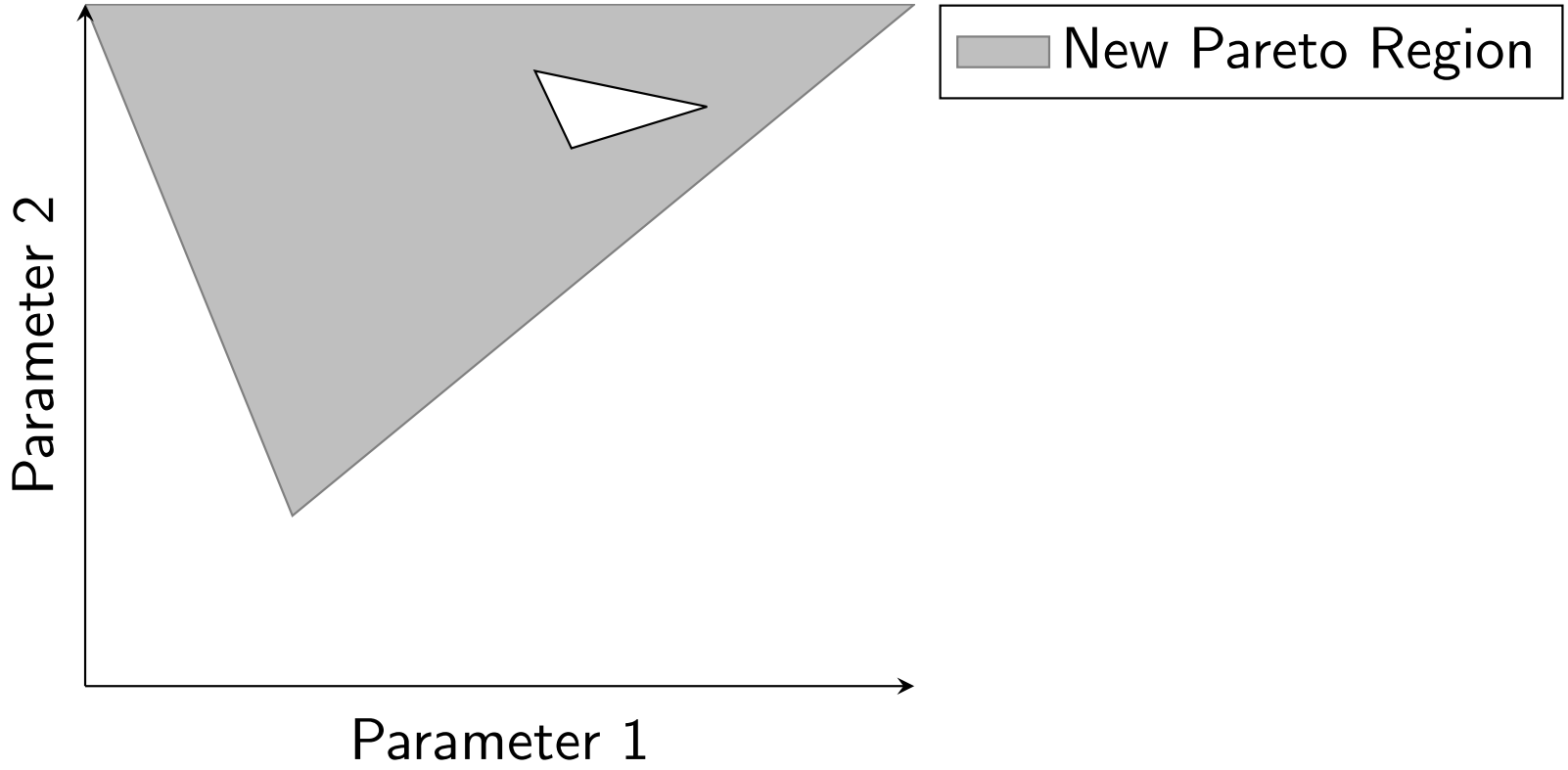


Pruning Illustration



Dominated Area is Removed from Pareto Region

Pruning Illustration



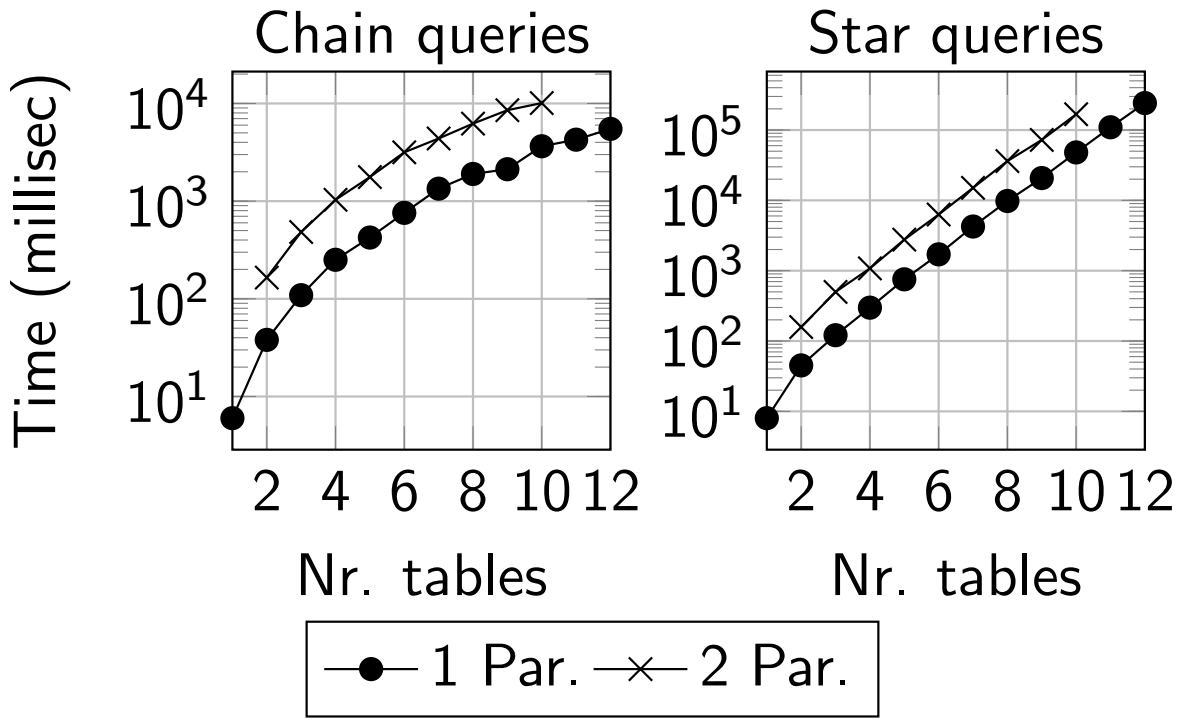
Complexity Analysis

Query Optimization Variant	**Nr. Pareto Plans per Relation**
Traditional	1
Multi-Objective	$2^{nrMetrics}$
Multi-objective Parametric	$2^{(nrParams+1)\cdot nrMetrics}$

Experimental Setup

- **Benchmark**
  - Randomly generated queries
  - Different query structures
  - Cloud scenario with two objectives
- **Algorithms**
  - Multi-objective parametric optimizer
- **Comparison**
  - Measure optimization time
  - Vary number of query tables

Experimental Results (Extract)



Conclusion

- Generalized query optimization model
- Presented first MPQ algorithm
- Future work: heuristics, approximation, ...

Thanks for your Attention!

Questions?

Slides/More on Query Optimization:
`www.itrummer.org`