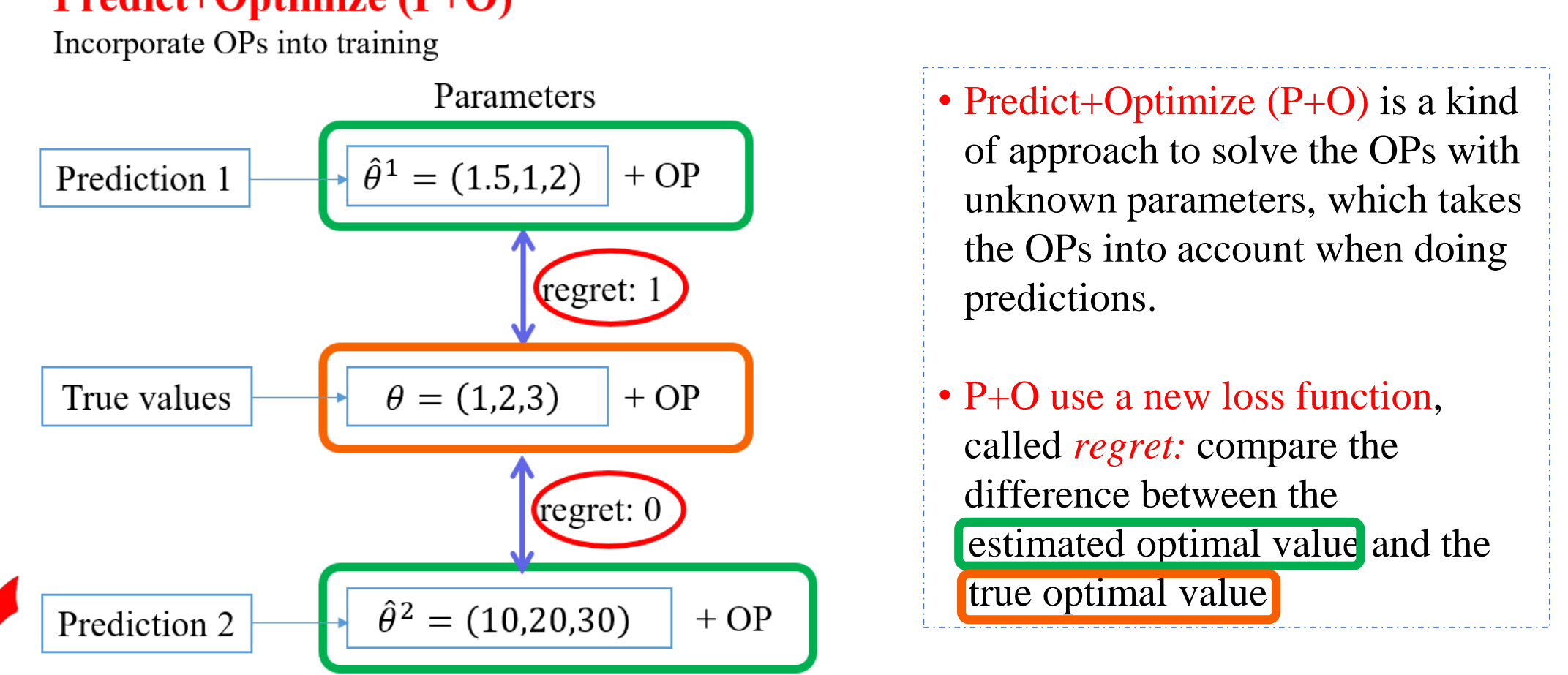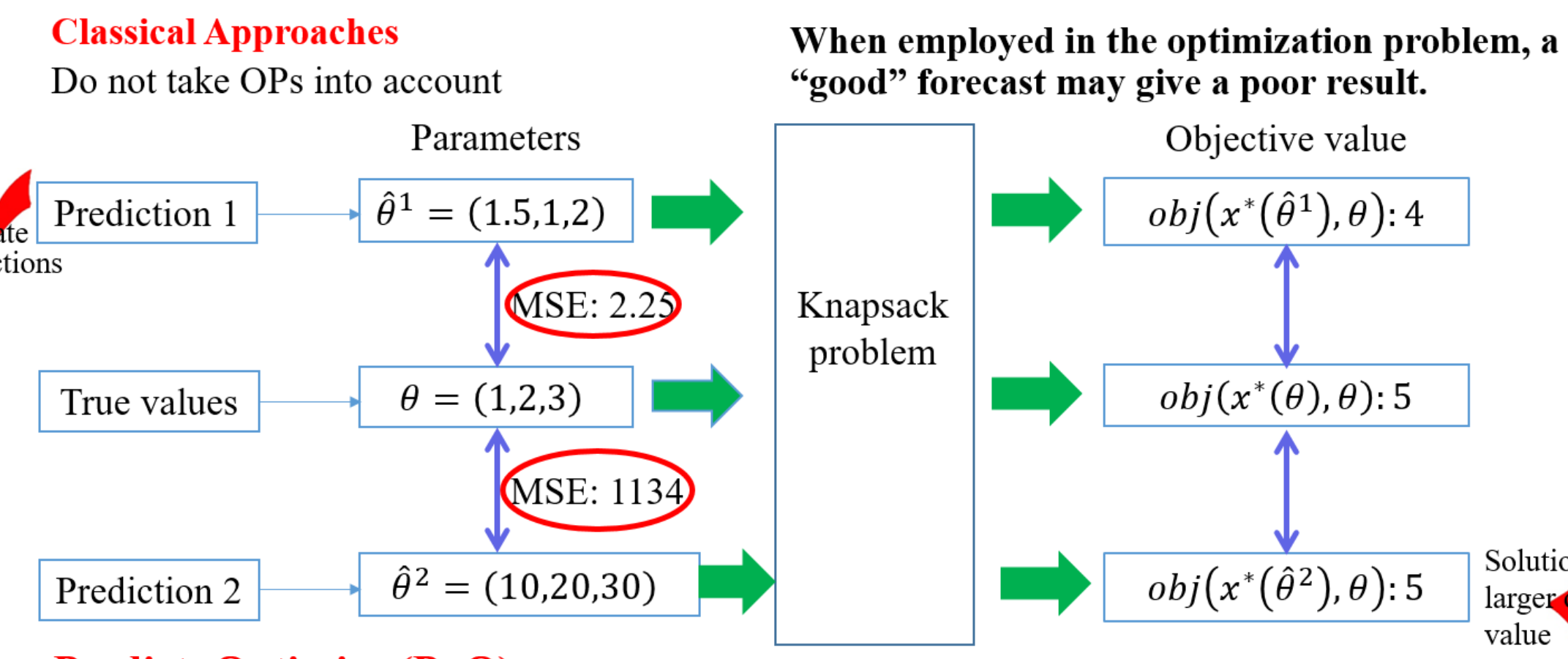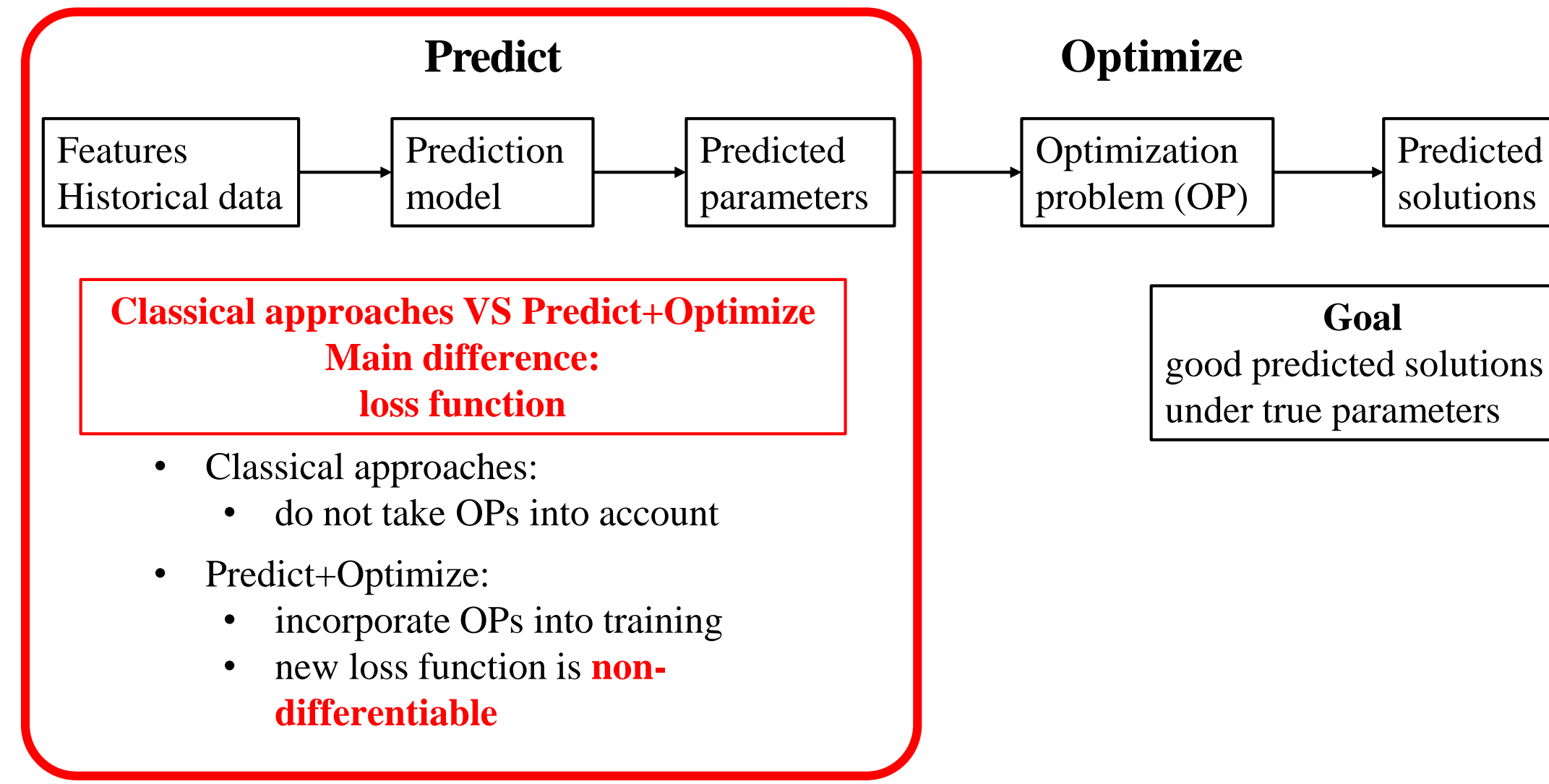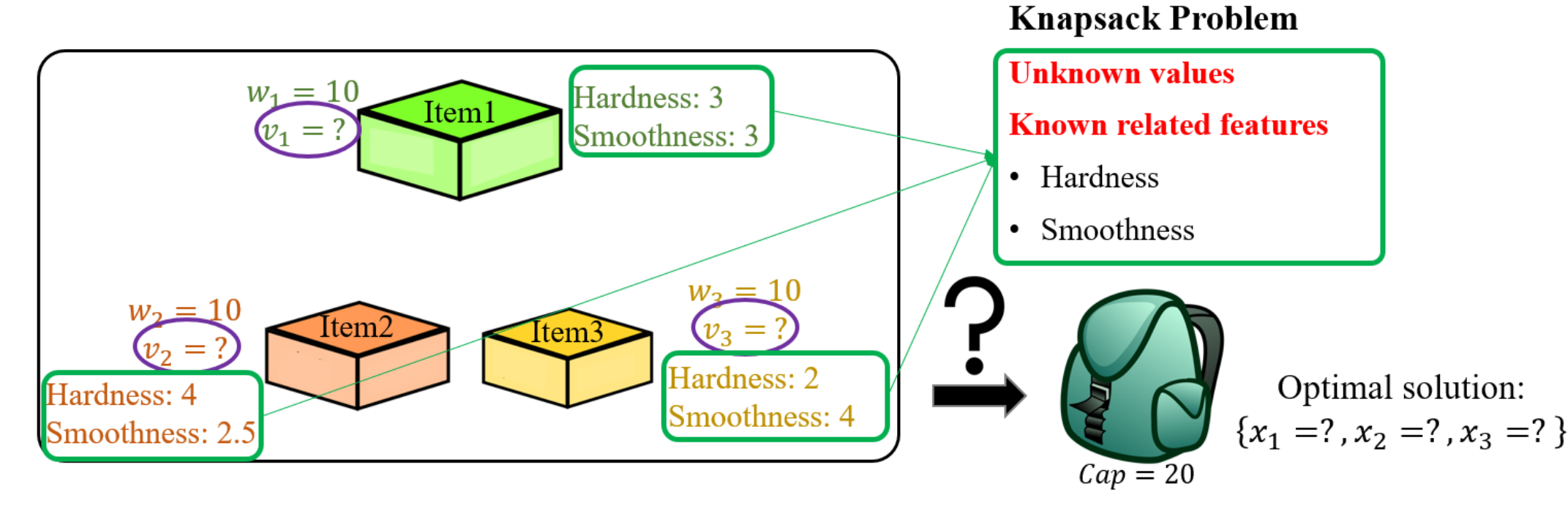# Predict+Optimize for Packing and Covering LPs with Unknown Parameters in Constraints

Xinyi Hu [1]   Jasper C.H. Lee [2]   Jimmy H.M. Lee [1]
[1] The Chinese University of Hong Kong
[2] University of Wisconsin–Madison

Motivation

In practice, optimization problems (OPs) usually contain parameters, some of which may be unknown, but with related features and historical data.

Knapsack Problem

Item1: $w_1 = 10$, $v_1 = ?$, Hardness: 3, Smoothness: 3
Item2: $w_2 = 10$, $v_2 = ?$, Hardness: 4, Smoothness: 2.5
Item3: $w_3 = 10$, $v_3 = ?$, Hardness: 2, Smoothness: 4

Unknown values
Known related features
- Hardness
- Smoothness

$Cap = 20$

Optimal solution: $\{x_1 = ?, x_2 = ?, x_3 = ?\}$

Predict → **Optimize**

Features / Historical data → Prediction model → Predicted parameters → Optimization problem (OP) → Predicted solutions

Goal
good predicted solutions under true parameters

Classical approaches VS Predict+Optimize
Main difference:
loss function

- Classical approaches:
  - do not take OPs into account
- Predict+Optimize:
  - incorporate OPs into training
  - new loss function is **non-differentiable**

Classical Approaches
Do not take OPs into account

When employed in the optimization problem, a "good" forecast may give a poor result.

Parameters → Objective value

More accurate predictions

Prediction 1 → $\hat{\theta}^1 = (1.5, 1, 2)$ → $obj(x^*(\hat{\theta}^1), \theta): 4$

True values → $\theta = (1, 2, 3)$ → $obj(x^*(\theta), \theta): 5$

Prediction 2 → $\hat{\theta}^2 = (10, 20, 30)$ → $obj(x^*(\hat{\theta}^2), \theta): 5$

MSE: 2.25
MSE: 1134

Knapsack problem

Solution with larger objective value

Predict+Optimize (P+O)
Incorporate OPs into training

Parameters

Prediction 1 → $\hat{\theta}^1 = (1.5, 1, 2)$ + OP
regret: 1
True values → $\theta = (1, 2, 3)$ + OP
regret: 0
Prediction 2 → $\hat{\theta}^2 = (10, 20, 30)$ + OP

- Predict+Optimize (P+O) is a kind of approach to solve the OPs with unknown parameters, which takes the OPs into account when doing predictions.

- P+O use a new loss function, called *regret:* compare the difference between the estimated optimal value and the true optimal value
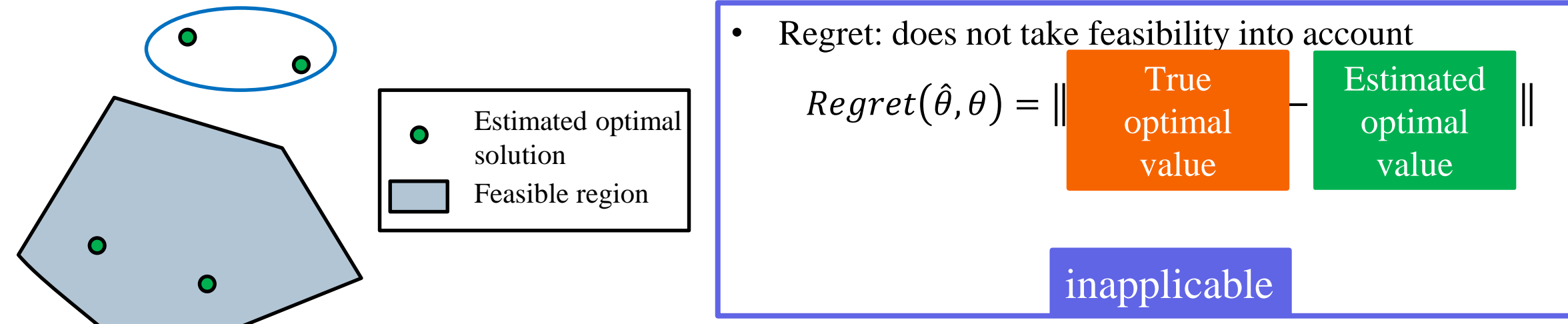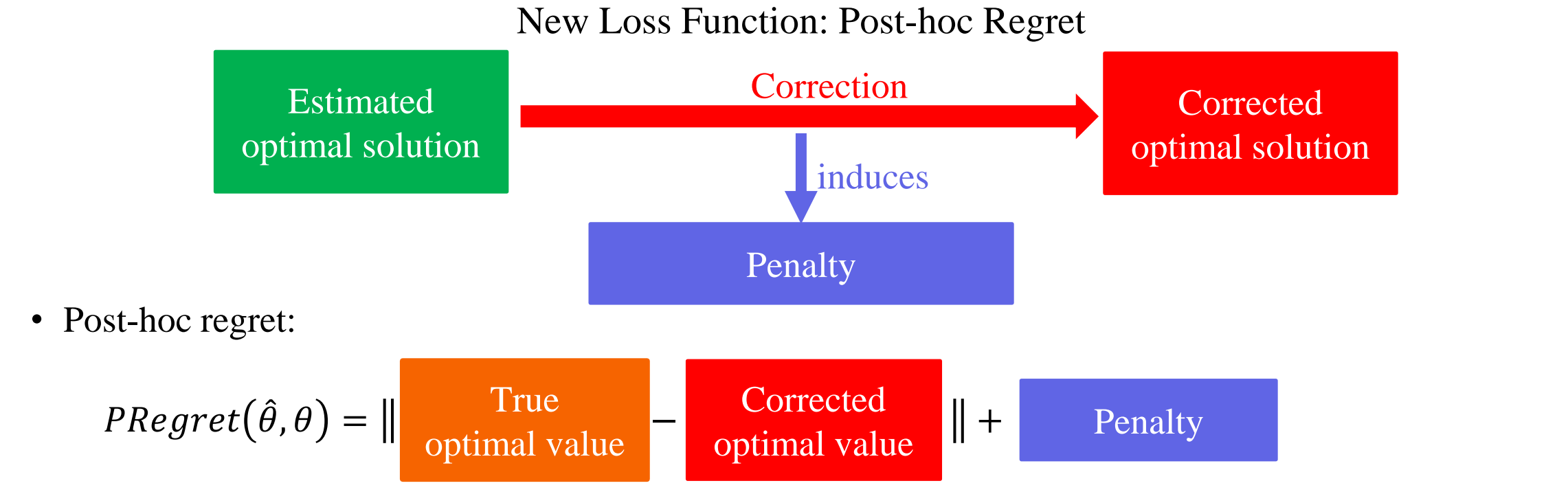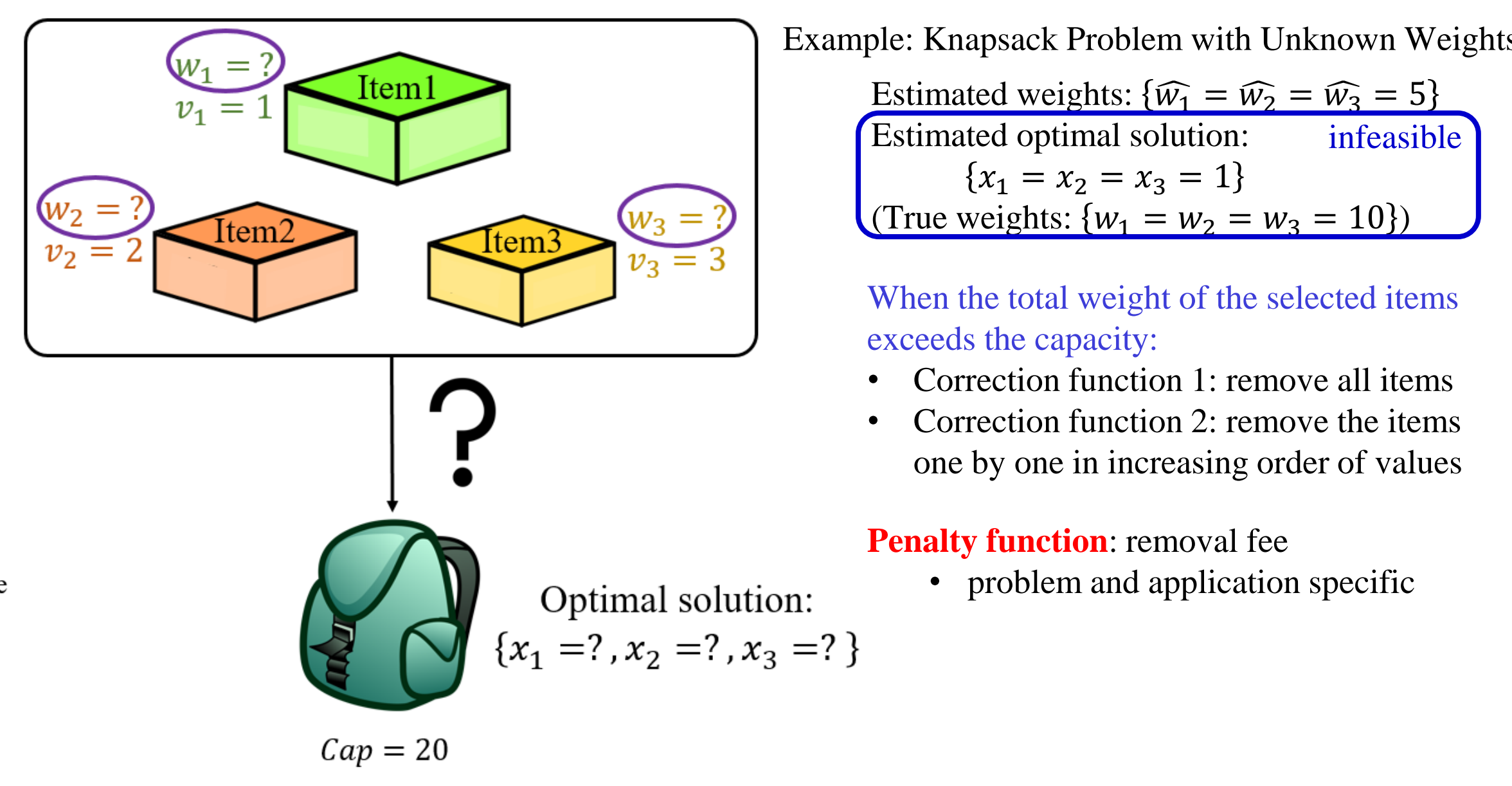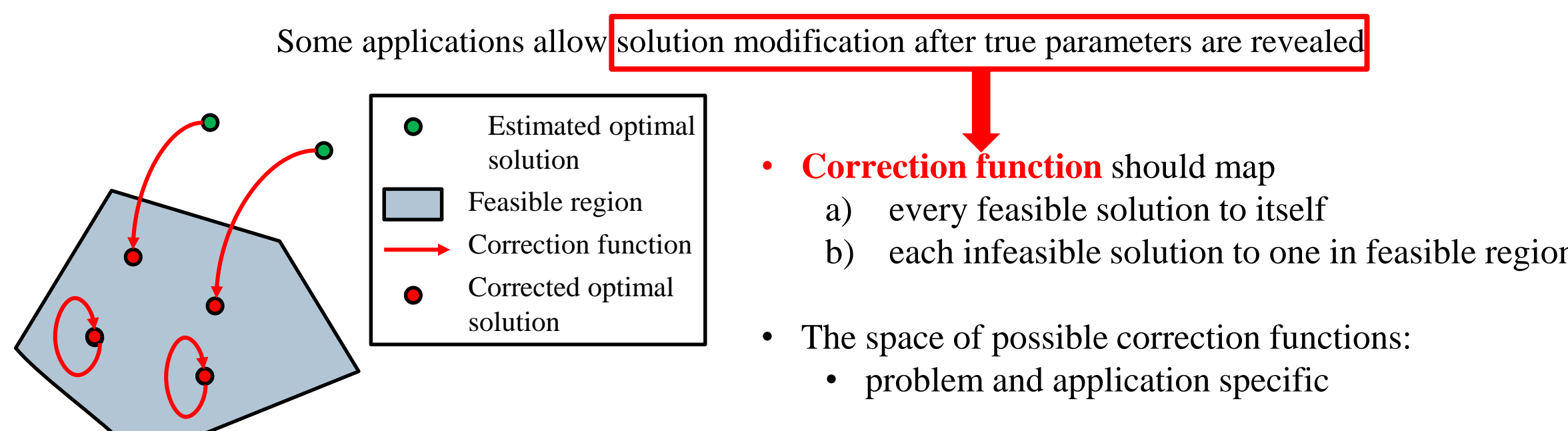
Limitations of Prior Works

- All focus on unknown parameters **only in the objective**

Challenge

Unknown parameters appearing in constraints (more complex)
- the estimated optimal solution may be out of the true solution space
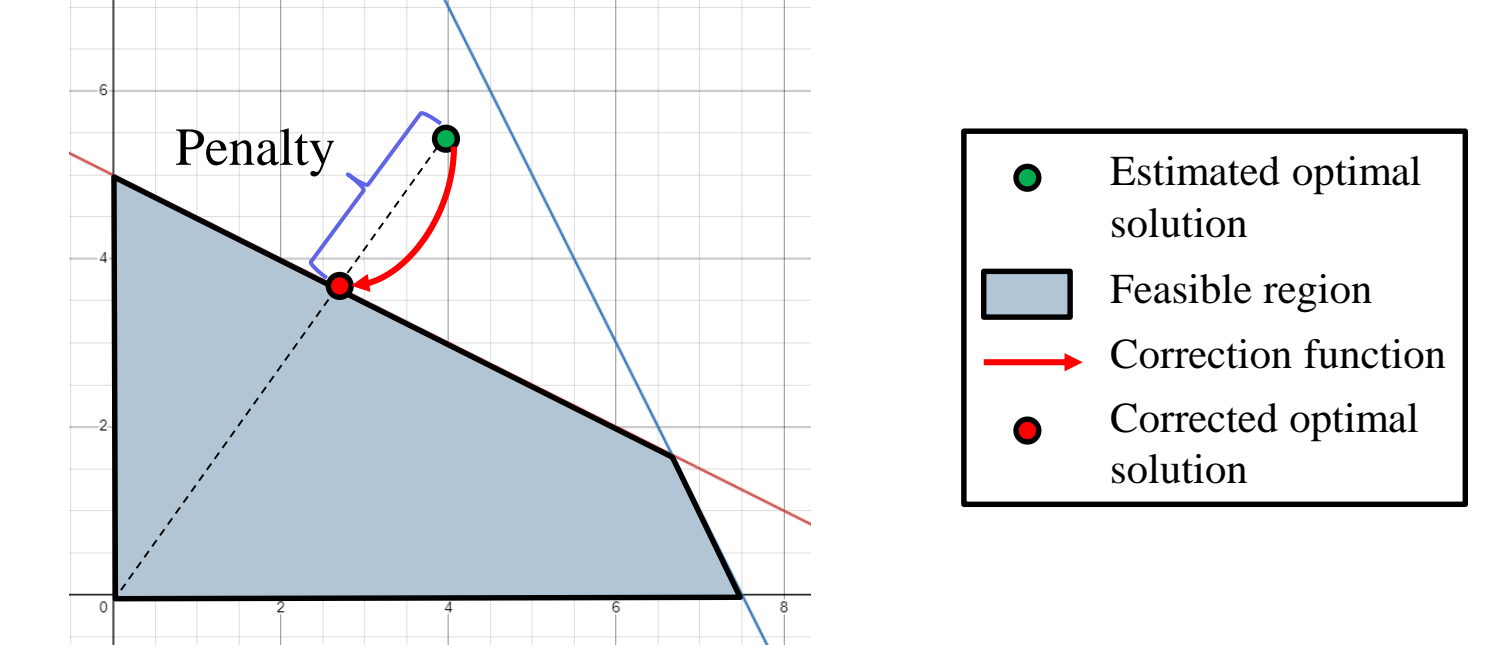
Estimated optimal solution
Feasible region

- Regret: does not take feasibility into account

$$Regret(\hat{\theta}, \theta) = \| \text{True optimal value} - \text{Estimated optimal value} \|$$

inapplicable

Contribution 1: P+O Framework for Unknown Constraints

Some applications allow solution modification after true parameters are revealed

Estimated optimal solution
Feasible region
Correction function
Corrected optimal solution

- **Correction function** should map
  a) every feasible solution to itself
  b) each infeasible solution to one in feasible region

- The space of possible correction functions:
  - problem and application specific

Example: Knapsack Problem with Unknown Weights

Item1: $w_1 = ?$, $v_1 = 1$
Item2: $w_2 = ?$, $v_2 = 2$
Item3: $w_3 = ?$, $v_3 = 3$

$Cap = 20$

Optimal solution: $\{x_1 = ?, x_2 = ?, x_3 = ?\}$

Estimated weights: $\{\hat{w}_1 = \hat{w}_2 = \hat{w}_3 = 5\}$
Estimated optimal solution: **infeasible**
$\{x_1 = x_2 = x_3 = 1\}$
(True weights: $\{w_1 = w_2 = w_3 = 10\}$)

When the total weight of the selected items exceeds the capacity:
- Correction function 1: remove all items
- Correction function 2: remove the items one by one in increasing order of values

Penalty function: removal fee
- problem and application specific

New Loss Function: Post-hoc Regret

Estimated optimal solution → Correction → Corrected optimal solution
induces → Penalty

- Post-hoc regret:

$$PRegret(\hat{\theta}, \theta) = \| \text{True optimal value} - \text{Corrected optimal value} \| + \text{Penalty}$$

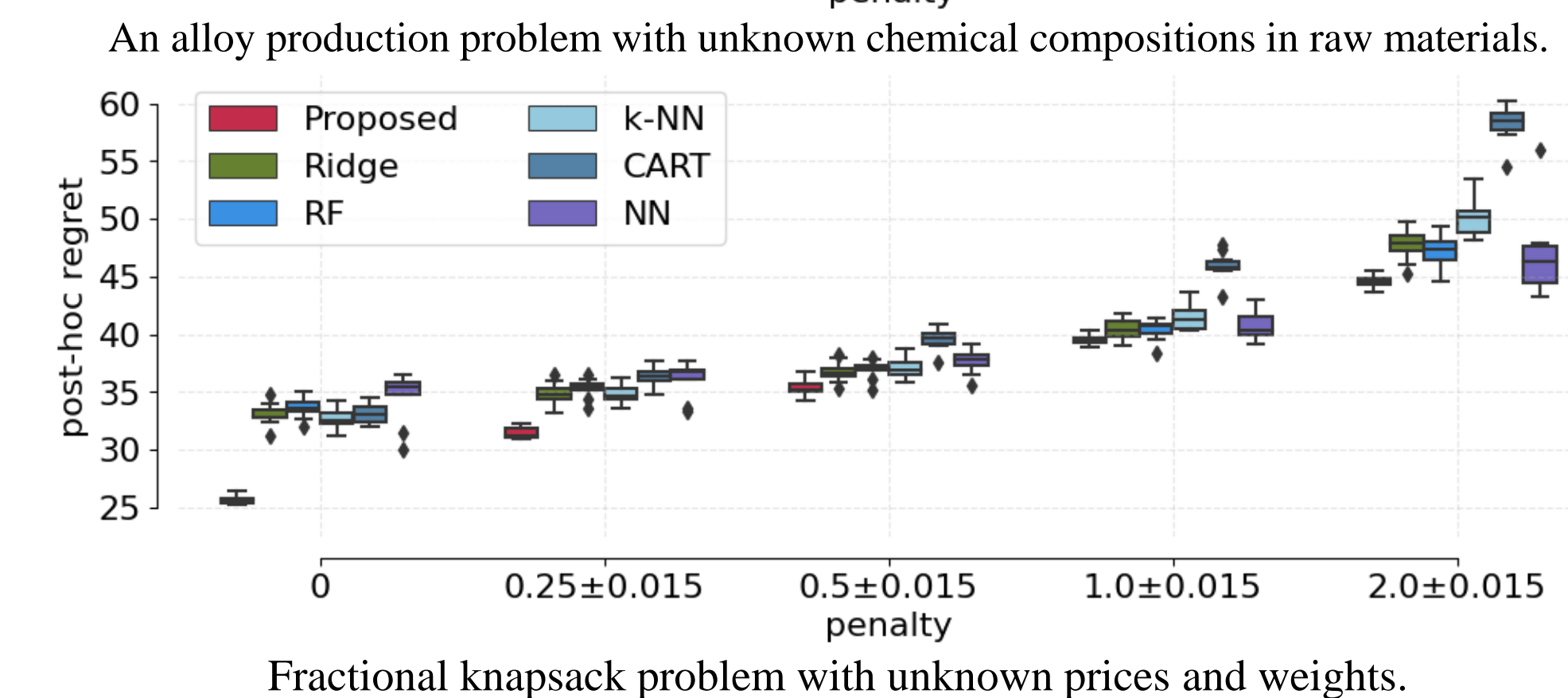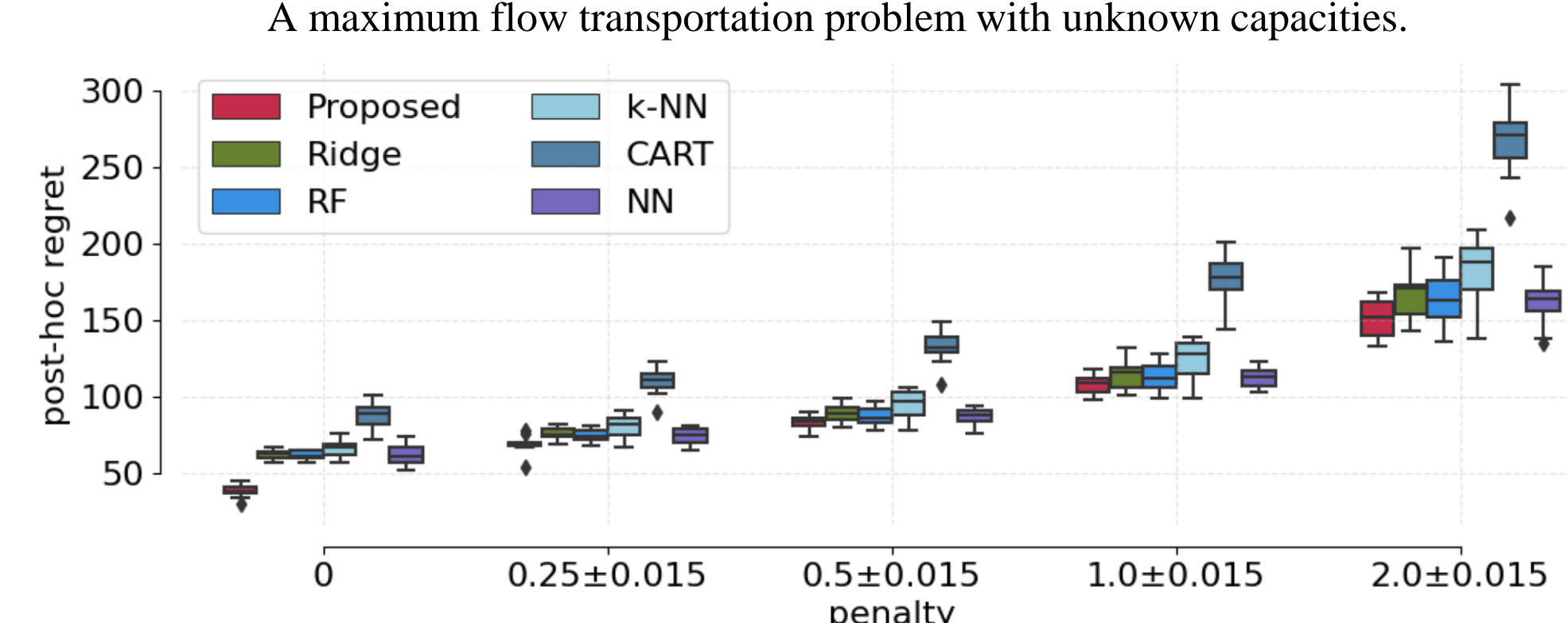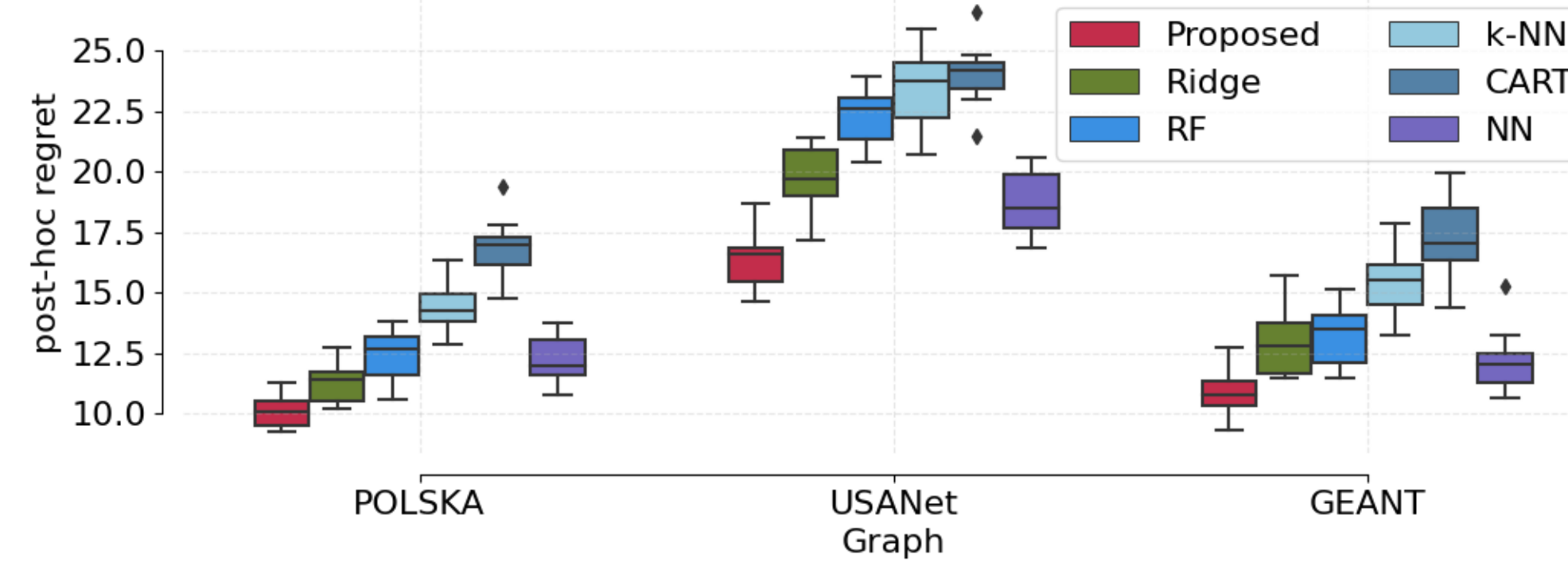## Contribution 2: P+O on Packing/Covering LPs

Consider a packing LP in the standard form:
$$x^* = \arg\max_x c^T x \quad s.t. \ Gx \le h, x \ge 0$$

Penalty

Estimated optimal solution
Feasible region
Correction function
Corrected optimal solution

Selected Experimental Results

Solution Quality:
- The proposed method achieves the best performance over classical approaches.

Legend: Proposed, Ridge, RF, k-NN, CART, NN

Graphs: POLSKA, USANet, GEANT (post-hoc regret)

A maximum flow transportation problem with unknown capacities.

An alloy production problem with unknown chemical compositions in raw materials.

Fractional knapsack problem with unknown prices and weights.

Conclusion

- The **first** Predict+Optimize framework for unknown constraints in general
- A specific training method for covering/packing LPs as a proof of concept