Boosting Power System Operation Economics via Closed-Loop Predict-and-Optimize

Xianbang Chen, Ph.D. Candidate, xchen130@stevens.edu
Lei Wu, Anson Wood Burchard Chair Professor, lei.wu@stevens.edu

Stevens Institute of Technology

June 28, 2023

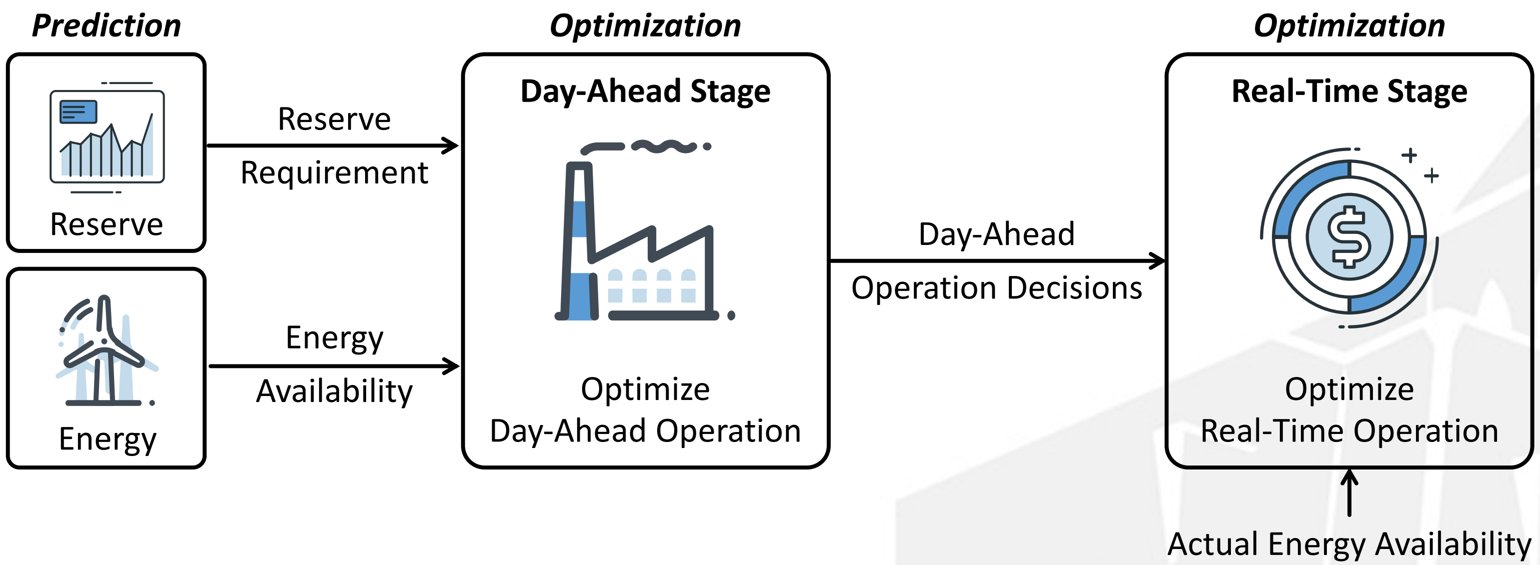
For the 2023 FERC Software Conference

Content

Boosting
Power System
Operation Economics
via
Closed-Loop
Predict-and-Optimize
(C-PO)

Background: Power System Operations in O-PO

- ***Operations in Open-Loop Predict-then-Optimize (O-PO) Framework***



- ***Operators' Goal: Minimum Operation Cost***

Background: Power System Operation Models

- ***Day-Ahead Unit Commitment (Mixed-Integer Programming)***

$$z(\hat{\boldsymbol{w}}) = \min_{\boldsymbol{x}} \boldsymbol{c}^\top \boldsymbol{x}$$
$$s.t.\ \boldsymbol{x} \in \mathcal{X}(\hat{\boldsymbol{w}}, \hat{\boldsymbol{r}})$$
$$\boldsymbol{x} \in \{0,1\}^m \times \mathbb{R}^n$$

Decision $\boldsymbol{x}$: On/off state, base-point generation, reserve deployment of units.

$\hat{\boldsymbol{w}}$ and $\hat{\boldsymbol{r}}$ are vectors of renewable and reserve predictions.

- ***Real-Time Economic Dispatch (Mixed-Integer Programming)***

$$z^{ed}(\boldsymbol{x}^*, \tilde{\boldsymbol{w}}) = \min_{\boldsymbol{y}} \boldsymbol{d}^\top \boldsymbol{y}$$
$$s.t.\ \boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x}^*, \tilde{\boldsymbol{w}})$$
$$\boldsymbol{y} \in \{0,1\}^p \times \mathbb{R}^q$$

Decision $\boldsymbol{y}$: On/off state of quick-start units, generation adjustment, reserve usage.

$\tilde{\boldsymbol{w}}$ is vector of actual energy availability.

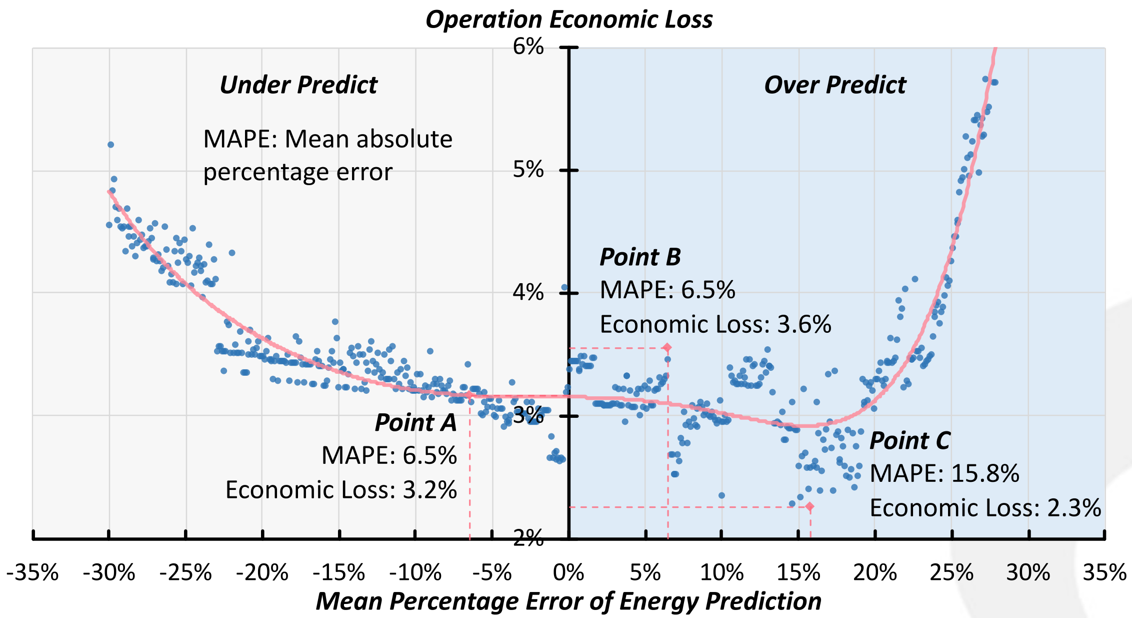
- ***Evaluate Operation Economics***

After solving the problems, the actual operation cost is available. A lower operation cost indicates a better operation economics.

Motivations: Flaws in Open-Loop Predict-then-Optimize (O-PO)

- ***Reserve Requirement is Redundant***
- ***Statistically More Accurate Prediction $\nRightarrow$ Better Operation Economics***

Samples with different prediction errors on an IEEE 118-bus system



Point A vs Point B

The **same prediction error (MAPE = 6.5%)** enable **different economic losses (3.2% vs 3.6%)**.

Point B vs Point C

A **worse prediction error (MAPE = 15.8%)** enables a **lower economic loss (2.3%)**.

Why?

Power systems are complex and nonlinear.

$$Economic\ Loss = \frac{Operation\ Cost^{With\ Error} - Operation\ Cost^{Error-Free}}{Operation\ Cost^{Error-Free}} \times 100\%$$

Motivations: Flaws in Open-Loop Predict-then-Optimize

*"In many real-world applications, the **ultimate goal** is not to make good predictions, but rather to use the often noisy predictions to **make good decisions**."*

---- Yoshua Bengio

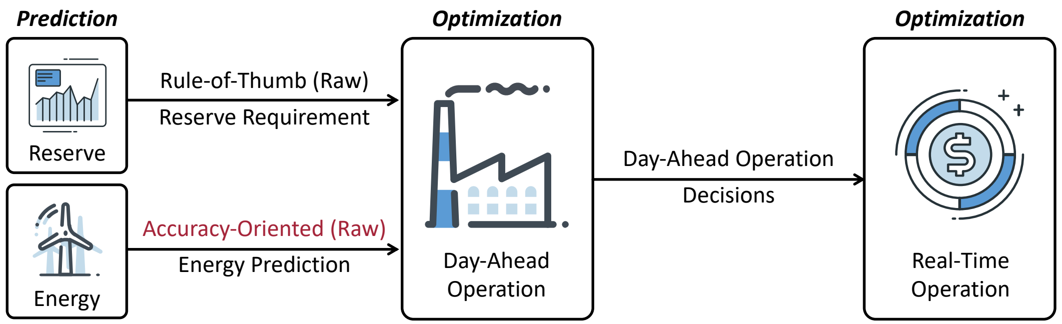
in *Using a Financial Training Criterion Rather than a Prediction Criterion*, 1997

Content

Boosting Power System Operation Economics via Closed-Loop Predict-and-Optimize (C-PO)

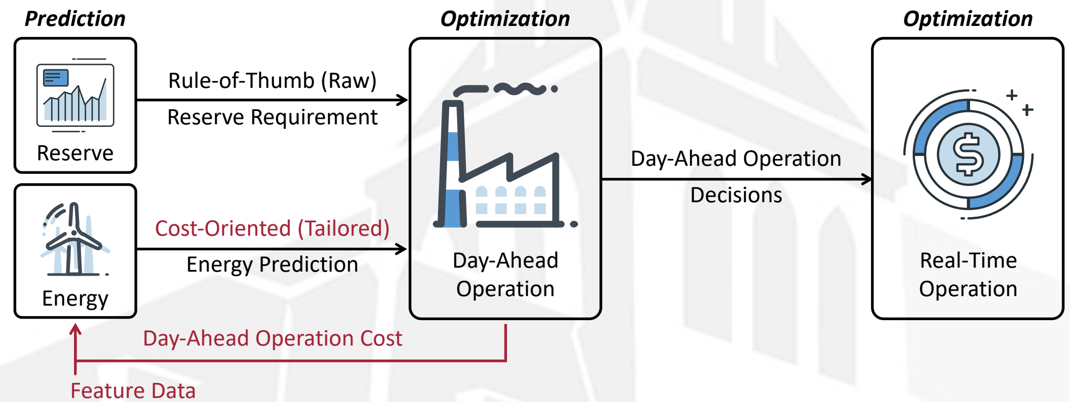
O-PO vs Feature-Driven C-PO

Open-Loop Predict-then-Optimize (O-PO)



- Train predictor with statistical accuracy criterion (Open-loop and accuracy-oriented)
- Sequentially predict and optimize (Predict-then-optimize)

Feature-Driven Closed-Loop Predict-and-Optimize (C-PO)



- Train predictor with operation cost criterion (Closed-loop and cost-oriented)
- Simultaneously predict and optimize (Predict-and-optimize)

Three Modules in Feature-Driven C-PO

- ***Data-Processing Module***
  - Feature selection
  - Training scenario selection
- ***Cost-Oriented Modeling-and-Training Module*** ☜
  - Model a cost-oriented empirical risk minimization problem
  - Solve the problem to get a well-trained predictor
- ***Closed-Loop Predict-and-Optimize Module***
  - Embed the predictor into the unit commitment model
  - Use the new model to predict and optimize simultaneously

Cost-Oriented Modeling-and-Training Module

- ***Smart "Predict-then-Optimize" (SPO) Loss Function for Training***

  $\ell oss^{SPO}(\widehat{\boldsymbol{w}},\widetilde{\boldsymbol{w}}) \coloneqq |z^\star(\widehat{\boldsymbol{w}}) - z^\star(\widetilde{\boldsymbol{w}})|$ Gap between imperfect UC cost $z^\star(\widehat{\boldsymbol{w}})$ (induced by prediction with error $\widehat{\boldsymbol{w}}$) and perfect UC cost $z^\star(\widetilde{\boldsymbol{w}})$ (induced by error-free prediction $\widetilde{\boldsymbol{w}}$).

- ***Cost-Oriented Empirical Risk Minimization (ERM) Problem based on SPO***

  $\text{UC} \coloneqq \min_{\boldsymbol{x}} \boldsymbol{c}^\top \boldsymbol{x}$ $\quad$ $\text{ERM} \coloneqq \min_{\boldsymbol{x},\boldsymbol{H}} \frac{1}{k}\sum_{k=1}^{K} \ell oss_k^{SPO}(\widehat{\boldsymbol{w}}_k,\widetilde{\boldsymbol{w}}_k)$ Minimize in-sample SPO loss

  $s.t.\ \boldsymbol{x} \in \mathcal{X}(\widehat{\boldsymbol{w}},\widehat{\boldsymbol{r}})$ $\Rightarrow$ $s.t.\ \boldsymbol{x}_k \in \mathcal{X}(\boldsymbol{H}\boldsymbol{f}_k,\widehat{\boldsymbol{r}}); k = 1,\dots,K.$ $\quad$ $\boldsymbol{f}_k$ is the feature vector of sceanrio k

  $\Downarrow$

  Essence: A linear regression problem.

  Intractability: All training scenarios are coupled by predictor matrix $\boldsymbol{H}$.

- ***Handle the Intractability***

  Lagrangian decomposition to enable parallel computation.

- ***After Solving the Empirical Risk Minimization Problem***

  Get a trained cost-oriented predictor $\boldsymbol{H}^\star$ tailored for operations, then embed it into UC to form a prescriptive UC.

Original Unit Commitment vs Prescriptive Unit Commitment

Original Unit Commitment Model

$$
\begin{aligned}
z(\hat{\boldsymbol{w}}) = \min_{\boldsymbol{x}} \ & \boldsymbol{c}^\top \boldsymbol{x} \\
s.t. \quad & \boldsymbol{x} \in \mathcal{X}(\hat{\boldsymbol{w}}, \hat{\boldsymbol{r}}) \\
& \boldsymbol{x} \in \{0,1\}^m \times \mathbb{R}^n
\end{aligned}
$$

- Predict-then-optimize
- Accuracy-oriented prediction

Prescriptive Unit Commitment Model

$$
\begin{aligned}
z(\boldsymbol{f}) = \min_{\boldsymbol{x}, \hat{\boldsymbol{w}}} \ & \boldsymbol{c}^\top \boldsymbol{x} \\
s.t. \quad & \boldsymbol{x} \in \mathcal{X}(\boldsymbol{H}^\star \boldsymbol{f}, \hat{\boldsymbol{r}}) \\
& \boldsymbol{x} \in \{0,1\}^m \times \mathbb{R}^n
\end{aligned}
$$

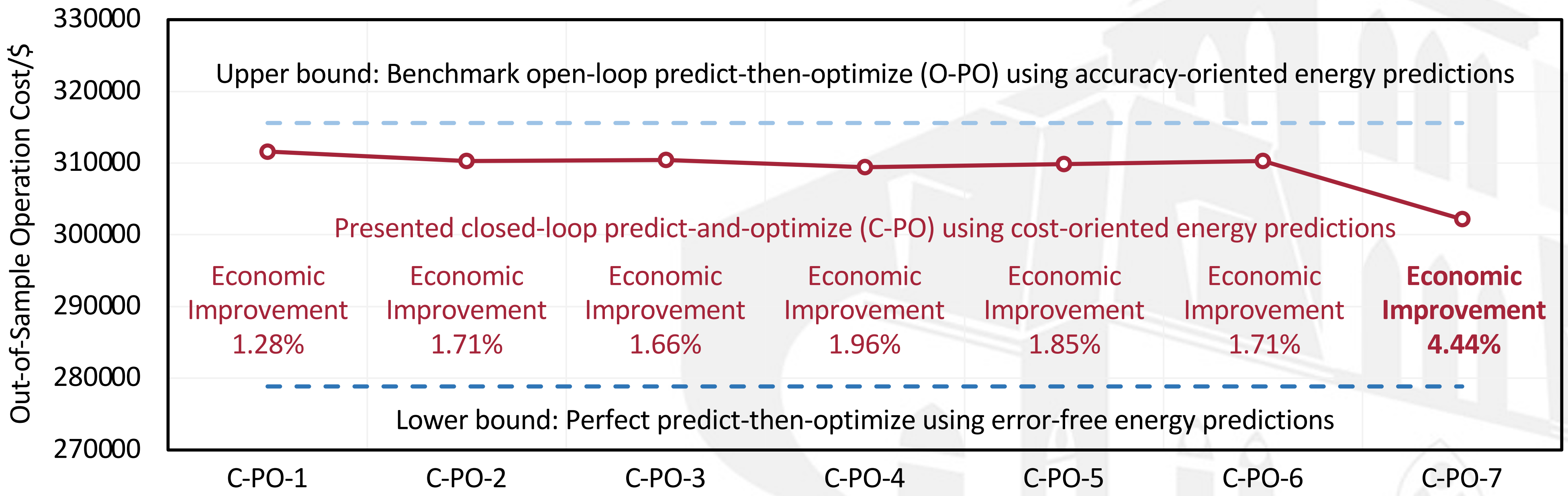
- Predict-and-optimize (i.e., prescription)
- Cost-oriented prediction driven by feature $\boldsymbol{f}$

Why is prescription?

Description	Prediction	Prescription
What does the data look like?	What will the data look like?	What decisions should be made from the data?

Case on Median-Size System: Operation Economics

- ***Benchmark: Predictions (Accuracy-Oriented) from a Belgian ISO***
- ***One-Year-Long Dataset for Out-of-Sample Testing (01/01/2020 to 12/31/2020)***
- ***7 C-POs with Different Configurations***



Case on Median-Size System: Prediction Accuracy

- ***Interesting Observations on Cost-Oriented Predictions***
  - Less accurate from a statistical perspective.
  - Lower MOPE and higher MUPE, i.e., more conservative.
- ***Why Conservative?***
  Predictors learn that an over-predicted energy availability may trigger expensive quick-start units, so they prefer to be under-predicted.

	Mean Absolute Error (MAE)/MW	Root Mean Square Error (RMSE)/MW	Mean Over-Prediction Percentage Error (MOPE)	Mean Under-Prediction Percentage Error (MUPE)
Accuracy-Oriented Prediction	104	130	34%	6%
Cost-Oriented Prediction	123	149	21%	12%

Case on Large-Size 5655-Bus System

- ***Does the Lagrangian Algorithm Work?***
  Enables training acceleration without optimality loss, showing potential for large-size system.

Average Training Time/second		**Average Optimality Gap**	
Solving Empirical Risk Minimization Directly	Solving Empirical Risk Minimization with Lagrangian Algorithm	Solving Empirical Risk Minimization Directly	Solving Empirical Risk Minimization with Lagrangian Algorithm
1604.2	930.2	0.65%	0.72%

42% faster ***High-quality solution***

Thinking

Feature-driven closed-loop predict-and-optimize performs well!

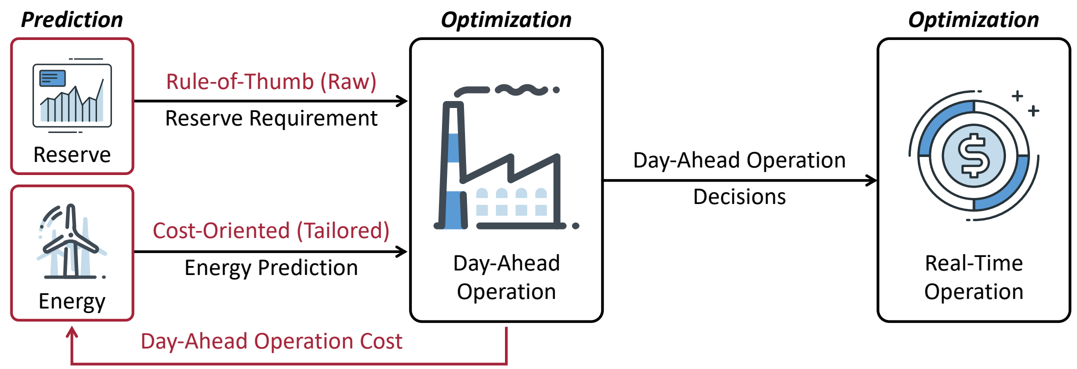
Can we make it better?

Content

Boosting
Power System
Operation Economics
via
Closed-Loop
Predict-and-Optimize
(C-PO)

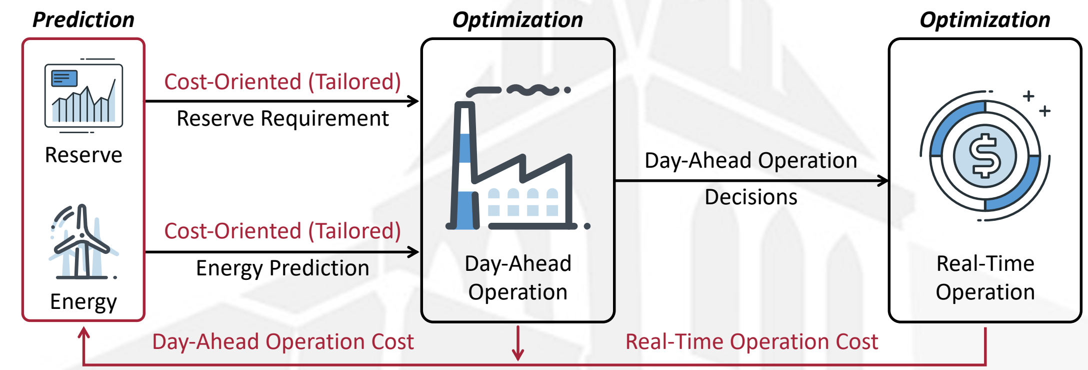
Feature-Driven C-PO vs Bilevel C-PO

Feature-Driven
Closed-Loop Predict-and-Optimize (C-PO)



- Feed the day-ahead cost back
- Tailor energy prediction
- Single-level empirical risk minimization (ERM)

Bilevel
Closed-Loop Predict-and-Optimize



- Feed the day-ahead and real-time costs back
- Jointly tailor energy prediction and reserve requirements
- Bilevel empirical risk minimization (ERM)

ERM Problem with Bilevel Mixed-Integer Programming Form

Previous Single-Level ERM

$$\min_{\boldsymbol{x},\boldsymbol{H}} \frac{1}{k}\sum_{k=1}^{K} \ell oss_k^{SPO}(\widehat{\boldsymbol{w}}_k, \widetilde{\boldsymbol{w}}_k)$$
$$s.t.\ \widehat{\boldsymbol{w}}_k = \boldsymbol{H}\boldsymbol{f}_k;\ \forall k$$
$$\boldsymbol{x}_k \in \mathcal{X}(\widehat{\boldsymbol{w}}_k, \widehat{\boldsymbol{r}});\ \forall k$$

- Training and in-sample operations have to share the same objective.
- In-sample decisions $\boldsymbol{x}$ may NOT follow the least-cost principle, i.e., overfitting.

Upgraded Bilevel ERM

Upper Level (Predictor Training)

$$\min_{\boldsymbol{H},\boldsymbol{R}} \frac{1}{K}\sum_{k=1}^{K}(\boldsymbol{b}^\top \boldsymbol{x}_k + \boldsymbol{d}^\top \boldsymbol{y}_k)$$
$$\widehat{\boldsymbol{w}}_k = \boldsymbol{H}\boldsymbol{f}_k,\quad \widehat{\boldsymbol{r}}_k = \boldsymbol{R}\boldsymbol{f}_k;\ \forall k$$

$\widehat{\boldsymbol{w}}_k$

$\widehat{\boldsymbol{r}}_k$

Lower Level (Day-Ahead Unit Commitment)

$$\boldsymbol{x}_k \in \underset{\boldsymbol{x}_k \in \mathcal{X}(\widehat{\boldsymbol{w}}_k, \widehat{\boldsymbol{r}}_k)}{\operatorname{argmin}}\ \boldsymbol{c}^\top \boldsymbol{x}_k;\ \forall k$$

$\boldsymbol{x}_k$

$\boldsymbol{x}_k$

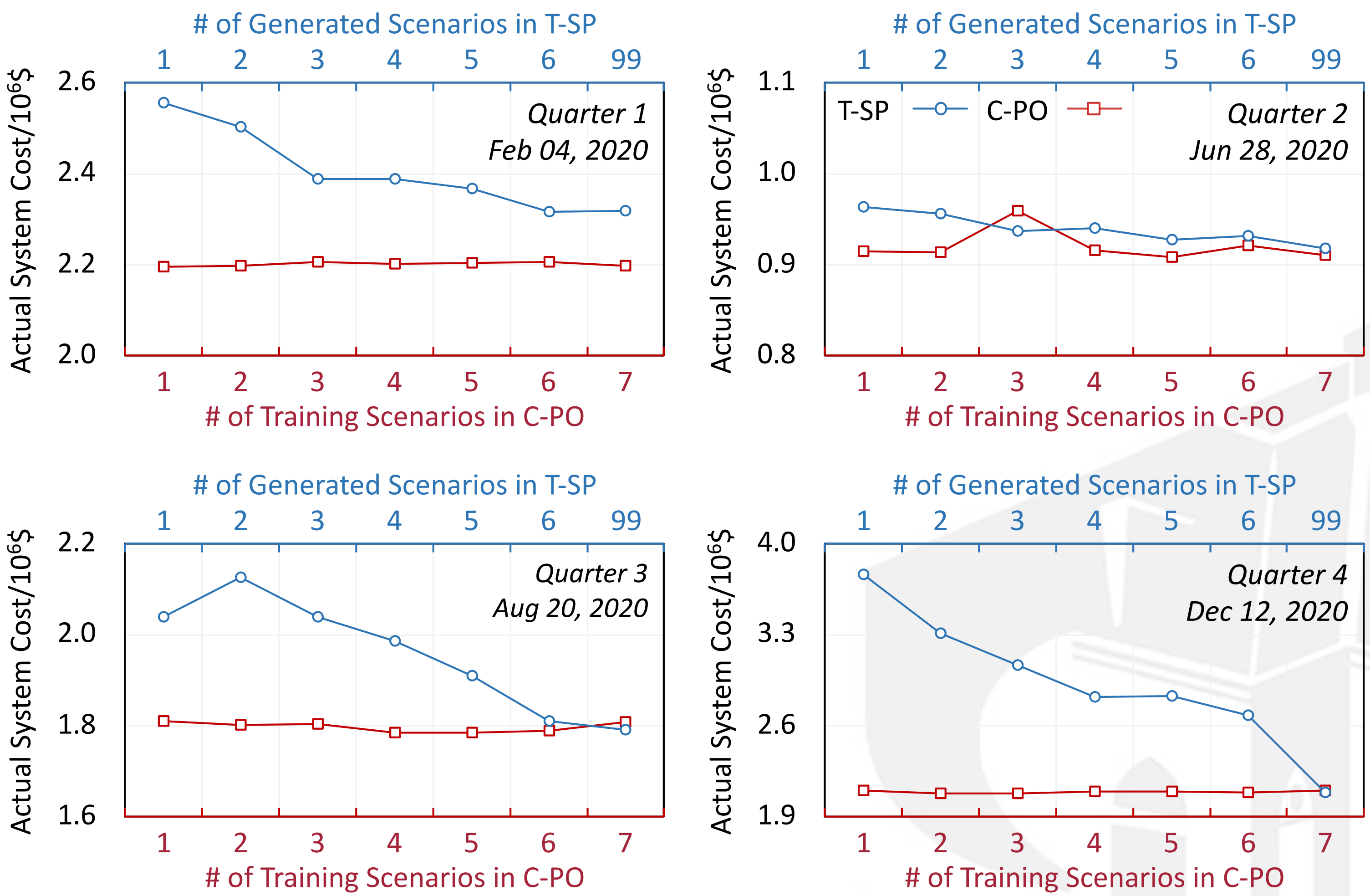
Lower Level (Real-Time Economic Dispatch)

$$\boldsymbol{y}_k \in \underset{\boldsymbol{y}_k \in \mathcal{Y}(\boldsymbol{x}_k, \widetilde{\boldsymbol{w}}_k)}{\operatorname{argmin}}\ \boldsymbol{d}^\top \boldsymbol{y}_k;\ \forall k$$

$\boldsymbol{y}_k$

- The training and operations use their own objectives.
- More robust when scenarios are limited.

Case Study: C-PO vs Two-Stage Stochastic Programming (T-SP)



Observations

- C-PO is **economically competitive**
- C-PO needs **fewer scenarios**
- C-PO is **more stable**

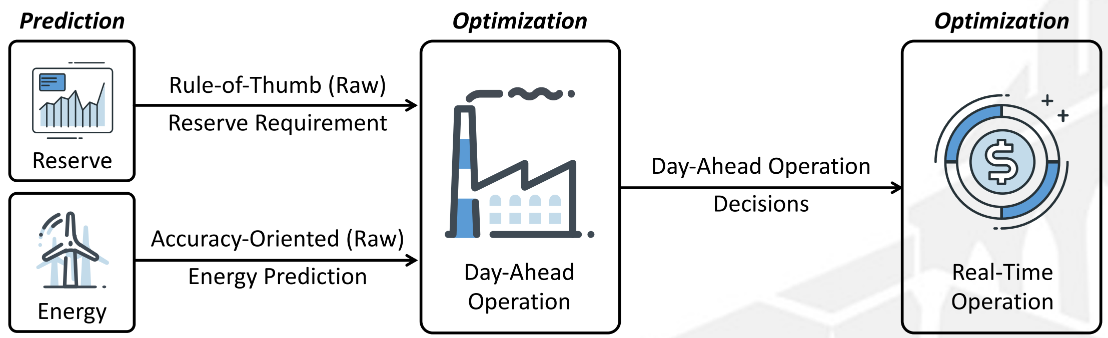
Content

Boosting
Power System
Operation Economics
via
Closed-Loop
Predict-and-Optimize
(C-PO)

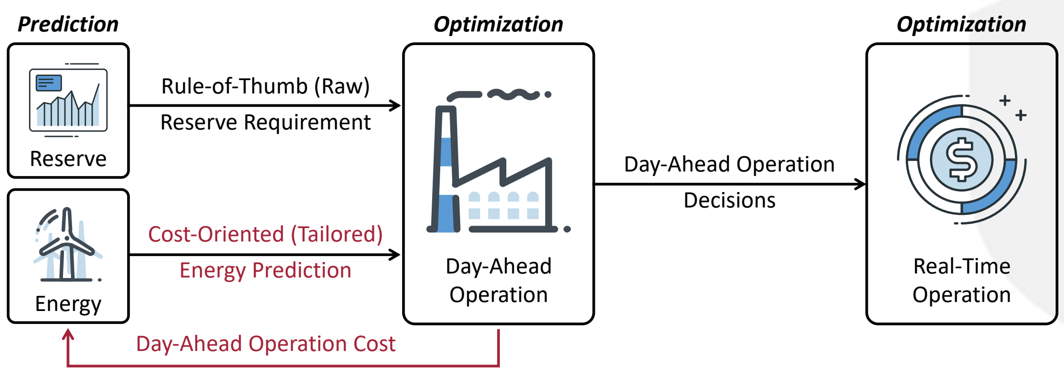
Summary

- ***Two Approaches Towards Improving Operation Economics***

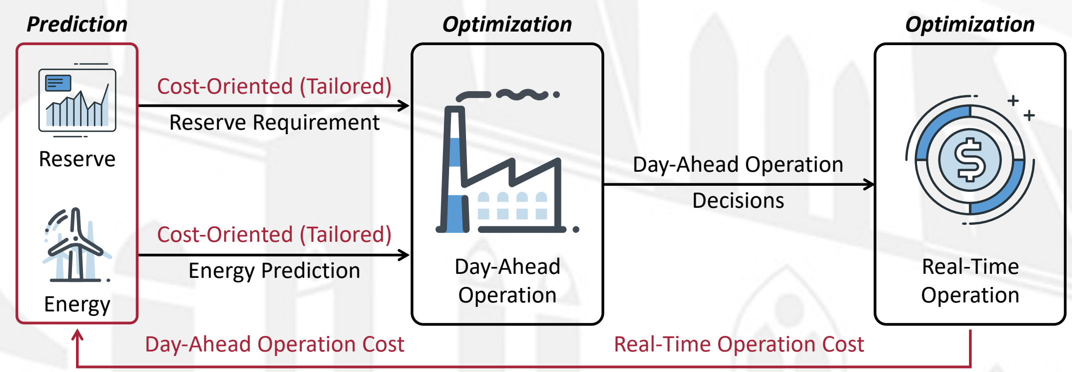
Traditional Open-Loop Predict-then-Optimize



Feature-Driven Closed-Loop Predict-and-Optimize



Bilevel Closed-Loop Predict-and-Optimize



Major Conclusions

- ***Conclusions***
  - The C-PO shows potential in **improving the system operation economics** by generating **cost-oriented predictions tailored for system operations**.
  - A cost-oriented energy prediction **may be slightly worse in accuracy**.
  - The bilevel empirical risk minimization problem enables the C-PO to lower down the operation cost using limited scenarios.

Summary

- ***Back to Essence: Just an Idea***
  Closed-loop predict-and-optimize is basically **an idea that feeds the optimization back to the prediction** for improving the optimization performance.

- ***Potential Tools to Realize This Idea***
  Deep learning, reinforcement learning, and so on.

- ***References***

1. Adam N. Elmachtoub and Paul Grigas, "Smart "Predict, then Optimize"," *Management Science*, 2022.

2. Xianbang Chen, Yafei Yang, Yikui Liu, Lei Wu, "Feature-Driven Economic Improvement for Network-Constrained Unit Commitment: A Closed-Loop Predict-and-Optimize Framework," *IEEE Transactions on Power Systems*, 2022.

3. Xianbang Chen, Yikui Liu, Lei Wu, "Towards Improving Operation Economics: A Bilevel MIP-Based Closed-Loop Predict-and-Optimize Framework for Prescribing Unit Commitment," *arXiv:2208.13065*, 2023.

Thank you!

For 2023 FERC Software Conference

xchen130@stevens.edu
Lei.wu@stevens.edu