A Day Ahead Market Bid-Aware Joint Forecasting Method for Photovoltaic-Energy Storage System

Chenghan Li, Xuan Wei, Yinliang Xu
Tsinghua Shenzhen International Graduate School
Tsinghua University
Shenzhen, China

Jinfeng Wang
Electric Power Research Institute
Guangdong Power Grid Co., Ltd
Guangzhou, China

Jeremy Watson
Department of Electrical and Computer Engineering
University of Canterbury
Canterbury, New Zealand

***Abstract*—With the increasing participation of photovoltaic-energy storage system (PV-ESS) power stations in the electricity market trading, accurate forecasting of both PV generation and electricity prices has become critically important. Moreover, because PV-ESS stations need to make a day-ahead market bid plan based on day-ahead forecasted values, a discrepancy inevitably arises between prediction and bid-making. To address this issue, we propose a two-stage bid-aware joint forecasting method based on Partially Input Convex Neural Networks (PICNN). Specifically, we adopt a two-stage training framework: firstly, the forecasting model is pre-trained to capture the fluctuation characteristics of PV generation and electricity prices, and then PICNN is employed as a surrogate loss function for decision-focused fine-tuning to enhance the bidding revenue of PV-ESS. We construct a PV-ESS bidding model using real-world datasets to validate the effectiveness of the proposed method. Experimental results demonstrate that the proposed method improves profit by 2.5% compared to baseline approaches, thereby validating its effectiveness.**

***Index Terms*—PV-ESS, Joint-Forecasting, PICNN, Decision-Focused**

I. Introduction

As the penetration of PV generation in power systems continues to rise, the inherent randomness and volatility of PV output pose challenges to renewable energy integration and the safe, stable operation of the grid. To enhance PV curtailment mitigation and reduce the impact of its variability, PV stations are now increasingly required to be equipped with energy storage systems, forming integrated solar-plus-storage facilities that can effectively participate in electricity markets [1] [2] [3].

Bidding in electricity markets can typically be modeled as an optimization problem. Past studies have proposed various optimization-based bidding strategies [4] [5] [6] [7] for applications such as wind farms, electric vehicle aggregators, and storage operators. Notably, Xu et al. [8] developed a stochastic programming approach to optimize the bidding power of overseas photovoltaic plants for maximum profit. However, existing optimization-based bidding methods typically approximate prediction uncertainty by simply adding gaussian noise, without adequately considering the true error characteristics and dynamic evolution patterns inherent in forecasting models. This may lead to inconsistencies between predictions and optimization bids, resulting in reduced revenue.

To address the inconsistency between prediction and bid optimization, existing research has proposed Decision-Focused Learning(DFL) [9] [10], which aims to unify prediction models with optimization decision-making processes. The mainstream approaches can be divided into two categories: the first embeds optimization problems as a differentiable optimization layer that directly participates in model training. Sang et al. [11] constructed a second-order cone programming (SOCP) layer, with a weighted sum of power loss and voltage violation as the optimization objective and propagated the decision error backward to the prediction module through chain differentiation. Amos et al. [12] combined implicit differentiation and a bilevel optimization framework to encapsulate a quadratic programming (QP) problem into a differentiable optimization layer for end-to-end gradient optimization. However, these approaches face the challenge that many optimization problems are non-differentiable and non-convex, preventing them from directly participating in prediction model training. The second approach involves finding surrogate loss functions to approximate the optimization process, making it suitable for training prediction models. Shah et al. [13] proposed a surrogate loss function for sample weights, which can model general optimization problems. However, it requires assigning a weight to each sample, leading to high computational costs. Zharmagambetov et al. [14] adopted local modeling as a replacement for global modeling, which reduces computational costs, but the problem of selecting local initialization parameters remains.

As discussed above, although significant progress has been achieved in DFL, two critical challenges remain unresolved. First, there is a shortage of efficient surrogate function modeling approaches, as existing methods often depend on effective initialization or involve high computational costs. Second, the interrelation between decision optimality and prediction feasibility has not been sufficiently addressed—specifically, the optimal decision point may fall outside the actual feasible region. Based on the above, this paper investigates a bid-aware joint forecasting problem for PV generation and DA electricity prices in PV-ESS systems within the DA market. First, a bid surrogate loss function is modeled using PICNN. Subsequently, a two-stage training method is employed: the forecasting model is initially trained with mean squared error(MSE) loss, followed by fine-tuning using the PICNN-

Corresponding author:xu.yinliang@sz.tsinghua.edu.cn

979-8-3315-8138-1

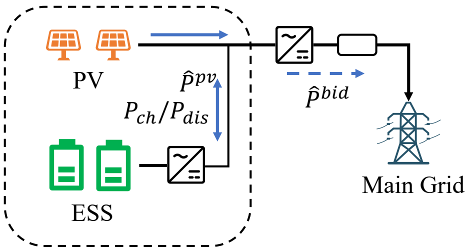

Fig. 1: Illustration of PV-ESS Bid

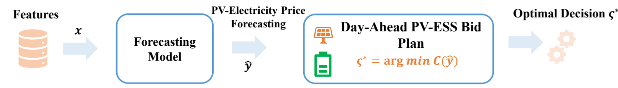

Fig. 2: Block Diagram of the Sequential Prediction-and-Optimization Framework

based surrogate loss. The detailed contributions of this paper are as follows:

1) We propose a novel approach to constructing surrogate loss functions based on the PICNN model. In contrast to previous studies [13] [12], the proposed method is not restricted by the structure of the downstream optimization problem and the requirement for precise initialization. In contrast to prior work that relies on predefined functional forms or requires precise problem-specific initialization, our PICNN-based surrogate loss is learned end-to-end from data. This allows it to capture complex, non-linear mappings between forecast errors and downstream regret without being constrained by the structure of the underlying optimization problem
2) We propose a two-stage learning framework. In the first stage, the forecasting model is trained to ensure that the predictions remain within physically feasible bounds. In the second stage, the model is fine-tuned using PICNN as a surrogate loss function to maximize decision-making revenue while maintaining prediction feasibility. Based on a case study constructed from real-world datasets, the proposed framework achieves a 2.5% increase in revenue compared with baselines, demonstrating its effectiveness.

The remainder of this paper is organized as follows. Section II introduces the PV–ESS bidding model and the formulation of the DFL framework. Section III describes the proposed methodology, including the construction of the surrogate loss function and two-stage training way. Section IV presents a case study to validate the effectiveness of the proposed approach. Finally, Section V concludes the paper.

II. Problem Formulation

A. Photovoltaic-ESS Bid

This paper considers a PV-ESS comprising PV arrays and ESS (see in Figure 1), which aims to maximize revenue by optimizing its DA market bidding strategy and coordinating energy storage operations. The station is assumed to be a price taker—that is, it has no influence on market prices. For simplicity, only participation in the DA electricity market is considered. The station must determine its bidding strategy based on DA electricity price forecasts and predicted PV power generation, as formulated in the following equation:

$$\max_{\boldsymbol{\xi}} R = \sum_{t=1}^{T} \left(\hat{\lambda}_t^{price} P_t^{bid} - c^{deg} (P_t^{cha} + P_t^{dis}) \right) \tag{1}$$

$$P_t^{bid} = \hat{P}_t^{pv} + P_t^{dis} - P_t^{cha}, \quad \forall t \in \{1, \ldots, T\} \tag{2}$$

$$0 \leq P_t^{cha} \leq u_t^{cha} P_{\max}^{cha}, \quad \forall t \tag{3}$$

$$0 \leq P_t^{dis} \leq u_t^{dis} P_{\max}^{dis}, \quad \forall t \tag{4}$$

$$u_t^{cha} + u_t^{dis} \leq 1, \quad \forall t \tag{5}$$

$$E_{t+1} = E_t + P_t^{cha} \eta_{cha} - \frac{P_t^{dis}}{\eta_{dis}}, \quad \forall t \tag{6}$$

$$E_0 = E_T \tag{7}$$

where $\boldsymbol{\xi} = \{P_t^{cha}, P_t^{dis}, u_t^{cha}, u_t^{dis}, E_t, P_t^{bid}\}_{t=1}^{T}$ denotes the decision variables, including charging/discharging power, their binary states, energy level, and bidding power. η_{cha} and η_{dis} denote charging and discharging efficiencies, and the cyclic constraint ensures $E_T = E_0$. $\hat{\lambda}_t^{price}$ and $\hat{P}_t^{pv}$ are forecasting values for DA price and PV generation.

The above problem can be cast as an optimization problem R, where the optimal operational decision $\boldsymbol{\xi}^*$ is given by

$$\boldsymbol{\xi}^* = \arg\max_{\boldsymbol{\xi} \in \mathcal{U}} R\big(\hat{\lambda}_t^{price}, \hat{P}_t^{pv}\big) \tag{8}$$

where $\mathcal{U}$ denotes the feasible set defined by operational constraints.

As illustrated above, the PV-ESS bidding problem entails a sequential predictive and optimization framework(see Figure 2): first, DA forecasts of electricity market prices and PV generation are generated; subsequently, these forecasts are used to formulate a bidding plan. However, this sequential approach inevitably leads to a misalignment between the prediction and the ultimate decision-making, resulting in inconsistency between the forecasting and optimization stages. To address this inconsistency, a framework is required that explicitly integrates the bid-making objective into the predictive modeling process, thereby aligning forecast generation with downstream bid optimization objectives.

B. Decision Focused Learning

Consider a dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$ consisting of historical DA electricity prices and PV generation data, where x_i denotes the historical inputs and y_i represents the corresponding target values for DA price and PV generation $(\hat{\lambda}_t^{price}, \hat{P}_t^{pv})$. The objective is to learn the optimal model parameters $\boldsymbol{\theta}$

979-8-3315-8138-1

that minimize the expected regret $\mathcal{R}$ of the PV-ESS revenue, formulated as

$$\min_{\boldsymbol{\theta}} \quad \mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\mathcal{R}(\hat{y},y)\right] \tag{9}$$

$$\text{s.t.} \quad \mathcal{R}(\hat{y},y) = R\big(\boldsymbol{\xi}^*(\hat{y})\big) - R\big(\boldsymbol{\xi}^*(y)\big) \tag{10}$$

$$\hat{y} = M_{\boldsymbol{\theta}}(x) \tag{11}$$

where $\mathcal{R}$ denotes the regret to be minimized, $\hat{y}$ is the prediction generated by the forecasting model $M_{\boldsymbol{\theta}}$, and $\boldsymbol{\xi}^*(\cdot)$ represents the optimal decision within a feasible region. Note that the second term $R\big(\boldsymbol{\xi}^*(y)\big)$ in Eq (10) corresponds to the optimal revenue under the true values and thus remains constant with respect to $\boldsymbol{\theta}$. Consequently, the optimization in Eq (9) is equivalent to minimizing $R\big(\boldsymbol{\xi}^*(\hat{y})\big)$.

Proposition 1 (Chain Rule)**.** For a composite differentiable function $\mathcal{L}(\boldsymbol{\theta}) = f(g(h(\boldsymbol{\theta})))$, the gradient with respect to $\boldsymbol{\theta}$ is given by

$$\nabla_{\boldsymbol{\theta}}\mathcal{L} = \frac{\partial f}{\partial g}\cdot\frac{\partial g}{\partial h}\cdot\frac{\partial h}{\partial \boldsymbol{\theta}}. \tag{12}$$

By Proposition 1, the gradient of the regret with respect to the model parameters admits the following decomposition:

$$\frac{\partial \mathcal{R}(\hat{y},y)}{\partial \boldsymbol{\theta}} = \frac{\partial \mathcal{R}(\hat{y},y)}{\partial \hat{y}}\cdot\frac{\partial \hat{y}}{\partial \boldsymbol{\theta}}. \tag{13}$$

However $\frac{\partial \mathcal{R}(\hat{y},y)}{\partial \hat{y}}$ may be undefined when the downstream optimizer is non-differentiable. To address this, we need to find a differentiable surrogate loss function $\mathcal{L}(\hat{y},y)$ that captures the regret direction of the gradient.

$$\frac{\partial \mathcal{R}(\hat{y},y)}{\partial \boldsymbol{\theta}} \approx \frac{\partial \mathcal{L}(\hat{y},y)}{\partial \hat{y}}\cdot\frac{\partial \hat{y}}{\partial \boldsymbol{\theta}}. \tag{14}$$

Although a differentiable surrogate loss enables end-to-end decision-focused optimization, it may produce predictions that fall outside the set of physically or economically feasible outcomes. By prioritizing alignment with decision-relevant gradients, the surrogate loss may yield predictions that violate feasibility constraints.

$$\hat{y}^*(x) = \arg\min_{\hat{y}} \mathcal{L}(\hat{y},y), \quad \hat{y}^*(x) \notin \mathcal{U}. \tag{15}$$

This observation suggests that directly applying end-to-end decision-focused learning may result in an illegal outcome, as the predictions could lie outside the feasible domain. In contrast, a two-stage learning framework ensures both prediction feasibility and decision optimality. Figure 3 provides an illustrative example.

III. Method

Building upon the problem formulation presented in the previous section, we propose a two-stage bid-aware forecasting framework, as illustrated in Figure 4. To simultaneously ensure decision optimality and prediction feasibility, the proposed framework consists of two modules. First, a base forecasting module is trained to produce predictions that lie within the feasible set $\mathcal{U}$. Subsequently, predictions are fine-tuned within $\mathcal{U}$ using a decision-aware surrogate loss to enhance downstream decision performance. Detailed descriptions of each component are provided in the following subsections.

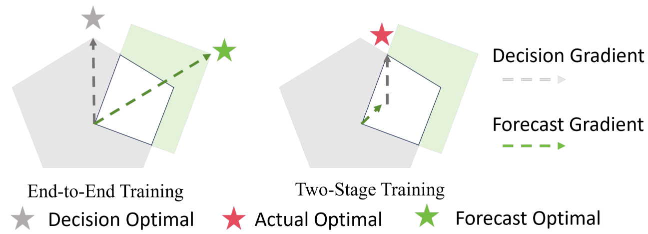

Fig. 3: End-to-End Learning vs Two-Stage Learning

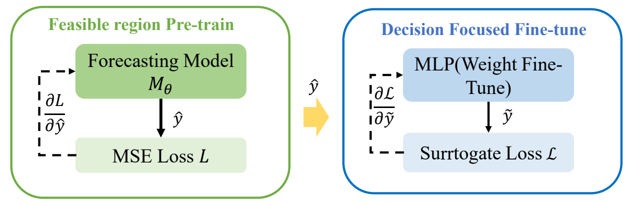

Fig. 4: Structure of Our Framework

A. Partial Input Convex Neural Network Surrogate Function

To construct a robust surrogate model for the mapping $(\hat{y},y) \mapsto \mathcal{R}$, we adopt the PICNN architecture [15], which guarantees convexity with respect to a subset of inputs—specifically, the framework can model flexible nonlinear relationships between the regret value $\mathcal{R}$ and the tuple $(y,\hat{y})$.

PICNN as illustrated in Figure 5, consists of two parallel streams: one processing the prediction $\hat{y}$ and the other handling the ground truth y. The $\hat{y}$-stream produces intermediate representations $u_1, u_2, \ldots, u_{k-1}$ through standard feedforward layers. Simultaneously, the y-stream computes $z_1, z_2, \ldots, z_k$, where each layer z_i receives not only the preceding hidden state z_{i-1} but also all prior outputs from the $\hat{y}$-stream $(u_1, \ldots, u_{i-1})$ as well as the original input y. Formally, the recurrence is defined as:

$$u_0 = \hat{y} \tag{16}$$

$$u_i = \sigma(W_i^u u_{i-1} + b_i^u) \tag{17}$$

$$z_0 = y \tag{18}$$

$$z_i = \sigma\left(W_i^{z,y} z_{i-1} + W_i^{z,u} u_{i-1} + W_i^{z,y} y + b_i^z\right) \tag{19}$$

where $\sigma(\cdot)$ denotes a non-linear activation function (e.g., ReLU or SOFTPLUS), and the final output $f(\hat{y},y) = z_k = \mathcal{R}$ represents the learned scalar-valued surrogate function.

To comprehensively model the relationship among $(y,\hat{y})$, and the regret $\mathcal{R}$, we synthesize a diverse training dataset by applying multi-scale scaling and gaussian noise injection to pairs of $(\hat{y},y)$ and then compute regret $\mathcal{R}$ to training PICNN.

979-8-3315-8138-1

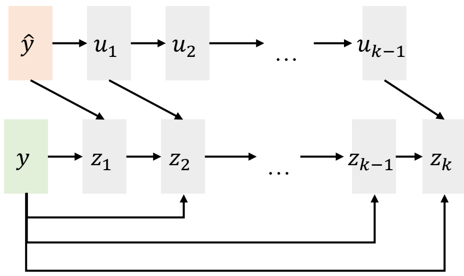

Fig. 5: Structure of PICNN

B. Two-Stage Training

We employ a two-stage training procedure for the bid-aware forecasting framework, as summarized in Algorithm 1. In the first stage, the predictive model M_θ is pre-trained using MSE loss to ensure that its outputs lie within the physically feasible space. In the second stage, a linear layer is appended to the pre-trained model, and fine-tuned using a PICNN surrogate loss $\mathcal{L}$, enabling predictions that directly optimize downstream bid-aware outcomes $\tilde{y}$. This two-stage approach ensures both prediction feasibility and alignment with decision objectives.

Algorithm 1 Two-Stage Training of Bid-Aware Forecasting Framework

Initialize model parameters θ for forecasting model M_θ.
Initialize surrogate network $\mathcal{L}_{PICNN}$ with random weights.
Set learning rate η, batch size B, and total epochs N_1, N_2.
Initialize linear decision layer parameters θ_{decision}.
Stage 1: Pre-training with MSE Loss
for each epoch $e_1 = 1, \ldots, N_1$ **do**
  for each batch $\{x_t, y_t\}_{t=1}^{B}$ **do**
    Generate predictions $\hat{y}_t = M_\theta(x_t)$.
    Compute $\mathcal{L}_{MSE} = \frac{1}{B}\sum_{t=1}^{B} \|\hat{y}_t - y_t\|^2$.
    Update $\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_{MSE}$.
  end for
end for
Freeze the pre-trained encoder layers of M_θ.
Append a linear decision layer with parameters $MLP_{\theta_{\text{decision}}}$ for downstream optimization alignment.
Stage 2: Fine-tuning with Decision-Aware Surrogate Loss
for each epoch $e_2 = 1, \ldots, N_2$ **do**
  for each batch $\{x_t, y_t\}_{t=1}^{B}$ **do**
    Generate predictions $\hat{y}_t = M_\theta(x_t)$.
    Generate decision predictions $\tilde{y} = MLP(\hat{y})$
    Evaluate surrogate regret loss $\mathcal{L}_{PICNN} = PICNN(\tilde{y}_t, y_t)$.
    Update decision layer parameters $\theta_{\text{decision}} \leftarrow \theta_{\text{decision}} - \eta \nabla_{\theta_{\text{decision}}} \mathcal{L}_{PICNN}$.
  end for
end for
return Optimized forecasting model M_θ^* and decision layer parameters $\theta_{\text{decision}}^*$ for PV-ESS bidding.

IV. Case Study

To validate the effectiveness of the proposed method, we utilize real-world datasets of PV generation and DA electricity prices (see Figure 6) [16], and construct a PV–ESS bidding optimization model using Pyomo. CNN and MLP are selected as baseline models, and performance is evaluated using two metrics: revenue ($), defined as the profit obtained when the predicted bidding strategy is settled under real prices, and MSE. All forecasting models are implemented in PyTorch with CUDA acceleration, while the optimization model is solved using Gurobi. The proposed framework is forecasting-model-agnostic, enabling compatibility with various time-series forecasting models. In this study, DLinear [17] is adopted due to its superior predictive performance for DA electricity price forecasting tasks.

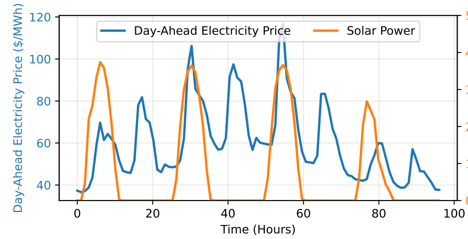

Fig. 6: Visualization of PV Generation and DA Electricity Price

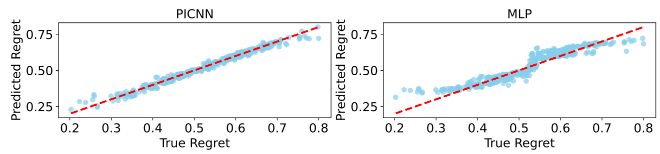

Fig. 7: Comparison of Fitting Results (PICNN vs MLP)

As shown in Table I, PICNN achieves the lowest MSE on the test set, demonstrating its superior capability as a surrogate loss function compared with MLP and CNN. Furthermore, as illustrated in Figure 7, the PICNN model exhibits a tighter correlation between predicted and true regret values, with data points clustering more closely around the diagonal dashed line, indicating higher prediction accuracy and improved generalization relative to the MLP model.

TABLE I: Performance comparison of different surrogate models

Model	PICNN	MLP	CNN
MSE	2e−3	8e−3	5e−3

The proposed bid-aware joint forecasting framework, which integrates a PICNN-based surrogate loss with DLinear feature extraction, demonstrates notable advantages for PV–ESS DA market participation. As shown in Table II, our approach sacrifices forecasting accuracy to achieve its primary objective: a 2.5% increase in revenue ($485.69 vs. $473.65). This indicates a misalignment in the gradient descent directions between prediction and decision-making. The training curves shown in Figure 8 validate the effectiveness of the two-stage strategy: initial DLinear pre-training ensures stable feature

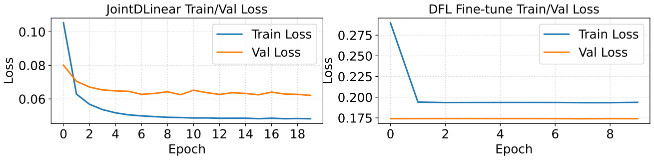

Fig. 8: Two-Stage of Training Loss

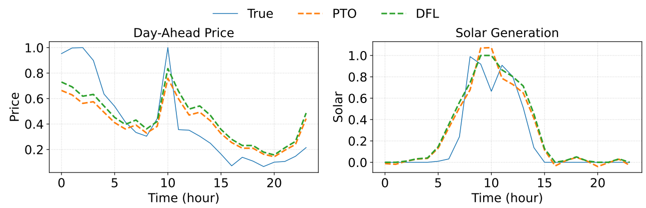

Fig. 9: Comparision of Forecasting Results(PTO vs Our)

extraction, and during the subsequent fine-tuning phase, the PICNN loss rapidly converges to a lower validation error, indicating effective decision-focused learning. This acceleration occurs because only the appended linear decision layer (θ_{decision}) is optimized, while the pre-trained feature extractor (M_θ) remains frozen, drastically reducing the parameter space. Figure 9 shows although our method exhibits slightly lower overall prediction accuracy than PTO, its forecasts better align with the dynamic trends of true values during peak hours of DA electricity price and PV generation, demonstrating that decision-oriented optimization enhances revenue in critical periods and validating the effectiveness of the proposed approach.

TABLE II: Performance comparison between PTO and Our method.

Method	MSE (DA Price)	MSE (PV)	Loss	Day Revenue($)
PTO	2.58e−2	2.75e−2	1.51e−1	473.65
CNN	3.07e−2	2.84e−2	2.05e−2	478.41
MLP	2.92e−2	2.87e−2	4.21e−2	470.20
Our	3.03e−2 (↓17.7%)	2.92e−2 (↓6.2%)	1.79e−2	485.69 (↑2.5%)

V. Conclusion

This paper proposes a novel bid-aware joint forecasting framework for PV–ESS, which explicitly integrates downstream bidding optimization objectives into the forecasting model via a PICNN-based surrogate loss function and a two-stage training strategy, thereby achieving coordination between prediction accuracy and decision optimality. Experimental results demonstrate that, compared to the conventional PTO approach, the proposed method yields a 2.5% improvement in revenue, validating its effectiveness in bridging the coupling gap between forecasting and PV–ESS bidding. Future work will extend this framework to multi-day or real-time market participation scenarios and incorporate more complex operational constraints, such as ramping limits and transmission congestion. Furthermore, we will explore generalizing this decision-oriented learning paradigm to other energy assets—including wind farms and demand response programs—to establish a scalable architecture for optimizing diverse energy trading strategies under uncertainty.

References

[1] Y. Xia and Y. Xu, "Learning-based optimal power bidding of an overseas photovoltaic-energy storage plant in singapore electricity market," *Available at SSRN 5221646*.

[2] Y. Wang, X. Wang, S. Li, X. Ma, Y. Chen, and S. Liu, "Optimization model for harmonic mitigation based on pv-ess collaboration in small distribution systems," *Applied Energy*, vol. 356, p. 122410, 2024.

[3] C. Liu, T. Wu, J. Zhu, R. Wei, Y. Wu, and A. Tian, "A transient reactive power control strategy of pv-ess enhances the system transient stability," *Journal of Energy Storage*, vol. 92, p. 112302, 2024.

[4] D. Yang, S. He, M. Wang, and H. Pandžić, "Bidding strategy for virtual power plant considering the large-scale integrations of electric vehicles," *IEEE Transactions on Industry Applications*, vol. 56, no. 5, pp. 5890–5900, 2020.

[5] S. A. Hosseini, J.-F. Toubeau, Z. De Greve, and F. Vallée, "An advanced day-ahead bidding strategy for wind power producers considering confidence level on the real-time reserve provision," *Applied energy*, vol. 280, p. 115973, 2020.

[6] M. Bruninx, T. Verstraeten, J. Kazempour, and J. Helsen, "Day-ahead bidding strategies for wind farm operators under a one-price balancing scheme," in *Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems*, 2025, pp. 719–726.

[7] S. Li, C. Chen, and L. Tong, "Convexifying regulation market clearing of state-of-charge dependent bid," in *2024 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2024, pp. 1–5.

[8] Y. Xu, "Optimal power bidding of overseas pv plants in singapore wholesale electricity market," *Energy Conversion and Economics*, vol. 5, no. 2, pp. 73–77, 2024.

[9] M. Huang and V. Gupta, "Decision-focused learning with directional gradients," *Advances in Neural Information Processing Systems*, vol. 37, pp. 79 194–79 220, 2024.

[10] G. Chen and J. Qin, "On the choice of loss function in learning-based optimal power flow," in *2024 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2024, pp. 1–5.

[11] L. Sang, Y. Xu, H. Long, and W. Wu, "Safety-aware semi-end-to-end coordinated decision model for voltage regulation in active distribution network," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 1814–1826, 2022.

[12] B. Amos and J. Z. Kolter, "Optnet: Differentiable optimization as a layer in neural networks," in *International conference on machine learning*. PMLR, 2017, pp. 136–145.

[13] S. Shah, K. Wang, B. Wilder, A. Perrault, and M. Tambe, "Decision-focused learning without decision-making: Learning locally optimized decision losses," *Advances in Neural Information Processing Systems*, vol. 35, pp. 1320–1332, 2022.

[14] A. Zharmagambetov, B. Amos, A. Ferber, T. Huang, B. Dilkina, and Y. Tian, "Landscape surrogate: Learning decision losses for mathematical optimization under partial information," *Advances in Neural Information Processing Systems*, vol. 36, pp. 27 332–27 350, 2023.

[15] H. Jeon, H. Bae, M. Park, C. Kim, and W. C. Kim, "Locally convex global loss network for decision-focused learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 25, 2025, pp. 26 805–26 812.

[16] Curyyfaust, "Pv-ess-dfl," https://github.com/Curyyfaust/PV-ESS-DFL, 2025, accessed: 2025-11-08.

[17] A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are transformers effective for time series forecasting?" in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 9, 2023, pp. 11 121–11 128.

979-8-3315-8138-1