Data-Driven Dimension Reduction for Industrial Load Modeling Using Inverse Optimization

Ruike Lyu, Hongye Guo, Goran Strbac, Chongqing Kang, *Fellow, IEEE*

***Abstract*—Harnessing the flexibility of industrial loads is crucial for maintaining a cost-effective supply-demand balance of the power system. However, the intricate mixed-integer constraints in industrial load models not only pose challenges for their direct integration into economic dispatch or market clearing processes but also render current analytical dimension-reduction methods ineffective. Unlike analytical methods based on original constraint forms and parameters, we propose a novel data-driven dimension-reduction approach for industrial load modeling, which uses the optimal energy usage data from industrial loads to train a dimension-reduced model that best fits the original constraints. Our approach, implemented by the adjustable load fleet model parameterized through inverse optimization, outperformed analytical methods across three industrial load datasets.**

***Index Terms*—industrial load, dimension reduction, inverse optimization, data-driven, load modeling**

I. Introduction

To harness industrial users' flexibility through grid-load interactions while mitigating adverse effects, it is crucial to mathematically model the constraints of their production processes [1]. The complex physics of industrial processes poses a key challenge, requiring integer variables or nonlinear functions for representation [2]. Therefore, precise industrial load models encounter obstacles in integrating into economic dispatch or market clearing. To address this issue, we explore dimension reduction techniques to approximate the intricate mixed-integer constraints of industrial loads with simplified lower-dimensional constraints.

Existing dimension reduction studies rely on analytical methods, which exhibit limited adaptability and compatibility with integer variables [3]. Mathematically, dimension reduction involves a projection process, which can be computationally challenging with high-dimensional temporally coupled constraints Therefore, approximations become necessary. However, current template-based approximation methods lack adaptability across various industrial loads [4]. Moreover, current analytical methods require convexity in the original constraints, which is incompatible with industrial load constraints involving integer variables [5].

In this paper, we propose a data-driven dimension reduction (D3R) framework that leverages optimal energy consumption data of industrial loads on the basis of the original constraints to train a reduced-constraint model through inverse optimization (Fig. 1). Compared with analytical methods, D3R exhibits enhanced adaptability, as it is not constrained by the specific form of the original constraints, making it compatible with constraints involving integer variables.

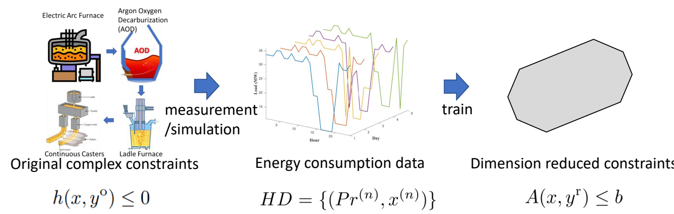

Fig. 1. Proposed data-driven dimension reduction framework for industrial load modeling, illustrated with a steel plant example.

II. Problem Description

Without loss of generality, our objective is to replace the original constraints of a given industrial load (original constraints, OCs), denoted as $h(x, y^{\rm o}) \le 0$, with a set of dimension-reduced linear constraints (reduced constraints, RCs) $A(x, y^{\rm r}) \le b$. Here, $h(\cdot)$ is the original constraint function, $x \in \mathbb{R}^T$ represents the hourly energy consumption (continuous), T is the number of time intervals, $y^{\rm o}$ represents other variables of the OCs (continuous or integer), $y^{\rm r}$ represents other variables of the RCs (continuous), and A and b are the parameter matrix of and the right-hand side of the RCs, respectively.

If we require strict equivalence between the RCs and OCs, then this involves projecting the variable space associated with the OCs onto the RCs, which means the following:

$$\forall x : \exists y^{\rm r}, A(x, y^{\rm r}) \le b \Leftrightarrow \exists y^{\rm o}, h(x, y^{\rm o}) \le 0. \qquad (1)$$

However, strict projection may be computationally infeasible in cases where $h(\cdot)$ has high dimensionality, $y^{\rm o}$ contains integer variables, or T is large. Moreover, in the current practice of grid-load interactions, small deviations in the demand response are permissible. Taking these practical factors into account, we relax Equ. (1) and introduce a loss function $J(h(\cdot), A, b)$ to quantify the error of the RCs relative to the OCs. Therefore, our objective transforms into determining appropriate A and b values to minimize the approximation error $J(h(\cdot), A, b)$.

Specifically, we do not derive A and b analytically on the basis of the form and parameters of the constraint $h(\cdot)$. Instead, we observe the historical energy usage of industrial users and parameterize (train) A and b in a data-driven manner. In doing so, we assume access to historical data of industrial users $HD = \{(Pr^{(n)}, x^{(n)})\}$, where $Pr^{(n)}$ and $x^{(n)}$ represent the hourly electricity price and the user's hourly electricity consumption on day n, respectively, $n = 1, ..., N$. We also assume that $x^{(n)}$ are optimized on the basis of the OCs to minimize the energy cost $Pr^{(n)\intercal} x^{(n)}$. Note that HD can be realistic historical data or simulated data. If we lack historical data from industrial users under fluctuating electricity prices or if the historical data are not optimal, then the D3R framework remains applicable; however, it initially needs to simulate optimal energy usage data under varying electricity prices via the OCs.

Under the above assumptions, the approximation error can be expressed as the distance between the optimal energy consumption under historical electricity prices given by the OCs and RCs, and the training problem is formulated as follows:

$$\min_{A,b}. J = \frac{1}{N}\sum_{n=1}^{N} ||x^{(n)} - x_n||^2 \tag{2a}$$

$$\text{s.t.} \quad x_n = \text{argmin.}\{Pr^{(n)\intercal} x_n \ : \ A(x_n, y^{\text{r}}) \leq b\}, \ \forall n \tag{2b}$$

where x_n is the variable for the optimization problem regarding day n based on the RCs to distinguish it from $x^{(n)}$ in the dataset. (2) is a data-driven inverse optimization problem [3], which will be further elaborated in Section III.

III. Methodology

In principle, D3R allows for the exploration of suitable parameter matrix sizes and element values within the space A/b. However, the high-dimensional search space and the high-dimensional nonlinear constraints of the optimality conditions in (2b) may lead to unaffordable computational costs. To address this challenge, our approach is twofold: first, we selected an appropriate form of the parameter matrix on the basis of the energy consumption characteristics of industrial users (Section III-A); and second, we devised a solution algorithm on the basis of zeroth-order stochastic gradient descent (ZOSGD) [6] to iteratively solve the inverse optimization problem, with a transformation of the optimality conditions to reduce the computational complexity at each iteration (Section III-B).

A. ALF-based Reduced Constraints

We chose an adjustable load fleet (ALF) composed of multiple adjustable loads (ALs) as the form of the RC to be trained. ALs are commonly used in power systems, have low computational complexity, and are compatible with existing economic dispatch models. The RCs in the form of an ALF are then:

$$x_t = \sum_{i=1}^{I} p_{t,i}\Delta t, \quad \underline{P}_i \leq p_{t,i} \leq \overline{P}_i \ : \ \underline{\mu}^{\text{P}}_{t,i}, \overline{\mu}^{\text{P}}_{t,i}, \ \forall t \tag{3a}$$

$$\underline{E}_i \leq \sum_{t=1}^{T} p_{t,i} \leq \overline{E}_i \ : \ \underline{\mu}^{\text{E}}_i, \overline{\mu}^{\text{E}}_i \tag{3b}$$

where the subscript n is omitted, I is the number of ALs as a hyperparameter, and x_t is the net energy consumption at time t, which is the summation of $p_{t,i}\Delta t$, representing the average consumption power of AL i at time t multiplied by the length of time interval Δt. $\theta = \{\underline{P}_i, \overline{P}_i, \underline{E}_i, \overline{E}_i\}$ include the power and energy limits of AL i, respectively, which are the parameters to be fitted. The Lagrange multipliers $\underline{\mu}^{\text{P}}_{t,i}, \overline{\mu}^{\text{P}}_{t,i}, \underline{\mu}^{\text{E}}_i, \overline{\mu}^{\text{E}}_i$ are the dual variables associated with the power and energy limits, respectively. The above linear constraints can be easily transformed into the form of $A(x, y^{\text{r}}) \leq b$, where $x = [x_t]$ and $y^{\text{r}} = [p_{t,i}]$.

We chose an ALF because intuitively, in commonly used state-task networks (STNs) [1] and the derived resource-task networks (RTNs) [2], as well as other industrial load models, the production goal constraints can be captured by the energy usage constraint (3b). Moreover, the equipment composition relationships in industrial processes can be captured via the parallel operation of multiple ALs.

B. Training of the ALF model

Substituting the form of the ALF, the optimality conditions (Karush-Kuhn-Tucker conditions) represented by Equation (2b) for day n include primal feasibility (3), stationarity (4a), dual feasibility (4b), and complementary slackness (4c, 4d):

$$Pr^{\text{E}}_t - \underline{\mu}^{\text{P}}_{t,i} + \overline{\mu}^{\text{P}}_{t,i} - \underline{\mu}^{\text{E}}_i + \overline{\mu}^{\text{E}}_i = 0, \ \forall t, \forall i \tag{4a}$$

$$(\underline{\mu}^{\text{P}}_{t,i}, \overline{\mu}^{\text{P}}_{t,i}, \underline{\mu}^{\text{E}}_i, \overline{\mu}^{\text{E}}_i) \geq 0, \ \forall t, \forall i \tag{4b}$$

$$\underline{\mu}^{\text{P}}_{t,i}(\underline{P}_i - p_{t,i}) = 0, \quad \overline{\mu}^{\text{P}}_{t,i}(p_{t,i} - \overline{P}_i) = 0, \quad \forall t, i \tag{4c}$$

$$\underline{\mu}^{\text{E}}_i(\underline{E}_i - \sum_{t=1}^{T} p_{t,i}) = 0, \quad \overline{\mu}^{\text{E}}_i(\sum_{t=1}^{T} p_{t,i} - \overline{E}_i) = 0, \quad \forall i \tag{4d}$$

The complementary slackness (4c, 4d) conditions involve bilinear constraints, which can be transformed via the Fortuny-Amat transformation [7] into a more easily solvable form. For example, the first term can be replaced by the following:

$$\underline{\mu}^{\text{P}}_{t,i} \leq M(1 - z_{t,i}), \quad p_{t,i} - \underline{P}_i \leq M z_{t,i}, \quad \forall t, i \tag{5}$$

where $z_{t,i}$ is a binary variable and M is a large positive number. The transformed conditions are used in inverse optimization (2), forming a mixed-integer linear programming (MILP) problem that can be solved by commercial solvers.

However, considering that the number of integer variables is proportional to the size of the dataset N, the computational complexity of direct solving becomes too high when data from multiple days are used. Instead, we employ a zeroth-order stochastic gradient descent (ZOSGD) [6] algorithm for iterative solving. This means that in each iteration, a batch (one day) of data is randomly selected and that the parameters are updated via estimated gradients:

1: Initialize $j = 0$ and $\theta^{(0)} = \{\underline{P}_i, \overline{P}_i, \underline{E}_i, \overline{E}_i\} = [0]$.

2: Randomly select day n and solve the batch problem: $\min_{\theta}. J = ||x^{(n)} - x_n||^2$ s.t.$(3) - (4)$ to obtain θ^*.

3: Update $\theta^{(j+1)} = (1-\alpha)\theta^{(j)} + \alpha\theta^*$; if $\theta^{(j+1)}$ converges, break; otherwise, set $j = j + 1$ and go to step 2.

where α is the learning rate, which can be adaptively adjusted. The ZOSGD algorithm can be easily implemented by using commercial solvers, and the computational complexity is independent of the number of days in the dataset.

IV. Numerical Results

a) Dataset and Task Description: We tested the D3R framework on three datasets from a cement plant [8], a steel powder plant [1], and a steelmaking plant [2]. The first two were originally modeled via the STN, whereas the steelmaking plant was modeled by using the RTN, both are high-dimensional models with integer variables, which were

TABLE I
NORMALIZED RMSE OF REDUCED CONSTRAINTS

Method	Cement plant	Steel powder manufactory	Steelmaking plant
OVB	25.4%	36.7%	N/A
SAL	10.2%	23.0%	8.3%
D3R-1	**8.9%**	17.2%	**3.6%**
D3R-2	9.5%	**10.3%**	4.2%

TABLE II
NUMBER OF VARIABLES IN THE MODELS

Continuous (integer)	Cement plant	Steel powder manufactory	Steelmaking plant
Original Model	196(288)	490(720)	0(10208)
OVB	48(0)	48(0)	N/A
SAL	24(0)	24(0)	24(0)
D3R-1	24(0)	24(0)	24(0)
D3R-2	48(0)	48(0)	48(0)

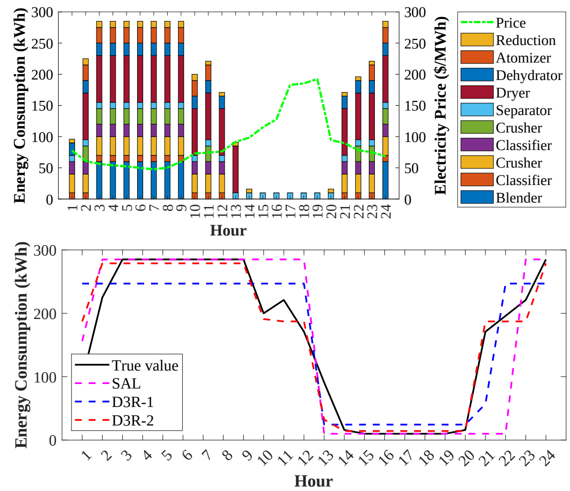

Fig. 2. Top: Optimal energy consumption of the production processes in the steel powder plant under typical electricity prices (July 23) with original high-dimensional constraints (true value). Bottom: Optimal energy consumption results provided by the reduced constraints under the same electricity prices.

used to simulate the optimal energy consumption results for training the ALF.

We used historical hourly electricity market prices from PJM in July 2022 for the simulations. The data from July 1st to 21st contained the training set, whereas the data from July 22nd to 31st were used to test the accuracy of the reduced constraints. The detailed settings and codes can be found in [9]. We employed Gurobi (V11.0.0) with YALMIP in MATLAB to solve the optimization problems on a workstation equipped with an Intel Core i9-10900X CPU (3.7 GHz) and 128 GB of RAM.

For complex industrial load models, there are currently no mature methods for dimension reduction. Therefore, we compared the accuracy of D3R with that of:

1) the simple AL model (SAL), which assigns values to θ based on the maximum and minimum values of the corresponding quantities in the historical load curves. For instance, the maximum power value from historical data is directly taken as the $P^{\max}$ for the SAL model.

2) the optimal virtual battery model (OVB) [10], an analytical inner-approximation method.

b) Results and comparison: The proposed D3R achieved the best performance on all three datasets. This means that the dimensionality reduction constraints obtained through this method yielded the lowest normalized root mean square error (RMSE) compared with the results based on the original constraints in the test set (Table I). Here, D3R-1(2) represents D3R via the ALF model with 1(2) ALs. The results from OVB based on the analytical inner approximation method were significantly inferior. Moreover, because OVB can be applied only to linear constraints, it cannot be applied to steelmaking plants. Fig. 2 visually demonstrates the effectiveness of D3R, where it reduces the industrial users' complex constraints and variables in the optimization problem with an error ranging from 4% to 10% (Table II).

V. CONCLUSION

Diverging from bottom-up analytical approaches, we propose D3R, a data-driven dimension reduction framework that trains a low-dimensional constraint model to fit the optimal solutions derived from the original constraints. D3R was tested on three industrial load datasets, and the results demonstrated that D3R reduced the high-dimensional constraints without losing much optimality, outperforming the analytical methods in terms of accuracy. D3R provides a new perspective for dimension reduction in industrial load modeling and other tasks with complex constraints to reduce, enhancing adaptability and compatibility with integer variables.

REFERENCES

[1] R. Lu, R. Bai, Y. Huang, Y. Li, J. Jiang, and Y. Ding, "Data-driven real-time price-based demand response for industrial facilities energy management," *Appl. Energy*, vol. 283, p. 116291, Feb. 2021.

[2] X. Zhang, G. Hug, and I. Harjunkoski, "Cost-Effective Scheduling of Steel Plants With Flexible EAFs," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 239–249, Jan. 2017.

[3] R. Lyu, H. Guo, and Q. Chen, "Approximating Energy-Regulation Feasible Region of Virtual Power Plants: A Data-driven Inverse Optimization Approach," in *2024 IEEE Power & Energy Society General Meeting*.

[4] F. L. Muller, J. Szabo, O. Sundstrom, and J. Lygeros, "Aggregation and Disaggregation of Energetic Flexibility From Distributed Energy Resources," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1205–1214, Mar. 2019.

[5] Y. Wen, Z. Hu, S. You, and X. Duan, "Aggregate Feasible Region of DERs: Exact Formulation and Approximate Models," *IEEE Trans. Smart Grid*, vol. 13, no. 6, pp. 4405–4423, Nov. 2022.

[6] S. Liu, P.-Y. Chen, B. Kailkhura, G. Zhang, A. Hero, and P. K. Varshney, "A Primer on Zeroth-Order Optimization in Signal Processing and Machine Learning," Jun. 2020, arXiv:2006.06224.

[7] J. Fortuny-Amat and B. McCarl, "A Representation and Economic Interpretation of a Two-Level Programming Problem," *J. Oper. Res. Soc*, vol. 32, no. 9, pp. 783–792, Sep. 1981.

[8] H. Golmohamadi, R. Keypour, B. Bak-Jensen, J. R. Pillai, and M. H. Khooban, "Robust Self-Scheduling of Operational Processes for Industrial Demand Response Aggregators," *IEEE Trans. Ind. Electron.*, vol. 67, no. 2, 2020.

[9] R. Lyu. [Online]. Available: https://github.com/Rick10119/Data-Driven-Dimension-Reduction

[10] Z. Tan, A. Yu, H. Zhong, X. Zhang, Q. Xia, and C. Kang, "Optimal virtual battery model for aggregating storage-like resources with network constraints," *CSEE J. Power Energy Syst.*, vol. 10, no. 4, pp. 1843–1847, 2024.