

Problem-Driven Scenario Reduction Framework for Power System Stochastic Operation

Yingrui Zhuang, *Student Member, IEEE*, Lin Cheng, *Senior Member, IEEE*, Ning Qi, *Member, IEEE*,
Mads R. Almassalkhi, *Senior Member, IEEE*, Feng Liu, *Senior Member, IEEE*

Abstract—Scenario reduction (SR) aims to identify a small yet representative scenario set to depict the underlying uncertainty, which is critical to scenario-based stochastic optimization (SBSO) of power systems. Existing SR techniques commonly aim to achieve statistical approximation to the original scenario set. However, SR and SBSO are commonly considered as two distinct and decoupled processes, which cannot guarantee a superior approximation of the original optimality. Instead, this paper incorporates the SBSO problem structure into the SR process and introduces a novel problem-driven scenario reduction (PDSR) framework. Specifically, we project the original scenario set in distribution space onto the mutual decision applicability between scenarios in problem space. Subsequently, the SR process, embedded by a distinctive problem-driven distance metric, is rendered as a mixed-integer linear programming formulation to obtain the representative scenario set while minimizing the optimality gap. Furthermore, *ex-ante* and *ex-post* problem-driven evaluation indices are proposed to evaluate the SR performance. Numerical experiments on two two-stage stochastic economic dispatch problems validate the effectiveness of PDSR, and demonstrate that PDSR significantly outperforms existing SR methods by identifying salient (e.g., worst-case) scenarios, and achieving an optimality gap of less than 0.1% within acceptable computation time.

Index Terms—Problem-driven, scenario reduction, stochastic optimization, worst-case scenario, risk management

NOMENCLATURE

A. Abbreviations

ADN	Active distribution network
DDSR	Distribution-driven scenario reduction
ES	Energy storage
MILP	Mixed-integer linear programming
OG	Optimality gap
PDD	Problem-driven distance
PDSR	Problem-driven scenario reduction

This work is supported in part by the National Key Research and Development Program of China (No. 2021YFE0191000), National Natural Science Foundation of China (No. 52037006), and China Postdoctoral Science Foundation special funded project (No. 2023TQ0169). Mads graciously recognizes support from the National Science Foundation (NSF) Award ECCS-2047306. Paper No. TPWRS-XXXXXX-2024. (Corresponding author: Yingrui Zhuang.)

Yingrui Zhuang, Lin Cheng and Feng Liu are with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: zyr21@mails.tsinghua.edu.cn).

Ning Qi is with the Department of Earth and Environmental Engineering, Columbia University, NY 10027, USA (e-mail: nq2176@columbia.edu).

Mads R. Almassalkhi is with the Department of Electrical and Biomedical Engineering, University of Vermont, Burlington VT 05405, USA (e-mail: malmassa@uvm.edu).

RES	Renewable energy sources
SBSO	Scenario-based stochastic optimization
SE	Scenario effectiveness
SoC	State of charge
SPDD	Sum of problem-driven distance
SR	Scenario reduction
TSSO	Two-stage stochastic optimization
B. Sets	
ξ	Original scenario set
ζ	Representative scenario set
Ω_B	Set of nodes
$\Omega_E/\Omega_R/\Omega_L$	Set of nodes with ES/RES/Load
C. Parameters	
$\Delta t/T$	Scheduling time interval/number of time steps in a scheduling cycle
η_j^c/η_j^d	Charge/discharge efficiency of j -th ES
γ_i	Probability of scenario ξ_i
\bar{V}_i/V_i	Upper/lower bound of squared voltage magnitude
\bar{P}^T	Upper bound of trading power
\bar{P}_j^E	Upper bound of charging/discharging power of j -th ES
$\pi^{R,c}/\pi^{L,s}$	Penalty price of RES curtailment/load shedding
π_j^E	Capacity procurement price of j -th ES
$\pi_{s,t}^{T+}/\pi_{s,t}^{T-}$	Intraday up-regulation/down-regulation price
$\pi_{s,t}^T$	Day-ahead trading price
$\underline{SoC}/\bar{SoC}$	Lower/upper bound of SoC
N	Number of original scenarios
$P_{i,s,t}^R/P_{i,s,t}^L$	RES generation/load demand power at node i
r_{ij}/x_{ij}	Line resistance/reactance of line ij
D. Variables	
β	Trade-off factor
$V_{i,s,t}$	Squared voltage magnitude at node i
ω_k	Probability of scenario ζ_k
$D_{s,t}^T$	Intraday balancing state of ADN
$D_{j,s,t}^E$	Charge/discharge state of j -th ES
E_j	Procured energy capacity of j -th ES
I_{ij}	Current of line ij
K	Number of representative scenarios
$p_{i,s,t}/q_{i,s,t}$	Active/reactive injection power at node i

$P_{i,s,t}^{R,c}/P_{i,s,t}^{L,s}$	RES curtailment/load shedding power at node i
$P_{ij,s,t}/P_{ij,s,t}$	Active/reactive power of line ij
$P_{j,s,t}^{E,c}/P_{j,s,t}^{E,d}$	Charge/discharge power of j -th ES
$P_{s,t}^{T+}/P_{s,t}^{T-}$	Intraday up-regulation/down-regulation power
P_t^T	Day-ahead trading power
z_{ξ}^*/z_{ζ}^*	Optimal solution of the original problem with ξ /reduced problem with ζ

E. Functions

$d(\cdot)$	Problem-driven distance function
$F(\cdot)$	Objective of TSSO problem

I. INTRODUCTION

THE rapid integration of renewable energy sources (RES) and new loads into the power systems has led to increased variability and uncertainty in operations. Thus, effective decision-making in power system operations must account for these uncertainties to manage risks [1]. With complete information of uncertainties (e.g., a known probability distribution), chance-constrained optimization [2], [3], robust optimization [4], [5] and distributionally robust optimization [6], [7] have been verified to be effective approaches. In cases where incomplete information of uncertainties [8] (e.g., historical, forecasted scenarios), a common practice is to employ scenario-based stochastic optimization (SBSO) [9]–[11], where a finite scenario set is utilized to approximate the probability distribution of uncertainties [12]. However, scenario-based techniques typically struggle with the “curse of dimensionality”, which becomes more pronounced as the variety and number of uncertainties increase [13]. To reduce this complexity, scenario reduction (SR) can be used to identify a smaller representative scenario set to replace the original scenario set for decision making while maintaining an acceptably robust optimal solution. However, three critical questions exist for SR: (i) how to quantify the similarity between scenarios? (ii) how to define the representativeness of the reduced scenarios prospectively? (iii) how to perform SR to yield a representative scenario set that optimally reflects the full original formulation?

Most SR methods implicitly assume that statistically *better* representations of the original scenario set in the distribution space necessarily yield *better* optimal solutions of SBSO. We refer to these methods as distribution-driven scenario reduction (DDSR) methods. The overview of DDSR methods is summarized in Fig. 1(a). DDSR methods generally construct the original scenario set using raw scenario data [14], deep features extracted by machine learning [15], [16], and relevant problem properties manually selected based on engineering experience (e.g., power ramping [17], network power flow [18], and investment cost [19]). Moreover, to account for the impacts of worst-case scenarios, the original scenario set is often split into “normal” and “worst-case” subsets and SR is performed separately on each subset [20]. However, the definition of “worst-case” scenarios varies across different problem formulations (e.g., economic dispatch and resilience-oriented dispatch [21]) and are often difficult to explicitly define. Subsequently, distribution-driven distance metrics, such as Euclidean

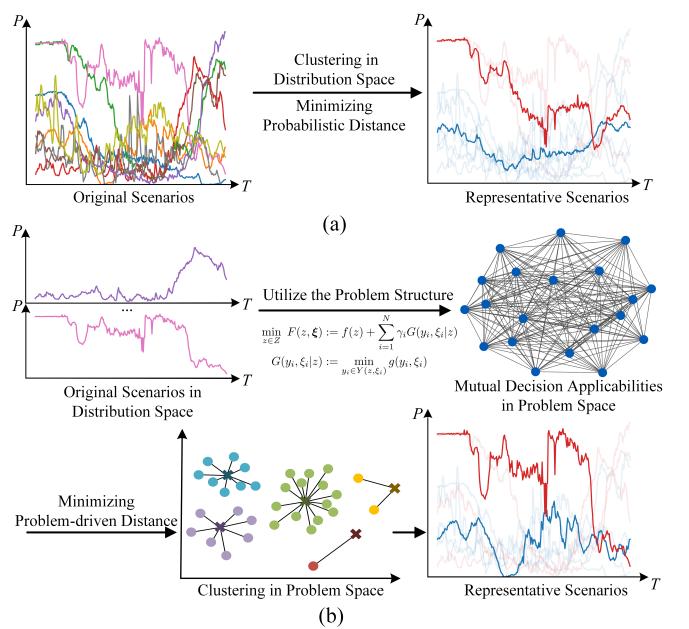


Fig. 1. Diagrams of scenario reduction methodologies: (a) distribution-driven scenario reduction and (b) problem-driven scenario reduction.

distance [14], Wasserstein distance [22], and dynamic time warping distance [23] are frequently used to measure the similarity between scenarios. Furthermore, clustering techniques, such as hierarchical clustering (HC) [24], K -means [25], K -medoids [26], and Gaussian mixture model [27], are employed to cluster the original scenario set into a representative scenario set. However, these methods generally rely on a myriad of hyper-parameters (e.g., random initialization and iterative adjustments). Finally, statistical indices based on distribution-driven distance metrics (e.g., Davis-Bouldin index) are used to validate the clustering performance. Unfortunately, higher statistical similarities between reduced and original scenario sets may not guarantee a better optimality approximation, which is particularly the case in optimization of power systems [28]. A simple example is that a slight increase in load demand beyond the safety limit can result in additional operational adjustments and penalty costs for system reliability losses [29]. Moreover, DDSR methods tend to generate the same representative scenario set for two different problems that share the same original scenario set. That is, since DDSR methods generally consider SR and SBSO as two distinct and decoupled processes, they suffer from a critical oversight: an inability to consider the impacts of the reduced scenario set on the optimal solution to the original SBSO.

To address this gap, SR methods should re-evaluate *representativeness* of scenarios. Specifically, the efficacy of SR should be gauged by the performance of the representative scenarios with consideration of the SBSO problem structure, i.e., in *problem space*, as illustrated in Fig. 1(b). *Problem space* is a projection of the distribution space that takes into account the SBSO problem structure, wherein the scenarios are positioned by the repulsive forces formed by their mutual decision impacts on the problem outcomes, rendering a spatial structure based on objective and decision elements. This paper,

therefore, focuses on identifying scenarios with high *decision applicability*, which refers to their relatively large impacts on the optimal objective value and decision-making. Recently, literature has integrated decision applicability into the SR process and referred the framework problem-dependent SR. In [30], a problem-dependent methodology is proposed for SR that relies on a computationally complex Wasserstein distance metric and alternating minimization algorithm, which limits scalability and potentially yields suboptimal solutions. A symmetric opportunity cost is employed in [31] as the distance metric to measure the decision applicability between scenarios, and heuristic methods are used to perform SR, which limits the ability to characterize the optimality gap. Ref. [32] interestingly develops a problem-driven scenario clustering method with asymmetric distance metric where the representative scenarios are selected based on their average decision applicability in the entire cluster, which limits scalability of the method to SBSOs with low dimensionality (e.g., few scenarios, simple problem structure) and is impractical in power system applications.

In this paper, we propose a novel problem-driven scenario reduction (PDSR) framework for solving SBSO problems to near-optimality and case studies illustrate improvements of up to ten times in terms of scalability and optimality gap over nine state-of-the-art methods, which enables PDSR applications to power systems for the first time. Specifically, our contributions are as follows:

- 1) *Problem-Driven Scenario Reduction Framework*: We propose a novel PDSR framework for general SBSO problems by defining the concept of optimality gap (OG) for SR and analytically characterizing the impacts of SR on the SBSO problem outcomes.
- 2) *Problem-Driven Distance Metric*: To evaluate the representativeness of scenarios, we introduce a provably effective problem-driven distance (PDD) metric that quantifies the mutual decision applicability between scenarios. We show that the OG can be bounded above by minimizing the sum of PDD within clusters (SPDD). Furthermore, we use the PDD and OG to introduce new *ex-ante* and *ex-post* problem-driven SR evaluation indices.
- 3) *Clustering Methodology*: Based on the PDD, we convert the original scenario set partitioning and representative scenario selection process into a mixed-integer linear program (MILP). The MILP objective balances minimization of SPPD within clusters and the reduction degree reflected by the total number of clusters (i.e., representative scenarios).
- 4) *Simulation-Based Analysis*: Two case studies validate the proposed PDSR framework by applying it to stochastic two-stage economic optimization problems within the context of active distribution networks and unit commitment, which co-optimize day-ahead decisions and intraday decision adjustments. Simulation results demonstrate PDSR's ability to identify salient scenarios that achieve an SR optimality gap up to ten times smaller than the nine state-of-the-art SR methods.

The remainder of the paper is organized as follows. Sec-

tion II introduces the proposed PDSR framework. Formulation of a scenario-based stochastic economic dispatch problem for an active distribution network is presented in Section III. Numerical studies based on real-world data are provided in Section IV to illustrate comparative performance. Finally, conclusions are summarized in Section V.

II. PROBLEM-DRIVEN SCENARIO REDUCTION FRAMEWORK

In this section, we detail the novel PDSR framework within the context of a general two-stage stochastic optimization (TSSO) problem, which represents a rich set of power system problems. Note that the PDSR framework can also be adapted to single-stage and multi-stage SBSO problems.

A. Formulation of Two-Stage Stochastic Optimization

Two-stage stochastic optimization is an effective formulation in stochastic optimization to address uncertainties due to its “here-and-now” and “wait-and-see” characteristics, which respectively represent the decisions to be made before and after the uncertainty is revealed. The general formulation of the TSSO built on the original scenario set is

$$\min_{z \in Z} F(z, \xi) := f(z) + \sum_{i=1}^N \gamma_i G(y_i, \xi_i | z) \quad (1a)$$

$$G(y_i, \xi_i | z) := \min_{y_i \in Y(z, \xi_i)} g(y_i, \xi_i), \quad (1b)$$

where $F(z, \xi)$ is the objective function of the TSSO, including the objective of the first stage and the expected objective of the second stage. $f(z)$ is the objective function of the first stage with the decision variable $z \in Z$. Here, $Z \subseteq \mathbb{R}^n$ denotes the bounded feasible set. The original set of N scenarios is denoted as $\xi = \{\xi_1, \xi_2, \dots, \xi_N\}$. The probability of scenario ξ_i is $\gamma_i > 0$, which satisfies $\sum_{i=1}^N \gamma_i = 1$. The second-stage problem is $G(y_i, \xi_i | z)$ under the uncertainty ξ_i with $y_i \in Y(z, \xi_i)$, $Y \subseteq \mathbb{R}^m$ as the decision variable and z as the parameter, since z remains constant across all second-stage problems. $g(y_i, \xi_i)$ is the objective function of the second stage. We denote the optimal solution of (1) as $z_{\xi}^* = \arg \min_{z \in Z} F(z, \xi)$. Additionally, we denote the scenario specific subproblem $F(z, \xi_i) := f(z) + G(y_i, \xi_i)$ with $z_{\xi_i}^* = \arg \min_{z \in Z} F(z, \xi_i)$, we have $F(z, \xi) = \sum_{i=1}^N \gamma_i F(z, \xi_i)$. We reasonably require the TSSO to satisfy the assumption of *relatively complete recourse*, implying that there exists a solution to TSSO for any $z \in Z$ and $\xi \in \xi$. This assumption is common in stochastic optimization [33], ensuring sufficient resources to handle potential risks, even costly.

To address the challenges of computation complexity when N is large, SR is employed to significantly reduce the computation complexity while maintaining the problem optimality approximation accuracy at an acceptable level. The SR process can be denoted as $C(\xi, K) = \{\{C_1, \dots, C_K\} : C_i \neq \emptyset, \forall i; C_i \cap C_j = \emptyset, \forall i \neq j; \cup_i C_i = I, I = \{1, 2, \dots, N\}\}$. The original scenario set ξ is partitioned into K ($K \ll N$) clusters and is reduced to a representative scenario set $\zeta = \{\zeta_1, \zeta_2, \dots, \zeta_K\}$. Each scenario cluster C_k is represented

by the representative scenario ζ_k with corresponding weight $\omega_k = \sum_{i \in C_k} \gamma_i$, which satisfies $\sum_{k=1}^K \omega_k = 1$. In this paper, we concentrate on selecting ζ as a subset of ξ , instead of generating new scenarios. The TSSO formulated on the representative scenario set ζ is given by

$$\min_{z \in Z} F(z, \zeta) := f(z) + \sum_{k=1}^K \omega_k G(y_k, \zeta_k | z) \quad (2a)$$

$$G(y_k, \zeta_k | z) := \min_{y_k \in Y(z, \zeta_k)} g(y_k, \zeta_k). \quad (2b)$$

The optimal solution of the reduced problem (2) is denoted as $z_{\zeta}^* = \arg \min_{z \in Z} F(z, \zeta)$. Specifically, we would like to understand how SR affects the optimality of SBSO. Thus, we seek to define an SR optimality gap metric next.

B. SR Optimality Gap

SR aims to minimize the optimality gap from using K representatives (ζ) vs. N scenarios (ξ). Towards this purpose, the OG can be defined as

$$OG := F(z_{\zeta}^*, \xi) - F(z_{\xi}^*, \xi), \quad (3)$$

where $F(z_{\zeta}^*, \xi)$ means solving (1) with $z = z_{\zeta}^*$. Compared to [32], we present a distinct and rigorous derivation of an upper bound on OG .

Since $F(z, \xi) \geq F(z_{\xi}^*, \xi)$ for all $z \in Z$, we have $OG \geq 0$. A smaller OG indicates a more accurate problem optimality approximation of ζ to ξ . Since $F(z, \zeta) \geq F(z_{\zeta}^*, \zeta)$ for all $z \in Z$, we can derive an upper bound of OG as

$$\begin{aligned} OG &\leq F(z_{\zeta}^*, \xi) - F(z_{\xi}^*, \xi) + (F(z_{\xi}^*, \zeta) - F(z_{\zeta}^*, \zeta)) \\ &= (F(z_{\zeta}^*, \xi) - F(z_{\zeta}^*, \zeta)) - (F(z_{\xi}^*, \xi) - F(z_{\xi}^*, \zeta)). \end{aligned} \quad (4)$$

Note that both the first and the last pair of terms share a common expression of $F(z, \xi) - F(z, \zeta)$, which can be further reformulated as

$$\begin{aligned} F(z, \xi) - F(z, \zeta) &= \sum_{i=1}^N \gamma_i F(z, \xi_i) - \sum_{k=1}^K \omega_k F(z, \zeta_k) \\ &= \sum_{k=1}^K \sum_{i \in C_k} \gamma_i F(z, \xi_i) - \sum_{k=1}^K \sum_{i \in C_k} \gamma_i F(z, \zeta_k) \\ &= \sum_{k=1}^K \sum_{i \in C_k} \gamma_i (F(z, \xi_i) - F(z, \zeta_k)). \end{aligned} \quad (5)$$

Combining (4) and (5), the upper bound of OG can be further expanded as

$$\begin{aligned} OG &\leq \sum_{k=1}^K \sum_{i \in C_k} \gamma_i (F(z_{\zeta}^*, \xi_i) - F(z_{\zeta}^*, \zeta_k)) \\ &\quad - \sum_{k=1}^K \sum_{i \in C_k} \gamma_i (F(z_{\xi}^*, \xi_i) - F(z_{\xi}^*, \zeta_k)) \\ &\leq \sum_{k=1}^K \sum_{i \in C_k} \gamma_i (|F(z_{\zeta}^*, \xi_i) - F(z_{\zeta}^*, \zeta_k)| \\ &\quad + |F(z_{\xi}^*, \xi_i) - F(z_{\xi}^*, \zeta_k)|). \end{aligned} \quad (6)$$

Given that z_{ξ}^* and z_{ζ}^* are not known, we use a result from [34] to derive an upper bound on (6) based on (7), which states that for a locally Lipschitz continuous function $F(z, \xi)$, there exists a continuous symmetric function $d(\cdot)$ and a non-decreasing function $h(\cdot)$, such that for each $z \in Z$ and $\xi_i, \zeta_k \in \xi$, we have

$$|F(z, \xi_i) - F(z, \zeta_k)| \leq h(\|z\|)d(\xi_i, \zeta_k), \quad (7)$$

where $d(\xi_i, \zeta_k)$ is required to satisfy the following properties:

- C1) *Consistency*: $d(\zeta_k, \xi_i) = 0 \Leftrightarrow \zeta_k = \xi_i$;
- C2) *Symmetricity*: $d(\zeta_k, \xi_i) = d(\xi_i, \zeta_k), \forall \zeta_k, \xi_i \in \xi$;
- C3) *Convergence*: $\sup\{d(\zeta_k, \xi_i) : \zeta_k, \xi_i \in \xi, \|\zeta_k - \xi_i\| \leq \delta\}$ tends to 0 as $\delta \rightarrow 0$;
- C4) *Triangle inequality*: \exists measurable, bounded function $\lambda(\cdot)$, where $d(\zeta_k, \xi_i) < \lambda(\zeta_k) + \lambda(\xi_i)$.

Finally, combining (6) and (7) begets

$$OG \leq (h(\|z_{\xi}^*\|) + h(\|z_{\zeta}^*\|)) \sum_{k=1}^K \sum_{i \in C_k} \gamma_i d(\xi_i, \zeta_k). \quad (8)$$

Now, the primary challenge to apply the above result lies in defining an appropriate distance metric in the problem space, $d(\xi_i, \zeta_k)$, that satisfies properties C1)-C4). Note that since Z is bounded, $\|z\|$ is well-defined, i.e., $\exists M \gg 1$, $\|z\| \leq M \forall z \in Z$. Next, we project the distribution space into the *problem space* and then define an appropriate metric $d(\xi_i, \zeta_k)$ within the problem space.

C. Problem Space Projection

In scenario-based problem formulations, each $z_{\zeta_k}^*$ is usually implemented within its respective scenario clusters, where $z_{\zeta_k}^* = \arg \min_{z \in Z} F(z, \zeta_k)$. Motivated by this, PDSR projects the original scenario set in the distribution space onto the problem space, which is constructed by the decision applicability between scenarios. The projection process can be denoted as $\xi \rightarrow F$, where $F := \{F_{ij} = F(z_{\xi_i}^*, \xi_j) \mid i, j \in I\}$. Each F_{ij} represents a scenario-specific problem and is bounded under the condition of *relatively complete recourse*. Note that, if $z_{\xi_i}^* = \arg \min_{z \in Z} F(z, \xi_i)$ has multiple optimal solutions, we select the one that minimizes $\sum_{\xi \in \xi} F(z, \xi)$, indicating better decision applicability to the original scenario set. In this projection, we can directly quantify the impacts of uncertainty and systematically incorporate the inherent characteristics of the SBSO problem into the SR process. Of course, this approach necessitates solving N^2 optimization problems to determine F , which may be computationally intensive for large N . However, since each problem is independent and can be solved in parallel, the absolute time required to find F can be reduced significantly. The algorithmic efficiency is discussed in IV-C.

D. Problem-Driven Distance Metric

Opportunity cost can be utilized to describe the decision applicability, which pertains to the trade-off wherein the selection of a particular action necessitates the relinquishment of potential benefits associated with the best alternative. In

the context of SR, when scenario ζ_k is chosen to represent scenario ξ_i , the opportunity cost is defined as

$$c(\zeta_k, \xi_i) := F(z_{\zeta_k}^*, \xi_i) - F(z_{\xi_i}^*, \xi_i). \quad (9)$$

However, $c(\zeta_k, \xi_i)$ does not satisfy all properties C1)-C4) and cannot serve as a distance metric $d(\cdot)$. Instead, we consider the following problem-driven distance metric between scenarios ξ_i and ζ_k :

$$\begin{aligned} d(\xi_i, \zeta_k) &:= c(\zeta_k, \xi_i) + c(\xi_i, \zeta_k) \\ &= F(z_{\zeta_k}^*, \xi_i) - F(z_{\xi_i}^*, \xi_i) + F(z_{\xi_i}^*, \zeta_k) - F(z_{\zeta_k}^*, \zeta_k). \end{aligned} \quad (10)$$

Proposition 1. *The distance metric $d(\xi_i, \zeta_k)$ in (10) satisfies all four properties C1)-C4) for (7).*

Proof. Please see proof in Appendix A. \square

The PDD in (10) effectively quantifies the mutual decision applicability between two scenarios. That is, a small $d(\xi_i, \zeta_k)$ implies that scenario ζ_k serves as an accurate representation of scenario ξ_i .

Combining (8) and (10) begets

$$\begin{aligned} OG &\leq 2h(M) \sum_{k=1}^K \sum_{i \in C_k} \gamma_i (F(z_{\zeta_k}^*, \xi_i) - F(z_{\xi_i}^*, \xi_i) \\ &\quad + F(z_{\xi_i}^*, \zeta_k) - F(z_{\zeta_k}^*, \zeta_k)). \end{aligned} \quad (11)$$

Note that while the above analysis can be extended to non-convex formulations, the bounds on OG are conditioned on finding the global optimum. Convex formulations, as well as certain simple non-convex formulations in (2) are the focus of this paper, as they can be solved to global optimality by existing commercial solvers. Next, we select salient scenarios that minimize the upper bound in (11).

E. MILP Reformulation of Clustering

Since $2h(M)$ in (11) is a constant, minimizing the sum of PDD within clusters achieves the lowest upper bound of OG . To achieve this, the process of scenario partitioning and representative scenario selection can be rendered as the following MILP formulation.

$$\min_{v, u, l, K} \sum_{j=1}^N l_j + \beta K/N \quad (12a)$$

$$\text{s.t. } \sum_{i=1}^N \gamma_i v_{ij} (F_{ji} - F_{ii} + F_{ij} - F_{jj}) \leq l_j, \quad \forall j \in I \quad (12b)$$

$$v_{ij} \leq u_j, \quad v_{jj} = u_j, \quad \forall i, j \in I \quad (12c)$$

$$\sum_{j=1}^N v_{ij} = 1, \quad \forall i \in I \quad (12d)$$

$$\sum_{j=1}^N u_j = K, \quad \forall j \in I \quad (12e)$$

$$v_{ij} \in \{0, 1\}, \quad u_j \in \{0, 1\}, \quad \forall i, j \in I. \quad (12f)$$

In (12a), $\sum_{j=1}^N l_j$ describes the SPDD, and K/N describes the reduction degree. β is the trade-off factor to achieve a balance between the SPDD and reduction degree, while

simultaneously deciding the optimal clustering number K . The binary variable u_j indicates whether scenario ξ_j is selected as a representative scenario of a cluster, while the binary variable v_{ij} determines whether scenario ξ_i is included in the cluster represented by scenario ξ_j . Constraint $v_{ij} \leq u_j$ ensures that ξ_i can only be assigned to a cluster that has a designated representative, while $v_{jj} = u_j$ enforces that ξ_i must be assigned to its respective cluster if it's a representative scenario. Constraint (12d) ensures that each scenario ξ_j can only be assigned to one cluster, while (12e) guarantees that exactly K clusters are formed. The weight of cluster C_k is calculated as $\omega_k = \sum_{i=1}^N v_{ik} \gamma_i$.

F. Problem-Driven Evaluation Indices

In this section, two types of problem-driven evaluation indices are introduced: *ex-ante* and *ex-post* indices. *Ex-ante* indices emphasize the SR's ability in partitioning and representing the original scenario set in the problem space before solving the reduced problem. *Ex-post* indices focus on the impacts of SR on the outcomes of the TSSO after solving the reduced problem. For the following indices, the first two indices are *ex-ante* indices, while the last two indices are *ex-post* indices.

1) Sum of PDD within clusters (SPDD):

$$\text{SPDD} := \sum_{j=1}^N l_j = \sum_{k=1}^K \sum_{i \in C_k} \gamma_i (F_{ki} - F_{ii} + F_{ik} - F_{kk}). \quad (13)$$

SPDD measures the dispersion between scenarios and their respective clusters. A smaller SPDD value indicates a tighter clustering result in the problem space.

2) Problem-driven Davies-Bouldin Index: Based on the Davies-Bouldin Index, we introduce the Problem-driven Davies-Bouldin Index (PDDBI) utilizing the PDD:

$$\text{PDDBI} := \frac{1}{K} \sum_{m=1}^K \max_{1 \leq n \neq m \leq K} \left(\frac{D_m + D_n}{d(\zeta_m, \zeta_n)} \right) \quad (14a)$$

$$D_m := \sum_{i \in C_m} \frac{\gamma_i}{\omega_m} d(\zeta_m, \xi_i). \quad (14b)$$

A smaller PDDBI value indicates a better quality of the balance between the within-cluster compactness and the between-cluster separation.

3) Optimality gap: After solving (1) and (2), we apply z_{ξ}^* and z_{ζ}^* to ξ , and calculate the percentage value of OG as

$$OG_{\zeta}(\%) := \frac{F(z_{\zeta}^*, \xi) - F(z_{\xi}^*, \xi)}{F(z_{\xi}^*, \xi)}. \quad (15)$$

This metric indicates the percentage deviation of the approximated optimality from the optimality of the original problem, and is desired to be close to zero ¹.

¹The feasibility for unexpected scenario realizations or outliers is ensured by the assumption of *relatively complete recourse* and the availability of corrective measures in the second stage.

4) *Representative scenario effectiveness*: For the SBSO based on the representative scenarios, identifying the relative importance of individual representative scenario is crucial for comprehending the problem structure and making reasonable decisions. We introduce the concept of “Scenario Effectiveness”, measuring the significance of a given representative scenario in the problem space. The scenario effectiveness of scenario ζ_k , denoted as $SE_{\zeta_k}(\%)$, is characterized by the changes in the percentage OG upon its removal from ζ :

$$SE_{\zeta_k}(\%) := OG_{\zeta_{-k}}(\%) - OG_{\zeta}(\%), \quad (16)$$

where $\zeta_{-k} = \zeta \setminus \{\zeta_k\}$. A higher value of $SE_{\zeta_k}(\%)$ signifies that the removal of ζ_k induces more substantial changes in the OG . This indicates that ζ_k holds greater significance in influencing the reduced problem outcomes.

G. Algorithm of PDSR

The proposed PDSR framework is illustrated in **Algorithm 1**. This algorithm consists of two steps:

Step 1 projects the distribution space onto the problem space by calculating the mutual decision applicability between scenarios, and constructs the problem space matrix F . In the projection process, the correlations between scenarios are inherently accounted for by leveraging the problem structure, making complex correlation analysis unnecessary. Notably, with stronger correlations between the original scenarios, the number of reduced scenarios is reduced.

Step 2 partitions the original scenario set into clusters and selects representative scenarios by first tuning β from *ex-ante* indices, and then solving the MILP in (12). Note that β depends on the specific problem formulation and the original scenario set. Thus, simple tuning is suggested to obtain an acceptable K .

Algorithm 1: Problem-Driven Scenario Reduction

Input: Scenario set ξ of N scenarios and weights γ .

Output: Scenario set ζ of K scenarios and weights ω .

Step 1 - Projection in Problem Space

```

Step 1.1: Initialize problem space matrix  $F = \mathbf{0}$ .
Step 1.2: Project the distribution space onto
          problem space by:
for  $i = 1$  to  $N$  do
  Solve the scenario-specific TSSO problem
   $F(z, \xi_i)$  and obtain the optimal decision  $z_{\xi_i}^*$ .
  Set  $F_{ii} = F(z_{\xi_i}^*, \xi_i)$ .
  for  $j = 1$  to  $N$ ,  $i \neq j$  parallel do
    Solve the single-stage and deterministic
    problem  $G(y_j, \xi_j | z_{\xi_i}^*)$  in (1b).
    Set  $F_{ij} = F(z_{\xi_i}^*, \xi_j)$ .
  end
end

```

Step 2 - Clustering

```

Step 2.1: Select  $\beta$  in (12a) from ex-ante indices.
Step 2.2: Solve MILP in (12) to obtain the
          representative scenario set  $\zeta$  with weights  $\omega$ .

```

To validate the effectiveness of the proposed PDSR framework, we consider applying PDSR to the following two-stage stochastic economic dispatch problem.

III. TWO-STAGE STOCHASTIC ECONOMIC DISPATCH FOR ACTIVE DISTRIBUTION NETWORKS

In this section, we consider an optimal stochastic economic dispatch of an active distribution network (ADN) that trades with the transmission system. The ADN's assets include wind turbines (WT), photovoltaic systems (PV) and energy storage (ES) facilities. We focus on uncertainties from WT, PV, loads and two electricity markets: day-ahead and intraday price [35], which engenders the two stages for dispatch. In the day-ahead stage, uncertainties are described using the representative scenario set and the ADN needs to sign contracts for power trading with the transmission system operator and procure ES capacity from the ES owner to address intraday uncertainties. In the intraday stage, due to the uncertainties, there will exist deviations between the scheduled day-ahead power trading and the actual intraday power demand, and even violations of safety constraints especially in worst-case scenarios. Therefore, the ADN decision-maker is risk-averse and prefers to make strategy adjustments to limit constraint violations, which are mainly considered as voltage magnitude constraint violations.

A. Objective Function

The objective is to minimize the total operation cost including the day-ahead trading cost and the expected intraday balancing cost and the penalty cost. The day-ahead trading cost in (17b) includes the cost of trading power with transmission system and the procurement of ES capacity. The intraday cost includes the balancing cost in the intraday balancing market in (17c), and the penalty cost of load shedding and RES curtailment in (17d).

$$\min C^{\text{DA}} + C^{\text{IN,im}} + C^{\text{IN,p}} \quad (17a)$$

$$C^{\text{DA}} = \sum_{s=1}^S \omega_s \sum_{t=1}^T \pi_{s,t}^{\text{T}} P_t^{\text{T}} \Delta t + \sum_{j \in \Omega_E} \pi_j^{\text{E}} E_j \quad (17b)$$

$$C^{\text{IN,im}} = \sum_{s=1}^S \omega_s \sum_{t=1}^T \Delta t (\pi_{s,t}^{\text{T+}} P_{s,t}^{\text{T+}} + \pi_{s,t}^{\text{T-}} P_{s,t}^{\text{T-}}) \quad (17c)$$

$$C^{\text{IN,p}} = \sum_{s=1}^S \omega_s \sum_{t=1}^T \Delta t \left(\sum_{j \in \Omega_R} \pi_{j,s,t}^{\text{R,c}} P_{j,s,t}^{\text{R,c}} + \sum_{j \in \Omega_L} \pi_{j,s,t}^{\text{L,s}} P_{j,s,t}^{\text{L,s}} \right). \quad (17d)$$

$\pi_{s,t}^{\text{T}}$ and P_t^{T} are the trading electricity price and trading power between the ADN and transmission system in the day-ahead market, respectively. π_j^{E} and E_j are the procurement price and procured ES capacity at node j in the day-ahead market. T and Δt are time periods and time interval for scheduling. S is the number of scenarios and ω_s is the weight of scenario s . $\pi_{s,t}^{\text{T+}} / \pi_{s,t}^{\text{T-}}$ are the imbalancing price of up-regulation and down-regulation in the intraday balancing market under scenario s and time t . $P_{s,t}^{\text{T+}} / P_{s,t}^{\text{T-}}$ are the imbalanced purchasing and selling power in the intraday balancing market.

In the intraday balancing market, the ADN can only purchase balancing energy at a higher price than in the day-ahead market, while selling electricity at a lower price. $P_{j,s,t}^{\text{R,c}}/P_{j,s,t}^{\text{L,s}}$ are the power of RES curtailment and load shedding at node j . $\Omega_{\text{R}}/\Omega_{\text{L}}/\Omega_{\text{E}}$ refer to the set of nodes of RES, load, and ES. $\pi^{\text{R,c}}/\pi^{\text{L,s}}$ are the penalty costs of RES curtailment and load shedding.

B. Operational Constraints

1) *Power flow constraints*: The linear version of the *Dist-Flow* model, i.e., *LinDistFlow* [36] is used in this paper to approximate nodal voltage magnitudes and active/reactive line flows in the ADN with the assumption that line losses can be neglected. We denote Ω_{N} as the set of nodes. $\forall j \in \Omega_{\text{N}}$, we have

$$\mathcal{V}_{j,s,t} = \mathcal{V}_{i,s,t} - 2(r_{ij}P_{ij,s,t} + x_{ij}Q_{ij,s,t}) \quad (18a)$$

$$p_{j,s,t} = P_{ij,s,t} - \sum_{l:j \rightarrow l} P_{jl,s,t} \quad (18b)$$

$$q_{j,s,t} = Q_{ij,s,t} - \sum_{l:j \rightarrow l} Q_{jl,s,t} \quad (18c)$$

$$\underline{\mathcal{V}}_j \leq \mathcal{V}_{j,s,t} \leq \overline{\mathcal{V}}_j, \quad (18d)$$

where r_{ij}/x_{ij} are the line resistance/reactance of line ij , respectively. $\mathcal{V}_{j,s,t}$ denotes the squared voltage magnitude at node j for scenario s at time t . $P_{ij,s,t}/Q_{ij,s,t}$ are the line active/reactive power of line ij , respectively. $p_{j,s,t}/q_{j,s,t}$ are the active/reactive injection power at node j . (18a) describes the voltage drop over line ij . (18b) and (18c) represent the active and reactive power balance at node j . (18d) describes voltage magnitude limits at node j , with $\overline{\mathcal{V}}_j/\underline{\mathcal{V}}_j$ as the upper/lower bound of squared voltage magnitude.

2) *RES curtailment and load shedding constraints*: $\forall j \in \Omega_{\text{R}}$, we have

$$0 \leq P_{j,s,t}^{\text{R,c}} \leq P_{j,s,t}^{\text{R}} \quad (19a)$$

$$0 \leq P_{j,s,t}^{\text{L,s}} \leq P_{j,s,t}^{\text{L}} \quad (19b)$$

where $P_{j,s,t}^{\text{R}}/P_{j,s,t}^{\text{L}}$ are the RES injection and active power consumption at node j for scenario s at time t , respectively.

3) *ES operation constraints*: $\forall j \in \Omega_{\text{E}}$, we have

$$SoC_{j,s,t+1} = SoC_{j,s,t} + \Delta t(P_{j,s,t}^{\text{E,c}}\eta_j^{\text{c}} - P_{j,s,t}^{\text{E,d}}/\eta_j^{\text{d}})/E_j \quad (20a)$$

$$\sum_{t=1}^T (P_{j,s,t}^{\text{E,c}}\eta_j^{\text{c}} - P_{j,s,t}^{\text{E,d}}/\eta_j^{\text{d}})\Delta t = 0 \quad (20b)$$

$$0 \leq P_{j,s,t}^{\text{E,c}} \leq (1 - D_{j,s,t}^{\text{E}})\overline{P_j^{\text{E}}} \quad (20c)$$

$$0 \leq P_{j,s,t}^{\text{E,d}} \leq D_{j,s,t}^{\text{E}}\overline{P_j^{\text{E}}} \quad (20d)$$

$$\underline{SoC} \leq SoC_{j,s,t} \leq \overline{SoC} \quad (20e)$$

$$D_{j,s,t}^{\text{E}} \in \{0, 1\}, \quad (20f)$$

where $P_{j,s,t}^{\text{E,c}}/P_{j,s,t}^{\text{E,d}}$ are the charge/discharge power of ES at node j . $\eta_j^{\text{c}}/\eta_j^{\text{d}}$ are the charge/discharge efficiency, respectively. $\overline{P_j^{\text{E}}}$ is the maximum charging/discharging power. $D_{j,s,t}^{\text{E}}$ is a binary variable indicating the charging/discharging state. $\underline{SoC}/\overline{SoC}$ denote the maximum/minimum SoC, respectively. Constraints (20a), (20b) and (20e) are related to the state of

charge (SoC). Constraint (20b) guarantees that the capacity at the last time period is equal to the initial capacity. Constraints (20c) and (20d) impose restrictions on the maximum charging/discharging power and charging state of ES.

4) Trading constraints with transmission system:

$$0 \leq P_{s,t}^{\text{T+}} \leq (1 - D_{s,t}^{\text{T}})\overline{P^{\text{T}}} \quad (21a)$$

$$0 \leq P_{s,t}^{\text{T-}} \leq D_{s,t}^{\text{T}}\overline{P^{\text{T}}} \quad (21b)$$

$$-\overline{P^{\text{T}}} \leq P_t^{\text{T}} \leq \overline{P^{\text{T}}} \quad (21c)$$

$$-\overline{P^{\text{T}}} \leq P_t^{\text{T}} + P_{s,t}^{\text{T+}} - P_{s,t}^{\text{T-}} \leq \overline{P^{\text{T}}} \quad (21d)$$

$$D_{s,t}^{\text{T}} \in \{0, 1\}, \quad (21e)$$

where $\overline{P^{\text{T}}}$ is the maximum trading power between ADN and transmission system. (21a) and (21b) indicate that the ADN can only be in one balancing state at one time, with the binary variable $D_{s,t}^{\text{T}}$ as the balancing state of ADN.

5) Energy balancing constraints: $\forall j \in \Omega_{\text{N}}$, we have

$$p_{j,s,t} = P_{j,s,t}^{\text{E,c}} - P_{j,s,t}^{\text{E,d}} + P_{j,s,t}^{\text{L}} - P_{j,s,t}^{\text{L,s}} - (P_{j,s,t}^{\text{R}} - P_{j,s,t}^{\text{R,c}}) \quad (22a)$$

$$q_{j,s,t} = Q_{j,s,t}^{\text{L}} - Q_{j,s,t}^{\text{L,s}}, \quad (22b)$$

where $Q_{j,s,t}^{\text{L}}/Q_{j,s,t}^{\text{L,s}}$ are the reactive power consumption and load shedding at node j . We assume that all the RESs are of the unity power factor and the power factor of load demand remains the same after load shedding. $\forall s, t$, if $j \notin \Omega_{\text{E}}$, $P_{j,s,t}^{\text{E,c}} = P_{j,s,t}^{\text{E,d}} = 0$. Similarly, if $j \notin \Omega_{\text{R}}$, $P_{j,s,t}^{\text{R}} = P_{j,s,t}^{\text{R,c}} = 0$.

C. Overall Problem Formulation

Finally, the optimal day-ahead economic dispatch problem is formulated as (23), which is a mixed-integer linear problem, and can be solved by commercial solvers.

$$\begin{aligned} \min_{\Xi_1, \Xi_2} \quad & C^{\text{DA}} + C^{\text{IN,im}} + C^{\text{IN,p}} \\ \text{s.t.} \quad & (18) - (22), \end{aligned} \quad (23)$$

where $\Xi_1 = [P_t^{\text{T}}, E_j]$ are the decision variables of first stage (i.e., day-ahead stage). $\Xi_2 = [P_{s,t}^{\text{T+}}, P_{s,t}^{\text{T-}}, P_{j,s,t}^{\text{E,c}}, P_{j,s,t}^{\text{E,d}}, P_{j,s,t}^{\text{R,c}}, P_{j,s,t}^{\text{L,s}}, D_{j,s,t}^{\text{T}}]$ are the decision variables of second stage (i.e., intraday stage).

IV. NUMERICAL CASE STUDIES

In this section, we conduct case studies on two TSSO problems to illustrate PDSR's priority. We mainly concentrate on the first case study, which involves a two-stage stochastic economic dispatch of a modified IEEE 33-node ADN introduced in Section III. Then, we provide a remark of applying PDSR to mixed integer problems such as unit commitment.

A. Problem Description

In the first case study on the modified IEEE 33-node ADN, one WT is located at node 10 and two PVs are located at nodes 16 and 24, respectively. Beside of the uncertain RES power output, the active load at nodes 10, 16, and 24 is also considered as random variables, while the active load at other nodes is assumed to have a fixed changing curve for

simplicity. An ES is located at node 13. The voltage magnitude is restricted as $|V_i| \in [0.90, 1.10]$ (p.u.), $\forall i \in \Omega_N$. The time step is set as $\Delta t = 15$ min with $T = 96$ steps. The original scenario set ξ comprises N scenarios. It is worth mentioning that, though PDSR does not require the specific scenario generation method, ξ should be generated with consideration of the SBSO problem. For example, historical observations can be used for long-term problems (e.g., system planning), while forecasting methods are more suitable for short-term problems (e.g., economic operation). Additionally, machine learning techniques (e.g., generative adversarial networks [37]) can be employed to enhance data quality. In this paper, we generate ξ from the forecasting results of a real distribution network. Additionally, to validate the performance of proposed PDSR framework in uncovering salient scenarios, particularly worst-case scenarios, we construct ξ by randomly selecting from the forecasting result, while also ensuring that it contains a specified number of bad scenarios [38]. Each individual scenario $\xi \in \mathbb{R}^{7 \times T}$, $\xi \in \xi$ is a multi-variable high-dimensional vector characterizing 7 sources of uncertainty, including the power of 1 WT, 2 PVs, 3 loads and the day-ahead electricity price. The capacities of WT and PVs are normalized to 1MW, 1.2MW and 1MW. For simplicity, the intraday balancing market price is set as $\pi_{s,t}^{T+} = 1.3\pi_{s,t}^T$ and $\pi_{s,t}^{T-} = 0.7\pi_{s,t}^T$. The power rating of ES is set as 0.4MW, and ES capacity procurement is limited by $E \leq 0.8$ MWh. The initial *SoC* of ES is 0.5 and $\eta^c = \eta^d = 0.95$. The penalty costs of load shedding and RES curtailment are set as \$1000/MWh and \$280/MWh, respectively. The TSSO problem built on ξ is used as the Benchmark. The optimization is coded in Python with the Cvxpy interface and solved by Gurobi 11.0 solver. The programming environment is Intel Core i9-13900HX @ 2.30GHz with RAM 16 GB.

B. Performance of PDSR

First, we consider $N = 400$ scenarios and construct the F matrix by solving N^2 scenario-specific deterministic problems. Then, we utilize the MILP formulation in (12) to decide K through a comprehensive analysis of the normalized *ex-ante* indices under different β . Notably, the simple tuning process of β can be completed within a few optimizations, each requires less than 3 seconds. The results are shown in Fig. 2. It is observed that $\beta \in [120, 140]$ and $K = 6$ correspond to a local minimum in PDDBI, and the balance between reduction degree K/N and SPDD is also achieved. This indicates that 6 representative scenarios can provide the most favorable clustering structure for the dataset under analysis.

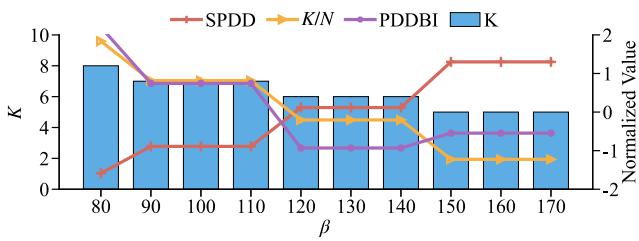


Fig. 2. The *ex-ante* validity of different β .

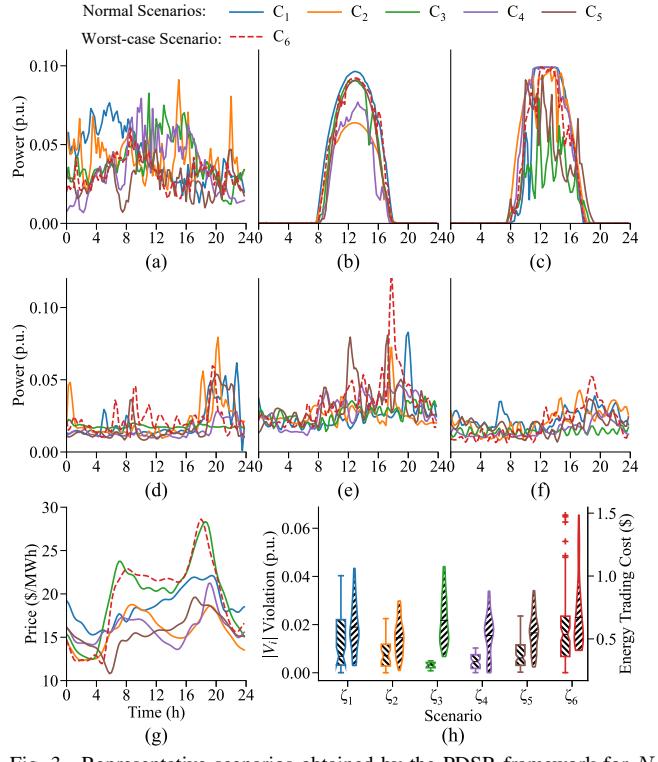


Fig. 3. Representative scenarios obtained by the PDSR framework for $N = 400$ and $K = 6$: (a) WT at node 10, (b) PV at node 16, (c) PV at node 24, (d) load at node 10, (e) load at node 16, (f) load at node 24, (g) day-ahead electricity price and (h) un-optimization voltage magnitude violations (boxplot, left) and energy trading cost (violinplot, right) of the 6 scenario clusters from AC power flow.

The obtained $K = 6$ representative scenario clusters of the proposed PDSR framework are presented in Fig. 3(a) to Fig. 3(g). Fig. 3(h) indicates the un-optimization voltage magnitude violations and energy trading cost of the 6 scenario clusters from AC power flow. This highlights the severity of these scenarios and underscores the importance of implementing reasonable dispatching practices. Curves in the same color belong to the same cluster. The corresponding weights of each cluster are 0.1575, 0.165, 0.1825, 0.175, 0.1425 and 0.1775, respectively. The representative scenarios in Fig. 3 illustrate PDSR's effectiveness in identifying salient features from a large set of uncertainties in the system. Moreover, PDSR includes one worst-case scenario (red ζ_6) in the reduced set. Specifically, worst-case scenarios are identified by the outliers in the problem space, which are detected by the condition $\varrho = \Delta^2(\text{sort}(\{\sum_{j=1}^N F(z_{\xi_j}^*, \xi_i)\}_{i=1}^N)) > \text{bound}$. $\text{sort}(\cdot)$ is the ascending sorting function, $\Delta^2(\cdot)$ is the second-order difference function, and bound is the threshold value, which can be set as 2 in this case. $\sum_{j=1}^N F(z_{\xi_j}^*, \xi_i)$ describes the decision adaptability level of all other potential solutions to scenario ξ_i . Scenario ζ_6 achieves a high value of ϱ and is identified as a worst-case scenario. Besides, its significant volatility and high values, and significant voltage magnitude violations and trading cost also verify it as a worst-case scenario. This observation is critical because incorporating too many worst-case scenarios into the representative scenario set may introduce conservatism, potentially leading to reduced

economic efficiency.

The optimality gap $OG(\%)$ is 0.09%, which suggests a high level of approximation accuracy. Notably, for all F_{ij} in \mathbf{F} , the MIP-gap is $< 10^{-4}$ (i.e., default Gurobi *MIPGap*), which ensures that the optimal solution is found despite the adopted case study being a non-convex MILP problem. Finding the optimal solution contributes to achieving high approximation accuracy and low OG. We utilize the evaluation indices of representative scenario effectiveness to further validate the performance of the proposed PDSR framework. The results are illustrated in Table I.

TABLE I
EVALUATION RESULTS OF REPRESENTATIVE SCENARIO EFFECTIVENESS
FOR $N = 400$ AND $K = 6$

	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5	ζ_6
SE_{ζ_k}	0.29%	0.25%	0.27%	0.28%	0.02%	0.33%

Table I suggests that each component of the representative scenario set has a relatively large value of SE_{ζ_k} , indicating their significant impact on the reduced problem outcomes. Notably, the scenario effectiveness of representative scenario ζ_6 , corresponding to the red curve in Fig. 3, is highlighted. This is consistent with the earlier analysis of the scenario being a worst-case scenario and an important component of the representative scenario set. This result demonstrates the effectiveness of the proposed PDSR framework in capturing the salient features of the original scenario set.

C. Comparison with State-of-the-Art SR Methods

To further demonstrate the benefit of the proposed PDSR framework, we conduct a comparative analysis. DDSR methods, including HC using Wasserstein distance (HC-W), K -means using Euclidean distances (KM-E), K -means using dynamic time warping distance (KM-D), K -medoids using Euclidean distances (KD-E), and Gaussian mixture model using Mahalanobis distance (G-M) are compared. For DDSR methods incorporating relevant problem properties, we include K -means based on network power flow (KM-pf) [18], and HC based on operational cost (HC-c) [19] in comparison. Regarding other problem-dependent SR research, we also include the method developed in [31] with graph clustering (GC), and the method developed in [30] using Wasserstein distance metric and alternating minimization algorithm (AM-W) for comparison. Additionally, the method of constructing the representative scenario set based on worst-case statistical indicators in the distribution space (WS) is also included for comparison.

The comparative indices include the *ex-post* SR performance indices and the computation efficiency indices. In this comparison, we concentrate on how the representative scenario set impact the problem outcomes. Thus, the considered SR performance indices include: the number of worst-case scenarios captured in the representative scenario set (κ), which indicates the SR method's effectiveness in identifying “worst-case” scenarios; the $OG(\%)$ defined in (15), which measures the

problem optimality approximation accuracy; the ES capacity procurement E from solving (2), which measures the ability of SR methods in capturing the underlying risk characteristics in the original scenario set and making risk provisions; the average verification penalty cost $\tilde{C}^{IN,P} = 1/N \sum_{\xi \in \mathcal{E}} C_{\xi}^{IN,P}$, where the penalty cost $C_{\xi}^{IN,P}$ is derived from $F(z_{\zeta}^*, \xi)$. The considered computation efficiency indices include: the time required to process the data input for clustering (τ_p), and the time required to solve the clustering problem (τ_c). The comparison results are presented in Table II.

TABLE II
COMPARING SR METHODS FOR $N = 400$ AND $K = 6$

Method	κ	$OG(\%)$	$E(\text{MWh})$	$\tilde{C}^{IN,P}(\$)$	$\tau_p(s)$	$\tau_c(s)$
Benchmark	12	0	0.17	155.1	—	—
PDSR	1	0.09	0.16	162.8	< 0.3	2.73
AM-W [30]	1	0.51	0.07	199.3	< 0.3	22.9
GC [31]	1	0.57	0.06	206.9	< 0.3	0.11
WS	3	0.83	0.58	67.53	0.01	—
HC-c [19]	0	1.70	0	250.6	< 0.8	0.01
KM-pf [18]	0	0.84	0.04	221.3	< 0.8	0.59
G-M	0	1.59	0.01	247.5	—	4.62
KM-D	0	1.71	0	250.6	—	266.5
KM-E	0	0.98	0.03	227.8	—	0.31
KD-E	0	0.77	0.04	217.7	—	1.46
HC-W	0	0.66	0.04	214.9	19.54	0.03

Optimality Gap: The observations from Table II suggest that the PDSR framework, with a small value of $OG(\%) = 0.09\%$, significantly outperforms other state-of-the-art SR methods. Besides, the ES capacity procurement E and $\tilde{C}^{IN,P}$ of PDSR also best approximate the results of Benchmark. The DDSR methods, seeking minimum statistical difference, fail to capture worst-case scenarios in their representative scenario sets, which leads to a neglect of potential risks during the operation, resulting in small procurements of ES capacity and an inability to cope with uncertainties, thus achieving relatively high penalty costs and $OG(\%)$. For instance, in heavy load situations, the node voltage might drop below safety requirements. Without the support of ES, the ADN must resort to lots of load shedding to prevent violating the voltage safety constraints, thereby incurring substantial penalty costs. WS selects 6 statistical worst-case scenarios in the distribution space, but only 3 of them are real worst-case scenarios in the problem space. This discrepancy highlights that severity in statistical metrics does not necessarily equate to severity in problem outcomes. Moreover, focusing only on worst-case scenarios may result in overly conservative decisions and unnecessarily high costs. WS procures too much ES capacity as $E = 0.58\text{MWh}$, achieving a low penalty cost, while also leading to a relatively high $OG(\%) = 0.83\%$. Regarding other problem-dependent SR research, The GC method in [31] captures 1 worst-case scenario with $OG(\%) = 0.51\%$, while the AM-W method in [30] also captures 1 worst-case scenario with $OG(\%) = 0.57\%$. Those results are much lower than the

DDSR methods. This indicates that measuring the difference between scenarios by the symmetric opportunity cost can enhance the SR performance. However, GC relies on heuristic methods to derive representative scenarios, whereas AM-W utilizes the alternating minimization algorithm, both of which can potentially yield suboptimal outcomes. Additionally, the method developed in [32] fails to solve the clustering process within 3 hours as their clustering methodology does not scale well with high-dimensional power system problems with complex problem structure and large scenario sets. Compared to the above methods, the proposed PDSR framework efficiently considers the potential impacts of the scenarios on the problem, and include one reasonable worst-case scenario in the representative scenario set, as analyzed in Section IV-B. These comparative findings suggest that PDSR exhibits superior accuracy in representing the original scenario set, thereby offering more reliable information for decision-making in power system energy management under uncertainty.

Representative Scenario Effectiveness: In Fig. 4, the comparative results of representative scenario effectiveness are depicted. For DDSR methods, removing some of the representative scenarios does not significantly impact the problem outcomes, as evidenced by the minimal changes in $OG(\%)$. While in PDSR, such a removal can significantly alter the problem outcomes. This indicates that the proposed PDSR framework can effectively capture the salient scenarios with significant impacts on the SBSO problem. Furthermore, all $OG_{\zeta-k}(\%)$ results of PDSR are much lower than the DDSR methods, indicating that statistically proximity in the distribution space does not equal to strategy closeness and better solution approximation in the problem space. Besides, the comparison between GC, AM-W and PDSR indicates that the proposed MILP clustering methodology is more effective than the graph clustering employed in GC and the alternating minimization algorithm in AM-W.

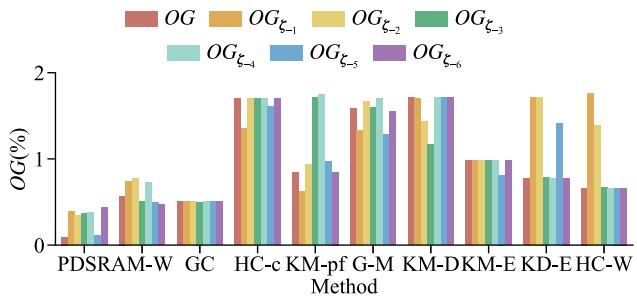


Fig. 4. Comparative results of representative scenario effectiveness for $N = 400$ and $K = 6$.

Computational Efficiency: The Benchmark struggles with computational complexity when N is large. For example, for $N = 400$ and $N = 600$, the Benchmark takes nearly 51 and 75 mins, respectively. Moreover, for $N > 600$, computation times become impractical for the Benchmark. The DDSR methods, however, overcome computation bottlenecks, but at the price of potentially large $OG(\%)$ values. Problem-dependent SR research, GC, AM-W and PDSR decompose the original SBSO problem into mostly parallelizable and simple scenario-specific subproblems. In this way, the computation

complexity is greatly reduced. Specifically, all GC, AM-W and PDSR involve N parallel computations followed by $N(N-1)$ parallel computations to determine matrix \mathbf{F} . In this paper, the computation time for each scenario-specific subproblem is $\tau_s < 0.3s$ (with LinDistFlow model). Theoretically, with $N(N-1)$ parallel processes available, the computation time for calculating the \mathbf{F} matrix can be reduced to $2 \times \tau_s = 0.6s^2$. For $N = 400$ and $K = 6$, after calculating the \mathbf{F} matrix, the MILP clustering problem is solved consistently with MIP-gap with $\tau_c < 3s$. We define $\tau_o(K)$ as the time required to solve (2) with K representative scenarios and $\tau_o(6) < 3s$. Generally, with W parallel processors, the total computation time of the PDSR is

$$\begin{aligned} \tau &:= \tau_p + \tau_c + \tau_o(K) \\ &= ([N/W] + [N(N-1)/W]) \tau_s + \tau_c + \tau_o(K). \end{aligned}$$

For example, with $N = 400$ and a practical $W = 100$ processors, $\tau \approx 8$ mins (or 15.7% of the Benchmark). Furthermore, as long as the original scenario set and SBSO problem formulation remain unchanged, matrix \mathbf{F} can be stored and re-used. In conclusion, the PDSR framework can effectively reduce computational complexity while achieving a low SR optimality gap.

Scalability: We further compare the $OG(\%)$ results under different N and K , as illustrated in Fig. 5. In Figs. 5(a,b), we observe that the proposed PDSR framework outperforms DDSR methods and GC across all values of N and K , as evidenced by its consistently lower $OG(\%)$. This underscores the distinct superiority of the PDSR framework in identifying salient scenarios. Interestingly, the PDSR results (in red) with $N = 100$ have smaller $OG(\%)$ than the DDSR methods even for $N = 400$, which means that to attain a comparable level of OG, the PDSR framework requires far fewer scenarios. Besides, as expected, increasing K decreases the $OG(\%)$ for PDSR. However, for DDSR methods, GC and AM-W, this inherent regular benefit is not present. Furthermore, we conduct the performance comparison with a large original scenario set, using $N = 1000$ and $K = 10$ (i.e., a 99% reduction) as an example in Fig. 5(c). In this case, the original SBSO problem in (1) is computationally intractable. Therefore, we use $F(z_{\zeta}^*, \xi)$ as the performance evaluation metric, instead of $OG(\%)$. A smaller value of $F(z_{\zeta}^*, \xi)$ indicates a better approximation to the original SBSO optimality. Notably, PDSR outperforms other SR methods with the lowest $F(z_{\zeta}^*, \xi)$. Based on the representative scenarios from PDSR, the day-ahead SBSO problem obtains superior operational strategies with high adaptability and effectiveness across different scenarios. Moreover, for $N = 1000$ and $K = 10$, we have $\tau_c = 45.8s$, indicating that the MILP formulation can be solved efficiently. In conclusion, the above comparative analysis further emphasizes the benefit of the proposed PDSR framework.

Remark. [PDSR for mixed-integer problems] Consider an example of a two-stage stochastic unit commitment problem applied to the IEEE 14-bus system, where nearly half of the

²Replacing LinDistFlow with its more accurate second-order conic relaxation, \mathbf{F} can still be calculated in parallel within $2 \times \tau_s = 3.2s$.

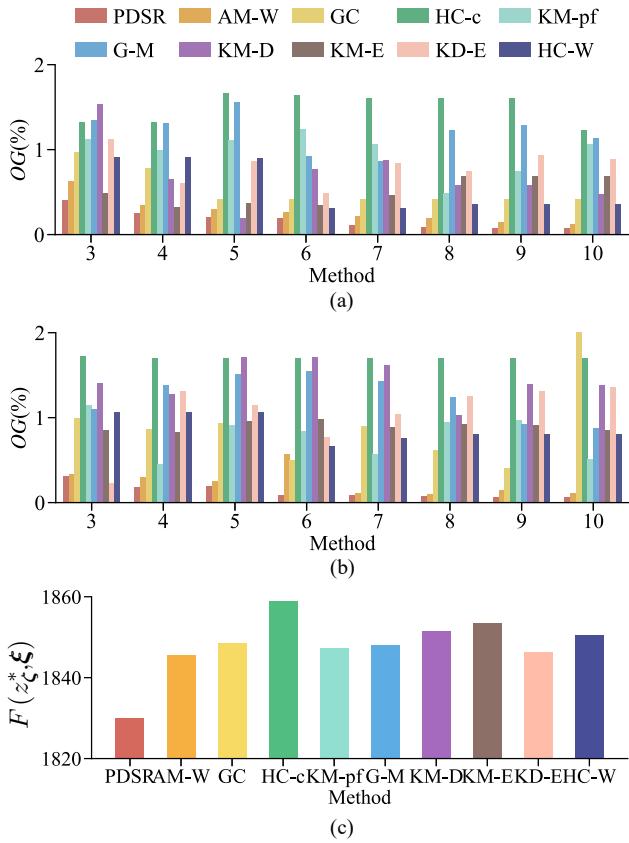


Fig. 5. Comparing results for different N and K : (a) $N = 100$, (b) $N = 400$ and (c) $N = 1000$ and $K = 10$.

decision variables are binary. In this example, five generators are available and one WT is located at bus 10, and two PVs are located at buses 12 and 14. The time step is set as $\Delta t = 1$ hour with $T = 24$ steps. The uncertainties stem from the RES power output and the load demand at the RES nodes, resulting in six distinct sources of uncertainty. The original scenario set ξ contains $N = 400$ scenarios and is reduced to $K = 5$ representative scenarios. The detailed problem formulation is presented in Appendix B. To be concise and comprehensive, we mainly show the comparison results of PDSR with other SR methods. The comparison indices include the number of worst-case scenarios captured κ , the optimality gap $OG(\%)$, the intraday operational cost $\tilde{C}^{IN}(M\$) = 1/N \sum_{\xi \in \xi} (\tilde{C}^{IN,g} + \tilde{C}^{IN,p})$ ($\tilde{C}^{IN,g}$ and $\tilde{C}^{IN,p}$ are from $F(z^*_{\xi}, \xi)$), as well as the preprocessing time $\tau_p(s)$, and clustering time $\tau_c(s)$. The comparative results are summarized in Table III.

Table III indicates that the proposed PDSR framework achieves the lowest $OG(\%) = 0.20\%$, which is significantly lower than the other SR methods. Additionally, the intraday operational cost \tilde{C}^{IN} of PDSR closely approximates the Benchmark results, demonstrating PDSR's effectiveness in facilitating the decision-making process for generating robust and cost-effective operational strategies. Moreover, τ_p with parallel setting of PDSR is within acceptable limits. τ_c is quite different from the result in Table II but is still acceptable, which is because that the different problem structure lead to different solving time of (12). The superiority of PDSR on the

mixed integer problem further demonstrates the effectiveness of the proposed PDSR framework.

TABLE III
COMPARING SR METHODS FOR MIXED INTEGER UC PROBLEM
($N = 400$ AND $K = 5$)

Method	κ	$OG(\%)$	$\tilde{C}^{IN}(M\$)$	$\tau_p(s)$	$\tau_c(s)$
Benchmark	31	0	773.3	—	—
PDSR	2	0.20	712.3	< 0.1	18.62
AM-W [30]	2	0.36	671.6	< 0.1	1.26
GC [31]	1	0.42	623.1	< 0.1	0.36
WS	3	1.52	783.4	0.01	—
HC-c [19]	0	0.52	650.6	< 0.7	0.01
KM-pf [18]	0	0.44	697.9	< 0.7	4.78
G-M	0	0.48	654.6	—	0.80
KM-D	0	0.69	650.5	—	8.94
KM-E	0	0.71	650.9	—	2.87
KD-E	0	0.69	651.6	—	1.23
HC-W	0	0.82	653.4	8.24	0.02

D. Summarizing Discussion on PDSR Framework

For SBSO problems, the definition of “representativeness” is essential to construct the representative scenario set and yield effective strategies. In this paper, we demonstrate, both through theoretical analysis and numerical validation, that the “representativeness” should be defined as the decision applicability of the representative scenario to its represented scenario cluster. The advantages of the proposed PDSR framework lie in the following aspects:

- Representativeness:** With the problem space constructed from decision applicability, the PDSR framework successfully achieves a low optimality gap, demonstrating a significant level of representativeness.
- Efficiency:** The PDSR framework successfully captures the scenarios with significant impacts on the problem, especially the worst-case scenarios, enhancing the robustness and reliability. Moreover, in comparison to other SR methods, the PDSR framework effectively reduces the number of required scenarios with the same level of optimality gap, which is beneficial in cases with limited monitoring and data accumulation.
- Determinateness:** Instead of using heuristic methods, the PDSR framework transforms the process of scenario partitioning and representative scenario selection into a MILP formulation, which attains deterministic and optimal outcomes with commercial solvers.
- Generality:** The proposed PDSR framework operates without reliance on probability distribution, and only has limited and reasonable assumptions on the SBSO problem structure. As a result, the PDSR framework can be applied to a broad range of SBSO problems, ensuring high generality and scalability.

Meanwhile, the potential limitations for PDSR lie in the following aspects:

- (i) *Scalability*: Large-scale power systems present scalability challenges for PDSR due to the curse of dimensionality. Firstly, to compute \mathbf{F} , PDSR requires $N + N(N - 1)$ parallel computations, which can be a significant computation burden. Second, the resulting N -by- N \mathbf{F} begets a large-scale MILP with $N(N + 1)$ binary variables (albeit with a sparse structure). Potential techniques for solving large-scale MILP include warm-start, neural networks [39], [40], and decomposition methods [41], [42]. Third, complex optimization frameworks (e.g., multi-stage stochastic optimization), further increase the computational complexity of scenario-specific subproblems.
- (ii) *Tractability*: PDSR depends on the resolution of scenario-specific subproblems to construct the \mathbf{F} matrix. Consequently, issues that cannot be resolved through a single scenario, such as the Conditional Value-at-Risk (CVaR) component within the objective function, are beyond PDSR's scope of solvability.

V. CONCLUSION

In this paper, a novel problem-driven scenario reduction (PDSR) framework is proposed for power system SBSO problems, which fully incorporates the problem structure into the SR process. Specifically, we utilize the mutual decision applicability to construct the problem space as the input for SR, and propose the problem-driven distance metric to measure the similarity of scenarios in problem space. That is, PDSR decomposes the original large-scale and complex optimization problem into independent and simpler scenario-specific subproblems, thereby significantly decreasing computational complexity. Thus, the presented PDSR framework obtains near-optimal approximation accuracy with just a few salient representative scenarios within acceptable computation time, as illustrated by extensive case studies: one ADN dispatch problem balancing operational energy storage capacity and economic costs, and another addressing a mixed-integer unit commitment problem. Moreover, a comprehensive comparative analysis with other SR methods is provided for different N and K values and demonstrates broadly the superior performance of our PDSR framework.

Future work will focus on further scaling Algorithm 1 by filtering the original scenario set for PDSR to reduce the size of the \mathbf{F} matrix necessary to guarantee the desired optimality gap. Additionally, PDSR will benefit from extending the analysis to characterize the impacts of local solutions arising from non-convex problems, such as with the AC optimal power flow. Lastly, we are interested in extending the PDSR framework to multi-stage stochastic problems and other SBSO problems (e.g., CVaR) relevant to power engineering.

APPENDIX

A. Proof of Proposition 1

In this part, we prove that the proposed problem-driven distance metric in (10) satisfies the required properties of (7).

Proof. C1) Consistency: First we notice that $\zeta_k = \xi_i \Rightarrow d(\zeta_k, \xi_i) = 0$. Conversely, given that both $c(\zeta_k, \xi_i)$, $c(\xi_i, \zeta_k)$ are nonnegative, $d(\zeta_k, \xi_i) = 0$ implies $c(\zeta_k, \xi_i) = c(\xi_i, \zeta_k) = 0$, thus $z_{\zeta_k}^* = z_{\xi_i}^*$. We reasonably require the problem to satisfy the assumption that $z_{\zeta_k}^* = z_{\xi_i}^* \Rightarrow \zeta_k = \xi_i$. This hypothesis rests on the premise that $F(z, \xi)$ is highly sensitive to variations in ξ at particular z , suggesting that identical solutions imply identical scenarios. This assumption depends on the problem structure and can be restrictive. For certain problems dissatisfy this assumption, we can adjust the PDD by simply incorporating a regularized scaled norm component ($\mu > 0$) as

$$\begin{aligned} \tilde{d}(\xi_i, \zeta_k) &= F(z_{\zeta_k}^*, \xi_i) - F(z_{\xi_i}^*, \xi_i) \\ &\quad + F(z_{\xi_i}^*, \zeta_k) - F(z_{\zeta_k}^*, \zeta_k) + \mu \|\zeta_k - \xi_i\|_2. \end{aligned} \quad (24)$$

In this case, $\tilde{d}(\xi_i, \zeta_k) = 0 \Rightarrow \zeta_k = \xi_i$ holds for any μ . In this paper, we set $\mu = 0$ and continue to use $d(\zeta_k, \xi_i)$ for brevity, but all the proofs and algorithms can be adapted for $\tilde{d}(\xi_i, \zeta_k)$.

C2) Symmetricity: From definition, $d(\zeta_k, \xi_i) = d(\xi_i, \zeta_k)$.

C3) Convergence: As $\delta \rightarrow 0$, ζ_k and ξ_i become arbitrarily close. Given the Lipschitz continuity of $F(z, \xi)$ with respect to ξ , it follows that $F(z_{\zeta_k}^*, \xi_i) \rightarrow F(z_{\zeta_k}^*, \zeta_k)$ and $F(z_{\xi_i}^*, \xi_i) \rightarrow F(z_{\xi_i}^*, \zeta_k)$. Consequently, $d(\zeta_k, \xi_i)$ tends to 0.

C4) Triangle inequality: Let $\lambda(\xi_i) = 2 \sup_{z \in Z} |F(z, \xi_i)|$, given $F(z, \xi)$ is bounded for all z and ξ . We have $F(z_{\zeta_k}^*, \xi_i) - F(z_{\xi_i}^*, \xi_i) < |F(z_{\zeta_k}^*, \xi_i)| + |F(z_{\xi_i}^*, \xi_i)| < \lambda(\xi_i)$, and similar for $\lambda(\zeta_k)$. Thus, we have $d(\zeta_k, \xi_i) < \lambda(\zeta_k) + \lambda(\xi_i)$. \square

B. Formulation of the Two-Stage Stochastic Unit Commitment Problem

The objective of the two-stage stochastic unit commitment problem is to minimize the total operation cost, which includes the day-ahead generation cost $C^{\text{DA,g}}$, intraday generation regulation cost $C^{\text{IN,g}}$, and intraday punishment cost $C^{\text{IN,p}}$ of RES curtailment and load shedding.

$$\min_{\Xi_1, \Xi_2} C^{\text{DA,g}} + C^{\text{IN,g}} + C^{\text{IN,p}} \quad (25a)$$

$$C^{\text{DA,g}} = \sum_{t=1}^T \sum_{g=1}^{N_G} (C_g^{\text{PG}} P_{g,t} + C_g^{\text{NL}} u_{g,t} + C_g^{\text{SC}}) \quad (25b)$$

$$C^{\text{IN,g}} = \sum_{s=1}^S \omega_s \sum_{t=1}^T \sum_{g=1}^{N_G} (C_g^+ P_{g,s,t}^+ + C_g^- P_{g,s,t}^-) \quad (25c)$$

$$C^{\text{IN,p}} = \sum_{t=1}^T \Delta t \left(\sum_{r=1}^{N_R} C_r^{\text{R,c}} P_{r,s,t}^{\text{R,c}} + \sum_{l=1}^{N_L} C_l^{\text{L,s}} P_{l,s,t}^{\text{L,s}} \right), \quad (25d)$$

where $P_{g,t}$ is the power output of generator g at time t , $u_{g,t}$ is the binary variable indicating whether the generator is on ($u_{g,t} = 1$) or off ($u_{g,t} = 0$). $P_{g,s,t}^+ / P_{g,s,t}^-$ are the up/down regulation power of generator g in scenario s at time t , $C_g^{\text{PG}} / C_g^{\text{NL}}$ are the linear generation cost coefficients of generator g , C_g^+ / C_g^- are the generator up/down regulation costs, C_g^{SC} is the start-up and shut-down cost. $P_{r,s,t}^{\text{R,c}} / C_r^{\text{R,c}}$ are the curtailed power of RES farm r and the associated penalty cost. $P_{l,s,t}^{\text{L,s}} / C_l^{\text{L,s}}$ are the load shedding of load demand l and

the associated penalty cost. ω_s is the probability of scenario s . $N_G/N_L/N_R$ are the number of generators/regular loads/RES farms.

The constraints are as follows:

$$P_{ij,s,t} = B_{ij}(\theta_{i,s,t} - \theta_{j,s,t}) \quad (26a)$$

$$(P_{g,t} + P_{g,s,t}^+ - P_{g,s,t}^-) + (P_{r,s,t}^R - P_{r,s,t}^{R,c}) - (P_{l,s,t}^L - P_{l,s,t}^{L,s}) \quad (26b)$$

$$= -\left(\sum_{j \rightarrow i} P_{ji,s,t} - \sum_{i \rightarrow k} P_{ik,s,t}\right) \quad (26b)$$

$$\theta_{1,s,t} = 0, -1/3\pi \leq \theta_{i,s,t} \leq 1/3\pi \quad (26c)$$

$$P_{g,s,t} = P_{g,t} + P_{g,s,t}^+ - P_{g,s,t}^- \quad (26d)$$

$$v_{g,t} = u_{g,t} - u_{g,t-1} \quad (26e)$$

$$P_{g,s,t}^+ \leq \bar{P}_g u_{g,t}, P_{g,s,t}^- \leq \bar{P}_g u_{g,t} \quad (26f)$$

$$P_{g,s,t}^+ \leq \bar{P}_g D_{g,s,t}, P_{g,s,t}^- \leq \bar{P}_g (1 - D_{g,s,t}) \quad (26g)$$

$$u_{g,t} \bar{P}_g \leq P_{g,t} \leq u_{g,t} \bar{P}_g \quad (26h)$$

$$u_{g,t} \bar{P}_g \leq P_{g,s,t} \leq u_{g,t} \bar{P}_g \quad (26i)$$

$$RD_g \leq P_{g,s,t} - P_{g,s,t-1} \leq RU_g \quad (26j)$$

$$RD_g \leq P_{g,t} - P_{g,t-1} \leq RU_g \quad (26k)$$

$$C_{g,t}^{SC} \geq v_{g,t} C_g^{ON}, C_{g,t}^{SC} \geq -v_{g,t} C_g^{OFF} \quad (26l)$$

$$\sum_{\tau=t+1}^{t+T_g^{on}} u_{g,\tau} \geq v_{g,t} T_g^{on} \quad (26m)$$

$$\sum_{\tau=t+1}^{t+T_g^{off}} (1 - u_{g,\tau}) \geq -v_{g,t} T_g^{off} \quad (26n)$$

$$0 \leq P_{r,s,t}^{R,c} \leq P_{r,s,t}^R \quad (26o)$$

$$0 \leq P_{l,s,t}^{L,s} \leq P_{l,s,t}^L \quad (26p)$$

$$D_{g,s,t} \in \{0, 1\}, u_{g,t} \in \{0, 1\}, \quad (26q)$$

where (26a)-(26c) are the DC power flow equations, (26e), (26m), (26n) are the generation on/off status and time constraints, (26d), (26f)-(26i) are the generation power range constraints, (26j)-(26k) are the generation ramp rate constraints, (26o)-(26p) are the RES curtailment and load shedding constraints. $P_{ij,s,t}$ is the power flow of line ij at scenario s time t , B_{ij} is the susceptance of line ij , $\theta_{i,s,t}$ is the phase angle. $P_{r,s,t}^R/P_{l,s,t}^L$ are the power of RES farm and load demand, respectively. $v_{g,t}$ is binary and indicates the on/off status of generator g at time t . \bar{P}_g/P_g are the maximum/minimum generator power output. $D_{g,s,t}$ is the binary variable indicating the up/down generator regulation status. RD_g/RU_g are the generator ramp down and ramp up limits. C_g^{ON}/C_g^{OFF} are the generator start-up and shutdown cost. T_g^{on}/T_g^{off} are the minimum up and down time of generator g . $\Xi_1 = [P_{g,t}, u_{g,t}]$ is the set of decision variables at the first stage. $\Xi_2 = [D_{g,s,t}, P_{g,s,t}^+, P_{g,s,t}^-, P_{r,s,t}^{R,c}, P_{l,s,t}^{L,s}]$ is the set of decision variables at the second stage.

REFERENCES

- [1] L. A. Roald, D. Pozo, A. Papavasiliou *et al.*, "Power systems optimization under uncertainty: A review of methods and applications," *Electric Power Systems Research*, vol. 214, p. 108725, 2023.
- [2] Z. Guo, P. Pinson, S. Chen *et al.*, "Chance-constrained peer-to-peer joint energy and reserve market considering renewable generation uncertainty," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 798–809, 2020.
- [3] N. Qi, P. Pinson, M. R. Almassalkhi *et al.*, "Chance-constrained generic energy storage operations under decision-dependent uncertainty," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 4, pp. 2234–2248, 2023.
- [4] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.
- [5] Y. Zhang, F. Liu, Z. Wang *et al.*, "Robust scheduling of virtual power plant under exogenous and endogenous uncertainties," *IEEE Transactions on Power Systems*, vol. 37, no. 2, pp. 1311–1325, 2021.
- [6] Y. Guo, K. Baker, E. Dall'Anese *et al.*, "Data-based distributionally robust stochastic optimal power flow—part i: Methodologies," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1483–1492, 2018.
- [7] Y. Qiu, Q. Li, Y. Ai *et al.*, "Two-stage distributionally robust optimization-based coordinated scheduling of integrated energy system with electricity-hydrogen hybrid energy storage," *Protection and Control of Modern Power Systems*, vol. 8, no. 2, pp. 1–14, 2023.
- [8] G. Liu, Y. Xu, and K. Tomsovic, "Bidding strategy for microgrid in day-ahead market based on hybrid stochastic/robust optimization," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 227–237, 2016.
- [9] J. Shi, Y. Guo, L. Tong *et al.*, "A scenario-oriented approach to energy-reserve joint procurement and pricing," *IEEE Transactions on Power Systems*, vol. 38, no. 1, pp. 411–426, 2023.
- [10] K. Antoniadou-Plytaria, D. Steen, L. A. Tuan *et al.*, "Scenario-based stochastic optimization for energy and flexibility dispatch of a microgrid," *IEEE Transactions on Smart Grid*, vol. 13, no. 5, pp. 3328–3341, 2022.
- [11] F. Fang, S. Yu, and X. Xin, "Data-driven-based stochastic robust optimization for a virtual power plant with multiple uncertainties," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 456–466, 2022.
- [12] W. Yin, S. Feng, and Y. Hou, "Stochastic wind farm expansion planning with decision-dependent uncertainty under spatial smoothing effect," *IEEE Transactions on Power Systems*, 2022.
- [13] Y. Liu, D. Liu, and H. Zhang, "Stochastic unit commitment with high-penetration offshore wind power generation in typhoon scenarios," *Journal of Modern Power Systems and Clean Energy*, 2023.
- [14] N. Qi, L. Cheng, H. Li *et al.*, "Portfolio optimization of generic energy storage-based virtual power plant under decision-dependent uncertainties," *Journal of Energy Storage*, vol. 63, p. 107000, 2023.
- [15] L. Liu, X. Hu, J. Chen *et al.*, "Embedded scenario clustering for wind and photovoltaic power, and load based on multi-head self-attention," *Protection and Control of Modern Power Systems*, vol. 9, no. 1, pp. 122–132, 2024.
- [16] M. Sun, Y. Wang, F. Teng *et al.*, "Clustering-based residential baseline estimation: A probabilistic perspective," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6014–6028, 2019.
- [17] N. Huang, W. Wang, S. Wang *et al.*, "Incorporating load fluctuation in feature importance profile clustering for day-ahead aggregated residential load forecasting," *IEEE Access*, vol. 8, pp. 25 198–25 209, 2020.
- [18] D. Z. Fitiwi, F. de Cuadra, L. Olmos *et al.*, "A new approach of clustering operational states for power network expansion planning problems dealing with res (renewable energy source) generation operational variability and uncertainty," *Energy*, vol. 90, pp. 1360–1376, 2015.
- [19] M. Sun, F. Teng, X. Zhang *et al.*, "Data-driven representative day selection for investment decisions: A cost-oriented approach," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2925–2936, 2019.
- [20] J. Xu, B. Wang, Y. Sun *et al.*, "A day-ahead economic dispatch method considering extreme scenarios based on wind power uncertainty," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 2, pp. 224–233, 2019.
- [21] X. Liu, M. Shahidehpour, Z. Li *et al.*, "Microgrids for enhancing the power grid resilience in extreme conditions," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 589–597, 2017.
- [22] X. Dong, Y. Sun, S. M. Malik *et al.*, "Scenario reduction network based on wasserstein distance with regularization," *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 4–13, 2024.
- [23] T. Teeraratkul, D. O'Neill, and S. Lall, "Shape-based approach to household electric load curve clustering and prediction," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5196–5206, 2018.
- [24] Y. Liu, R. Sioshansi, and A. J. Conejo, "Hierarchical clustering to find representative operating periods for capacity-expansion modeling," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3029–3039, 2017.
- [25] S. Lin, C. Liu, Y. Shen *et al.*, "Stochastic planning of integrated energy system via frank-copula function and scenario reduction," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 202–212, 2021.
- [26] J. Pei, J. Wang, Z. Wang *et al.*, "Precise recovery of corrupted synchronphasors based on autoregressive bayesian low-rank factorization and adaptive k-medoids clustering," *IEEE Transactions on Power Systems*, vol. 38, no. 6, pp. 5834–5848, 2023.
- [27] Z. Shi, L. Wu, and Y. Zhou, "Predicting household energy consumption in an aging society," *Applied Energy*, vol. 352, p. 121899, 2023.

- [28] H. Teichgraeb and A. R. Brandt, "Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison," *Applied Energy*, vol. 239, pp. 1283–1293, 2019.
- [29] B. Ghosh, A. K. Chakraborty, and A. R. Bhowmik, "Reliability and efficiency enhancement of a radial distribution system through value-based auto-recloser placement and network remodelling," *Protection and Control of Modern Power Systems*, vol. 8, no. 1, pp. 1–14, 2023.
- [30] D. Bertsimas and N. Mundru, "Optimization-based scenario reduction for data-driven two-stage stochastic optimization," *Operations Research*, vol. 71, no. 4, pp. 1343–1361, 2022.
- [31] M. Hewitt, J. Ortmann, and W. Rei, "Decision-based scenario clustering for decision-making under uncertainty," *Annals of Operations Research*, vol. 315, no. 2, pp. 747–771, 2022.
- [32] J. Keutchan, J. Ortmann, and W. Rei, "Problem-driven scenario clustering in stochastic optimization," *Computational Management Science*, vol. 20, no. 1, p. 13, 2023.
- [33] J. L. Higle, "Stochastic programming: optimization when uncertainty matters," in *Emerging theory, methods, and applications*. Informs, 2005, pp. 30–53.
- [34] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, "Scenario reduction in stochastic programming," *Mathematical Programming*, vol. 95, no. 3, pp. 493–511, 2003.
- [35] M. Almassalkhi, S. Brahma, N. Nazir *et al.*, "Hierarchical, grid-aware, and economically optimal coordination of distributed energy resources in realistic distribution systems," *Energies*, vol. 13, no. 23, 2020.
- [36] M. Baran and F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 735–743, 1989.
- [37] I. Goodfellow, J. Pouget-Abadie, M. Mirza *et al.*, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [38] "Original scenarios." [Online]. Available: <https://github.com/Yingrui-Z/Original-Scenarios>
- [39] V. Nair, S. Bartunov, F. Gimeno *et al.*, "Solving mixed integer programs using neural networks," *arXiv preprint arXiv:2012.13349*, 2020.
- [40] J. Qin and N. Yu, "Solve large-scale unit commitment problems by physics-informed graph learning," *arXiv preprint arXiv:2311.15216*, 2023.
- [41] R. Vujanic, P. Mohajerin Esfahani, P. J. Goulart *et al.*, "A decomposition method for large scale milps, with performance guarantees and a power system application," *Automatica*, vol. 67, pp. 144–156, 2016.
- [42] A. Camisa, I. Notarnicola, and G. Notarstefano, "Distributed primal decomposition for large-scale milps," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 413–420, 2022.



Yingrui Zhuang (Student Member, IEEE) was born in 1999. He received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2021. He is currently pursuing the Ph.D. degree in electrical engineering with the State Key Laboratory of Power System Operation and Control, Department of Electrical Engineering, Tsinghua University. His current research interests include risk analysis of distribution systems with high renewable energy resources penetration and data-driven distributed resources coordination and optimization.



Lin Cheng (Senior Member, IEEE), was born in 1973. He received the B.S. degree in electrical engineering from Tianjin University, Tianjin, China, in 1996 and the Ph.D. degree from Tsinghua University, Beijing, China, in 2001. Currently, he is a tenured Professor in the Department of Electrical Engineering at Tsinghua University, serving as the deputy director of the State Key Laboratory of Power System Operation and Control, and is also a Fellow of IET. His research interests include operational reliability evaluation and application of power systems, operation optimization of distribution systems with flexible resources, and perception and control of uncertainty in wide-area measurement systems.



Ning Qi (Member, IEEE) was born in 1996. He received a B.S. degree in Electrical Engineering from Tianjin University, Tianjin, China, in 2018 and the Ph.D. degree in Electrical Engineering from Tsinghua University, Beijing, China, in 2023. He is currently a postdoctoral research scientist in Earth and Environmental Engineering at Columbia University. Before joining Columbia University, he was a visiting scholar at the Technical University of Denmark in 2022. He was a research associate in Electrical Engineering at Tsinghua University in 2024. His current research focuses on data-driven modeling, optimization under uncertainty, and market design for power systems with generalized energy storage. He is currently serving as the Youth Editorial Board Member Power System Protection and Control, and the guest editor of Processes and Frontier in Energy Research.



Mads R. Almassalkhi (M'13, SM'18) received his dual degree in electrical engineering and applied mathematics from the University of Cincinnati, Ohio, in 2008, and the Ph.D. degree in electrical engineering: systems from the University of Michigan in 2013. He is currently the L. Richard Fisher Associate Professor with the Department of Electrical and Biomedical Engineering at The University of Vermont. He also holds a joint appointment with Pacific Northwest National Laboratory as Chief Scientist and co-founded clean-tech startup company, Packetized Energy. Before joining the University of Vermont, he was with another energy startup company Root3 Technologies. His research interests lie at the intersection of power systems, mathematical optimization, and control systems. He is currently serving as Chair of the IEEE CSS Technical Committee on Smart Grids and Associate Editor of IEEE Transactions on Power Systems.



Feng Liu (M'10–SM'18) received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1999 and 2004, respectively. He is currently an Associate Professor of Tsinghua University. From 2015 to 2016, he was a visiting associate at California Institute of Technology, CA, USA. Dr. Feng Liu's research interests include stability analysis, optimal control, robust dispatch and game theory based decision-making in energy and power systems. Dr. Liu is an IET Fellow. He is an associated editor of several international journals including IEEE Transactions on Power Systems, IEEE Transactions on Smart Grid, and Control Engineering Practice. He also served as a guest editor of IEEE Transactions on Energy Conversion.