An Iterative Problem-Driven Scenario Reduction Framework for Stochastic Optimization with Conditional Value-at-Risk

Yingrui Zhuang[1], Lin Cheng[1], Ning Qi[2*], Mads R. Almassalkhi[3], Feng Liu[1]
[1]*Department of Electrical Engineering, Tsinghua University, Beijing, China*
[2]*Department of Earth and Environmental Engineering, Columbia University, New York, USA*
[3]*Department of Electrical and Biomedical Engineering, University of Vermont, Burlington VT, USA*
[*]*Email: nq2176@columbia.edu*

***Abstract*—Scenario reduction (SR) alleviates the computational complexity of scenario-based stochastic optimization with conditional value-at-risk (SBSO-CVaR) by identifying representative scenarios to depict the underlying uncertainty and tail risks. Existing distribution-driven SR methods emphasize statistical similarity but often exclude extreme scenarios, leading to weak tail-risk awareness and insufficient problem-specific representativeness. Instead, this paper proposes an iterative problem-driven scenario reduction framework. Specifically, we integrate the SBSO-CVaR problem structure into SR process and project the original scenario set from distribution space onto the problem space. Subsequently, to minimize the SR optimality gap with acceptable computation complexity, we propose a tractable iterative problem-driven scenario reduction (IPDSR) method that selects representative scenarios that best approximate the optimality distribution of the original scenario set while preserving tail risks. Furthermore, the iteration process is rendered as a mixed-integer program to enable scenario partitioning and representative scenario selection. And *ex-post* problem-driven evaluation indices are proposed to evaluate the SR performance. Numerical experiments show that IPDSR significantly outperforms existing SR methods by achieving an optimality gap of less than 1% within an acceptable computation time.**

***Index Terms*—Conditional value-at-risk, problem-driven, risk-averse, scenario reduction, stochastic optimization.**

I. Introduction

The increasing integration of renewable energy sources (RES) and emerging flexible loads introduces substantial uncertainties and operational risks into power systems [1]. Effective decision-making under uncertainty requires risk-averse optimization. Compared to robust optimization [2], chance-constrained optimization [3], and distributionally robust optimization approaches [4], stochastic optimization with conditional value-at-risk (CVaR) [5], [6] is widely adopted in power systems for its convexity, coherency, and explicit tail-risk characterization. Accurate CVaR computation requires precise modeling of uncertainty distributions, particularly in the tail regions. When exact distributional information is unavailable, scenario-based stochastic optimization with conditional value-at-risk (SBSO-CVaR) [7] offers an effective way to manage tail risks. However, SBSO-CVaR relies on a large number of scenarios to accurately capture low-probability, high-loss events, which leads to substantial computational complexity. To alleviate computational burden, scenario reduction (SR) is employed to identify a smaller representative scenario set to replace the original scenario set while maintaining an acceptably optimal solution [8]. Specifically, SR for SBSO-CVaR presents two main challenges: *i)* how to identify salient and representative scenarios that enable substantial scenario reduction, and *ii)* how to ensure that the reduced scenario set retains the ability to accurately capture tail risk behavior and approximate the optimal solution of the original problem.

Existing SR methods for SBSO problems can be broadly classified into two categories: *distribution-driven scenario reduction* (DDSR) and *problem-driven scenario reduction* (PDSR). DDSR methods aim to preserve the statistical properties of the original scenario set, assuming that a more faithful representation of the underlying distribution will lead to a more accurate approximation of the original optimal solution. Certain distribution-driven similarity metrics (e.g., Euclidean distance [9], Wasserstein distance [10]) are commonly adopted by DDSR methods to measure the similarity between scenarios. Distribution-driven clustering methods (e.g., K-means [11], hierarchical clustering [12], and Gaussian mixture model [13]) are adopted to perform scenario clustering. That is, DDSR methods generally consider SR and SBSO as two distinct and decoupled processes, where the SR process is solely driven by the statistical characteristics of scenarios. However, since the same scenario may yield different solutions under different problem formulations (e.g., economic dispatch and resilience-oriented dispatch), statistical similarity does not necessarily imply a similarity in decision-making [14]. Therefore, DDSR methods suffer from a critical oversight: an inability to consider the impacts of the reduced scenario set on optimality. To address this limitation, our recent work [15] proposed a PDSR framework for risk-neutral SBSO problems. The key idea of PDSR is to project the original scenario set from the distribution space onto the problem space by explicitly incorporating the optimization problem

This work is supported in part by National Natural Science Foundation of China (No. 52037006) and China Postdoctoral Science Foundation special funded project (No. 2023TQ0169).

structure. SR is then performed directly in the problem space by identifying scenarios that significantly influence decision-making optimality. This idea aligns with the emerging literature on decision-focused learning [16], [17], which emphasizes preserving decision-relevant structures to minimize the final downstream decision loss instead of prioritizing statistical approximation. The PDSR framework can be viewed as a specific realization of this philosophy within the field of scenario reduction. The effectiveness of PDSR has been demonstrated through its ability to closely approximate the original optimal solution while maintaining computational efficiency. However, the PDSR framework in [15] primarily focuses on risk-neutral problems and has not yet been extended to risk-averse settings.

Unlike risk-neutral SBSO, the CVaR objective makes SBSO-CVaR structurally more complex, necessitating precise tail-risk modeling [18]. Despite this additional complexity, most existing studies addressing SBSO-CVaR still employ traditional DDSR methods [19], [20]. Such methods discard rare yet influential tail scenarios since they typically prioritize scenarios near the distribution center and use averaging to derive representative scenarios. Moreover, DDSR methods fail to account for the problem-specific characteristics of SBSO-CVaR or the sensitivity of tail risk measures to individual scenarios during SR. Consequently, these methods fail to capture the weighted marginal contributions of scenarios to tail losses, resulting in a poor representation of the tail risk distribution. This may lead to different decisions from those of the original problem, and ultimately, inadequate effectiveness to tail risks.

Existing PDSR methods for SBSO-CVaR focus exclusively on the CVaR term, addressing tail risks but neglecting other salient features of the underlying distribution. For example, the method in Ref. [21] iteratively evaluates all tail scenarios, but results in a large number of scenarios and overlooks the impact of SR on the optimal solution. Consequently, this method suffers from scalability and suboptimality issues. In Ref. [22], tail-region scenarios capture the CVaR value well, but all remaining scenarios are aggregated into a single scenario. Such aggregation requires knowledge of the optimal solution under the original scenario set, which is generally unavailable in practice.

Our previous work on PDSR focused on risk-neutral SBSO problems [15]. Directly extending our risk-neutral PDSR [15] to risk-averse SBSO-CVaR problems is computationally intractable due to the non-smooth, tail-dependent nature of CVaR. To overcome this, we propose a novel iterative problem-driven scenario reduction (IPDSR) framework that efficiently achieves near-optimal solutions with significantly reduced computational burden. Specifically, our contributions are as follows:

1) *IPDSR Framework:* We propose an IPDSR framework that provides a tractable extension of [15] to incorporate SBSO-CVaR. IPDSR incorporates the problem structure of SBSO-CVaR (especially the emphasis on tail risks) into the SR process. By analytically minimizing the optimality gap (OG), IPDSR iteratively selects the representative scenarios and ensures the reduced scenario set accurately approximates the original optimality.
2) *Solution Methodology:* We iteratively identify the most representative scenarios within the projected problem space. In each iteration, the optimization process is reformulated as a mixed-integer programming (MIP) model that jointly performs scenario partitioning and representative scenario selection. Additionally, to efficiently handle large-scale problems, we propose an objective aggregation technique that significantly reduces the computational complexity of the MIP.
3) *Simulation Analysis:* We evaluate the effectiveness of the proposed IPDSR framework using a risk-averse offering problem within the context of virtual power plant (VPP) operations. Simulation results demonstrate that IPDSR outperforms state-of-the-art SR methods by effectively identifying salient scenarios and achieving an optimality gap of less than 1%.

The remainder of this paper is organized as follows. Section II introduces the proposed IPDSR framework. Section III presents the problem formulation of a risk-averse SBSO-CVaR problem. Section IV presents case studies to validate the effectiveness of IPDSR. We conclude the paper in Section V.

II. Iterative Problem-Driven Scenario Reduction Framework

A. Formulation of SBSO-CVaR

SBSO-CVaR represents a rich set of risk-averse power system problems. The general formulation of the SBSO-CVaR problem built on an original scenario set $\boldsymbol{\xi}$ is

$$\min_{z\in Z} F(z,\boldsymbol{\xi}) := \sum_{i=1}^{N}\gamma_i f(z,\xi_i) + \lambda\left(v^{\alpha}_{\boldsymbol{\xi}} + \frac{1}{1-\alpha}\sum_{i=1}^{N}\gamma_i [f(z,\xi_i) - v^{\alpha}_{\boldsymbol{\xi}}]_+\right), \tag{1}$$

where $F(z,\boldsymbol{\xi})$ is the objective function of the SBSO-CVaR problem with decision variable $z\in Z$, including the expected cost across $\boldsymbol{\xi}$ (the first term) and the CVaR term (the second term) with $\alpha\in(0,1)$ as the confidence level. The risk aversion parameter is $\lambda\geq 0$, where $\lambda=0$ corresponds to the risk-neutral case and a larger λ indicates higher risk aversion. $v^{\alpha}_{\boldsymbol{\xi}}$ is the VaR under confidence level α. The positive part of x is denoted $[x]_+ := \max\{x,0\}$. The inclusion of the CVaR term emphasizes the conditional expected loss in the tail distribution, transforming a risk-neutral optimization problem into a risk-averse one, which is inherently different from our previous work [15]. The original set of N scenarios is denoted as $\boldsymbol{\xi} = \{\xi_1,\ldots,\xi_N\}$. The probability of scenario ξ_i is $\gamma_i > 0$, which satisfies $\sum_{i=1}^{N}\gamma_i = 1$. $f(z,\xi_i)$ denotes the scenario-specific cost function associated with scenario ξ_i. We denote the optimal solution of the original problem (1) as $z^*_{\boldsymbol{\xi}} = \operatorname{argmin}_{z\in Z} F(z,\boldsymbol{\xi})$. Similar to [15], we also assume that the SBSO-CVaR problem satisfies the assumption of *relatively complete recourse*.

Scenario reduction is employed to reduce the computational complexity by clustering large N into $K \ll N$ clusters that preserve the accuracy of approximating optimality of the original problem. The SR process can be denoted as $\boldsymbol{C}(\boldsymbol{\xi},K) = \{\{C_1,\ldots,C_K\} : C_i \neq \emptyset, \forall i; C_i \cap C_j = \emptyset, \forall i \neq j; \cup_i C_i = I, I = \{1,2,\ldots,N\}\}$. That is, the original scenario set $\boldsymbol{\xi}$ is partitioned into K clusters and is reduced to a representative scenario set $\boldsymbol{\zeta}=\{\zeta_1,...,\zeta_K\}$. 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 focus on selecting $\boldsymbol{\zeta}\subseteq\boldsymbol{\xi}$, instead of generating synthetic scenarios. The reduced problem formulated on the reduced scenario set $\boldsymbol{\zeta}$ can be formulated as (2) and the corresponding optimal solution is denoted as $z^*_{\boldsymbol{\zeta}}=\underset{z\in Z}{\operatorname{argmin}}\ F(z,\boldsymbol{\zeta})$.

$$\min_{z\in Z}F(z,\boldsymbol{\zeta}):=\sum_{k=1}^{K}\omega_k f(z,\zeta_k)+\lambda\left(v^{\alpha}_{\boldsymbol{\zeta}}+\frac{1}{1-\alpha}\sum_{k=1}^{K}\omega_k\Big[f(z,\zeta_k)-v^{\alpha}_{\boldsymbol{\zeta}}\Big]_+\right). \tag{2}$$

CVaR emphasizes tail risk characterization, inducing non-smooth piecewise convex structures that markedly increase computational complexity. As a result, SR for SBSO-CVaR poses substantial challenges beyond the risk-neutral case. In what follows, we extend the framework to accommodate the SBSO-CVaR setting.

B. Optimality Gap of Scenario Reduction

SR aims to minimize the optimality gap (OG) from using K representatives ($\boldsymbol{\zeta}$) vs. N scenarios ($\boldsymbol{\xi}$). Towards this purpose, the OG can be defined as

$$OG:=F(z^*_{\boldsymbol{\zeta}},\boldsymbol{\xi})-F(z^*_{\boldsymbol{\xi}},\boldsymbol{\xi}), \tag{3}$$

where $F(z^*_{\boldsymbol{\zeta}},\boldsymbol{\xi})$ means solving (1) with $z=z^*_{\boldsymbol{\zeta}}$. Since $F(z,\boldsymbol{\xi})\geq F(z^*_{\boldsymbol{\xi}},\boldsymbol{\xi})$ for all $z\in Z$, we have $OG\geq 0$. A smaller OG indicates a more accurate problem optimality approximation of $\boldsymbol{\zeta}$ to $\boldsymbol{\xi}$.

Notably, the OG is defined relative to the original scenario set $\boldsymbol{\xi}$ rather than the true distribution. Under Sample Average Approximation (SAA) theory, a sufficiently large $\boldsymbol{\xi}$ acts as a tractable, high-confidence proxy for the true distribution. Thus, the objective is to minimize information loss when compressing $\boldsymbol{\xi}$ to $\boldsymbol{\zeta}$. However, like all SAA-based methods, the ultimate performance inherently relies on the representativeness of $\boldsymbol{\xi}$.

Next, we present two methods to minimize the OG:

i) PDSR: we extend the framework in [15] from risk-neutral formulations to risk-averse formulations with CVaR, and derive a theoretical upper bound of the OG leveraging relaxation techniques, as detailed in Section II-C.

ii) IPDSR: we iteratively select the representative scenario set $\boldsymbol{\zeta}$ to minimize the OG, as detailed in Section II-D.

C. Problem-Driven Scenario Reduction

This method provides a risk-averse formulation with CVaR, and further incorporates the problem structure of SBSO-CVaR to devise a relaxation strategy. We first project the distributional space onto the problem space and define a problem-driven distance metric, upon which the SR process is reformulated as an MIP model aiming to minimize the relaxed theoretical upper bound on the OG. The detailed derivations are presented in Appendix A.

However, the MIP formulation involves numerous binary variables and constraints, making it computationally intractable for standard solvers. Therefore, we introduce an alternative approach to minimize the OG and greatly reduce the SR problem complexity. Since $z^*_{\boldsymbol{\xi}}$ remains unknown in practice and $F(z^*_{\boldsymbol{\xi}},\boldsymbol{\xi})$ is a constant, the OG can be minimized by minimizing $F(z^*_{\boldsymbol{\zeta}},\boldsymbol{\xi})$. This transformation implies that the OG can be minimized without solving the computationally intensive original N-scenario problem (i.e., knowledge of $z^*_{\boldsymbol{\xi}}$ is not required). We next develop an IPDSR framework to identify a high-quality $\boldsymbol{\zeta}$ and corresponding $z^*_{\boldsymbol{\zeta}}$ that minimize the OG.

D. Iterative Problem-Driven Scenario Reduction

IPDSR consists of two main steps: initialization and iterative scenario partitioning.

1) Initialization: First, $\boldsymbol{\zeta}^0$ and $z^*_{\boldsymbol{\zeta}^0}$ are initialized. We denote κ as the iteration index, initialized as $\kappa=0$. We validate $z^*_{\boldsymbol{\zeta}^0}$ in $\boldsymbol{\xi}$ and calculate the validation value of $F(z^*_{\boldsymbol{\zeta}^0},\boldsymbol{\xi})$ and the initial VaR value $v^{\alpha,0}_{\boldsymbol{\xi}}$. Note that this initialization can be performed by applying traditional DDSR methods like K-means. Furthermore, multiple initializations can be performed to obtain a superior initial $\boldsymbol{\zeta}^0$.

2) Iterative Scenario Reduction: At step κ, we have $\boldsymbol{\zeta}^\kappa$, $z^*_{\boldsymbol{\zeta}^\kappa}$, $F(z^*_{\boldsymbol{\zeta}^\kappa},\boldsymbol{\xi})$ and $v^{\alpha,\kappa}_{\boldsymbol{\xi}}$. We seek to select a new reweighted representative scenario set $\boldsymbol{\zeta}^{\kappa+1}\subset\boldsymbol{\xi}$ with weights $\omega^{\kappa+1}$ such that $F(z^*_{\boldsymbol{\zeta}^\kappa},\boldsymbol{\zeta}^{\kappa+1})$ can best approximate $F(z^*_{\boldsymbol{\zeta}^\kappa},\boldsymbol{\xi})$, which can be formulated as the following optimization problem:

$$\begin{aligned}
\boldsymbol{\zeta}^{\kappa+1}&=\underset{\boldsymbol{\zeta}\subset\boldsymbol{\xi}}{\operatorname{argmin}}\Big|F(z^*_{\boldsymbol{\zeta}^\kappa},\boldsymbol{\zeta})-F(z^*_{\boldsymbol{\zeta}^\kappa},\boldsymbol{\xi})\Big|=\underset{\boldsymbol{\zeta}\subset\boldsymbol{\xi}}{\operatorname{argmin}}\\
&\Bigg|\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i(f(z^*_{\boldsymbol{\zeta}^\kappa},\zeta_k)+\lambda(v^{\alpha,\kappa}_{\boldsymbol{\zeta}}+\frac{1}{1-\alpha}[f(z^*_{\boldsymbol{\zeta}^\kappa},\zeta_k)-v^{\alpha,\kappa}_{\boldsymbol{\zeta}}]_+))\\
&-\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i(f(z^*_{\boldsymbol{\zeta}^\kappa},\xi_i)+\lambda(v^{\alpha,\kappa}_{\boldsymbol{\xi}}+\frac{1}{1-\alpha}[f(z^*_{\boldsymbol{\zeta}^\kappa},\xi_i)-v^{\alpha,\kappa}_{\boldsymbol{\xi}}]_+))\Bigg|\\
&=\underset{\boldsymbol{\zeta}\subset\boldsymbol{\xi}}{\operatorname{argmin}}\Bigg|\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i\Big(f(z^*_{\boldsymbol{\zeta}^\kappa},\xi_i)-f(z^*_{\boldsymbol{\zeta}^\kappa},\zeta_k)+\lambda(v^{\alpha,\kappa}_{\boldsymbol{\xi}}-v^{\alpha,\kappa}_{\boldsymbol{\zeta}})\\
&+\frac{\lambda}{1-\alpha}([f(z^*_{\boldsymbol{\zeta}^\kappa},\xi_i)-v^{\alpha,\kappa}_{\boldsymbol{\xi}}]_+-[f(z^*_{\boldsymbol{\zeta}^\kappa},\zeta_k)-v^{\alpha,\kappa}_{\boldsymbol{\zeta}}]_+)\Big)\Bigg|.
\end{aligned} \tag{4}$$

For convenience of notation, we omit the superscript κ in the following discussion. We sort the scenarios in a non-decreasing manner with respect to the validated objective function $f(z^*_{\boldsymbol{\zeta}},\cdot)$, i.e., $f(z^*_{\boldsymbol{\zeta}},\xi_{(1)})\leq f(z^*_{\boldsymbol{\zeta}},\xi_{(2)})\leq\cdots\leq f(z^*_{\boldsymbol{\zeta}},\xi_{(N)})$, where $\xi_{(i)}$ is the ordered i-th scenario with corresponding probability $\gamma_{(i)}$. Let the integer index $N^{\alpha}_{\boldsymbol{\xi}}$ be defined such that $\sum_{i=1}^{N^{\alpha}_{\boldsymbol{\xi}}-1}\gamma_{(i)}<\alpha\leq\sum_{i=1}^{N^{\alpha}_{\boldsymbol{\xi}}}\gamma_{(i)}$. Then we have $v^{\alpha}_{\boldsymbol{\xi}}=f(z,\xi_{(N^{\alpha}_{\boldsymbol{\xi}})})$, and for $i>N^{\alpha}_{\boldsymbol{\xi}}$, we have $\Big[f(z,\xi_{(i)})-v^{\alpha}_{\boldsymbol{\xi}}\Big]_+=f(z,\xi_{(i)})-v^{\alpha}_{\boldsymbol{\xi}}$, and is 0 otherwise. Similarly for $\boldsymbol{\zeta}$, we have $K^{\alpha}_{\boldsymbol{\zeta}}$ such that $v^{\alpha}_{\boldsymbol{\zeta}}=f(z^*_{\boldsymbol{\zeta}},\zeta_{(K^{\alpha}_{\boldsymbol{\zeta}})})$.

In this way, we project the original scenario set $\boldsymbol{\xi}$ from the distribution space onto the problem space of SBSO-CVaR by explicitly incorporating the problem structure. Note that the problem space here is a collapsed space since we have the validation solution $z^*_{\boldsymbol{\zeta}}$. The projection process can be denoted as $\boldsymbol{\xi}\rightarrow\boldsymbol{F}$, where $\boldsymbol{F}:=\{F_i=f(z^*_{\boldsymbol{\zeta}},\xi_{(i)})\}$. In this projection, we can directly quantify the impacts of uncertainty and systematically incorporate the inherent characteristics of the SBSO-CVaR problem into the SR process.

Then, we transform (4) into a MIP as (5), which conducts scenario partitioning and representative scenario selection in the problem space simultaneously.

$$\min \left|\sum_{j=1}^{N}\Big(\sum_{i=1}^{N}v_{ij}\gamma_i(F_i-F_j+\frac{\lambda}{1-\alpha}([F_i-v_{\boldsymbol{\xi}}^{\alpha}]_+-[F_j-v_{\boldsymbol{\zeta}}^{\alpha}]_+))\right.$$

$$\left.+\lambda(v_{\boldsymbol{\xi}}^{\alpha}-v_{\boldsymbol{\zeta}}^{\alpha})\right| \tag{5a}$$

$$\text{s.t. } v_{ij}\le u_j,\ v_{jj}=u_j,\ n_j\le u_j,\ h_j\le u_j, \tag{5b}$$

$$\sum\nolimits_{j=1}^{N}v_{ij}=1,\ \sum\nolimits_{j=1}^{N}u_j=K, \tag{5c}$$

$$\sum\nolimits_{j=1}^{N}n_j=1,\ \sum\nolimits_{i=1}^{N}n_iF_i=v_{\boldsymbol{\zeta}}^{\alpha}, \tag{5d}$$

$$h_i=u_i(1-\sum\nolimits_{p=1}^{i}n_p), \tag{5e}$$

$$\sum\nolimits_{j=1}^{N}h_j\sum\nolimits_{i=1}^{N}v_{ij}\gamma_i\le\alpha, \tag{5f}$$

$$\sum_{j=1}^{N}h_j\sum_{i=1}^{N}v_{ij}\gamma_i+(u_k-h_k)\sum_{i=1}^{N}v_{ik}\gamma_i\ge\alpha(u_k-h_k), \tag{5g}$$

$$v_{ij}\in\{0,1\},\ u_j\in\{0,1\},\ n_j\in\{0,1\},\ h_j\in\{0,1\}, \tag{5h}$$

where u_j indicates whether scenario ξ_j is selected as a representative scenario of a cluster. v_{ij} determines whether scenario ξ_i is included in the cluster represented by scenario ξ_j. n_j indicates whether scenario ξ_j is the scenario corresponding to the VaR value (i.e., whether $j=K_{\boldsymbol{\zeta}}^{\alpha}$), h_j indicates $\mathbf{1}_{j\le K_{\boldsymbol{\zeta}}^{\alpha}}$, and (5e) is only valid for sorted scenarios. Constraint $v_{ij}\le u_j$ ensures that ξ_i can only be assigned to a cluster that has a designated representative, while $v_{jj}=u_j$ enforces that ξ_j must be assigned to its respective cluster if it's a representative scenario. Constraint $n_j\le u_j$ ensures that the VaR scenario of $\boldsymbol{\zeta}$ can only be selected from the representative scenarios. Constraint $\sum_{j=1}^{N}v_{ij}=1$ ensures that each scenario ξ_i can only be assigned to one cluster, while constraint $\sum_{j=1}^{N}u_j=K$ guarantees that exactly K clusters are formed. Constraint $\sum_{j=1}^{N}n_j=1$ ensures that there is only one VaR scenario in the reduced scenario set $\boldsymbol{\zeta}$. Constraints (5f) and (5g) are used to ensure the confidence level of the VaR scenario in the reduced scenario set $\boldsymbol{\zeta}$.

Note that the MIP model is nonlinear due to the presence of absolute values and variable products. Auxiliary variables can be introduced to partially linearize the model to make it solvable by Gurobi. By solving the MIP model in (5), we can obtain the optimal representative scenario set $\boldsymbol{\zeta}^{\kappa+1}$ and the associated weights $\omega_k^{\kappa+1}=\sum_{i=1}^{N}v_{ik}^{\kappa+1}\gamma_{(i)}$.

E. *Objective Aggregation for MIP Simplification*

Notably, the MIP model in (5) is computationally intensive with numerous binary variables, especially when N is large. To address this issue, we propose to precede the solution process of (5) with an objective aggregation step. Specifically, $\boldsymbol{F}=\{F_i\}_{i=1}^{N}$ is aggregated into $\bar{\boldsymbol{F}}:=\{\bar{F}_j\}_{j=1}^{N'}$, and the corresponding probability weights $\{\gamma_i\}_{i=1}^{N}$ are aggregated into $\{\bar{\gamma}_j\}_{j=1}^{N'}$, where $N'\ll N$ is the number of aggregated values.

$$\bar{F}_j=\frac{\sum_{i\in\mathcal{I}_j}\gamma_iF_i}{\sum_{i\in\mathcal{I}_j}\gamma_i},\quad \bar{\gamma}_j=\sum_{i\in\mathcal{I}_j}\gamma_i, \tag{6}$$

where $\mathcal{I}_j$ is the index set of the aggregated cluster j. These aggregated values are then used in the MIP model (5) in place of the original ones, effectively reducing the model size while preserving the key characteristics of the scenario distribution. Then, the results of the MIP model are re-mapped to the original scenario set $\boldsymbol{\xi}$.

F. *Algorithm of IPDSR*

The overall IPDSR algorithm is summarized in Algorithm 1.

Algorithm 1: IPDSR

Input: Scenario set $\boldsymbol{\xi}$ of N scenarios and weights $\boldsymbol{\gamma}$.
Output: Scenario set $\boldsymbol{\zeta}$ of K scenarios and weights $\boldsymbol{\omega}$.
Step 1 - Initialization
  Initialize $\boldsymbol{\zeta}^0$ with $z_{\boldsymbol{\zeta}^0}^*$ and $\kappa=0$.
Step 2 - Iterative Scenario Reduction
  while $\kappa<$ *max_iter* **do**
    Project the scenario set $\boldsymbol{\xi}$ from the distribution space onto the problem space by computing $\boldsymbol{F}$.
    Perform objective aggregation on $\boldsymbol{F}$ and generate $\bar{\boldsymbol{F}}$ for MIP simplification.
    Construct and solve the MIP model in (5) on $\bar{\boldsymbol{F}}$ to obtain $\boldsymbol{\zeta}^{\kappa+1}$ with weights $\boldsymbol{\omega}^{\kappa+1}$.
    Update $z_{\boldsymbol{\zeta}^{\kappa+1}}^*=\underset{z\in Z}{\operatorname{argmin}}\ F(z,\boldsymbol{\zeta}^{\kappa+1})$.
    Compute the validation value $F(z_{\boldsymbol{\zeta}^{\kappa+1}}^*,\boldsymbol{\xi})$.
    Set $\kappa=\kappa+1$.
  end
  Select $\boldsymbol{\zeta}^{\kappa+1}$ and corresponding $\boldsymbol{\omega}^{\kappa+1}$ with the smallest $F(z_{\boldsymbol{\zeta}^{\kappa+1}}^*,\boldsymbol{\xi})$ as the final output.

Unlike the original PDSR formulation in Appendix A, which is strongly NP-hard and computationally intractable for standard solvers due to intricate nonlinear binary variable products, by structurally decoupling the problem via alternating optimization, the proposed IPDSR framework transforms this intractable exponential complexity into a sequence of rapidly solvable, small-scale linear MIPs. Note that two factors are required for the computational efficiency of IPDSR: 1) optimizing $z_{\boldsymbol{\zeta}^{\kappa}}^*$ on the reduced set $\boldsymbol{\zeta}$ ($K\ll N$) significantly shrinks the problem scale thereby reducing solution time; and 2) evaluating $F(z_{\boldsymbol{\zeta}^{\kappa}}^*,\xi_i)$ on the full set $\boldsymbol{\xi}$ via straightforward, parallelizable calculations under fixed decisions, yields significant time savings.

G. *Problem-Driven Evaluation Indices*

In this part, we introduce three *Ex-post* indices focusing on the impacts of SR on the optimality of SBSO-CVaR after solving the reduced problem.

1) Validation Distribution Similarity: The distribution similarity between $\{f(z^*_{\boldsymbol{\zeta}},\xi_{(i)})\}_{i=1}^N$ and $\{f(z^*_{\boldsymbol{\xi}},\xi_{(i)})\}_{i=1}^N$ can be evaluated by the Wasserstein distance (WD):

$$WD(\boldsymbol{\zeta},\boldsymbol{\xi})=\sum\nolimits_{i=1}^N \gamma_i \left|f(z^*_{\boldsymbol{\zeta}},\xi_{(i)})-f(z^*_{\boldsymbol{\xi}},\xi_{(i)})\right|, \quad (7)$$

where $\xi_{(i)}$ is the i-th scenario in the ordered set. A smaller WD implies a closer match between the validation outcomes, indicating that the reduced scenario set yields a solution with better applicability to the original scenarios.

2) Optimality Gap: After solving (1) and (2), we apply $z^*_{\boldsymbol{\xi}}$ and $z^*_{\boldsymbol{\zeta}}$ to $\boldsymbol{\xi}$, and calculate the percentage value of OG as

$$OG_{\boldsymbol{\zeta}}(\%):=\frac{F(z^*_{\boldsymbol{\zeta}},\boldsymbol{\xi})-F(z^*_{\boldsymbol{\xi}},\boldsymbol{\xi})}{F(z^*_{\boldsymbol{\xi}},\boldsymbol{\xi})}. \quad (8)$$

This metric indicates the percentage deviation of the approximated optimality from the optimality of the original problem, and should ideally be close to zero.

3) Representative Scenario Effectiveness: For the SBSO with representative scenarios, we introduce Scenario Effectiveness ($SE_{\zeta_k}(\%)$) to quantify the importance of scenario ζ_k. It is measured by the change in the OG (in percentage) when ζ_k is removed from $\boldsymbol{\zeta}$:

$$SE_{\zeta_k}(\%):=OG_{\boldsymbol{\zeta}_{-k}}(\%)-OG_{\boldsymbol{\zeta}}(\%), \quad (9)$$

where $\boldsymbol{\zeta}_{-k}=\boldsymbol{\zeta}\setminus\zeta_k$. A larger $SE_{\zeta_k}(\%)$ indicates a greater influence of ζ_k on the reduced problem outcomes. More details of this metric can be found in [15].

III. Day-Ahead Risk-Averse Offering Problem Formulation of a Virtual Power Plant

We consider a day-ahead risk-averse offering problem for a virtual power plant (VPP) trading with the upper grid. The VPP aggregates renewable energy sources (RES) and energy storage (ES) facilities, managing uncertainties from net load demand and market price fluctuations, which are described using a representative scenario set. The VPP makes risk-averse decisions on day-ahead power trading while considering intraday adjustments for deviations between the scheduled day-ahead power trading and the actual intraday power demand.

The VPP day-ahead risk-averse offering problem is formulated as (10).

$$\min_{\Xi}\sum\nolimits_{i=1}^N \gamma_i C_i+\lambda(v_\alpha+\frac{1}{1-\alpha}\sum\nolimits_{i=1}^N \gamma_i[C_i-v_\alpha]_+) \quad (10a)$$

$$C_i=\Delta t\sum\nolimits_{t=1}^T(\pi^{\mathrm{T}}_{i,t}P^{\mathrm{T}}_t+1.3\pi^{\mathrm{T}}_{i,t}P^{\mathrm{T}+}_{i,t}-0.7\pi^{\mathrm{T}}_{i,t}P^{\mathrm{T}-}_{i,t}) \quad (10b)$$

$$SoC_{i,t+1}=SoC_{i,t}+\Delta t(P^{\mathrm{E,c}}_{i,t}\eta^{\mathrm{c}}-P^{\mathrm{E,d}}_{i,t}/\eta^{\mathrm{d}})/E \quad (10c)$$

$$\sum\nolimits_{t=1}^T(P^{\mathrm{E,c}}_{i,t}\eta^{\mathrm{c}}-P^{\mathrm{E,d}}_{i,t}/\eta^{\mathrm{d}})\Delta t=0 \quad (10d)$$

$$0\le P^{\mathrm{E,c}}_{i,t}\le(1-D^{\mathrm{E}}_{i,t})\overline{P^{\mathrm{E}}},\ 0\le P^{\mathrm{E,d}}_{i,t}\le D^{\mathrm{E}}_{i,t}\overline{P^{\mathrm{E}}} \quad (10e)$$

$$\underline{SoC}\le SoC_{i,t}\le\overline{SoC} \quad (10f)$$

$$P^{\mathrm{T}}_t\le\overline{P}^{\mathrm{T}}_t,\ -\overline{P}^{\mathrm{T}}_t\le P^{\mathrm{T}}_t+P^{\mathrm{T}+}_{i,t}-P^{\mathrm{T}-}_{i,t}\le\overline{P}^{\mathrm{T}}_t \quad (10g)$$

$$0\le P^{\mathrm{R,u}}_{i,t}\le P^{\mathrm{R}}_{i,t} \quad (10h)$$

$$-P^{\mathrm{R,u}}_{i,t}+P^{\mathrm{L}}_{i,t}+P^{\mathrm{E,c}}_{i,t}-P^{\mathrm{E,d}}_{i,t}=P^{\mathrm{T}}_t+P^{\mathrm{T}+}_{i,t}-P^{\mathrm{T}-}_{i,t} \quad (10i)$$

$$0\le P^{\mathrm{T}+}_{i,t}\le D^{\mathrm{T}}_{i,t}\overline{P}^{\mathrm{T}+},\ 0\le P^{\mathrm{T}-}_{i,t}\le(1-D^{\mathrm{T}}_{i,t})\overline{P}^{\mathrm{T}-} \quad (10j)$$

$$D^{\mathrm{E}}_{i,t},\ D^{\mathrm{T}}_{i,t}\in\{0,1\}. \quad (10k)$$

where C_i is the cost of scenario i, $\pi^{\mathrm{T}}_{i,t}$ is the day-ahead market price of scenario i at time step t, P^{T}_t is the day-ahead trading power, $P^{\mathrm{T}+}_{i,t}/P^{\mathrm{T}-}_{i,t}$ are the up/down intraday balancing market power, $P^{\mathrm{E,c}}_{i,t}/P^{\mathrm{E,d}}_{i,t}$ are the charging/discharging power of the ES, $E/SoC_{i,t}$ are the energy capacity/state of charge of the ES, $\eta^{\mathrm{c}}/\eta^{\mathrm{d}}$ are the charging/discharging efficiency of the ES, $D^{\mathrm{E}}_{i,t}$ and $D^{\mathrm{T}}_{i,t}$ are binary variables indicating the ES charging state and the intraday balancing state, respectively. $\overline{P^{\mathrm{E}}}$ is the power rating of the ES, $\overline{P}^{\mathrm{T}}_t$ is the maximum power exchange with the upper network. $\underline{SoC}/\overline{SoC}$ are the minimum and maximum states of charge of the ES, respectively. Ξ is the set of decision variables. Eq. (10b) is the scenario-specific economic cost of the VPP, including the cost of day-ahead trading and intraday power balancing. Constraint (10c) limits the charging and discharging power and state of the ES. Constraint (10f) bounds the range of state of charge (SoC). Constraint (10d) guarantees that the SoC at the last time period is equal to the initial capacity. Constraint (10g) limits the day-ahead power trading and intraday balancing power. Constraint (10h) limits the renewable power curtailment. Constraint (10i) is the power balance equation of the VPP. Constraint (10j) ensures that the VPP cannot simultaneously buy and sell power in the intraday balancing market.

IV. Numerical Case Study

A. Set up

In the considered VPP operation problem, each individual scenario $\xi\in\mathbb{R}^{2\times T}$, $\xi\in\boldsymbol{\xi}$ is a multi-variable high-dimensional vector characterizing two sources of uncertainty (net load demand and market price). The time step is set to $\Delta t=15\mathrm{min}$ with $T=96$. The capacity of the wind power plant is set to $1000\mathrm{kW}$, the capacity of the ES is set to $200\mathrm{kW}$ with 2 hours of charging/discharging duration. The initial SoC of the ES is 0.5 and $\eta^{\mathrm{c}}=\eta^{\mathrm{d}}=0.95$. We set $\lambda=0.5$ and $\alpha=0.95$. The problem built on $\boldsymbol{\xi}$ is used as the Benchmark. The optimization is coded in Python with the Gurobipy interface and solved using Gurobi 12.0. The programming environment is Intel Core i9-13900HX @ 2.30GHz with RAM 16 GB.

B. Performance of IPDSR

We first visualize the nonlinear mapping between the distribution space and the problem space to highlight the effectiveness of IPDSR and the limitations of classical DDSR methods (e.g., hierarchical clustering, HC). For clarity, we consider a small-scale example with $N=100$ and $K=10$. We solve the original optimization problem (1) over $\boldsymbol{\xi}$ to obtain $z^*_{\boldsymbol{\xi}}$ and then compute the scenario-specific objective values $\{f(z^*_{\boldsymbol{\xi}},\xi_i)\}_{i=1}^N$. The histogram of $\{f(z^*_{\boldsymbol{\xi}},\xi_i)\}_{i=1}^N$ is shown as the light-colored bars in Fig. 1(a), with the VaR threshold indicated by a shaded bar. Fig. 1(a) illustrates the distribution of all scenarios in the problem space. Since the decision $z^*_{\boldsymbol{\xi}}$ is given, the problem space collapses to a one-dimensional space of values. Notably, there exists an extreme scenario in the right tail of the distribution, which is critical for SBSO-CVaR problems.

We then utilize the IPDSR framework to obtain the reduced scenario set $\boldsymbol{\zeta}$ and the associated weights $\boldsymbol{\omega}$. The values of $\{f(z^*_{\boldsymbol{\xi}},\zeta_k)\}_{k=1}^{K}$ are then mapped onto the same histogram as narrow blue bars, where scenarios falling into the same bin have their probabilities aggregated. For comparison with DDSR methods, we also mark the representative scenarios obtained by the HC method as narrow orange bars using the same procedure. Fig. 1(a) shows that scenarios obtained by IPDSR are more widely spread across the objective value distribution and effectively capture the objective value distribution features, especially in the right tail. In contrast, HC-selected scenarios are concentrated in the central region, failing to represent right-tail outcomes. This contrast stems from the methodological difference: DDSR performs clustering in the distribution space, while IPDSR clusters in the risk-averse problem space. Due to the nonlinear relationship between the distribution and problem spaces, similarity in distributional features does not guarantee effective decision-making in the problem space.

This point is further illustrated in Fig. 1(b), where we present four scenario samples. Although the blue and red scenarios exhibit significant differences in the distribution space, they fall into the same bin in the problem space. Conversely, the green and purple scenarios are distributionally similar but belong to different bins in the problem space. This clearly indicates that distributional similarity does not necessarily translate into decision-relevant similarity.

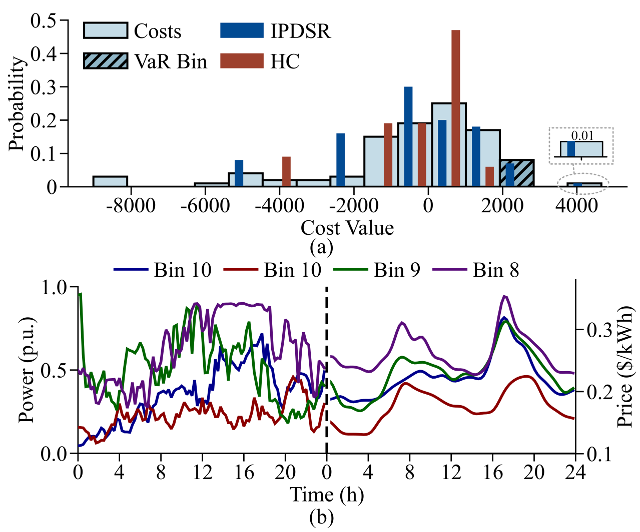

Fig. 1. Illustrative example of nonlinear mapping between distribution space and problem space: (a) scenarios in the problem space; (b) four example scenarios in the distribution space: left – net load scenarios; right – corresponding price scenarios.

Moreover, the optimality gap of IPDSR is only 0.50%, whereas that of HC is 6.83%, highlighting the superior performance of IPDSR. This is largely because IPDSR accounts for the SBSO-CVaR problem structure (particularly tail risk) in the SR process, leading to more adaptable representative scenarios and solutions. Overall, these results highlight the advantage of IPDSR in preserving tail-risk problem structures and improving the representativeness of reduced scenario sets in the problem space.

Fig. 2 uses a single iteration to illustrate the key steps and effects of IPDSR. Fig. 2(a) presents the initial distribution of the validation values $\{f(z^*_{\boldsymbol{\zeta}},\xi_i)\}_{i=1}^{N}$. Fig. 2(b) shows the results after effective objective aggregation, where the number of values is reduced from 100 to 50. This dimensionality reduction significantly alleviates the computational burden of the subsequent MIP problem (5), while preserving the essential characteristics of the original distribution. Based on the aggregated data in Fig. 2(b), the MIP (5) is then solved with an approximation loss of 0, and the results are shown by the orange bars in Fig. 2(c). Fig. 2(c) reveals that the cumulative distribution function (CDF) curve after the IPDSR iteration (orange line) closely aligns with the original distribution (blue line), particularly in the right-tail region beyond the VaR threshold. Notably, for clear illustration, the probabilities of *Costs-iter* are divided by 8. This indicates that IPDSR preserves critical tail structures during scenario selection, yielding risk-consistent solutions.

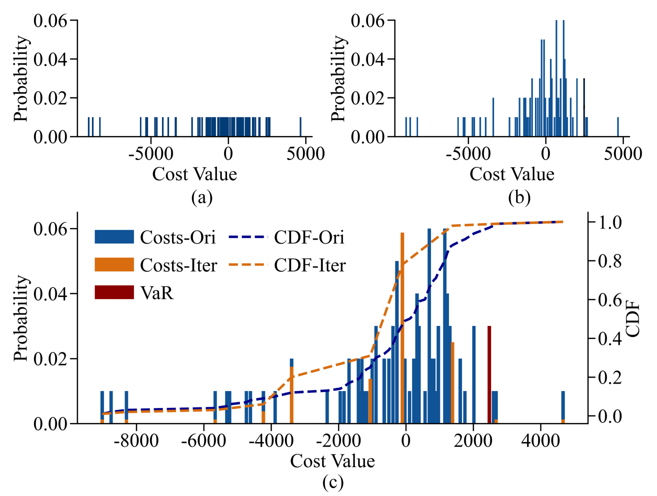

Fig. 2. Illustrative example of the iterative scenario reduction process: (a) initial values of $\{f(z^*_{\boldsymbol{\zeta}},\xi_i)\}_{i=1}^{N}$; (b) aggregated values of $\{f(z^*_{\boldsymbol{\zeta}},\xi_i)\}_{i=1}^{N}$; (c) reduced values of $\boldsymbol{\zeta}$ after one iterative process.

Moreover, we present the iterative curve of $OG(\%)$ in Fig. 3. The $OG(\%)$ value converges to a smaller value after a few iterations, demonstrating the effectiveness of the iteration process in IPDSR.

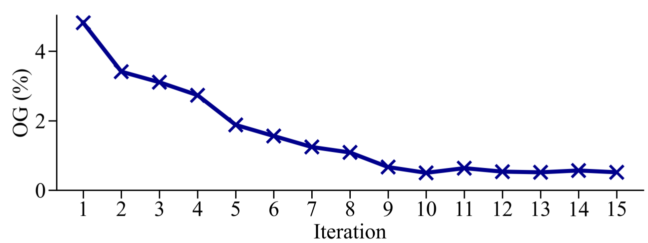

Fig. 3. The iterative curve of the OG value of IPDSR.

Next, we utilize the evaluation index of representative scenario effectiveness to further validate the performance of IPDSR, as illustrated in Table I. Table I suggests that some scenarios of the representative scenario set have relatively large values of $SE_{\zeta_k}(\%)$, indicating their significant impact on the outcomes of the reduced problem. This result demonstrates the effectiveness of the proposed IPDSR framework in capturing the salient features of the original scenario set.

TABLE I
EVALUATION RESULTS OF REPRESENTATIVE SCENARIO EFFECTIVENESS FOR $N=100$ AND $K=10$

	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5	ζ_6	ζ_7	ζ_8	ζ_9	ζ_{10}
ω_k	0.3	0.18	0.05	0.05	0.01	0.06	0.04	0.16	0.07	0.08
$SE_{\zeta_k}(\%)$	0.45	0.49	0.04	0.02	0.52	0.09	0.03	0.35	0.57	0.04

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

To further demonstrate the benefits of IPDSR, we conduct a comparative analysis with $N = 400$ and $K = 20$. DDSR methods including HC using Euclidean distance (HC-E), HC using Wasserstein distance (HC-W), K-means using Euclidean distance (KM-E), K-means using dynamic time warping distance (KM-D), and Gaussian mixture model using Mahalanobis distance (G-M) are compared. Regarding other PDSR research, we also include the method developed in [21] focusing on iterative tail scenario reduction (ITSR) for comparison. Recall that the PDSR in Section II-C is not tractable for existing standard solvers, it is not included for comparison.

The comparative indices include the SR performance indices and the computation efficiency indices. We define the values falling beyond the VaR threshold of $\{f(z^*_{\boldsymbol{\xi}}, \xi_i)\}_{i=1}^N$ as the "worst-case" scenarios. The considered SR performance indices include: the number of worst-case scenarios captured in the representative scenario set (ϱ), the WD defined in (7), and the $OG(\%)$ defined in (8). The considered computational 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 $K=20$

Method	ϱ	WD	$OG(\%)$	$\tau_{\mathrm{p}}(s)$	$\tau_{\mathrm{c}}(s)$
Benchmark	19	0	0	–	–
IPDSR	3	22.98	0.73	<0.3	<35
PDSR	Intractable for existing standard solvers				
ITSR [21]	9	162.87	8.86	<0.3	0.01
G-M	0	48.30	1.89	–	1.63
KM-D	0	36.34	2.74	–	12.7
KM-E	0	42.95	2.09	–	0.29
HC-E	0	33.70	1.37	–	0.02
HC-W	0	90.77	1.59	8.8	0.03

Optimality Gap: Table II suggests that IPDSR, with a small value of $OG(\%) = 0.73\%$, significantly outperforms other SR methods. The DDSR methods, seeking minimum statistical difference, fail to capture worst-case scenarios in their representative scenario sets. Such limitations lead to a neglect of potential risks during the operation, and insufficient ability to make risk-averse decisions for potential future uncertainties, thus yielding a relatively high $OG(\%)$. This discrepancy highlights that distributional similarity does not necessarily translate into approximated decision-making optimality in the problem space. The ITSR method [21] concentrates exclusively on tail scenarios and disregards the expected loss beyond the CVaR term, thereby incurring a significant OG bias and rendering it unsuitable for the SBSO-CVaR formulation considered in this paper. Compared to the above methods, IPDSR efficiently considers the potential impacts of the scenarios on the risk-averse problem, and includes three reasonable worst-case scenarios in the representative scenario set. These comparative findings suggest that IPDSR exhibits superior accuracy in representing the optimality of the original scenario set, thereby offering more reliable information for decision-making under uncertainty.

Computational Efficiency: The Benchmark struggles with computational complexity when N is large, especially for complex SBSO-CVaR problems. For example, for $N=400$ and $N=1000$ and a simple SBSO-CVaR problem as in Section III, the Benchmark takes nearly 4 minutes and 10 minutes, respectively. Moreover, for larger N and more complex SBSO-CVaR problems, the original problem becomes computationally intractable. The DDSR methods overcome computational bottlenecks, but at the price of potentially large $OG(\%)$ values. IPDSR involves N parallel computations of $\boldsymbol{F}=\{f(z^*_{\boldsymbol{\zeta}},\xi_i)\}_{i=1}^N$ in each iteration to obtain the validation values, which can be efficiently executed in parallel. In this paper, the computation time for each scenario-specific subproblem is $\tau_{\mathrm{s}}<0.3\ s$. For $N=400$ and $K=20$, after aggregating the $\boldsymbol{F}$ matrix into 100 values, the MIP clustering problem is solved consistently with MIP-gap with $\tau_{\mathrm{c}}<35\ s$.

D. Scalability of IPDSR

We further compare the $OG(\%)$ results of multiple SR methods under different N and K, as illustrated in Fig. 4. Note that the $OG(\%)$ of ITSR [21] is all above 8% and is thus not included. In Fig. 4, we observe that IPDSR outperforms DDSR methods across all values of N and K, as evidenced by its consistently lower $OG(\%)$. Besides, as expected, increasing K decreases the $OG(\%)$ for IPDSR, as more representative scenarios can better capture the distribution features, especially the right-tail scenarios. However, for DDSR methods, this inherent expected trend is not present, indicating that DDSR methods may capture some scenarios that are not impactful on the problem outcomes. This problem-driven reduction mechanism ensures that the retained scenarios are truly representative with respect to decision-making relevance. Based on the representative scenarios from IPDSR, solving the SBSO-CVaR problem yields superior operational strategies with high adaptability and effectiveness for potential future uncertainties. Moreover, unlike SR for the risk-neutral SBSO problem [15], the SBSO-CVaR formulation requires a markedly larger set of representative scenarios to achieve a satisfactory $OG(\%)$. This is attributable to its pronounced emphasis on tail risk, which can be adequately characterized only when the scenarios are sufficiently diverse.

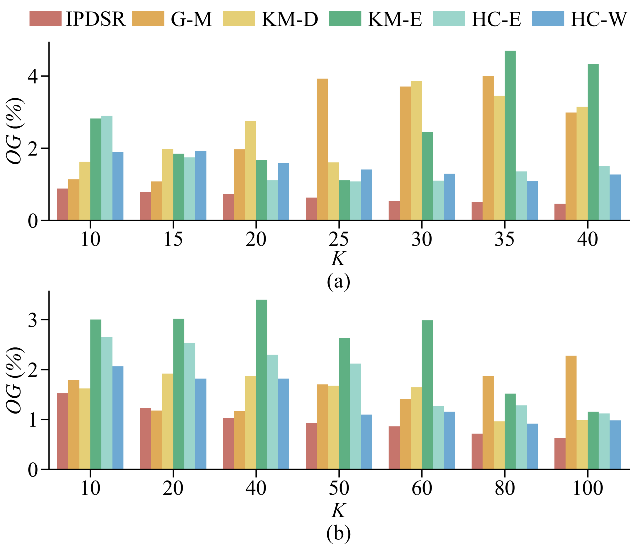

Fig. 4. Comparing results for different N and K: (a) $N=400$, (b) $N=1000$.

In practice, selecting K balances tail-risk accuracy against computational efficiency. A small K risks omitting α-defined tail scenarios (underestimating CVaR), whereas a large K diminishes SR's computational benefits. We recommend using the classic "Elbow Method" [23]. Leveraging IPDSR's rapid evaluation, practitioners can plot the OG trajectory via ex-ante tests (Fig. 4). The inflection point—where OG improvement becomes marginal—indicates the optimal K.

V. Conclusion

In this paper, we propose an IPDSR framework for power system SBSO-CVaR problems, which fully incorporates the risk-averse problem structure into the SR process. Specifically, we iteratively search for a set of salient representative scenarios that can best approximate the validation distribution of the original scenario set, thereby minimizing the optimality gap. Moreover, we propose an MIP formulation to conduct the scenario partitioning and representative scenario selection. In the case study, we visualize the nonlinear mapping relationship between the distribution space and the problem space. By conducting SR in the problem space, the presented IPDSR method obtains near-optimal approximation accuracy with just a few salient representative scenarios within an acceptable computational time. Moreover, a comprehensive comparative analysis with other SR methods is provided for different N and K values and broadly demonstrates the performance of our IPDSR method.

Future work will extend the method proposed in this paper to other types of risk measures such as standard deviation, and further improve the efficiency.

Appendix

In this appendix, we provide the detailed derivation process of the PDSR method in Section II-C, including the mathematical derivation process of the upper bound of the OG, the corresponding MIP formulation that minimizes the upper bound of the OG, and the algorithmic steps.

A. Derivation Process of the Upper Bound of OG

For brevity of mathematical notation, we denote $G(z,\xi_i)$ and $G(z,\zeta_k)$ as follows:

$$G(z,\xi_i)=f(z,\xi_i)+\frac{\lambda}{1-\alpha}[f(z,\xi_i)-v_{\boldsymbol{\xi}}^{\alpha}]_+ \tag{11a}$$

$$G(z,\zeta_k)=f(z,\zeta_k)+\frac{\lambda}{1-\alpha}[f(z,\zeta_k)-v_{\boldsymbol{\zeta}}^{\alpha}]_+ \tag{11b}$$

Thus, we have $F(z,\boldsymbol{\xi})=\sum_{i=1}^{N}\gamma_i G(z,\xi_i)+\lambda v_{\boldsymbol{\xi}}^{\alpha}$ and $F(z,\boldsymbol{\zeta})=\sum_{k=1}^{K}\omega_k G(z,\zeta_k)+\lambda v_{\boldsymbol{\zeta}}^{\alpha}$. Note that both the first and the last pair of (3) share a common expression of $F(z,\boldsymbol{\xi})-F(z,\boldsymbol{\zeta})$. Since $\sum_{i=1}^{N}\gamma_i=\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i=\sum_{k=1}^{K}\omega_k$, we have

$$\begin{aligned}
&|F(z,\boldsymbol{\xi})-F(z,\boldsymbol{\zeta})|\\
=&\left|\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i G(z,\xi_i)+\lambda v_{\boldsymbol{\xi}}^{\alpha}-\left(\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i G(z,\zeta_k)+\lambda v_{\boldsymbol{\zeta}}^{\alpha}\right)\right|\\
=&\left|\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i (G(z,\xi_i)-G(z,\zeta_k))+\lambda(v_{\boldsymbol{\xi}}^{\alpha}-v_{\boldsymbol{\zeta}}^{\alpha})\right|\\
\leq&\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i |G(z,\xi_i)-G(z,\zeta_k)|+\lambda|v_{\boldsymbol{\xi}}^{\alpha}-v_{\boldsymbol{\zeta}}^{\alpha}|
\end{aligned} \tag{12}$$

Similar to Section II-D2, we sort the scenarios in a non-decreasing manner with respect to the validated objective function $f(z,\cdot)$. Thus, we have

$$\begin{aligned}
&G(z,\xi_i)=f(z,\xi_i)+\frac{\lambda}{1-\alpha}(f(z,\xi_i)-f(z,\xi_{N_{\boldsymbol{\xi}}^{\alpha}}))\mathbf{1}_{i>N_{\boldsymbol{\xi}}^{\alpha}}\\
&G(z,\zeta_k)=f(z,\zeta_k)+\frac{\lambda}{1-\alpha}(f(z,\zeta_k)-f(z,\zeta_{K_{\boldsymbol{\zeta}}^{\alpha}}))\mathbf{1}_{k>K_{\boldsymbol{\zeta}}^{\alpha}}
\end{aligned} \tag{13}$$

where $\mathbf{1}_{\text{condition}}$ is the indicator function, which equals 1 if the condition is true and 0 otherwise.

We give a uniform formulation of $|G(z,\xi_i)-G(z,\zeta_k)|$ as:

$$\begin{aligned}
&|G(z,\xi_i)-G(z,\zeta_k)|\\
=&|a(f(z,\xi_i)-f(z,\zeta_k))+b(f(z,\xi_{N_{\boldsymbol{\xi}}^{\alpha}})-f(z,\zeta_{K_{\boldsymbol{\zeta}}^{\alpha}}))\\
&+c(f(z,\xi_i)-f(z,\xi_{N_{\boldsymbol{\xi}}^{\alpha}}))+d(f(z,\zeta_k)-f(z,\zeta_{K_{\boldsymbol{\zeta}}^{\alpha}}))|\\
\leq&|a||f(z,\xi_i)-f(z,\zeta_k)|+|b||f(z,\xi_{N_{\boldsymbol{\xi}}^{\alpha}})-f(z,\zeta_{K_{\boldsymbol{\zeta}}^{\alpha}})|\\
&+|c||f(z,\xi_i)-f(z,\xi_{N_{\boldsymbol{\xi}}^{\alpha}})|+|d||f(z,\zeta_k)-f(z,\zeta_{K_{\boldsymbol{\zeta}}^{\alpha}})|
\end{aligned} \tag{14}$$

where a, b, c, and d are coefficients summarized in Table III.

TABLE III
Values of a, b, c, and d for different cases

Case	a	b	c	d
$i\leq N_{\boldsymbol{\xi}}^{\alpha}$ and $k\leq K_{\boldsymbol{\zeta}}^{\alpha}$	1	0	0	0
$i> N_{\boldsymbol{\xi}}^{\alpha}$ and $k\leq K_{\boldsymbol{\zeta}}^{\alpha}$	1	0	$\frac{\lambda}{1-\alpha}$	0
$i\leq N_{\boldsymbol{\xi}}^{\alpha}$ and $k> K_{\boldsymbol{\zeta}}^{\alpha}$	$1+\frac{\lambda}{1-\alpha}$	$-\frac{\lambda}{1-\alpha}$	$-\frac{\lambda}{1-\alpha}$	0
$i> N_{\boldsymbol{\xi}}^{\alpha}$ and $k> K_{\boldsymbol{\zeta}}^{\alpha}$	$1+\frac{\lambda}{1-\alpha}$	$-\frac{\lambda}{1-\alpha}$	0	0

We introduce binary indicator variable g_i to indicate $\mathbf{1}_{i\le N_{\boldsymbol{\xi}}^{\alpha}}$, and h_k to indicate $\mathbf{1}_{k\le K_{\boldsymbol{\xi}}^{\alpha}}$. Then we have

$$\begin{aligned}&|G(z,\xi_i)-G(z,\zeta_k)|\\ &\le(h_k+(1+\frac{\lambda}{1-\alpha})(1-h_k))|f(z,\xi_i)-f(z,\zeta_k)|\\ &+\frac{\lambda}{1-\alpha}(1-h_k)|f(z,\xi_{N_{\xi}^{\alpha}})-f(z,\zeta_{K_{\zeta}^{\alpha}})|\\ &+\frac{\lambda}{1-\alpha}|h_k-g_i||f(z,\xi_i)-f(z,\xi_{N_{\xi}^{\alpha}})|\end{aligned}\tag{15}$$

Similar to our previous work [15], we define the problem-driven distance function $d(\xi_i,\zeta_k)$ as in (16) and have the relation as in (17). Please refer to [15] for more details about (16) and (17).

$$d(\xi_i,\zeta_k)=f(z^*_{\zeta_k},\xi_i)-f(z^*_{\xi_i},\xi_i)+f(z^*_{\xi_i},\zeta_k)-f(z^*_{\zeta_k},\zeta_k).\tag{16}$$

$$|f(z,\xi_i)-f(z,\zeta_k)|\le h(\|z\|)d(\xi_i,\zeta_k),\tag{17}$$

Then, we have

$$\begin{aligned}&|F(z,\boldsymbol{\xi})-F(z,\boldsymbol{\zeta})|\le\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i|G(z,\xi_i)-G(z,\zeta_k)|+\lambda|v^{\alpha}_{\boldsymbol{\xi}}-v^{\alpha}_{\boldsymbol{\zeta}}|\\ &\le h(\|z\|)(\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i((h_k+(1+\frac{\lambda}{1-\alpha})(1-h_k))d(\xi_i,\zeta_k)\\ &+(\frac{\lambda}{1-\alpha}(1-h_k))d(\xi_{N_{\boldsymbol{\xi}}^{\alpha}},\zeta_{K_{\boldsymbol{\xi}}^{\alpha}})+\frac{\lambda}{1-\alpha}|h_k-g_i|d(\xi_i,\xi_{N_{\boldsymbol{\xi}}^{\alpha}}))\\ &+\lambda d(\xi_{N_{\boldsymbol{\xi}}^{\alpha}},\zeta_{K_{\boldsymbol{\xi}}^{\alpha}}))\end{aligned}\tag{18}$$

Since Z is bounded, $\|z\|$ is well-defined, i.e., $\exists M\gg1$, $\|z\|\le M$, $\forall z\in Z$. Thus, we have the upper bound of OG as follows:

$$\begin{aligned}&OG\le|F(z^*_{\boldsymbol{\zeta}},\boldsymbol{\xi})-F(z^*_{\boldsymbol{\zeta}},\boldsymbol{\zeta})|+|F(z^*_{\boldsymbol{\xi}},\boldsymbol{\xi})-F(z^*_{\boldsymbol{\xi}},\boldsymbol{\zeta})|\\ &\le2h(M)(\sum_{k=1}^{K}\sum_{i\in C_k}\gamma_i((h_k+(1+\frac{\lambda}{1-\alpha})(1-h_k))d(\xi_i,\zeta_k)\\ &+(\frac{\lambda}{1-\alpha}(1-h_k))d(\xi_{N_{\boldsymbol{\xi}}^{\alpha}},\zeta_{K_{\boldsymbol{\xi}}^{\alpha}})+\frac{\lambda}{1-\alpha}|h_k-g_i|d(\xi_i,\xi_{N_{\boldsymbol{\xi}}^{\alpha}}))\\ &+\lambda d(\xi_{N_{\boldsymbol{\xi}}^{\alpha}},\zeta_{K_{\boldsymbol{\xi}}^{\alpha}}))\end{aligned}\tag{19}$$

B. MIP Formulation

Since $2h(M)$ is a constant, minimizing the following MIP formulation achieves the lowest upper bound of OG.

$$\min\sum_{j=1}^{N}l_j^{\mathrm{a}}+\sum_{j=1}^{N}l_j^{\mathrm{b}}+\sum_{j=1}^{N}l_j^{\mathrm{c}}+t\tag{20a}$$

$$\text{s.t.}\sum_{i=1}^{N}v_{ij}\gamma_i(h_j+(1+\frac{\lambda}{1-\alpha})(1-h_j))\tilde{F}_{ij}\le l_j^{\mathrm{a}},\tag{20b}$$

$$\sum_{i=1}^{N}v_{ij}\gamma_i(\frac{\lambda}{1-\alpha}(1-h_j))\sum_{p=1}^{N}\sum_{q=1}^{N}m_pn_q\tilde{F}_{pq}\le l_j^{\mathrm{b}},\tag{20c}$$

$$\sum_{i=1}^{N}v_{ij}\gamma_i\frac{\lambda}{1-\alpha}|h_j-g_i|\sum_{p=1}^{N}m_p\tilde{F}_{ip}\le l_j^{\mathrm{c}},\tag{20d}$$

$$\lambda\sum_{p=1}^{N}\sum_{q=1}^{N}m_pn_q\tilde{F}_{pq}\le t,\tag{20e}$$

$$v_{ij}\le u_j,\ v_{jj}=u_j,\ n_j\le u_j,\tag{20f}$$

$$h_j\le u_j,\ m_j\le g_j,\ n_j\le h_j,\tag{20g}$$

$$\sum\nolimits_{j=1}^{N}v_{ij}=1,\tag{20h}$$

$$\sum\nolimits_{j=1}^{N}u_j=K,\tag{20i}$$

$$\sum\nolimits_{j=1}^{N}m_j=1,\ \sum\nolimits_{j=1}^{N}n_j=1,\tag{20j}$$

$$\sum\nolimits_{i=1}^{N}g_i\gamma_i\le\alpha,\tag{20k}$$

$$\sum\nolimits_{i=1}^{N}g_i\gamma_i+\gamma_k(1-g_k)\ge\alpha(1-g_k),\tag{20l}$$

$$\sum\nolimits_{j=1}^{N}h_j\sum_{i=1}^{N}v_{ij}\gamma_i\le\alpha,\tag{20m}$$

$$\sum_{j=1}^{N}h_j\sum_{i=1}^{N}v_{ij}\gamma_i+(u_k-h_k)\sum_{i=1}^{N}v_{ik}\gamma_i\ge\alpha(u_k-h_k),\tag{20n}$$

$$\{v_{ij},\ u_j,\ m_j,\ n_j,\ h_k,\ g_i\}\in\{0,1\}.\tag{20o}$$

where $\tilde{F}_{ij}:=f_{ij}-f_{jj}+f_{ji}-f_{ii}$ and $f_{ij}=f(z^*_{\xi_i},\xi_j)$. m_j indicates whether scenario ξ_j is the VaR scenario of $\boldsymbol{\xi}$, and n_j indicates whether scenario ξ_j is the VaR scenario of $\boldsymbol{\zeta}$. Constraint $\sum_{j=1}^{N}m_j=1$ ensures that there is only one VaR scenario in the original scenario set $\boldsymbol{\xi}$, and $\sum_{j=1}^{N}n_j=1$ ensures that there is only one VaR scenario in the reduced scenario set $\boldsymbol{\zeta}$. Constraints (20k)-(20l) ensures that the selected VaR scenario is valid.

However, the MIP formulation involves numerous binary variables and intricate constraints, making it computationally intractable for standard solvers. We believe that this formulation can still provide valuable insights for SR in SBSO-CVaR problems, and can be used as a theoretical basis for further research.

C. Algorithm of PDSR

The PDSR method is illustrated in **Algorithm 1**.

Algorithm 2: Problem-Driven Scenario Reduction

Input: Scenario set $\boldsymbol{\xi}$ of N scenarios and weights $\boldsymbol{\gamma}$.
Output: Scenario set $\boldsymbol{\zeta}$ of K scenarios and weights $\boldsymbol{\omega}$.
Step 1 - Projection in Problem Space
  Step 1.1: Initialize problem space matrix $\boldsymbol{F}=\mathbf{0}$.
  Step 1.2: Project the distribution space onto the problem space by:
  for $i=1$ **to** N **do**
    Solve the scenario-specific SO 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**
      Set $f_{ij}=f(z^*_{\xi_i},\xi_j)$.
    end
  end
Step 2 - Clustering
  Solve MIP in (20) to obtain the representative scenario set $\boldsymbol{\zeta}$ with weights $\boldsymbol{\omega}$.

REFERENCES

[1] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo, "Power systems optimization under uncertainty: A review of methods and applications," *Electric Power Systems Research*, vol. 214, p. 108725, 2023.

[2] T. Tan, M. Yang, R. Xie, Y. Cao, and Y. Chen, "Adjustable robust optimization with decision-dependent uncertainty for power system problems: A review," *Energy Conversion and Economics*, vol. 6, no. 3, pp. 155–169, 2025.

[3] N. Qi, P. Pinson, M. R. Almassalkhi, L. Cheng, and Y. Zhuang, "Chance-constrained generic energy storage operations under decision-dependent uncertainty," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 4, pp. 2234–2248, 2023.

[4] N. Qi, P. Pinson, M. R. Almassalkhi, Y. Zhuang, Y. Su, and F. Liu, "Capacity credit evaluation of generalized energy storage considering strategic capacity withholding and decision-dependent uncertainty," *Applied Energy*, vol. 397, p. 126310, 2025.

[5] M. Wang, X. Li, C. Dong, Y. Mu, H. Jia, and F. Li, "Day-ahead optimal bidding for a retailer with flexible participation of electric vehicles," *IEEE Transactions on Smart Grid*, vol. 14, no. 2, pp. 1482–1494, 2023.

[6] J. Zhu, G. Li, Y. Guo, J. Chen, H. Liu, Y. Luo, and W. Liu, "Real-time risk-averse dispatch of an integrated electricity and natural gas system via conditional value-at-risk-based lookup-table approximate dynamic programming," *Protection and Control of Modern Power Systems*, vol. 9, no. 2, pp. 47–60, 2024.

[7] J. Hu, M. Shi, X.-p. Zhang, and Y. Jia, "Stochastic carbon footprint tracing for power systems with uncertainty," *Energy Conversion and Economics*, vol. 6, no. 2, pp. 101–110, 2025.

[8] H. Li, Z. Ren, M. Fan, W. Li, Y. Xu, Y. Jiang, and W. Xia, "A review of scenario analysis methods in planning and operation of modern power systems: Methodologies, applications, and challenges," *Electric Power Systems Research*, vol. 205, p. 107722, 2022.

[9] N. Qi, L. Cheng, H. Li, Y. Zhao, and H. Tian, "Portfolio optimization of generic energy storage-based virtual power plant under decision-dependent uncertainties," *Journal of Energy Storage*, vol. 63, p. 107000, 2023.

[10] X. Dong, Y. Sun, S. M. Malik, T. Pu, Y. Li, and X. Wang, "Scenario reduction network based on wasserstein distance with regularization," *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 4–13, 2024.

[11] S. Lin, C. Liu, Y. Shen, F. Li, D. Li, and Y. Fu, "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.

[12] A. M. Alonso, F. J. Nogales, and C. Ruiz, "Hierarchical clustering for smart meter electricity loads based on quantile autocovariances," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4522–4530, 2020.

[13] Z. Shi, L. Wu, and Y. Zhou, "Predicting household energy consumption in an aging society," *Applied Energy*, vol. 352, p. 121899, 2023.

[14] H. Teichgraeber 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.

[15] Y. Zhuang, L. Cheng, N. Qi, M. R. Almassalkhi, and F. Liu, "Problem-driven scenario reduction framework for power system stochastic operation," *IEEE Transactions on Power Systems*, vol. 40, no. 4, pp. 3232–3246, 2025.

[16] A. N. Elmachtoub and P. Grigas, "Smart "predict, then optimize"," *Management Science*, vol. 68, no. 1, pp. 9–26, 2022.

[17] J. Mandi, J. Kotary, S. Berden, M. Mulamba, V. Bucarey, T. Guns, and F. Fioretto, "Decision-focused learning: Foundations, state of the art, benchmark and future opportunities," *Journal of Artificial Intelligence Research*, vol. 80, pp. 1623–1701, 2024.

[18] S. Zhao, C. Shao, J. Ding, B. Hu, K. Xie, X. Yu, and Z. Zhu, "Unreliability tracing of power systems for identifying the most critical risk factors considering mixed uncertainties in wind power output," *Protection and Control of Modern Power Systems*, vol. 9, no. 5, pp. 96–111, 2024.

[19] S. M. Mohseni-Bonab, I. Kamwa, A. Rabiee, and C. Chung, "Stochastic optimal transmission switching: A novel approach to enhance power grid security margins through vulnerability mitigation under renewables uncertainties," *Applied Energy*, vol. 305, p. 117851, 2022.

[20] K. Li, F. Yang, L. Wang, Y. Yan, H. Wang, and C. Zhang, "A scenario-based two-stage stochastic optimization approach for multi-energy microgrids," *Applied Energy*, vol. 322, p. 119388, 2022.

[21] R. García-Bertrand and R. Mínguez, "Iterative scenario based reduction technique for stochastic optimization using conditional value-at-risk," *Optimization and Engineering*, vol. 15, no. 2, pp. 355–380, 2014.

[22] J. Fairbrother, A. Turner, and S. W. Wallace, "Problem-driven scenario generation: an analytical approach for stochastic programs with tail risk measure," *Mathematical Programming*, vol. 191, no. 1, pp. 141–182, 2022.

[23] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 1990.