PSCC 2026

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 Tsinghua University; 2 Columbia University; 3 University of Vermont

Thursday June 11th, 2026

Paper Number: EPSR-D-25-06631

Congrats to newly minted to *Dr.* Yingrui Zhuang

Decision-Focused Learning for Risk Management in the Coordinated Operation of Distribution Networks and Microgrids

Dissertation submitted to

Tsinghua University

in partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Zhuang Yingrui

Dissertation Supervisor: Professor Cheng Lin

May, 2026

University of Vermont

Two PDSR papers in one presentation!

3232 IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 40, NO. 4, JULY 2025

Problem-Driven Scenario Reduction Framework for Power System Stochastic Operation

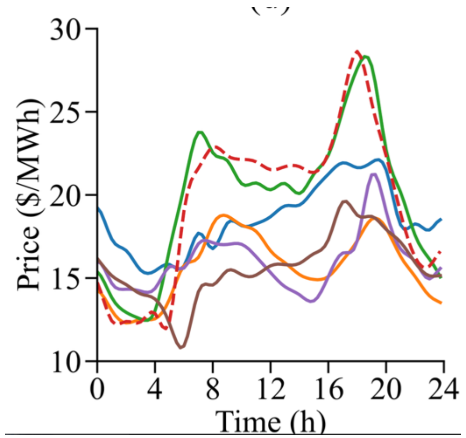
Yingrui Zhuang, *Graduate Student Member, IEEE*, Lin Cheng, *Senior Member, IEEE*,
Ning Qi, *Member, IEEE*, Mad

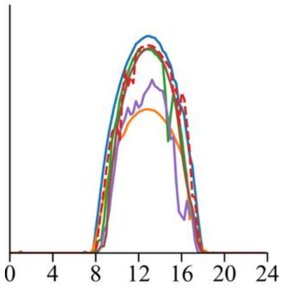
An Iterative Problem-Driven Scenario Reduction Framework for Stochastic Optimization with Conditional Value-at-Risk

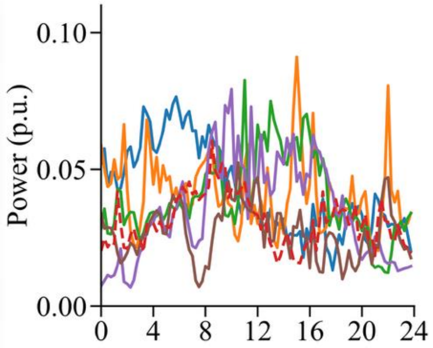
Yingrui Zhuang[1], Lin Cheng[1], Ning Qi[2*], Mads R. Almassalkhi[3], Feng Liu[1]
[1]*Department of Electrical Engineering, Tsinghua University, Beijing, China*
[2]*Department of Earth and Environmental Engineering, Columbia University, New York, USA*
[3]*Department of Electrical and Biomedical Engineering,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

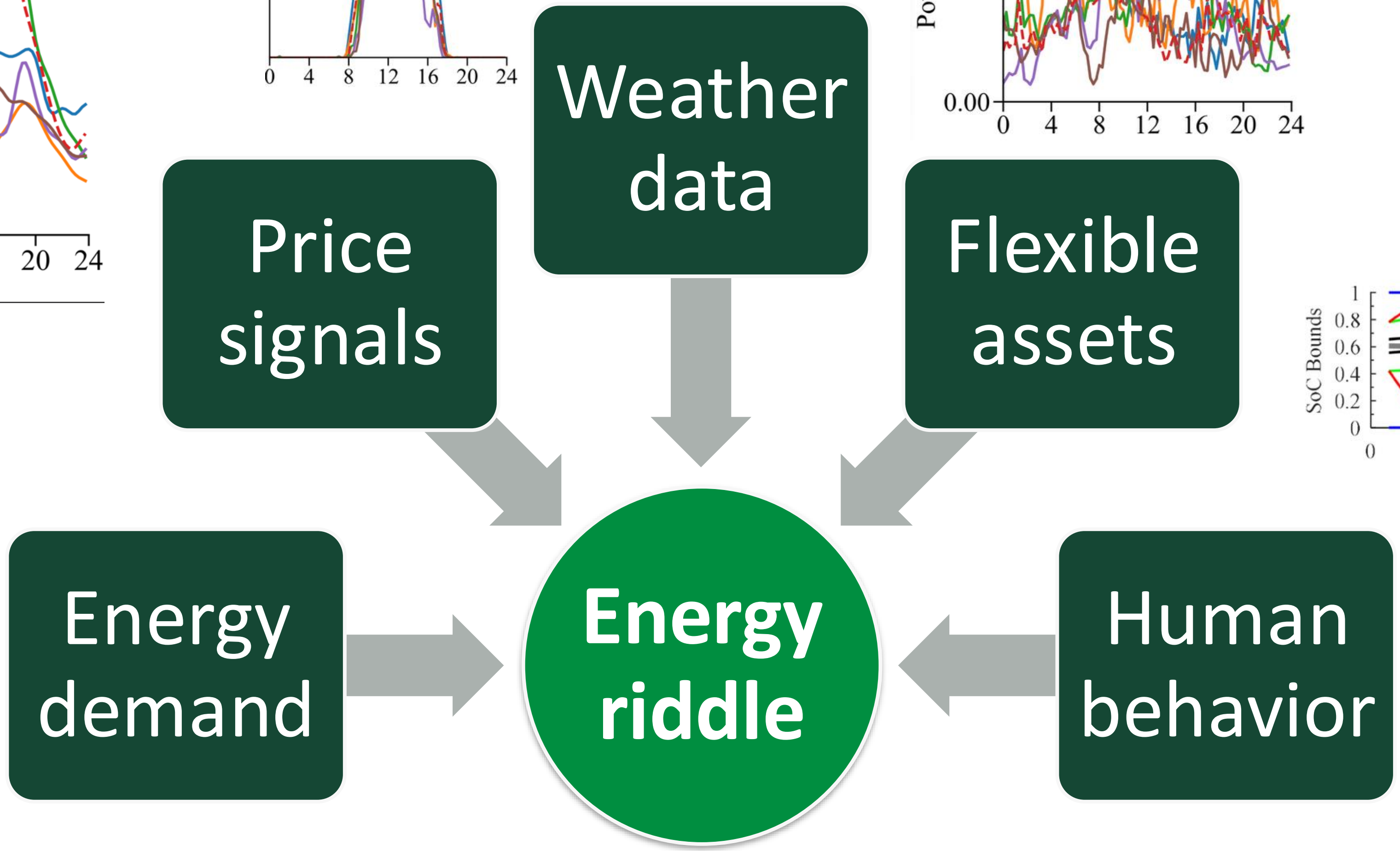
University of Vermont

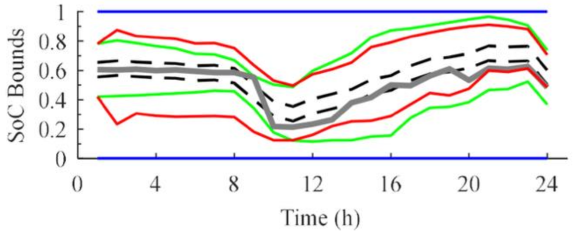
Solving energy riddles: optimization under uncertainty

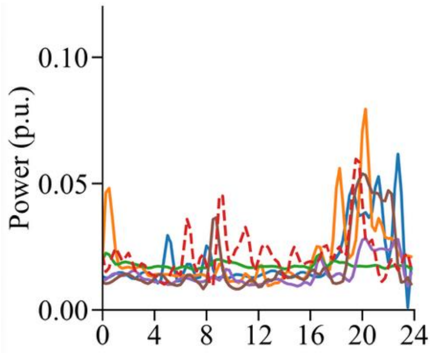












University of Vermont

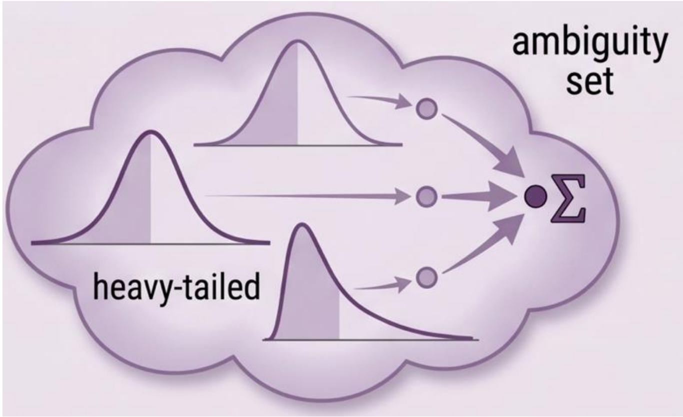
A. Khurram, Luis Duffaut Espinosa, Roland Malhamé, Mads Almassalkhi, "Identification of Hot Water End-use Process of EWHs from Energy Measurements," PSCC, 2020

M. Amini and M. Almassalkhi, "Corrective optimal dispatch of uncertain virtual energy resources," IEEE Transactions on Smart Grid, 2020.

N. Qi, P. Pinson, M. Almassalkhi, et al, "Chance Constrained Economic Dispatch of Generic Energy Storage under Decision-Dependent Uncertainty," IEEE TSE. 2023.

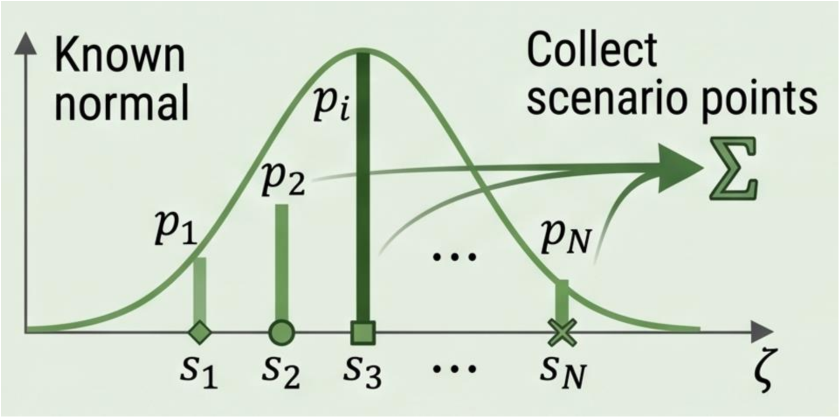
But all uncertainties are not the same

Distribution is (partially) known

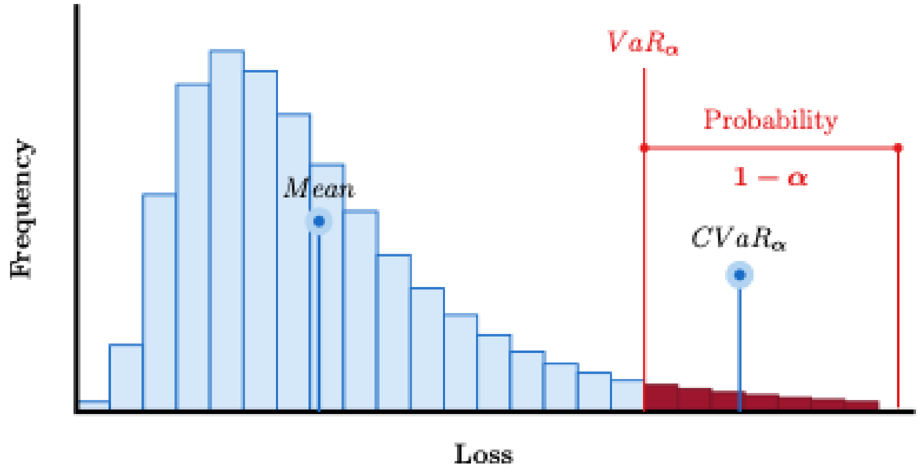


Distributionally Robust Optimization
(e.g., chance-constraint or CC)

Distribution is unknown



Scenario-based optimization
(e.g., forecasted scenarios, SAA)



Operators care about tail outcomes: low probability, high loss/costs.

VaR: quantile cost (defines CC)
CVaR: conditional value at risk, expected tail costs.

A working energy riddle example: day-ahead VPP offering

A VPP (renewables + storage) trades day-ahead, <u>then</u> adjusts intraday: *two stages*

symbol	role	in the VPP example
z	here-and-now decision	day-ahead trading schedule
ξ_i	uncertainty realization	net load + price trajectory i
y_i	wait-and-see recourse	intraday adjustments, storage dispatch
$f_1(z)$	first-stage cost	day-ahead trading cost
$g(y_i, \xi_i)$	recourse cost	imbalance / regulation cost

How are decisions, scenarios, and recourse variables related?

$$\underbrace{f(z,\xi_i) \;:=\; f_1(z)}_{\textit{Cost of decision } z} \;+\; \underbrace{Q(z,\xi_i), \qquad Q(z,\xi_i) \;:=\; \min_{y_i \in \mathcal{Y}(z,\xi_i)} g(y_i,\xi_i)}_{\textit{Cost of adjustment } y \textit{ under decision } z \textit{ for realization } i}$$

Assumption: for *all* decisions z and *any* scenario, there is a **feasible** adjustment y ➔ Q is finite (*relatively complete recourse*)

Two-stage stochastic optimization (TSSO) problem

Consider N=1 scenario, ξ_i: simple, deterministic problem!

$$\min_{z\in Z} f(z,\xi_i) = \min_{z\in Z}\left[f_1(z) + \min_{y\in\mathcal{Y}(z,\xi_i)} g(y,\xi_i)\right] = \min_{z\in Z,\ y\in\mathcal{Y}(z,\xi_i)} \left[f_1(z) + g(y,\xi_i)\right]$$

Wait-and-see **solution** ➔ $z^*_{\xi_i} = \arg\min_{z\in Z} f(z,\xi_i)$

What if z **is fixed/constant/data?**

$$\bar{z} \in Z \quad ➔ \quad f(\bar{z},\xi_i) = \underbrace{f_1(\bar{z})}_{\text{constant}} + \min_{y\in\mathcal{Y}(\bar{z},\xi_i)} g(y,\xi_i)$$

One simple recourse solve

Stack N scenarios: *one* decision serves all

N>>1 scenarios: $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_N\}$ **with weights** $\gamma_i > 0, \sum_{i=1}^{N} \gamma_i = 1$

Stacked optimization is harder now:

$$
\begin{aligned}
z^*_{\boldsymbol{\xi}} = \arg \min_{z,\, y_1, \ldots, y_N} \quad & \sum_{i=1}^{N} \gamma_i \big(f_1(z) + g(y_i, \xi_i)\big) \\
\text{s.t.} \quad & z \in Z, \\
& y_1 \in \mathcal{Y}(z, \xi_1) \\
& \vdots \\
& y_N \in \mathcal{Y}(z, \xi_N)
\end{aligned}
$$

If z is fixed/data, evaluation is trivial!

$$
\left.
\begin{aligned}
& \min_{y_1 \in \mathcal{Y}(\bar{z}, \xi_1)} g(y_1, \xi_1) \\
& \min_{y_2 \in \mathcal{Y}(\bar{z}, \xi_2)} g(y_2, \xi_2) \\
& \vdots \\
& \min_{y_N \in \mathcal{Y}(\bar{z}, \xi_N)} g(y_N, \xi_N)
\end{aligned}
\right\}
$$

N independent problems
➔ solve in parallel!
➔ begets $F(\bar{z}, \xi)$

Full SBSO objective on full scenario set can also add CVaR: <u>tail fidelity makes N LARGE!</u>

$$
F(z, \xi) := \underbrace{\sum_{i=1}^{N} \gamma_i f(z, \xi_i)}_{\text{expected cost}} + \lambda \underbrace{\Big(\underset{\downarrow \, \mathrm{VaR}_\alpha}{v^\alpha_\xi} + \tfrac{1}{1-\alpha} \sum_{i=1}^{N} \gamma_i \left[f(z, \xi_i) - v^\alpha_\xi \right]_+ \Big)}_{\mathrm{CVaR}_\alpha[f(z,\xi)]}, \qquad \lambda \geq 0,
$$

NB! $\lambda = 0 \Rightarrow TPWRS$
$\lambda > 0 \Rightarrow PSCC$

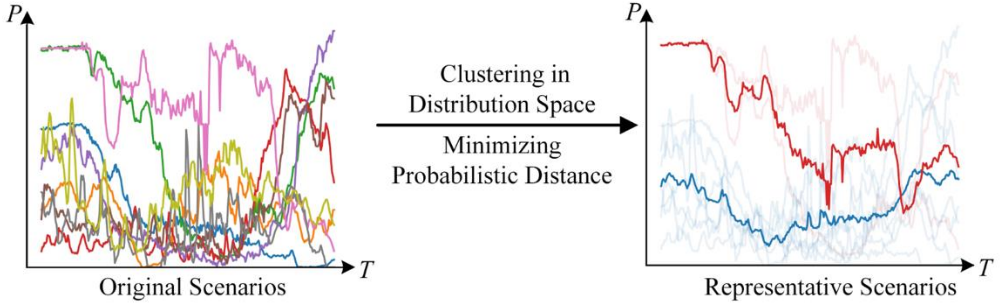
University of Vermont

The Philosophies of Scenario Reduction: DDSR vs. PDSR

Research Objective: reduce set of *N* scenarios, ξ, to a representative set of $K \ll N$ scenarios, $\zeta \subset \xi$, that minimizes the optimality gap (OG), $\mathrm{OG} := F(z^*_\zeta, \xi) - F(z^*_\xi, \xi) \geq 0$

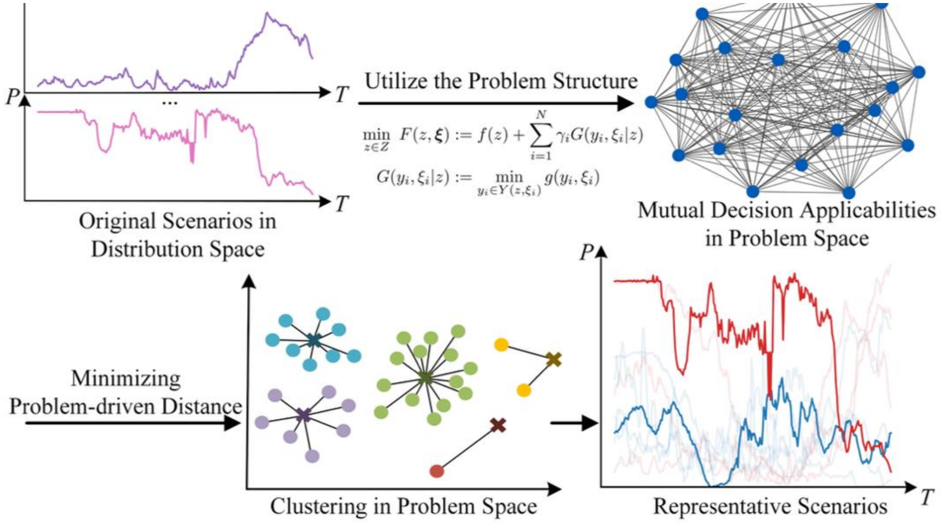
Distribution-driven SR (DDSR)

- Clusters scenarios in *distribution space*
  - K-means, GMM, etc.
- Blind to optimization problem.
- *Statistical similarity ≠ decision similarity!*
  - Only sees tails in distribution space.



Problem-driven SR (PDSR)

- Clusters scenarios in *problem space* by
- *Mutual decision applicability:* $f_{ij} := f(z^*_{\xi_i}, \xi_j)$
- SR based on optimization problem!
- *Prior PDSR Lit*: risk-neutral, no OG characterization, limited scalability.



Risk-neutral case: bounding the optimality gap

Reduced set K clusters: $\boldsymbol{\zeta} = \{\zeta_1, \ldots, \zeta_K\} \subset \boldsymbol{\xi}, \quad \omega_k = \sum_{i \in C_k} \gamma_i, \quad z^*_{\boldsymbol{\zeta}} = \arg\min_z F(z, \boldsymbol{\zeta}).$

1. **Augment OG**: $\mathrm{OG} \;\le\; F(z^*_\zeta, \xi) - F(z^*_\xi, \xi) + F(z^*_\xi, \zeta) - F(z^*_\zeta, \zeta)$
2. Convert *set-level diff* into ***pairwise intra-cluster*** **mismatches**

$$F(z, \boldsymbol{\xi}) - F(z, \boldsymbol{\zeta}) = \sum_{k=1}^{K} \sum_{i \in C_k} \gamma_i\, f(z, \xi_i) - \sum_{k=1}^{K} \sum_{i \in C_k} \gamma_i\, f(z, \zeta_k) = \sum_{k=1}^{K} \sum_{i \in C_k} \gamma_i \underbrace{\big(f(z, \xi_i) - f(z, \zeta_k)\big)}_{\text{scenario } i \text{ vs. its representative}}$$

3. **Apply to OG to get *set-level decisions* vs. scenarios**:

$$\mathrm{OG} \;\le\; \sum_{k=1}^{K} \sum_{i \in C_k} \gamma_i \Big(\big| f(z^*_{\boldsymbol{\zeta}}, \xi_i) - f(z^*_{\boldsymbol{\zeta}}, \zeta_k) \big| + \big| f(z^*_{\boldsymbol{\xi}}, \xi_i) - f(z^*_{\boldsymbol{\xi}}, \zeta_k) \big| \Big).$$

4. And $||z|| < M$ & metric $d(.)$ **bounds term:** $\big| f(z, \xi_i) - f(z, \zeta_k) \big| \;\le\; h(M)\, d(\xi_i, \zeta_k)$

5. **Combine** ➔ $\mathrm{OG} \;\le\; 2h(M) \cdot \underbrace{\sum_{k=1}^{K} \sum_{i \in C_k} \gamma_i\, d(\xi_i, \zeta_k)}_{\mathrm{SPDD}}$

Minimize SPDD
➔ minimize upper OG bound (decision-free!)

University of Vermont

Distance metric and Decoupling

Defining problem-driven distance (PDD) metric as *symmetric opp. costs*:

$$d(\xi_i, \zeta_k) := \underbrace{f(z^*_{\zeta_k}, \xi_i) - f(z^*_{\xi_i}, \xi_i)}_{\text{cost of using } \zeta_k\text{'s decision on } \xi_i} + \underbrace{f(z^*_{\xi_i}, \zeta_k) - f(z^*_{\zeta_k}, \zeta_k)}_{\text{and vice versa}}$$

How to pre-compute metric PDD? Need to compute $f_{ij} := f(z^*_{\xi_i}, \xi_j)$

- **Diagonals** ($i=j$) are N <u>single-scenario</u> joint *wait-and-see* solves for $(z_{\xi_i}{}^*, y_i)$
- **Off-diagonals** *($i\neq j$)* are $N(N\text{-}1)$ <u>fixed</u>-$z_{\xi i}{}^*$ simple, parallel *recourse* problems!
- **Optimality Gap becomes:**

Clustering selects which *f*-entries are summed

j = the scenario chosen as ζ_k ➔ *Optimize selection?*

$$\text{OG} \leq 2h(M) \cdot \sum_{k=1}^{K} \sum_{i \in C_k} \gamma_i \left(f_{ji} - f_{ii} + f_{ij} - f_{jj}\right)$$

Clustering as a MILP (TPWRS result)

Solving MILP selects K representatives *and* partitions N scenarios

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

$$\text{s.t.} \quad \sum_{i=1}^{N} \gamma_i v_{ij} \left(f_{ji} - f_{ii} + f_{ij} - f_{jj}\right) \leq l_j, \quad \forall j$$

$$v_{ij} \leq u_j, \quad v_{jj} = u_j, \quad \textstyle\sum_j v_{ij} = 1, \quad \sum_j u_j = K, \quad v_{ij}, u_j \in \{0,1\}.$$

Minimizes bound on OG

$u_j = 1$ ➔ ξ_j **is a representative scenario.** $v_{ij} = 1$ ➔ **scenario** ξ_i **is represented by** ξ_j

Does it "work"?

- **Case studies**: 1) ADN dispatch and 2) unit-commitment (N=400 scenarios)
- **Outperforms** 9 state-of-the-art SR methods (OG = 0.09% and 0.20%)
- **Compute**: clustering MILP is efficient 3-18s (even for N+N^2 binaries!)
  - Setting up MILP requires all NxN f_{ij}: Fully parallel = 0.6s! On 100 cores: 8 mins.
  - Note that Benchmark (full stack) takes 51mins

Clustering as a MILP (TPWRS result)

Solving MILP selects K representatives *and* partitions N scenarios

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

Minimizes bound on OG

Are we done? Only with **risk-neutral** case, where we leveraged structure: the OG bound is built from a **static, decision-free** matrix f_{ij}. ***CVaR destroys exactly that!***

- **Case studies**: 1) ADN dispatch and 2) unit-commitment (N=400 scenarios)
- **Outperforms** 9 state-of-the-art SR methods (OG = 0.09% and 0.20%)
- **Compute**: clustering MILP is efficient 3-18s (even for $N+N^2$ binaries!)
  - Setting up MILP requires all NxN f_{ij}: Fully parallel = 0.6s! On 100 cores: 8 mins.
  - Note that Benchmark (full stack) takes 51mins

Risk-averse case (PSCC '26): CVaR breaks structure!

$$F(z, \boldsymbol{\xi}) = \sum_i \gamma_i f(z, \xi_i) + \lambda \left(v_{\boldsymbol{\xi}}^{\alpha} + \frac{1}{1-\alpha} \sum_i \gamma_i \left[f(z, \xi_i) - v_{\boldsymbol{\xi}}^{\alpha} \right]_+ \right)$$

VaR threshold is depends on z and ordering of all scenarios (in problem-space) ➔ **No separation.**

Non-smooth operation: scenario's contribution depends on which side of VaR it lands

- "Distance" between two scenarios now *depends on everyone else* (order!)➔ complex f_{ij}***!***
- *Can we still get OG bound?* Yes! Direct extension, but clustering MIP is now MINLP and not tractable! ➔ See Appendix A. *(Not our proposed method, only motivation!)*

Iterative approach: fix decision, restore separability!

***Key observation:** $F(z_\xi{}^*, \xi)$ is constant w.r.t. reduced set ζ ➔ the OG can be minimized by minimizing $F(z_\zeta{}^*, \xi)$ ➔ No need for upper bound on OG.*

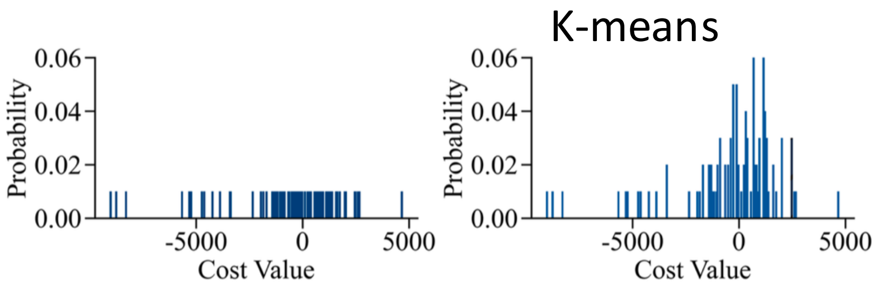
Proposed Iterative Approach:

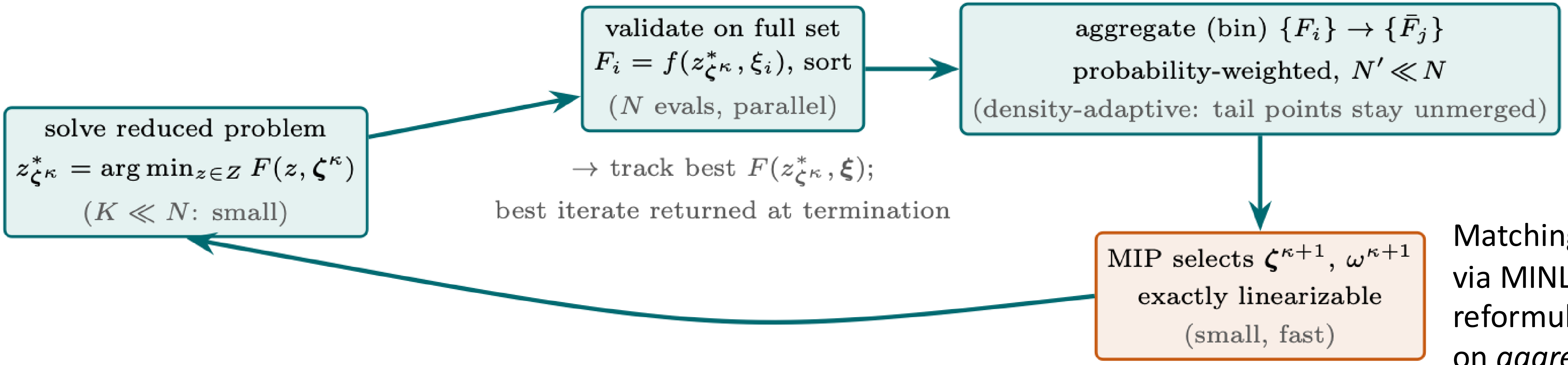
1. Take initial reduced guess ζ^0 (e.g., K-means); solve reduced problem $z_\zeta{}^{0*}$
2. **Freeze** $z = z^*_{\boldsymbol{\zeta}^\kappa}$ and validate on full set: $F_i := f(z^*_{\boldsymbol{\zeta}^\kappa}, \xi_i)$ ➔ N parallel evals.
   - Note that NxN matrix f_{ij} collapsed to N-vector for fixed z!
3. **Order** the validated costs: $f(z^*_{\boldsymbol{\zeta}^\kappa}, \xi_{(1)}) \leq \cdots \leq f(z^*_{\boldsymbol{\zeta}^\kappa}, \xi_{(N)})$ ➔ VaR pos known.
4. **Selection step:** $\boldsymbol{\zeta}^{\kappa+1} = \arg\min_{\boldsymbol{\zeta} \subset \boldsymbol{\xi}} \left| F(z^*_{\boldsymbol{\zeta}^\kappa}, \boldsymbol{\zeta}) - F(z^*_{\boldsymbol{\zeta}^\kappa}, \boldsymbol{\xi}) \right|$

Match the validated cost distribution, including its tail with a *re-weighted* K-point empirical distribution?

University of Vermont

Iterative approach that is I-PDSR





Matching is achieved via MINLP ➔ MILP reformulation solved on *aggregated* F

No global guarantee, no convergence guarantee, no monotonicity guarantee.

But it kinda works well in 10-20 iterations!

Results: Risk-averse VPP day-ahead offer (N=400; K=20; α = 0.95)

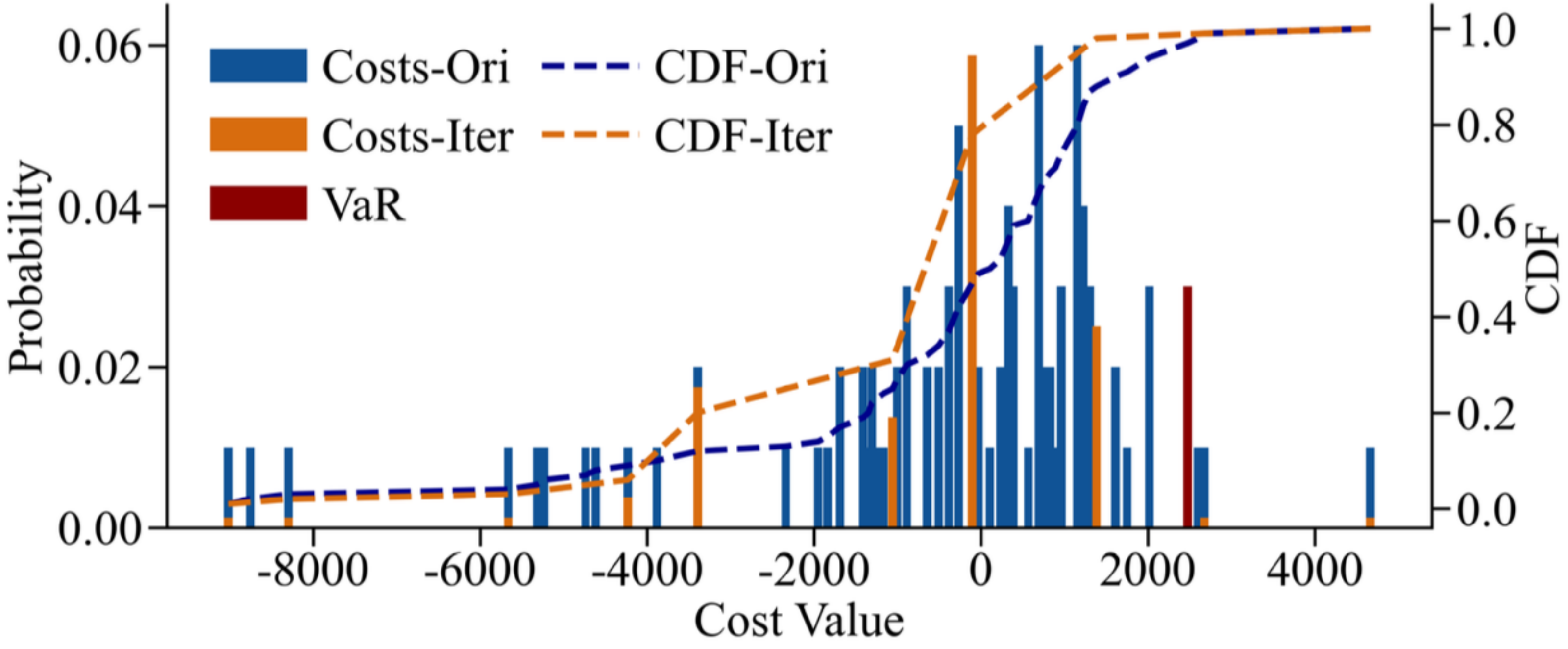
method	OG (%)
IPDSR (proposed)	**0.73**
best DDSR baselines	1.4 – 2.7

Performant

Solve times: ~35s per iteration, 10-20 iterations expected
- Objective aggregation compressed F to 100 values
- 5-10 minutes total time.

Tractable

Tail captured!



University of Vermont

Summarizing

SBSO is a large stacked mathematical program. Proposed two methods for SR in problem space: $\lambda = 0$ (risk-neutral, TPWRS '25) vs. $\lambda > 0$ (risk-averse CVaR, PSCC paper).

- At $\lambda = 0$, the OG bound is built from a static matrix of independent deterministic solves, so SR becomes a parallel pre-computation plus one (large, sparse) MILP.
- At $\lambda > 0$ CVaR breaks that separability in two ways because it **demands order!**
  - IPDSR restores tractability by freezing the decision, sorting validated costs, and iterating: a sequence of small linear MIPs that preserves the tail.

Overall results: <1% optimality gap with K = 5-20 clusters from N=400-1000 scenarios, in a tractable manner (order of seconds-minutes). **It kinda works!**

Thank you! Questions? Comments?

malmassa@uvm.edu   @theEnergyMads   https://madsalma.github.io

TPWRS 2025   **PSCC 2026**

Dr. Yingrui Zhuang
zyr21@mails.tsinghua.edu.cn

Dr. Ning Qi
nq2176@columbia.edu

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

Y. Zhuang, L. Cheng, N. Qi, M. R. Almassalkhi, and F. Liu, ``An Iterative Problem-Driven Scenario Reduction Framework for Stochastic Optimization with Conditional Value-at-Risk,'' accepted by *Electric Power Systems Research*, 2025