

# KDE Robust Satisficing for Optimal Load Shedding Under Renewable Uncertainty

Shu M X, Liu W, Chen F Y, Chen H X, Zhang Z Z, and Yu B

**Abstract**—Renewable-driven direct-current optimal load shedding (DC-OLS) requires a model that is interpretable to operators, data driven under continuous forecast errors, sensitive to severe security failures, and computationally tractable. This paper develops a budgeted KDE- $\phi$ -HMCR-RS-OLS framework for that purpose. Robust satisficing (RS) replaces ambiguity-radius tuning with an admissible shedding budget. A one-dimensional KDE reference family with  $\phi$ -divergence reweighting gives a smooth finite-sample representation of renewable uncertainty and a tractable blockwise reformulation. Higher-moment coherent risk (HMCR) distinguishes the severity of reserve exhaustion and line overload once violations occur. The paper derives a continuous-distribution dual template, the main finite-sample reformulation, and Kullback–Leibler and Pearson  $\chi^2$  specializations, and develops a stabilized bundle algorithm for benchmark-scale instances. Results on RBTS and RTS79 show that the proposed method improves security relative to deterministic DC-OLS while avoiding the full preventive cost of heavier robust baselines.

**Index Terms**—Optimal load shedding, robust satisficing, kernel density estimation, higher-moment coherent risk, renewable uncertainty.

## I. INTRODUCTION

**O**PTIMAL load shedding (OLS) remains a last-resort action, but in renewable-rich systems it can no longer be treated as a simple emergency switch. When forecast errors, reserve shortages, and transmission stress interact, the question is not only how to shed less load, but how to avoid an under-protected schedule without moving to an overly conservative one. Modern OLS therefore sits squarely at the economy–security tradeoff, and recent work has accordingly moved from fixed rules to optimization models that account explicitly for uncertainty, risk, and network limits [1]–[5], [27].

Existing uncertainty-aware models clarify this tradeoff, but they still leave a gap for renewable-driven DC-OLS. Chance-constrained models control violation probabilities, yet their operating meaning can become fragile when stressed renewable conditions make the underlying probability model or sample surrogate less trustworthy [6]. Classical robust models provide hard protection, but they may become excessively conservative when the uncertainty set is only loosely calibrated from data [7]. DRO offers a useful middle ground by hedging against a family of probability laws around a reference distribution, and recent studies have shown its value in dispatch and OPF-type settings [8], [9], [11]–[14]. Even so, renewable-driven DC-OLS still lacks a framework that makes the robustness parameter interpretable to operators, models renewable forecast errors by more than empirical point masses, distinguishes shallow and deep security failures, and remains solver-ready at benchmark scale [29].

This paper addresses that gap through a budgeted KDE- $\phi$ -HMCR-RS-OLS framework. The design is problem driven. First, ambiguity-radius tuning is statistically natural but operationally opaque, so robustness is organized through robust satisficing: following Long *et al.* [15], the model starts from an acceptable target and measures fragility around it. In OLS, that target becomes an admissible shedding budget. Second, renewable forecast errors are continuous and irregular, whereas empirical point masses can be too rough and repeated multivariate density estimation is too expensive. A one-dimensional KDE reference family therefore smooths the scalar violation severity entering each security block, and  $\phi$ -divergence turns distributional misspecification into a finite-dimensional reweighting problem with a tractable dual form [17], [21], [22]. Third, policies with similar violation probabilities can still have very different positive violation magnitudes once reserve exhaustion or line overload occurs; HMCR is used because it keeps the coherent convex structure of CVaR at  $p = 1$  while penalizing deeper tail excursions more strongly for  $p > 1$  [17], [19], [20].

These choices also clarify the paper’s relation to prior work. Relative to Long *et al.* [15], we translate target-based robust satisficing into an operator-facing budgeted formulation for network-constrained OLS. Relative to Liu *et al.* [16], we retain a reference-centered RS interpretation rather than introducing a separate inner ambiguity set. Relative to the KDE–HMCR portfolio model of Liu *et al.* [17], we transfer the KDE- $\phi$ -HMCR reformulation logic to renewable-driven DC-OLS and reorganize it around an admissible shedding budget and blockwise network-security constraints. The innovation is therefore the integration of operator-facing budgeting, continuous-reference uncertainty modeling, finite-dimensional perturbation geometry, and deep-tail risk discrimination in one executable OLS framework.

The paper has two linked aims. The first is to develop this budgeted KDE- $\phi$ -HMCR-RS-OLS model together with tractable reformulations. The second is to show that the resulting workflow remains numerically reliable through matched direct-solve checks and larger RBTS/RTS79 studies. The direct formulations serve as fidelity checks on matched stressed cases, whereas the stabilized bundle method is the benchmark-scale solver. Taken together, the theoretical and numerical results support one practical message: the proposed framework improves security relative to deterministic DC-OLS while avoiding the full preventive cost of heavier robust baselines.

The contributions of the paper are fourfold.

- **Formulation:** We formulate renewable-driven OLS as a budgeted KDE- $\phi$ -HMCR-RS-OLS model in which RS gives robustness an admissible-budget interpretation, HMCR measures deep-tail security severity, and KDE with  $\phi$ -divergence provides a smooth finite-sample uncertainty representation with tractable weight perturbations.
- **Reformulation:** We derive a continuous-distribution bridge result, the main finite-sample KDE reformulation, and solver-ready Kullback–Leibler and Pearson  $\chi^2$  specializations.
- **Computation:** We develop a stabilized reserve-aware bundle algorithm for benchmark-scale instances and show that it remains close to direct solves on matched stressed cases.
- **Evidence:** RBTS and RTS79 studies show that the integrated design improves security relative to deterministic DC-OLS while avoiding the full conservativeness of heavier robust baselines.

The remainder of the paper is organized as follows. Section II introduces the KDE reference family, the  $\phi$ -divergence perturbation geometry, the HMCR representation, and the RS budget interpretation. Section III develops tractable reformulations. Section IV embeds the risk model into the DC-OLS problem. Section V presents the computation strategy. Section VI reports the RBTS and RTS79 results. Section VII concludes the paper.

## II. KDE- $\phi$ -HMCR-RS INGREDIENTS

This section introduces the ingredients in the order used later: KDE,  $\phi$ -divergence, HMCR, and RS. Their roles are distinct: KDE with  $\phi$ -divergence makes continuous finite-sample uncertainty tractable, HMCR captures severity rather than probability alone, and RS gives robustness an operator-facing budget interpretation. Throughout,  $x$  denotes the deterministic decision vector,  $\tilde{\xi}$  the renewable forecast-error vector, and  $\{\xi^i\}_{i=1}^T$  its historical samples. We use  $q(x, \xi)$  for a generic scalar violation mapping, reserving  $q_j^U$ ,  $q_j^D$ ,  $q_\ell^U$ , and  $q_\ell^D$  for the reserve and line-security blocks in DC-OLS. The symbols  $\lambda$  and  $\nu$  denote KDE weights and scalar dual multipliers, while  $w^f$  and  $\bar{F}_\ell$  denote renewable point forecasts and line-flow limits. We use  $\hat{\mathbb{P}}$  for a reference law on the full uncertainty space and  $\hat{p}_\lambda$  for the one-dimensional KDE density induced by a scalar violation block.

*Remark 1.* Two divergence layers are used throughout the paper. Theorems and appendices are stated for a generic generator  $\phi$ , whereas computation is restricted to Kullback–Leibler and Pearson  $\chi^2$  choices. This keeps the theory reusable and the implementation explicit, as is common in the DRO literature [10], [23].

### A. KDE reference family and $\phi$ -divergence perturbations

We begin with a blockwise KDE reference family and its associated  $\phi$ -divergence reweighting geometry because this module resolves the uncertainty-representation bottleneck. Let  $\{\xi^i\}_{i=1}^T$  denote historical renewable forecast-error samples and let  $k(\cdot)$  be a kernel with bandwidth  $h > 0$ . In the present RS construction, KDE is used to build a

smooth data-based reference family rather than an explicit ambiguity set. A purely empirical law is often too rough and too brittle under reweighting, whereas blockwise one-dimensional KDE preserves continuous error structure, keeps the smoothed positive-part mapping  $\Psi_p$  convex, and yields the finite-dimensional reweighting model used later in Theorem 1. Once the reference law is written as a weighted KDE mixture, uncertainty enters through the mixture weights, so robustification becomes a reweighting step and  $\phi$ -divergence provides the perturbation geometry whose Fenchel conjugate makes the later reformulation tractable [12], [21], [22], [29].

For a generic scalar violation function  $q(x, \xi)$ , the weighted KDE family is defined by

$$\hat{p}_\lambda(z) = \sum_{i=1}^T \lambda_i h^{-1} k\left(\frac{z - q(x, \xi^i)}{h}\right), \quad (1)$$

where  $\lambda \in \mathbb{R}_+^T$ ,  $\mathbf{1}^\top \lambda = 1$ . Distributional perturbations are induced by reweighting the mixture components around the uniform reference weight  $\lambda^0 = T^{-1}\mathbf{1}$  through the penalty  $I_\phi(\lambda, \lambda^0) = (1/T) \sum_{i=1}^T \phi(T\lambda_i)$ , where  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}$  is closed, convex, and satisfies  $\phi(1) = 0$ . Thus the protection scope is the KDE-induced perturbation family generated by admissible sample weights, not the set of all probability laws on the support.

The following assumptions are standard.

**Assumption 1.** *The kernel  $k(\cdot)$  is nonnegative, bounded, measurable, and satisfies  $\int_{-\infty}^{\infty} k(t) dt = 1$ .*

**Assumption 2.** *The kernel is symmetric, i.e.,  $k(-t) = k(t)$ , and has finite  $p$ th moment for the chosen risk order  $p \geq 1$ .*

**Assumption 3.** *As  $T \rightarrow \infty$ , the bandwidth satisfies  $h(T) \rightarrow 0$  and  $Th(T) \rightarrow \infty$ .*

To simplify notation, define  $G_p(x) := \int_{-\infty}^x (x-t)^p k(t) dt$ ,  $\psi_p(x) := G_p(x)^{1/p}$ , and the perspective-smoothed positive-part operator  $\Psi_p(x, h) := h \psi_p(x/h)$ . Then the KDE-based HMCR of  $q(x, \xi)$  under weight vector  $\lambda$  becomes

$$\text{HMCR}_\gamma^{p, \lambda}(q) = \inf_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{1-\gamma} \left( \sum_{i=1}^T \lambda_i \cdot (\Psi_p(q(x, \xi^i) - \alpha, h))^p \right)^{1/p} \right\}. \quad (2)$$

**Proposition 1.** *Under Assumptions 1–2, the function  $\Psi_p(x, h)$  satisfies the following properties for  $p \geq 1$  and  $h > 0$ :*

- 1)  $\Psi_p(x, h)$  is nondecreasing in  $x$ ;
- 2)  $\Psi_p(x, h)$  is jointly convex in  $(x, h)$ ;
- 3)  $\Psi_p(x, h) \rightarrow [x]_+$  as  $h \downarrow 0$  for each fixed  $x$ , and  $\Psi_p(0, h) = \mathcal{O}(h)$ .

Proposition 1 matters because joint convexity preserves tractability when the argument is affine in the decision variables [26].

*Remark 2.* The one-dimensional KDE construction is deliberate: the reformulation only needs the scalar violation severity entering each risk block, not a full multivariate density estimate for every network state.

Therefore, this module supplies the continuous-reference, finite-sample uncertainty layer for the later blockwise reformulation.

### B. Higher-moment coherent risk

We next introduce HMCR as the tail-severity layer because the remaining bottleneck is not whether violations occur, but how severe they become once they occur. For a random loss  $Y$  with confidence level  $\gamma \in [0, 1)$  and moment order  $p \geq 1$ , the higher-moment coherent risk is defined as [19]

$$\text{HMCR}_\gamma^p(Y) := \inf_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{1-\gamma} \|[Y - \alpha]_+\|_p \right\}, \quad (3)$$

where  $[u]_+ = \max\{u, 0\}$  and  $\|\cdot\|_p$  is the  $L_p$  norm. When  $p = 1$ , (3) reduces to the conditional value-at-risk representation [20]. For  $p > 1$ , the measure becomes more sensitive to tail severity and therefore provides a more discriminative representation of extreme reserve or line-flow violations.

Using the standard perspective identity  $\zeta^{1/p} = \inf_{y>0} \{\zeta/(py^{p-1}) + (p-1)y/p\}$  for  $\zeta \geq 0$ , (3) can be rewritten as

$$\text{HMCR}_\gamma^p(Y) = \inf_{\alpha \in \mathbb{R}, y>0} \left\{ \alpha + \frac{p-1}{p(1-\gamma)}y + \mathbb{E} \left[ \frac{[Y - \alpha]_+^p}{p(1-\gamma)y^{p-1}} \right] \right\}. \quad (4)$$

The representation in (4) will be central in the tractable reformulation.

The choice of HMCR is deliberate. CVaR-type surrogates preserve convexity, but stressed reserve exhaustion or line overload can have similar tail probabilities while differing sharply in post-threshold magnitude. HMCR keeps the same convex structure, nests CVaR at  $p = 1$ , and penalizes deeper tail excursions more strongly for  $p > 1$ . Therefore, this module turns the security model from a probability-only rule into a severity-sensitive tail-risk layer.

### C. Robust satisficing under distributional uncertainty

We then place the resulting risk model inside a robust-satisficing budget interpretation because the last bottleneck is parameter interpretability. Let  $x \in \mathcal{X}$  denote the decision vector and let  $Z(x, \tilde{\xi})$  be a random loss induced by renewable forecast error  $\tilde{\xi}$ . In the RS paradigm, the decision maker specifies a satisfactory target  $\tau$  and seeks a decision with the smallest fragility against distributional misspecification. A generic RS constraint takes the form

$$\mathcal{R}_{\mathbb{P}}(Z(x, \tilde{\xi})) \leq \tau + \kappa \Delta(\mathbb{P}, \hat{\mathbb{P}}), \quad \forall \mathbb{P} \in \mathcal{P}_0, \quad (5)$$

where  $\mathcal{R}_{\mathbb{P}}(\cdot)$  is a generic risk functional under distribution  $\mathbb{P}$ ,  $\hat{\mathbb{P}}$  is the reference distribution estimated from data,  $\Delta(\cdot, \cdot)$  is a divergence or distance, and  $\kappa \geq 0$  is the fragility index. The operational meaning of (5) is immediate: the worst admissible growth of the risk relative to the target is penalized proportionally to distributional perturbation. In the remainder of the paper, the symbol  $\rho$  is reserved exclusively for the satisficing budget factor in (6); it is not used to denote a generic risk measure.

Equation (5) is the reference-distribution-centered RS form of Long *et al.* [15]. Unlike the GDRC formulation of Liu *et al.* [16], no inner ambiguity set is imposed here; the reference object is a data-estimated distribution, later specialized to a KDE family. In optimal load shedding, the target is naturally tied to an economically acceptable load-shedding cost. Let  $Z_0$  denote the optimal cost of a deterministic baseline DC-OLS model. We define

$$\tau_{\text{cost}} = (1 + \rho)Z_0, \quad (6)$$

where  $\rho \geq 0$  is a user-specified satisficing margin. This construction avoids the calibration difficulty of a purely statistical ambiguity radius and makes the robustness parameter interpretable from an operational budget perspective, complementing conventional probability-radius and ambiguity-radius calibrations used in robust and distributionally robust power-system models [8], [9], [11], [14], [24].

Canonical RS is often written for a single risk-aware loss. Here  $\tau_{\text{cost}}$  bounds planned shedding cost, whereas reserve and line-security constraints receive separate fragility indices with zero nominal targets. Therefore, this module supplies the operator-facing budget interpretation that organizes the full model.

### D. Regularity conditions for dual reformulation

The tractable reformulations below rely on mild regularity conditions that make the minimax interchange and the divergence duality fully explicit; such assumptions are standard in convex duality and modern DRO reformulations [10], [23].

**Assumption 4.** *The divergence generator  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{+\infty\}$  is proper, closed, convex, satisfies  $\phi(1) = 0$ , and is superlinear in the sense that  $\lim_{t \rightarrow +\infty} \phi(t)/t = +\infty$ .*

**Assumption 5.** *For every fixed decision vector  $x \in \mathcal{X}$ , the mapping  $q(x, \cdot)$  is Borel measurable and the function  $\ell_{\alpha, y}(\cdot)$  defined in (9) is  $\hat{\mathbb{P}}$ -integrable for every  $(\alpha, y) \in \mathbb{R} \times \mathbb{R}_{++}$ . Moreover, the map*

$$(\alpha, y) \mapsto \alpha + \frac{p-1}{p(1-\gamma)}y + \int \ell_{\alpha, y}(\xi) \hat{p}(\xi) d\xi$$

*is proper, lower semicontinuous, and coercive on  $\mathbb{R} \times \mathbb{R}_{++}$ .*

**Lemma 1.** *Under Assumptions 4 and 5, for every fixed  $x$  and every finite level  $\tau$ , the outer minimization variables  $(\alpha, y)$  in the continuous and finite-sample  $\phi$ -HMCR-RS reformulations can be restricted, without loss of optimality, to a compact convex subset of  $\mathbb{R} \times \mathbb{R}_{++}$ . Consequently, Sion's minimax theorem can be applied on the truncated domain and the original value is recovered by letting the truncation radius tend to infinity.*

## III. KDE- $\phi$ -HMCR-RS REFORMULATIONS

This section develops the reformulations used later, moving from a continuous  $\phi$ -HMCR-RS dual template to the finite-sample KDE model.

### A. A continuous $\phi$ -HMCR-RS dual template

We first record the  $\phi$ -HMCR-RS dual structure under a generic continuous reference law with density  $\widehat{p}(\xi)$  on  $\mathbb{R}^m$ . This object lives on the full uncertainty space and should be distinguished from the one-dimensional blockwise density  $\widehat{p}_\lambda$  introduced in Section II.

**Proposition 2.** *Let  $p > 1$ ,  $\gamma \in [0, 1)$ , and  $\kappa > 0$ . Under Assumptions 4 and 5, the  $\phi$ -HMCR-RS constraint*

$$\sup_{\mathbb{P} \in \mathcal{P}(\mathbb{R}^m)} \left\{ \text{HMCR}_{\gamma, \mathbb{P}}^p(q(x, \xi)) - \kappa D_\phi(\mathbb{P} \| \widehat{\mathbb{P}}) \right\} \leq \tau \quad (7)$$

is equivalent to

$$\inf_{\alpha \in \mathbb{R}, y > 0, \nu \in \mathbb{R}} \left\{ \alpha + \frac{p-1}{p(1-\gamma)} y + \nu + \kappa \int \phi^* \left( \frac{\ell_{\alpha, y}(\xi) - \nu}{\kappa} \right) d\widehat{\mathbb{P}}(\xi) \right\} \leq \tau \quad (8)$$

where

$$\ell_{\alpha, y}(\xi) := \frac{[q(x, \xi) - \alpha]_+^p}{p(1-\gamma)y^{p-1}}. \quad (9)$$

*Proof.* Appendix B proves the result in three explicit steps: truncation of the outer minimization via Lemma 1, minimax interchange on the truncated domain, and Fenchel–Lagrange dualization of the density-ratio maximization.  $\square$

*Remark 3.* Proposition 2 is a distribution-level bridge: it shows how the continuous  $\phi$ -HMCR-RS constraint collapses to a finite-dimensional dual form once the divergence penalty is conjugated. The finite-sample KDE model used in the executable DC-OLS formulation is developed next.

### B. A finite-sample KDE- $\phi$ -HMCR-RS reformulation

Proposition 2 gives a distribution-level  $\phi$ -HMCR-RS dual template. Replacing the generic reference law with the weighted mixture (1) yields the finite-data, blockwise model used in the rest of the paper.

**Theorem 1.** *Let  $p > 1$  and  $\kappa \geq 0$ . Under Assumptions 4 and 5, the finite-sample KDE- $\phi$ -HMCR-RS constraint*

$$\sup_{\lambda \in \mathbb{R}_+^T, \mathbf{1}^\top \lambda = 1} \left\{ \text{HMCR}_{\gamma}^{p, \lambda}(q) - \frac{\kappa}{T} \sum_{i=1}^T \phi(T\lambda_i) \right\} \leq \tau \quad (10)$$

is equivalent to the existence of variables  $\alpha \in \mathbb{R}$ ,  $\eta \in \mathbb{R}$ ,  $y > 0$ , and  $v_i \in \mathbb{R}$  such that

$$\alpha + \eta + \frac{p-1}{p(1-\gamma)} y + \frac{1}{T} \sum_{i=1}^T \kappa \phi^* \left( \frac{v_i - \eta}{\kappa} \right) \leq \tau, \quad (11a)$$

$$\frac{1}{p(1-\gamma)} y^{1-p} (\Psi_p(q(x, \xi^i) - \alpha, h))^p \leq v_i, \quad i = 1, \dots, T. \quad (11b)$$

When  $\kappa = 0$ , the convention  $0\phi^*(s/0) = 0$  for  $s \leq 0$  and  $+\infty$  otherwise recovers the nominal KDE-HMCR constraint.

*Proof.* Appendix C shows that the result follows from the same minimax argument as Proposition 2, followed by a finite-dimensional simplex dualization and an epigraph lift that exploits the monotonicity of  $\phi^*$ .  $\square$

*Remark 4.* Theorem 1 is the main tractability result. It replaces worst-case risk evaluation over perturbed KDE weights by a convex model whose extra complexity grows linearly with the sample size through  $\{v_i\}_{i=1}^T$ . All later solver-ready and case-study models follow by instantiating  $\phi^*$  in (11a).

### C. A Kullback–Leibler specialization of the KDE- $\phi$ -HMCR-RS reformulation

We first specialize the KDE- $\phi$ -HMCR-RS reformulation to the Kullback–Leibler case because it is the main direct-solve benchmark and yields the cleanest exponential-cone representation. For  $\phi(t) = t \log t - t + 1$ , the Fenchel conjugate is  $\phi^*(s) = e^s - 1$ , so (11a) reduces to the corresponding exponential-budget form.

**Corollary 1.** *Under  $\phi(t) = t \log t - t + 1$ , the KDE- $\phi$ -HMCR-RS constraint can be represented by (11b) together with auxiliary variables  $\zeta_i \geq 0$  satisfying*

$$\alpha + \eta + \frac{p-1}{p(1-\gamma)} y + \frac{1}{T} \sum_{i=1}^T \zeta_i \leq \tau, \quad (12a)$$

$$(v_i - \eta, \kappa, \kappa + \zeta_i) \in \mathcal{K}_{\text{exp}}, \quad i = 1, \dots, T, \quad (12b)$$

where  $\mathcal{K}_{\text{exp}} := \{(u, v, w) : v > 0, v \exp(u/v) \leq w\} \cup \{(u, 0, w) : u \leq 0, w \geq 0\}$  is the exponential cone.

Corollary 1 shows that the Kullback–Leibler case can be implemented directly in conic solvers such as MOSEK through standard exponential-cone constraints [26]. This is why KL is the main direct-solve formulation used in the paper.

### D. A quadratic form under the Pearson $\chi^2$ divergence

We next specialize the same KDE- $\phi$ -HMCR-RS reformulation to the Pearson  $\chi^2$  divergence, for which  $\phi(t) = (t-1)^2$  on  $t \geq 0$  and  $\phi^*(s) = -1$  for  $s < -2$  and  $\phi^*(s) = s + s^2/4$  for  $s \geq -2$ . Substituting this conjugate into (11a) yields a piecewise-quadratic robust-satisficing constraint.

**Corollary 2.** *Under the Pearson  $\chi^2$  divergence, the KDE- $\phi$ -HMCR-RS constraint is equivalent to (11b) together with auxiliary variables  $\zeta_i \in \mathbb{R}$  satisfying*

$$\alpha + \eta + \frac{p-1}{p(1-\gamma)} y + \frac{1}{T} \sum_{i=1}^T \zeta_i \leq \tau, \quad (13a)$$

$$\zeta_i \geq \kappa \phi^* \left( \frac{v_i - \eta}{\kappa} \right), \quad i = 1, \dots, T. \quad (13b)$$

On the active branch  $v_i - \eta \geq -2\kappa$ , this becomes the convex quadratic epigraph  $\zeta_i \geq (v_i - \eta) + (v_i - \eta)^2/(4\kappa)$ , so the Pearson  $\chi^2$  case leads to a QCQP/SOCP-friendly implementation [26].

*Remark 5.* The roles of the two instantiated divergences are different. The Kullback–Leibler divergence is the main direct-solve formulation used in the numerical checks because of its smooth entropy-type penalty and mature exponential-cone support. The Pearson  $\chi^2$  divergence is retained as a second direct-solve comparison, exposing how the same generic reformulation behaves under a lighter-tailed, second-order ambiguity penalty.

#### IV. APPLICATION TO DC OPTIMAL LOAD SHEDDING

This section embeds the reformulation into the blockwise DC-OLS model used later. We specify the deterministic baseline, the reserve and line-security violation blocks, and the resulting tractable template.

##### A. Deterministic baseline model

Consider a power system with bus set  $\mathcal{B}$ , generator set  $\mathcal{G}$ , transmission line set  $\mathcal{L}$ , and renewable unit set  $\mathcal{W}$ . Let  $g \in \mathbb{R}^{|\mathcal{G}|}$  denote conventional generation,  $r^U$  and  $r^D$  denote up/down reserve capacities,  $s \in \mathbb{R}^{|\mathcal{B}|}$  denote curtailed load, and  $\beta \in \mathbb{R}^{|\mathcal{G}|}$  denote the affine participation factors that balance the aggregate renewable forecast error. Let  $w^f \in \mathbb{R}^{|\mathcal{W}|}$  denote the renewable point forecast. The deterministic benchmark DC-OLS problem is

$$\min_{g, r^U, r^D, s, \beta} c_L^\top s \quad (14a)$$

$$\text{s.t. } g + r^U \leq g^{\max}, \quad (14b)$$

$$g - r^D \geq g^{\min}, \quad (14c)$$

$$\mathbf{1}^\top g + \mathbf{1}^\top w^f = \mathbf{1}^\top (l - s), \quad (14d)$$

$$0 \leq s \leq l, \quad (14e)$$

$$\mathbf{1}^\top \beta = 1, \quad \beta \geq 0, \quad (14f)$$

where  $c_L$  is the nodal interruption-cost vector and  $l$  is the demand vector. Model (14) is a compact emergency-dispatch surrogate of standard DC network-constrained scheduling models [6], [7], [30]. Let its optimal value be  $Z_0$ .

##### B. Random security functions

Let  $\tilde{\xi} \in \mathbb{R}^{|\mathcal{W}|}$  be the renewable forecast-error vector. Under affine balancing, the realized generator output of unit  $j$  is  $g_j(\tilde{\xi}) = g_j - \beta_j \mathbf{1}^\top \tilde{\xi}$ . Define the deterministic decision block as  $x := (g, r^U, r^D, s, \beta)$ . The following scalar violation functions are exactly the standardized forms of the original chance constraints in the source manuscript.

$$q_j^U(x, \xi) = -\beta_j \mathbf{1}^\top \xi - r_j^U, \quad j \in \mathcal{G}, \quad (15a)$$

$$q_j^D(x, \xi) = \beta_j \mathbf{1}^\top \xi - r_j^D, \quad j \in \mathcal{G}, \quad (15b)$$

$$q_\ell^U(x, \xi) = (M_\ell^W - M_\ell^G \beta \mathbf{1}^\top) \xi + M_\ell^G g + M_\ell^W w^f - M_\ell^B (l - s) - \bar{F}_\ell, \quad \ell \in \mathcal{L}, \quad (15c)$$

$$q_\ell^D(x, \xi) = -(M_\ell^W - M_\ell^G \beta \mathbf{1}^\top) \xi - M_\ell^G g - M_\ell^W w^f + M_\ell^B (l - s) - \bar{F}_\ell, \quad \ell \in \mathcal{L}. \quad (15d)$$

where  $M^G$ ,  $M^W$ , and  $M^B$  are the shift-factor matrices and  $\bar{F}_\ell$  is the transmission limit of line  $\ell$ . Each violation block is affine in the decision variables for fixed  $\xi^i$ , matching standard convex DRO and chance-constrained dispatch constructions [8], [9], [11].

##### C. Cost-budgeted, constraint-level RS model for DC-OLS

We keep total load-shedding cost below  $\tau_{\text{cost}}$  and minimize security fragility. This yields a cost-budgeted, constraint-level extension of canonical RS: the budget handles the economic side, whereas reserve-up, reserve-down, and line-security blocks are protected separately. The scalar fragility parameter

$\kappa$  therefore becomes block specific, giving  $\kappa_{1j}, \kappa_{2j}, \kappa_{1\ell}, \kappa_{2\ell}$ , and the weights  $w_{1j}, w_{2j}, w_{1\ell}, w_{2\ell}$  control only the relative emphasis across blocks. Using Theorem 1 block by block, we obtain the following model.

$$\min_{\Xi} \sum_{j \in \mathcal{G}} (w_{1j} \kappa_{1j} + w_{2j} \kappa_{2j}) + \sum_{\ell \in \mathcal{L}} (w_{1\ell} \kappa_{1\ell} + w_{2\ell} \kappa_{2\ell}) \quad (16a)$$

$$\text{s.t. } c_L^\top s \leq \tau_{\text{cost}}, \quad (16b)$$

$$(14b) - (14f),$$

$$y_{1j}, y_{2j} \geq 0, \quad \kappa_{1j}, \kappa_{2j} \geq 0, \quad \forall j \in \mathcal{G}, \quad (16c)$$

$$y_{1\ell}, y_{2\ell} \geq 0, \quad \kappa_{1\ell}, \kappa_{2\ell} \geq 0, \quad \forall \ell \in \mathcal{L}, \quad (16d)$$

$$\alpha_{1j} + \eta_{1j} + \frac{p-1}{p(1-\gamma_{1j})} y_{1j} + \frac{1}{T} \sum_{i=1}^T \kappa_{1j} \phi^* \left( \frac{v_{1ji} - \eta_{1j}}{\kappa_{1j}} \right) \leq 0, \quad \forall j \in \mathcal{G}, \quad (16e)$$

$$\frac{1}{p(1-\gamma_{1j})} y_{1j}^{1-p} \times \left( \Psi_p(q_j^U(x, \xi^i) - \alpha_{1j}, h_j) \right)^p \leq v_{1ji}, \quad \forall i, j, \quad (16f)$$

$$\alpha_{2j} + \eta_{2j} + \frac{p-1}{p(1-\gamma_{2j})} y_{2j} + \frac{1}{T} \sum_{i=1}^T \kappa_{2j} \phi^* \left( \frac{v_{2ji} - \eta_{2j}}{\kappa_{2j}} \right) \leq 0, \quad \forall j \in \mathcal{G}, \quad (16g)$$

$$\frac{1}{p(1-\gamma_{2j})} y_{2j}^{1-p} \times \left( \Psi_p(q_j^D(x, \xi^i) - \alpha_{2j}, h_j) \right)^p \leq v_{2ji}, \quad \forall i, j, \quad (16h)$$

$$\alpha_{1\ell} + \eta_{1\ell} + \frac{p-1}{p(1-\gamma_{1\ell})} y_{1\ell} + \frac{1}{T} \sum_{i=1}^T \kappa_{1\ell} \phi^* \left( \frac{v_{1\ell i} - \eta_{1\ell}}{\kappa_{1\ell}} \right) \leq 0, \quad \forall \ell \in \mathcal{L}, \quad (16i)$$

$$\frac{1}{p(1-\gamma_{1\ell})} y_{1\ell}^{1-p} \times \left( \Psi_p(q_\ell^U(x, \xi^i) - \alpha_{1\ell}, h_\ell) \right)^p \leq v_{1\ell i}, \quad \forall i, \ell, \quad (16j)$$

$$\alpha_{2\ell} + \eta_{2\ell} + \frac{p-1}{p(1-\gamma_{2\ell})} y_{2\ell} + \frac{1}{T} \sum_{i=1}^T \kappa_{2\ell} \phi^* \left( \frac{v_{2\ell i} - \eta_{2\ell}}{\kappa_{2\ell}} \right) \leq 0, \quad \forall \ell \in \mathcal{L}, \quad (16k)$$

$$\frac{1}{p(1-\gamma_{2\ell})} y_{2\ell}^{1-p} \times \left( \Psi_p(q_\ell^D(x, \xi^i) - \alpha_{2\ell}, h_\ell) \right)^p \leq v_{2\ell i}, \quad \forall i, \ell. \quad (16l)$$

Here,  $\Xi$  collects all primal and auxiliary variables. The blockwise form makes RS operational for OLS: cost is

budgeted, whereas security is protected through separate fragility penalties on the physically relevant blocks.

Throughout the remainder of this section, we impose the following standing well-posedness conditions for the DC-OLS application: the deterministic feasible set induced by (14b)–(14f) is nonempty, all operating limits are finite, and the bandwidth parameters  $h_j$  and  $h_\ell$  used in the smoothed constraints are strictly positive.

**Proposition 3.** *Under these standing DC-OLS conditions, if each  $q_j^U$ ,  $q_j^D$ ,  $q_\ell^U$ , and  $q_\ell^D$  is affine in the decision variables for fixed  $\xi^i$ , then model (16) is a convex optimization problem. If, in addition,  $\phi(t) = t \log t - t + 1$ , then (16) admits an exponential-cone representation; if  $\phi(t) = (t - 1)^2$ , it admits a convex quadratic representation.*

*Proof.* The proof follows from Proposition 1, Theorem 1, and the affine structure of the violation functions. A concise derivation is included in Appendix D.  $\square$

*Remark 6.* These standing conditions simply exclude pathological cases in which the deterministic benchmark is infeasible or the smoothing parameters degenerate. Under them, (16b) controls admissible load-shedding cost, whereas the objective minimizes reserve and transmission fragility.

*Remark 7.* The results are proved for the DC model because the violation maps in (15) are affine, which keeps (16) convex after the HMCR and KDE transformations. The same budgeted-fragility logic can be extended to approximate-AC or other convexified models whenever the resulting surrogates remain affine or convex [9]; here we claim exact tractability only for DC-OLS.

## V. COMPUTATION AND REPRODUCIBILITY

This section explains how the executable model from Section IV is solved and checked. Direct formulations provide matched fidelity checks, whereas the stabilized bundle method delivers the benchmark-scale computations used later in Section VI.

### A. Direct formulations and proposed algorithm

Section III yields two computational routes. Corollary 1 and Corollary 2 give direct solver-ready formulations for smaller matched instances, whereas the larger RBTS and RTS79 suites are solved by a stabilized reserve-aware bundle algorithm. The direct models serve only as matched fidelity checks; the bundle method solves the same budget–fragility model at benchmark scale through a reserve-aware phase-I center, blockwise fragility oracles, and an acceptance rule for numerically reliable final policies.

### B. Stabilized reserve-aware bundle algorithm

The proposed algorithm is a tailored proximal-bundle scheme for a nonsmooth convex master objective assembled from blockwise fragility oracles [25]. It first solves the deterministic baseline (14) and forms  $\tau_{\text{cost}}$  through (6). Let  $\Xi_{\text{tr}} = \{\xi^i\}_{i=1}^T$  denote the retained training samples, and define  $Q^{\text{up}} := \max\{0, -\min_i \mathbf{1}^\top \xi^i\}$  and  $Q^{\text{dn}} := \max\{0, \max_i \mathbf{1}^\top \xi^i\}$ . A

**Algorithm 1** Stabilized reserve-aware bundle algorithm for the proposed model

- 1: Solve deterministic baseline and set  $\tau_{\text{cost}} = (1 + \rho)Z_0$ .
- 2: Build the training set, bandwidths, and aggregate error envelope  $(Q^{\text{up}}, Q^{\text{dn}})$ .
- 3: Solve the phase-I reserve-aware center problem to obtain  $x^0$ .
- 4: **for**  $k = 0, 1, \dots$  until stopping **do**
- 5: Evaluate all reserve and line fragility blocks at  $x^k$  by oracle calls.
- 6: If any oracle fails, reject the iterate; if any block is infinite, add the corresponding penalty cut.
- 7: Otherwise, collect bundle cuts and solve the stabilized master problem with (17).
- 8: Apply the serious/null-step test and update the proximal parameter.
- 9: **end for**
- 10: Accept the solution only if oracle failures are zero, all infinite-fragility flags vanish, and final  $\zeta$ -slacks are numerically negligible.

Fig. 1. Bundle algorithm used in the numerical experiments.

phase-I center then computes a reserve-aware feasible point by minimizing reserve slacks  $\sigma^{\text{up}}$  and  $\sigma^{\text{dn}}$  subject to the deterministic network constraints and  $\sum_{j \in \mathcal{G}} r_j^U + \sigma^{\text{up}} \geq Q^{\text{up}}$ ,  $\sum_{j \in \mathcal{G}} r_j^D + \sigma^{\text{dn}} \geq Q^{\text{dn}}$ . This prevents the method from starting from a cost-feasible but reserve-degenerate center.

Given a center  $x^k$ , the algorithm evaluates the fragility objective by solving blockwise KL oracles for  $q_j^U$ ,  $q_j^D$ ,  $q_\ell^U$ , and  $q_\ell^D$ . Each oracle returns a fragility value and subgradient when the block is finite, and an *infinite-fragility* flag otherwise. The master problem then updates the iterate through a standard proximal-bundle step [25], minimizing  $t + \omega_U \mathbf{1}^\top \zeta^U + \omega_D \mathbf{1}^\top \zeta^D$  subject to the deterministic DC-OLS constraints, the cost budget, accumulated bundle cuts for  $t$ , and the soft reserve-allocation links

$$r_j^U + \zeta_j^U \geq Q^{\text{up}} \beta_j, \quad r_j^D + \zeta_j^D \geq Q^{\text{dn}} \beta_j, \quad \zeta_j^U, \zeta_j^D \geq 0. \quad (17)$$

These links regularize intermediate iterates by aligning total reserve and affine participation factors. In the reported finite-fragility RBTS solutions,  $\sum_j \zeta_j^U$  and  $\sum_j \zeta_j^D$  are on the order of  $10^{-6}$  MW at termination, so the accepted solution is numerically indistinguishable from one of the original model.

### C. Workflow and scaling remarks

Direct KL or Pearson solves instantiate model (16) as a conic or quadratic-conic program. The bundle algorithm keeps the same data split, budget logic, and smoothed risk construction, but replaces the monolithic solve by phase-I initialization, blockwise fragility oracles, and a stabilized master. The auxiliary-variable burden remains  $\mathcal{O}(T(2N_G + 2N_L))$ , with one additional oracle call per reserve or line block at each iteration. In practice, this makes the method useful when direct runs stagnate, collapse reserves, or return numerically infinite fragility blocks [25], [26].

Table I summarizes the numerical studies: matched direct solves check numerical fidelity on smaller stressed cases, the

TABLE I  
NUMERICAL STUDIES IN THE PAPER

Study	Scope	Observed numerical behavior	Purpose	Note
Direct KL check	3 matched stressed RBTS hours	3/3 solver-attainable with Opt./FP terminations	Check numerical fidelity on matched stressed hours	Serves as the main direct-solve reference
Direct Pearson check	Same 3 matched RBTS hours	3/3 solver-feasible but all Slow./FP	Provide a second direct-solve comparison	Numerically weaker than KL on these instances
Five-row matched check	5 screened RBTS rows, matched workflow	3/5 strict common-reportable; 5/5 bridge-acceptable; full matches 5/5 for Matched-KDE-RS and 4/5 for Matched- $h \downarrow$ 0-RS	Test whether the implementation preserves stressed-hour directions	A diagnostic check rather than a main benchmark table
RBTS benchmark suite	8 RBTS benchmark snapshots plus 8 fixed-hard rows	Accepted only after zero oracle failures, zero infinite-fragility blocks, and negligible final stabilization slacks	Show the main RBTS economy–security tradeoff and budget ordering	Main benchmark evidence
RTS79 validation suite	9 regime/layout cases with 5–11 screened snapshots and 3 seeds	Stable multi-layout second-system evidence, with altB remaining the most screening-sensitive layout	Check that the same pattern survives on a larger network	Not a universal reportability guarantee

“Opt./FP” and “Slow./FP” denote OPTIMAL/FEASIBLE\_POINT and SLOW\_PROGRESS/FEASIBLE\_POINT. Accepted policies satisfy the Section V screen: zero oracle failures, no infinite-fragility blocks, and negligible stabilization slacks.

five-row comparison checks direction preservation, and the RBTS/RTS79 suites provide the main evidence.

## VI. RBTS AND RTS79 RESULTS

This section tests whether the model delivers the intended economy–security compromise and whether the solution workflow remains numerically reliable. We report the RBTS operating-point comparison, the budget scan, the matched direct checks, and the RTS79 validation in that order.

### A. Setup and evaluation protocol

The benchmark uses one RBTS MATPOWER case with bus-level loads, system-level loads, and one wind profile [30]–[32]. After timestamp alignment and duplicate-key averaging, 8734 hourly observations remain. The main benchmark applies the common stress ladder 1.25/0.75/0.80/0.90, retains eight stress-screened snapshots with at least 72 h spacing, uses the preceding 720 h for training and the subsequent 336 h for testing, and caps retained training samples at 168. A final-stage solution is reportable only when its retained robust margin is at least 0.10.

We compare deterministic DC-OLS, Empirical-CC, fixed-radius DRO, KDE-HMCR, and Proposed-RS on the same snapshots using planned shedding cost, additional out-of-sample EENS, total EENS, line-event rate, and reportability. We also run a fixed-hard RBTS scan on rows {3227, 4070, 7433, 7601, 7770, 8106, 8274, 8346} for  $\rho \in \{1, 2, 3, 4, 5, 8\}$ .

For numerical fidelity, we run matched direct checks on stressed hours 8275, 8106, and 7771. Because direct monolithic solves with 168 retained samples often stall, these checks use the matched  $h \downarrow 0$  setting with the same 720 h/336 h split, 12 retained samples,  $\gamma = 0.95$ , and  $\rho = 0.5$ . We also run a five-row matched comparison on rows {8298, 8106, 7434, 7842, 7594} with the same split, 12 retained samples, and  $\rho = 0.5$ , using matched-KDE and matched  $h \downarrow 0$  comparators. Appendix E re-screens the fixed-hard RBTS records for  $\theta \in \{0.05, 0.10, 0.15, 0.20\}$  without re-solving the models.

We further run a single-snapshot  $\rho$ -sensitivity experiment on the same RBTS case using the KL-based bundle implementation with  $p = 2$ ,  $\gamma = 0.95$ , load scale 1.15, one forced

TABLE II  
RBTS EIGHT-SNAPSHOT BENCHMARK AT THE REFERENCE OPERATING POINT  $\rho = 3.0$

Method	Report. snaps.	Cost	Add. EENS (MWh)	Total EENS (MWh)	Line event rate
Deterministic	8/8	1,187.80	7.496	8.684	0.1789
Empirical-CC	8/8	29,527.07	0.242	29.769	0.0019
Fixed-DRO	8/8	44,411.38	0.001	44.412	0.0000
KDE-HMCR	8/8	34,384.78	0.090	34.475	0.0004
<b>Proposed-RS</b>	7/8	<b>4299.99</b>	<b>4.603</b>	<b>8.903</b>	<b>0.0867</b>

Proposed-RS averages use the reportable subset at  $\rho = 3.0$  because one high-stress snapshot remains below the retained-margin screen.

generator outage, and a 30 MW wind installation. The grid is  $\rho \in \{0.1, 0.3, 0.5, 1.0\}$ , and each case is evaluated on the full aligned history using LOLP, LOLE, EDNS, EENS, and line-overload probability [34].

### B. Average performance at the reference operating point

Table II reports the reference operating point at  $\rho = 3.0$ . Deterministic DC-OLS is cheapest but also has the largest additional EENS and line-event rate. Empirical-CC, fixed-radius DRO, and KDE-HMCR nearly eliminate deterioration at much higher planned cost. Proposed-RS remains the intended middle-ground policy: more protective than deterministic scheduling and markedly less preventive than the heavier robust baselines.

One borderline snapshot remains below the retained-margin screen and is therefore excluded from the Proposed-RS average. Appendix E shows that changing the cutoff affects screening severity, not the budget ordering.

### C. Fixed-hard RBTS budget scan and reportability threshold

Table III freezes the hard stress setting and scans  $\rho \in \{1, 2, 3, 4, 5, 8\}$ . Reportability improves from 6/8 at  $\rho = 1.0$  to 7/8 at  $\rho \geq 2$ ; row 8106 becomes reportable at  $\rho \geq 2$ , whereas row 7770 remains the boundary case. Over the same grid, average additional EENS drops from 4.939 to 4.162 MWh and the line-event rate from 0.0928 to 0.0838. The budget therefore pushes the policy toward the finite-fragility, higher-reliability regime, though one difficult row remains.

### D. Matched direct-solve checks on stressed RBTS snapshots

Table IV reports three matched stressed RBTS hours. Direct KL solves reach OPTIMAL/FEASIBLE\_POINT on all three

TABLE III  
FIXED-HARD RBTS BUDGET SCAN UNDER THE COMMON RETAINED-MARGIN SCREEN

$\rho$	Reportable snaps.	Avg. budget $\tau_{\text{cost}}$	Avg. planned cost	Avg. add. EENS (MWh)	Avg. line event rate
1.0	6/8	2,471.35	1,961.79	4.939	0.0928
2.0	7/8	3,937.52	3,361.36	4.702	0.0897
3.0	7/8	5,250.02	4,444.04	4.519	0.0867
4.0	7/8	6,562.53	5,688.82	4.416	0.0850
5.0	7/8	7,875.03	6,933.59	4.371	0.0850
8.0	7/8	11,812.55	9,141.35	4.162	0.0838

Common hard RBTS stress 1.25/0.75/0.80/0.90 is applied to rows {3227, 4070, 7433, 7601, 7770, 8106, 8274, 8346}. No tested budget reaches full reportability; row 8106 becomes reportable from  $\rho = 2$ , whereas row 7770 remains the boundary case.

TABLE IV  
MATCHED DIRECT-SOLVE CHECKS ON THREE STRESSED RBTS SNAPSHOTS

Hour	Model	Status	Cost	Base shed (MW)	Up res. (MW)	Down res. (MW)	Reserve event rate
8275	Proxy-LP	Opt./FP	8,283.28	8.283	0.005	18.123	0.0030
8275	Exact-KL	Opt./FP	12,270.10	12.270	0.511	18.990	0.0000
8275	Exact-Pearson	Slow./FP	12,413.20	12.413	4.056	61.442	0.0000
8106	Proxy-LP	Opt./FP	1,811.76	1.812	0.005	18.072	0.0149
8106	Exact-KL	Opt./FP	2,708.51	2.709	0.885	18.920	0.0000
8106	Exact-Pearson	Slow./FP	2,709.21	2.709	0.887	51.025	0.0000
7771	Proxy-LP	Slow./FP	1,349.43	1.349	0.170	17.853	0.0030
7771	Exact-KL	Opt./FP	1,766.83	1.767	0.576	18.308	0.0000
7771	Exact-Pearson	Slow./FP	1,767.62	1.768	0.583	182.793	0.0000

All runs use the same 720 h/336 h split, 12 retained samples,  $\gamma = 0.95$ ,  $\rho = 0.5$ , and the  $h \downarrow 0$  matched setting.  $Z_0 = \{8,278.34, 1,806.82, 1,178.71\}$  and  $\tau_{\text{cost}} = \{12,417.5, 2,710.22, 1,768.06\}$ . Direct KL remains main-table eligible on all three snapshots, the proxy LP on two, and direct Pearson on none.

TABLE V  
FIVE-ROW MATCHED COMPARISON ON SCREENED RBTS ROWS

Comparator	Strict common reportable	Bridge common acceptable	Avg. policy match score	Avg. protection match score	Full matches
Matched-KDE-RS	3/5	5/5	1.000	1.000	5/5
Matched- $h \downarrow 0$ -RS	3/5	5/5	0.800	0.933	4/5

All bridge runs use rows {7434, 7594, 7842, 8106, 8298}, the same 720 h/336 h split, 12 retained samples, and  $\rho = 0.5$ . “Strict common reportable” applies  $\theta = 0.10$  to both methods; “bridge common acceptable” keeps solver-acceptable, slack-free policies slightly below that cutoff. Policy scores average cost and base-shedding directions; protection scores average additional-EENS, reserve-event, and line-event directions.

snapshots; direct Pearson solves stay feasible but terminate with SLOW\_PROGRESS/FEASIBLE\_POINT. Once the workflow is aligned, all methods move away from the lightly protected proxy policy toward a more protective one under the same budget logic, with KL clearly the more reliable direct formulation.

#### E. Five-row matched comparison on screened RBTS rows

We next compare five screened RBTS rows {8298, 8106, 7434, 7842, 7594} under the same 720 h/336 h split, 12 retained samples, and  $\rho = 0.5$ , using matched-KDE and matched  $h \downarrow 0$  variants.

Table V shows that the implemented workflow preserves the stressed-hour directions that matter for policy comparison: the matched KDE comparison attains 5/5 full matches and the matched  $h \downarrow 0$  comparison 4/5.

#### F. Budget sensitivity and reliability indicators

Table VI shows how the satisficing factor  $\rho$  moves the solution from a low-budget, numerically fragile regime to a finite-fragility one. For  $\rho \in \{0.1, 0.3, 0.5\}$ , multiple upward-reserve blocks remain infinite and the stabilization slacks

$\sum_j \zeta_j^U$  stay positive. At  $\rho = 1.0$ , all infinite-fragility flags disappear and both  $\sum_j \zeta_j^U$  and  $\sum_j \zeta_j^D$  collapse to numerical noise. Over the same range, LOLP falls from 0.3488 to 0.0029, LOLE from 3047.4 to 25.4 h/yr, and EENS from 13315.25 to 0.55 MWh/yr, at the price of more preventive curtailment and reserve commitment.

The down-reserve envelope remains around 20.22 MW, so the very large down-reserve value at  $\rho = 0.3$  is only a symptom of the low-budget, infinite-fragility regime. Once  $\rho$  reaches the finite-fragility regime, the optimized reserve levels track that envelope much more closely.

#### G. Strengthened RTS79 multi-wind second-system validation

We finally test whether the RBTS message survives on RTS79 under more demanding wind siting patterns and screened windows. In addition to the base placement (3, 6, 10, 14, 19, 18), we test altA (2, 5, 8, 13, 18, 21), altB (4, 7, 11, 15, 20, 23), and altC (2, 6, 9, 14, 17, 21). Hard cases use 1.25/0.75/0.80/0.90 with  $\rho = 3.0$ ; mild cases use 1.20/0.85/0.90/0.95 or 1.15/0.85/0.90/0.95 with  $\rho = 1.0$ . Each case starts from the same 240-hour candidate pool, applies the same fragility and temporal-dispersion screens, and is re-

TABLE VI  
RBTS STRESSED-HOUR  $\rho$  SENSITIVITY AT A REPRESENTATIVE STRESSED HOUR

$\rho$	Base shed (MW)	Up res. (MW)	Down res. (MW)	$n_\infty^U$ (blocks)	$\sum \zeta^U$ (MW)	LOLP	LOLE (h/yr)	EDNS (MW)	EENS (MWh/yr)
0.1	7.0097	0.6372	35.6290	10	5.6969	0.3488	3047.4	1.5242	13315.25
0.3	8.2842	1.9117	101.6660	8	4.4224	0.3286	2870.7	1.1014	9621.97
0.5	9.5587	3.1862	35.8256	8	3.1479	0.3198	2794.0	0.7509	6559.61
1.0	12.7435	6.3558	22.4889	0	$\approx 0$	0.0029	25.4	$6.3 \times 10^{-5}$	0.55

TABLE VII  
RTS79 VALIDATION SUMMARY AT THE REFERENCE BUDGETS (THREE SEEDS)

Case	Method	$\rho$	$N$	Cost ( $\times 10^6$ )	Add. EENS	eLOLP	Rep.
hard_base, altA, altC	Det.	0.0	33	0.106	278.406	0.562	0.394
	DRO	1.0	33	3.731	47.393	0.232	0.970
	RS	3.0	33	3.715	51.076	0.252	0.939
mild20_base, altA, altC	Det.	0.0	18	0.002	222.373	0.478	0.389
	DRO	1.0	18	2.409	36.568	0.164	0.944
	RS	1.0	18	1.956	57.379	0.206	0.944
mild15_base	Det.	0.0	18	0.002	185.371	0.422	0.333
	DRO	1.0	18	1.762	30.880	0.155	0.944
	RS	1.0	18	1.529	41.442	0.168	0.944
hard_altB	Det.	0.0	18	0.456	269.300	0.563	0.000
	DRO	1.0	18	3.826	48.486	0.240	0.222
	RS	3.0	18	3.826	48.972	0.245	0.500
mild20_altB	Det.	0.0	9	0.002	193.640	0.455	0.000
	DRO	1.0	9	2.401	16.916	0.141	0.222
	RS	1.0	9	2.224	24.470	0.174	0.333

Det., DRO, and RS denote deterministic DC-OLS, fixed-DRO, and Proposed-RS. Cases with identical results across base, altA, and altC are grouped to save space. eLOLP counts only hours with positive emergency shedding beyond the preventive schedule.

evaluated under three random seeds with the same 720 h/336 h protocol.

Table VII shows the same qualitative pattern. Across the base, altA, and altC layouts, Proposed-RS improves security relative to deterministic scheduling while remaining less conservative than fixed-DRO in planned cost. The main exception is altB, the most screening-sensitive layout. Thus the larger-system message matches RBTS: the method remains a useful middle-ground policy on the main layouts, while altB marks the regime where siting stress still dominates.

## VII. CONCLUSION

This paper develops a budgeted KDE- $\phi$ -HMCR-RS-OLS framework for renewable-driven DC optimal load shedding. Its contribution is the integration of four roles in one executable model: RS gives robustness an admissible-budget interpretation, KDE provides a smooth finite-sample reference law,  $\phi$ -divergence yields a tractable perturbation geometry, and HMCR makes protection sensitive to deep security failures. The resulting model is interpretable, data driven, tail sensitive, and solver ready.

The theory follows the same structure: Proposition 2 gives the continuous-distribution dual template, Theorem 1 gives the main finite-sample reformulation, and Corollaries 1–2 produce solver-ready KL and Pearson  $\chi^2$  forms. On both RBTS and RTS79, the proposed framework improves security relative to deterministic DC-OLS while avoiding the full preventive cost of heavier robust baselines, and the stabilized bundle algorithm remains close to direct solves on matched stressed instances. Natural next steps are stronger direct formulations, richer AC-aware approximations, and broader operating datasets.

The appendices retain only the steps needed for the main equivalence and convexity claims.

## APPENDIX A

### AUXILIARY LEMMA FOR MINIMAX TRUNCATION

*Proof of Lemma 1.* Fix  $x \in \mathcal{X}$  and define

$$F(\alpha, y, \mathbb{P}) := \alpha + \frac{p-1}{p(1-\gamma)}y + \mathbb{E}_{\mathbb{P}}[\ell_{\alpha, y}(\tilde{\xi})] - \kappa D_{\phi}(\mathbb{P} \parallel \hat{\mathbb{P}}). \quad (18)$$

Assumption 5 makes  $(\alpha, y) \mapsto \alpha + \frac{p-1}{p(1-\gamma)}y + \mathbb{E}_{\mathbb{P}}[\ell_{\alpha, y}(\tilde{\xi})]$  proper, lower semicontinuous, and coercive on  $\mathbb{R} \times \mathbb{R}_{++}$ . Since  $D_{\phi}(\mathbb{P} \parallel \hat{\mathbb{P}}) \geq 0$ , the same lower bound holds uniformly in  $\mathbb{P}$ . Hence, for every finite level  $\tau$ , there exist constants  $M_{\tau} > 0$  and  $0 < \underline{y}_{\tau} < \bar{y}_{\tau} < +\infty$  such that

$$F(\alpha, y, \mathbb{P}) > \tau \quad \text{whenever } |\alpha| > M_{\tau} \text{ or } y \notin [\underline{y}_{\tau}, \bar{y}_{\tau}], \quad (19)$$

$$\forall \mathbb{P} \ll \hat{\mathbb{P}}.$$

Therefore, every minimizer of the outer problem belongs to the compact convex set

$$\mathcal{C}_{\tau}(x) := [-M_{\tau}, M_{\tau}] \times [\underline{y}_{\tau}, \bar{y}_{\tau}]. \quad (20)$$

Sion's theorem applies on  $\mathcal{C}_{\tau}(x)$  because the truncated minimization set is compact and convex, and the probability or simplex feasible set is convex. Letting  $M_{\tau} \uparrow +\infty$ ,  $\underline{y}_{\tau} \downarrow 0$ , and  $\bar{y}_{\tau} \uparrow +\infty$  recovers the original value.  $\square$

## APPENDIX B

### PROOF OF PROPOSITION 2

*Proof of Proposition 2.* Define

$$F(\alpha, y, \mathbb{P}) := \alpha + \frac{p-1}{p(1-\gamma)}y + \mathbb{E}_{\mathbb{P}}[\ell_{\alpha, y}(\tilde{\xi})] - \kappa D_{\phi}(\mathbb{P} \parallel \hat{\mathbb{P}}), \quad (21)$$

where  $\ell_{\alpha, y}$  is given by (9). By the perspective representation (4), constraint (7) is equivalent to

$$\sup_{\mathbb{P} \in \mathcal{P}(\mathbb{R}^m)} \inf_{\alpha \in \mathbb{R}, y > 0} F(\alpha, y, \mathbb{P}) \leq \tau. \quad (22)$$

*Step 1.* By Lemma 1, the outer minimization in (22) can be restricted to a compact convex set  $\mathcal{C}_{\tau}(x)$ . For fixed  $(\alpha, y)$ ,  $\mathbb{P} \mapsto F(\alpha, y, \mathbb{P})$  is concave; for fixed  $\mathbb{P}$ ,  $(\alpha, y) \mapsto F(\alpha, y, \mathbb{P})$  is convex and lower semicontinuous by Assumption 5. Sion's theorem then gives

$$\sup_{\mathbb{P}} \inf_{(\alpha, y) \in \mathcal{C}_{\tau}(x)} F(\alpha, y, \mathbb{P}) = \inf_{(\alpha, y) \in \mathcal{C}_{\tau}(x)} \sup_{\mathbb{P}} F(\alpha, y, \mathbb{P}), \quad (23)$$

and Lemma 1 removes the truncation afterward.

*Step 2.* Since  $D_\phi(\mathbb{P}||\hat{\mathbb{P}}) = +\infty$  when  $\mathbb{P} \not\ll \hat{\mathbb{P}}$ , every optimizer admits a likelihood ratio  $L = d\mathbb{P}/d\hat{\mathbb{P}}$  with  $L \geq 0$  and  $\int L d\hat{\mathbb{P}} = 1$ . Hence, for fixed  $(\alpha, y)$ ,

$$\begin{aligned} \sup_{\mathbb{P}} F(\alpha, y, \mathbb{P}) &= \alpha + \frac{p-1}{p(1-\gamma)}y \\ &+ \sup_{L \geq 0, \int L d\hat{\mathbb{P}}=1} \int \left( \ell_{\alpha, y}(\xi) L(\xi) \right. \\ &\quad \left. - \kappa \phi(L(\xi)) \right) d\hat{\mathbb{P}}(\xi). \end{aligned} \quad (24)$$

The feasible point  $L \equiv 1$  satisfies the normalization constraint, so Fenchel–Rockafellar duality applies. *Step 3.* Introduce the multiplier  $\nu \in \mathbb{R}$  for  $\int L d\hat{\mathbb{P}} = 1$ . The Lagrangian is

$$\begin{aligned} \mathcal{L}(L, \nu) &= \alpha + \frac{p-1}{p(1-\gamma)}y + \nu \\ &+ \int \left[ (\ell_{\alpha, y}(\xi) - \nu)L(\xi) - \kappa \phi(L(\xi)) \right] d\hat{\mathbb{P}}(\xi). \end{aligned} \quad (25)$$

Taking the supremum pointwise in  $\xi$  gives

$$\begin{aligned} \sup_{L \geq 0} \int \left[ (\ell_{\alpha, y}(\xi) - \nu)L(\xi) - \kappa \phi(L(\xi)) \right] d\hat{\mathbb{P}}(\xi) \\ = \int \sup_{u \geq 0} \left[ (\ell_{\alpha, y}(\xi) - \nu)u - \kappa \phi(u) \right] d\hat{\mathbb{P}}(\xi) \\ = \kappa \int \phi^* \left( \frac{\ell_{\alpha, y}(\xi) - \nu}{\kappa} \right) d\hat{\mathbb{P}}(\xi), \end{aligned} \quad (26)$$

Substituting the conjugate identity into (24) gives

$$\begin{aligned} \sup_{\mathbb{P}} F(\alpha, y, \mathbb{P}) &= \alpha + \frac{p-1}{p(1-\gamma)}y \\ &+ \inf_{\nu \in \mathbb{R}} \left\{ \nu \right. \\ &\quad \left. + \kappa \int \phi^* \left( \frac{\ell_{\alpha, y}(\xi) - \nu}{\kappa} \right) d\hat{\mathbb{P}}(\xi) \right\}. \end{aligned} \quad (27)$$

Combining this with (23) gives (8).  $\square$

## APPENDIX C

### PROOF OF THEOREM 1

*Proof of Theorem 1.* For fixed sample weights  $\lambda$ , substitute (2) and the perspective identity introduced in Section II into (10). This yields

$$\begin{aligned} \sup_{\lambda \in \Delta_T} \inf_{\alpha \in \mathbb{R}, y > 0} \left\{ \alpha + \frac{p-1}{p(1-\gamma)}y + \sum_{i=1}^T \lambda_i A_i(\alpha, y) \right. \\ \left. - \frac{\kappa}{T} \sum_{i=1}^T \phi(T\lambda_i) \right\} \leq \tau \end{aligned} \quad (28)$$

where  $\Delta_T := \{\lambda \in \mathbb{R}_+^T : \mathbf{1}^\top \lambda = 1\}$  and

$$A_i(\alpha, y) := \frac{1}{p(1-\gamma)} y^{1-p} (\Psi_p(q(x, \xi^i) - \alpha, h))^p. \quad (29)$$

*Step 1.* The simplex  $\Delta_T$  is compact and convex. By Proposition 1, each  $A_i(\alpha, y)$  is convex in  $(\alpha, y)$ , and Assumption 5 provides coercivity of the outer minimization. Lemma 1 and Sion's theorem imply

$$\sup_{\lambda \in \Delta_T} \inf_{\alpha, y > 0} (\dots) = \inf_{\alpha, y > 0} \sup_{\lambda \in \Delta_T} (\dots). \quad (30)$$

*Step 2.* For fixed  $(\alpha, y)$ , the inner maximization is

$$\sup_{\lambda \in \Delta_T} \left\{ \sum_{i=1}^T \lambda_i A_i(\alpha, y) - \frac{\kappa}{T} \sum_{i=1}^T \phi(T\lambda_i) \right\}. \quad (31)$$

With multiplier  $\eta \in \mathbb{R}$  for  $\mathbf{1}^\top \lambda = 1$ , the Lagrangian is

$$\eta + \sum_{i=1}^T \left[ \lambda_i (A_i - \eta) - \frac{\kappa}{T} \phi(T\lambda_i) \right]. \quad (32)$$

Strong duality then gives

$$\begin{aligned} \sup_{\lambda \geq 0, \mathbf{1}^\top \lambda = 1} \left\{ \sum_{i=1}^T \lambda_i A_i - \frac{\kappa}{T} \sum_{i=1}^T \phi(T\lambda_i) \right\} \\ = \inf_{\eta \in \mathbb{R}} \left\{ \eta + \sum_{i=1}^T \sup_{\lambda_i \geq 0} \left[ \lambda_i (A_i - \eta) - \frac{\kappa}{T} \phi(T\lambda_i) \right] \right\}. \end{aligned} \quad (33)$$

With the substitution  $\mu_i = T\lambda_i$  and  $s_i := (A_i - \eta)/\kappa$ , each scalar supremum equals

$$\begin{aligned} \sup_{\mu_i \geq 0} \left[ \lambda_i (A_i - \eta) - \frac{\kappa}{T} \phi(T\lambda_i) \right] &= \frac{1}{T} \sup_{\mu_i \geq 0} \left[ \mu_i (A_i - \eta) \right. \\ &\quad \left. - \kappa \phi(\mu_i) \right] \\ &= \frac{\kappa}{T} \phi^*(s_i). \end{aligned}$$

Thus (10) is equivalent to the existence of  $(\alpha, \eta, y)$  satisfying

$$\begin{aligned} \alpha + \eta + \frac{p-1}{p(1-\gamma)}y + \frac{1}{T} \sum_{i=1}^T \kappa \phi^* \left( \frac{A_i(\alpha, y) - \eta}{\kappa} \right) \\ \leq \tau. \end{aligned} \quad (34)$$

*Step 3.* Since  $\phi^*$  is convex and nondecreasing on  $\mathbb{R}_+$ , (34) holds iff there exist variables  $v_i$  such that

$$A_i(\alpha, y) \leq v_i, \quad i = 1, \dots, T, \quad (35)$$

and

$$\alpha + \eta + \frac{p-1}{p(1-\gamma)}y + \frac{1}{T} \sum_{i=1}^T \kappa \phi^* \left( \frac{v_i - \eta}{\kappa} \right) \leq \tau. \quad (36)$$

If (34) holds, take  $v_i = A_i(\alpha, y)$ . Conversely, if  $A_i(\alpha, y) \leq v_i$ , monotonicity of  $\phi^*$  gives

$$\phi^* \left( \frac{A_i(\alpha, y) - \eta}{\kappa} \right) \leq \phi^* \left( \frac{v_i - \eta}{\kappa} \right), \quad (37)$$

so the lifted system implies (34). This proves (11) for  $\kappa > 0$ ; when  $\kappa = 0$ , the conjugate term reduces to the indicator of the nonpositive orthant and recovers the nominal KDE-HMCR constraint.  $\square$

## APPENDIX D

### PROOF OF PROPOSITION 3

*Proof of Proposition 3.* Under the standing DC-OLS well-posedness conditions stated in Section IV, the baseline constraints (14b)–(14f) and the budget constraint (16b) are affine. For each sample  $\xi^i$ , the functions  $q_j^U$ ,  $q_j^D$ ,  $q_\ell^U$ , and  $q_\ell^D$  are affine in the decision vector by construction. Proposition 1 shows that  $(x, \alpha, y) \mapsto y^{1-p} (\Psi_p(q(x, \xi^i) - \alpha, h))^p$  is convex on  $y > 0$ , and Theorem 1 expresses every worst-case risk constraint as the intersection of these convex epigraph inequalities with the convex conjugate term in (11a). Thus each block of constraints in (16e)–(16l) is convex, and the weighted sum of fragility indices in (16a) is linear. The model is therefore convex. If  $\phi(t) = t \log t - t + 1$ , Corollary 1 replaces every conjugate term by an exponential-cone epigraph, yielding an exponential-cone representation of (16).  $\square$

## APPENDIX E

### REPORTABILITY-THRESHOLD SENSITIVITY AND INTERPRETATION

Reportability is benchmarked at the retained-margin threshold  $\theta = 0.10$ . Because  $\mathbf{1}\{r_s \geq \theta\}$  is monotone in  $\theta$ , this appendix simply re-screens the same fixed-hard RBTS records at  $\theta \in \{0.05, 0.10, 0.15, 0.20\}$  without re-solving the models.

The ordering is unchanged. At  $\theta = 0.05$ , the suite is already at 7/8 reportable snapshots for  $\rho = 1$  to 5 and reaches 8/8 only at  $\rho = 8$ . At the benchmark screen  $\theta = 0.10$ , the counts are 6/8 at  $\rho = 1$  and 7/8 for  $\rho \geq 2$ . At tighter screens  $\theta \in \{0.15, 0.20\}$ , the counts drop to 3/8 at  $\rho = 1$ , 4/8 for  $\rho = 2$  to 5, and 5/8 at  $\rho = 8$ . Thus  $\theta = 0.10$  is a reporting convention rather than a tuned parameter. The same reading carries to RTS79, where altB remains the most screen-sensitive layout.

TABLE VIII  
FIXED-HARD RBTS REPORTABLE COUNTS UNDER ALTERNATIVE  
RETAINED-MARGIN THRESHOLDS

$\rho$	$\theta = 0.05$	$\theta = 0.10$	$\theta = 0.15$	$\theta = 0.20$
1.0	7/8	6/8	3/8	3/8
2.0	7/8	7/8	4/8	4/8
3.0	7/8	7/8	4/8	4/8
4.0	7/8	7/8	4/8	4/8
5.0	7/8	7/8	4/8	4/8
8.0	8/8	7/8	5/8	5/8

Counts re-screen the fixed-hard RBTS records from Table III; no models are re-solved.

#### DATA AVAILABILITY

The core RBTS and RTS79 input files, processed benchmark tables, and scripts used to generate the reported results will be released in a public repository upon acceptance. During review, the authors can provide the run settings, screened snapshot definitions, and core scripts for the reported tables and figures upon request.

#### REFERENCES

- [1] V. Terzija, "Adaptive underfrequency load shedding based on the magnitude of the disturbance estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1260–1266, Aug. 2006.
- [2] T. Shekari, F. Aminifar, and M. Sanaye-Pasand, "An analytical adaptive load shedding scheme against severe combinational disturbances," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4135–4143, Sep. 2016.
- [3] C. Li, Y. Wu, Y. Sun, H. Zhang, Y. Liu, Y. Liu, and V. Terzija, "Continuous under-frequency load shedding scheme for power system adaptive frequency control," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 950–961, Mar. 2020.
- [4] F. Elyasichamazkoti, S. Teimourzadeh, and F. Aminifar, "Optimal distribution of power grid under-frequency load shedding with security considerations," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 4110–4112, Sep. 2022.
- [5] Z. Luo, H. Liu, N. Wang, T. Zhao, and J. Tian, "Optimal adaptive decentralized under-frequency load shedding for islanded smart distribution network considering wind power uncertainty," *Appl. Energy*, vol. 370, Art. no. 123162, 2024.
- [6] H. Zhang and P. Li, "Chance constrained programming for optimal power flow under uncertainty," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2417–2424, Nov. 2011.
- [7] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4742–4751, Nov. 2013.
- [8] W. Xie and S. Ahmed, "Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1860–1867, Mar. 2018.
- [9] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu, "Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4924–4936, Sep. 2018.
- [10] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, nos. 1–2, pp. 115–166, Sep. 2018.
- [11] B. K. Poolla, A. R. Hota, S. Bolognani, D. S. Callaway, and A. Cherukuri, "Wasserstein distributionally robust look-ahead economic dispatch," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2010–2022, May 2021.
- [12] H. Yang, R. Liang, Y. Yuan, B. Chen, S. Xiang, J. Liu, H. Zhao, and E. Ackom, "Distributionally robust optimal dispatch in the power system with high penetration of wind power based on net load fluctuation data," *Appl. Energy*, vol. 313, Art. no. 118813, 2022.
- [13] W. Huang, T. Qian, W. Tang, and J. Wu, "A distributionally robust chance constrained optimization approach for security-constrained optimal power flow problems considering dependent uncertainty of wind power," *Appl. Energy*, vol. 383, Art. no. 125264, 2025.
- [14] J. Li, X. Wang, Z. Wang, and Z. Song, "Conditional distributionally robust dispatch for integrated transmission-distribution systems via distributed optimization," *Int. J. Electr. Power Energy Syst.*, vol. 169, Art. no. 110559, 2025.
- [15] D. Z. Long, M. Sim, and M. Zhou, "Robust satisficing," *Oper. Res.*, vol. 71, no. 1, pp. 61–82, Jan.-Feb. 2023.
- [16] F. Liu, Z. Chen, and S. Wang, "Globalized distributionally robust counterpart," *INFORMS J. Comput.*, vol. 35, no. 5, pp. 1120–1142, Sep.-Oct. 2023.
- [17] W. Liu, L. Yang, and B. Yu, "KDE distributionally robust portfolio optimization with higher moment coherent risk," *Ann. Oper. Res.*, vol. 307, pp. 363–397, 2021, doi: 10.1007/s10479-021-04171-4.
- [18] J. Liu and P. A. Krokhmal, "Higher moment coherent risk measures: Representation and properties," *Risk Anal.*, vol. 41, no. 12, pp. 2377–2395, Dec. 2021.
- [19] P. A. Krokhmal, "Higher moment coherent risk measures," *Quant. Finance*, vol. 7, no. 4, pp. 373–387, Aug. 2007.
- [20] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, no. 3, pp. 21–41, 2000.
- [21] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. London, U.K.: Chapman and Hall, 1986.
- [22] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, 2nd ed. Hoboken, NJ, USA: Wiley, 2015.
- [23] D. Kuhn, P. Mohajerin Esfahani, M. Sim, and S. Uryasev, "Distributionally robust optimization," *Acta Numer.*, vol. 34, pp. 1–164, 2025.
- [24] H. Liu, J. Qiu, and J. Zhao, "A data-driven scheduling model of virtual power plant using Wasserstein distributionally robust optimization," *Int. J. Electr. Power Energy Syst.*, vol. 137, Art. no. 107801, 2022.
- [25] C. Lemaréchal, A. Nemirovskii, and Y. Nesterov, "New variants of bundle methods," *Math. Program.*, vol. 69, pp. 111–148, 1995.
- [26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [27] C. Wang, Z. Wang, S. Chu, H. Ma, N. Yang, Z. Zhao, C. S. Lai, and L. L. Lai, "A two-stage underfrequency load shedding strategy for microgrid groups considering risk avoidance," *Appl. Energy*, vol. 367, Art. no. 123343, 2024.
- [28] M. Sarvarizadeh, M. Rajabdorri, E. Lobato, L. Sigrist, B. Kazemtabrizi, and M. Troffaes, "Under frequency load shedding aware unit commitment in island power systems," *Int. J. Electr. Power Energy Syst.*, vol. 170, Art. no. 110884, Sep. 2025.
- [29] Z. Wang, S. Bu, J. Wen, and C. Huang, "A comprehensive review on uncertainty modeling methods in modern power systems," *Int. J. Electr. Power Energy Syst.*, vol. 166, Art. no. 110534, May 2025.
- [30] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
- [31] R. Billinton, S. Kumar, N. Chowdhury, K. Chu, K. Debnath, L. Goel, E. Khan, P. Kos, G. Nourbakhsh, and J. Oteng-Adjei, "A reliability test system for educational purposes—basic data," *IEEE Trans. Power Syst.*, vol. 4, no. 3, pp. 1238–1244, Aug. 1989.
- [32] Reliability Test System Task Force of the Application of Probability Methods Subcommittee, "IEEE reliability test system," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 6, pp. 2047–2054, Nov.-Dec. 1979.
- [33] H. Qi, C. Li, C. Gai, R. Sun, and J. Yan, "Optimization of power system emergency load shedding with surrogate-assisted proliferation-reduction differential evolution algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 173, Art. no. 111358, Dec. 2025.
- [34] M. Tillmanns, J. Schöttler, and A. Praktiknjo, "A review of probabilistic resource adequacy assessments in power systems: Methods, applications, and future challenges," *Energy Policy*, vol. 209, pt. A, Art. no. 114924, Feb. 2026.