FROM MODEL COVERAGE TO CONTROL VALUE: AUDITING FEW-SHOT SAFETY FILTERS UNDER TOPOLOGY SHIFT

Anonymous authors
Paper under double-blind review

ABSTRACT

Safety filters are often judged by model error, yet a tighter model can be useless or uncertified: its error premise may fail, its reserve may empty the feasible set, or its rows may be inactive at the deployed decision. We introduce an optimizer-level, four-stage audit that separates model-premise validity, certificate availability, decision relevance, and control value. Under shared convex non-target constraints and a strictly convex objective, a single-solve certificate proves that every candidate constraint model whose feasible set contains a reference optimizer with strict target-row slack returns the same unique action; a computable perturbation radius extends the screen to a local model family. This certificate requires verified optimality and a shared problem, not voltage feasibility alone. We instantiate the audit for few-shot voltage shielding of electric-vehicle charging under topology shift using three complete target sensitivity queries. On three frozen feeders, interpolation upper-covers all 60/60 unqueried interior matrices and reduces final upper-bound overstatement by 38.74–61.96% relative to a same-budget baseline. The tighter model nevertheless changes no action in 34,560 comparisons: one of 15 base-feasible target–pressure layers fails zero-action voltage membership, and all remaining deployed actions have strictly inactive voltage rows. A prospectively frozen IEEE 8500 stress test covers only 3/20 matrices and therefore stops at the first gate; a state-binned reserve improves transfer coverage in 0/8 development comparisons and is not confirmed. The audit converts null and failed-transfer results into falsifiable diagnoses without treating finite-grid coverage as a population, arbitrary-topology, or nonlinear-AC guarantee.

1 INTRODUCTION

Safety filters make learned or heuristic controllers easier to deploy: an arbitrary proposal is projected into a modeled feasible set (Wabersich & Zeilinger, 2021; Amos & Kolter, 2017; Agrawal et al., 2019). But predictive fidelity and decision value are different objects. This is a central lesson of predict-then-optimize and decision-focused learning (Elmachtoub & Grigas, 2022; Wilder et al., 2019); safety filtering adds two failure modes absent from ordinary objective regret. An uncertainty premise can fail, and a valid reserve can remove every certifiable action before model accuracy reaches the optimizer. Under topology shift, however, policy transfer and constraint transfer are different problems. A policy can produce a dimensionally valid charging action on a new feeder while the downstream voltage model assigns the wrong sensitivity to a particular phase or station. The proposal transfers; the physical information contract may not.

A common few-shot response is to query a few target operating points, retain a source- or topology-derived nominal model as the center, and inflate an error set around it. This can waste the most expensive information in the protocol: if each query already returns a complete target sensitivity matrix, those matrices can update the center directly. The remaining uncertainty is an interpolation residual, not the full discrepancy from a coarse nominal model. The distinction is invisible when constraints are loose, but it determines the width and interpretation of the declared safe set.

We audit this distinction in phase-resolved EV charging and make availability an explicit part of the error contract. At load scales 0.25, 0.50, and 0.75, we measure the complete target matrix relating

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

eight station powers to all protected phase voltages. InterpBox linearly interpolates adjacent matrices and adds a one-sided local margin whose single ratio was frozen on four development feeders. We compare it with an elementwise box and a row-max box centered on the same topology model, using the same target queries, live no-EV phase voltages, and nonlinear reserve. We then separate coefficient, base-state, and nonlinear residual error and test whether a prespecified state-binned reserve transfers between seen feeders.

The study is organized around falsifiable evidence roles. Static confirmation tests whether a frozen upper coefficient model covers every entry of unqueried interior matrices. Closed-loop confirmation tests whether tighter models change actions or safety on new feeders and dates. A development-only IEEE 34 case supplies a mechanism stress test, but it is not allowed to substitute for held-out confirmation. Finally, certificate availability is reported separately from AC replay: a controller that returns zero because its reserve already exceeds base headroom has not demonstrated useful safe service.

Our contributions are:

- a reusable, optimizer-level four-stage evidence funnel—model- premise validity, certificate availability, decision relevance, and control value—with explicit failure semantics and a single-solve decision- invariance certificate as a conditional, per-decision diagnostic using one verified reference optimizer and row evaluations;
- an availability-aware error contract that separates a one-sided coefficient upper matrix, base-voltage overestimation, and nonlinear residual within a conditional convex certificate, together with a computable local model-change radius;
- frozen confirmation evidence on three held-out topologies showing 60/60 simultaneous matrix coverage and a post hoc 38.74–61.96% reduction in final upper-bound overstatement, followed by a diagnosed closed-loop null, a prospectively frozen IEEE 8500 premise-gate failure, and a failed reserve- transfer gate rather than outcome-selected positive claims.

The methodological output is a stopping rule, not a dominance claim: reject an unsupported premise, report an unavailable certificate, certify decision invariance, or proceed to paired control evaluation. Its empirical contract is deterministic and limited to named feeders and load intervals, not distribution-free coverage, arbitrary-topology transfer, or deployment authorization.

2 Related Work

Safety filters and conservatism. Predictive safety filters minimally modify proposed actions under modeled uncertainty (Wabersich & Zeilinger, 2021); differentiable optimization and completion/correction layers embed feasibility in learning pipelines (Amos & Kolter, 2017; Agrawal et al., 2019; Donti et al., 2021). Data-driven safe-set expansion targets filter conservatism (Bajelani & van Heusden, 2024). Our comparison holds proposal, horizon, target-query budget, and online voltage information fixed. It asks whether complete target measurements should update the center before uncertainty is added.

Power-system feasibility layers. Power-system methods project learned actions or predictions toward AC-feasible sets and construct robustly feasible predictors (Zhang et al., 2023; Han et al., 2024; Stratigakos et al., 2023). EV-charging work similarly combines policies with optimization or rule layers (Li et al., 2020; Zhang et al., 2024; Guan et al., 2024). Most directly, H-ACI combines a robust model-based constraint manifold with learned active- constraint refinement for DER/EV scheduling (Zhang et al., 2026), while SafePowerGraph benchmarks safety-aware GNNs across simulators and grid perturbations (Ma et al., 2026). These works rule out broad novelty claims for robust constraint guidance or safety-aware graph evaluation. Topology-aware GNNs also transfer OPF mappings across network states (Liu et al., 2023; Zhou et al., 2023; Yang et al., 2024). Our narrower question is different: once complete target sensitivities have already been measured, how should they center a conditional post-policy certificate, and when does a tighter certificate alter the deployed decision?

Calibration language under shift. Conformal safety methods make probabilistic statements under explicit exchangeability, adaptation, or shift assumptions (Strawn et al., 2023; Huriot et al., 2026;

Rahaman et al., 2026). CPSF calibrates future-agent trajectory regions, while ACoFi adapts a switching threshold from realized safety-value errors and provides a long-run soft guarantee. Our feeder, date, and load grids are deterministic. We report complete-matrix counts and paired date effects; we do not call the frozen maximum a confidence bound or a conformal quantile. In particular, the resulting finite-grid statement is not a conformal or population guarantee.

Decision-focused evaluation. Predict-then-optimize methods replace parameter error with downstream decision loss because a more accurate prediction need not induce a better decision (Elmachtoub & Grigas, 2022; Wilder et al., 2019). Our audit applies the same principle to uncertain constraints rather than predicted objectives. Here the chain has two additional gates: the uncertainty premise must hold and the reserved feasible set must exist before constraint refinement can have decision value. Constraint relevance itself is not new: active-set learning targets constraints that determine optimized solutions (Misra et al., 2022), and learned safety constraints need only be correct on the nominal policy's critical set (Massiani et al., 2021). Our distinction is operational: we separate invalid premises, unavailable certificates, and inactive deployed constraints under a fixed information budget, then audit the last gate from one reference solve and candidate-row feasibility checks. We do not train through the optimizer; we use frozen interventions and service outcomes to identify where the chain breaks.

3 Problem and Availability-Aware Shield

3.1 Charging, voltage, and one-step service viability

There are $M = 8$ charging stations and 96 fifteen-minute slots per day. At a slot, a proposal policy outputs nonnegative station power $z \in \mathbb{R}^M_+$. Earliest-deadline-first disaggregation maps station power to active sessions, while station, port, and transformer caps remain exact linear constraints. For a session with remaining energy E, n available slots including the current one, duration Δ, and port cap $\bar{p}$, the current power must satisfy

$$\underline{p} = \frac{\max\{0, E - \Delta\bar{p}(n-1)\}}{\Delta}, \qquad p \le \min\{\bar{p}, E/\Delta\}. \tag{1}$$

This interval is necessary and sufficient to leave a port-feasible completion schedule; future grid feasibility is a separate premise.

Let s denote information available before optimization, including the background-load scale λ. Write $u_0(s) \in \mathbb{R}^J$ for the true no-EV squared voltage and $\widehat{u}_0(s)$ for its online estimate. A local voltage-drop model is

$$u(P, s) = u_0(s) - A(s)P - r(P, s), \tag{2}$$

where each row is an individual protected phase voltage, $A \in \mathbb{R}^{J\times M}$ is the local sensitivity, and r contains nonlinear and discrete-control residuals. The endpoint is $u(P, s) \ge v^2_{\min}\mathbf{1}$ with $v_{\min} = 0.95$ p.u.

3.2 Interpolate the queried matrices and separate errors

We query $A_k = A(\lambda_k)$ at $\lambda_k \in \{0.25, 0.50, 0.75\}$. For $\lambda \in [\lambda_k, \lambda_{k+1}]$, let $t = (\lambda - \lambda_k)/(\lambda_{k+1} - \lambda_k)$ and define

$$\bar{A}(\lambda) = (1 - t)A_k + tA_{k+1}. \tag{3}$$

The local query-only scale is

$$S_{jm} = \max\{|(A_k)_{jm}|, |(A_{k+1})_{jm}|, \tau_k\}, \qquad \tau_k = 0.01\,\mathrm{median}_{S^0_{jm}>0} S^0_{jm}, \tag{4}$$

where $S^0_{jm} = \max\{|(A_k)_{jm}|, |(A_{k+1})_{jm}|\}$. The safety-relevant upper model is

$$U(\lambda) = \bar{A}(\lambda) + qS(\lambda), \qquad q = 0.0203262. \tag{5}$$

The ratio q is 1.025 times the largest normalized positive interpolation residual on four development feeders. It was frozen before the three confirmation feeders were measured. Negative coefficient residuals do not enlarge Eq. (5) because charging powers are nonnegative.

The availability-aware contract keeps three one-sided errors distinct:

$$A(s) \le U(s), \qquad \widehat{u}_0(s) - u_0(s) \le \eta_0(s), \qquad r(P,s) \le \rho(P,s). \tag{6}$$

The shield projects the proposal under Eq. (1), exact capacity rows, and

$$U(s)P + \eta_0(s) + \rho(P,s) \le \widehat{u}_0(s) - v_{\min}^2 \mathbf{1}. \tag{7}$$

EDF disaggregation gives an active-session proposal $\widetilde{z}$; the implemented shield minimizes $\frac{1}{2}\|p - \widetilde{z}\|_2^2$ over session powers p, whose station aggregate is P. Thus the deployed projection objective used below is strictly convex. For eight stations, a matrix requires one base solve and eight fixed-power finite differences; three probes cost 27 solves. The live no-EV phase-voltage snapshot is additional deployment information, not evaluation leakage. The primary frozen controller is the special case $\widehat{u}_0 = u_0$, $\eta_0 = 0$, and $\rho(P,s) = \varepsilon \mathbf{1}$ with $\varepsilon = 0.005$. If s is fixed before optimization and every $\rho_j(\cdot, s)$ is convex, Eq. (7) and the charging constraints define a convex feasible set. A global or state-binned constant reserve is the simplest such case.

Theorem 1 (Conditional phase-voltage feasibility)**.** *Suppose all three inequalities in Eq. (6) hold. Any $P \ge 0$ satisfying Eq. (7) has $u_j(P,s) \ge v_{\min}^2$ for every j.*

The proof is the one-sided box support argument in Appendix A. The result is conditional: finite development or confirmation data can audit the premises on named states but do not prove that empirical maxima cover a future population. If a numerical post-check permits row slack down to $-\tau_{\text{num}}$, the corresponding conclusion is only $u_j \ge v_{\min}^2 - \tau_{\text{num}}$; exact endpoint certification requires tightening the implemented right-hand side by that tolerance.

This motivates the evaluation chain

$$\text{premise validity} \implies \text{certificate availability} \implies \text{decision relevance} \implies \text{control value}. \tag{8}$$

The arrows are evidentiary preconditions for attributing *certified* control value to a model refinement, not physical or probabilistic implications: a method may deliver service after a premise failure, but that service is not certified by the declared contract, and passing one stage does not guarantee the next. Evaluate the gates in order and return the first supported diagnosis: `premise_failed` if a required error test fails; `unverified` if required error-bound evidence is missing; `unavailable` if the declared voltage set is proven empty; `unverified` if optimizer or shared-problem evidence is missing; and `decision_invariant` if one strict reference optimum is feasible under every candidate; otherwise return `effect_required`. These outcomes stop certified attribution, not empirical measurement. The screen is per decision. Trajectory invariance additionally requires its premises at every visited state, common initial conditions and exogenous inputs, and identical policy and state updates; otherwise paired evaluation remains necessary for a trajectory claim (Appendix A).

3.3 Same-budget baselines and certificate availability

Let H be a topology-derived nominal matrix. Within each query bracket, the element baseline E uses $H + C^E$, where C^E_{jm} is 1.025 times the largest absolute endpoint residual $|(A_k - H)_{jm}|$. The row baseline R replaces the entries of each row of C^E by their row maximum. A one-sided version U replaces the absolute endpoint residual by its positive part, isolating the cost of symmetric error modeling. I, E, R, and U use the same three complete target matrices, live u_0, and reserve. A current-matrix reference F uses $A(\lambda)$ at every evaluation point and is not query matched. N uses H with no coefficient bound and no nonlinear reserve.

E's feasible set contains R's whenever powers are nonnegative, but no set inclusion is asserted between I and E because their centers differ. More fundamentally, even the zero action satisfies the general voltage certificate only if

$$\widehat{u}_0(s) \ge v_{\min}^2 \mathbf{1} + \eta_0(s) + \rho(0,s). \tag{9}$$

For v0.6 this reduces to $u_0 \ge v_{\min}^2 + 0.005$. Thus a base state can exceed 0.95 p.u. yet fail availability. We report this state as uncertified, not as a safe zero action.

Proposition 1 (Single-solve decision-invariance certificate)**.** *Let p denote session powers and $P = Sp$ their station aggregation. Let C be a convex non-voltage set in session space, $f(p)$ a strictly convex objective, and $V_0 = \{p : B_0 p \le d_0\}$, where $B_0 = U_0 S$. If the optimizer $p^\star$ over $C \cap V_0$ satisfies*

$B_0 p^\star < d_0$ componentwise, then $p^\star$ is the unique optimizer over C. Consequently, every candidate voltage set V_k containing $p^\star$ has the same unique optimizer, even if a candidate row is active. One reference solve plus candidate-row feasibility checks can therefore certify action invariance across a finite model family.

Corollary 1 (Local model-change invariance radius)**.** *Under Proposition 1, let $P^\star = Sp^\star$ and b_{0j} be station-space rows of U_0. Define slack $s_j = d_{0j} - b_{0j}^\top P^\star > 0$. For perturbed rows $b'_j = b_{0j} + \Delta b_j$ and $d'_j = d_{0j} + \Delta d_j$, $p^\star$ remains the unique optimizer if, for every row, $\|\Delta b_j\|_\infty \|P^\star\|_1 + |\Delta d_j| < s_j$.*

Appendix A gives the proofs. After the reference solve, for K models including the reference, checking $K-1$ candidates with J rows and M stations costs $O((K-1)JM)$ arithmetic after aggregation. If a candidate changes the action, the reference rows cannot be strict or that action must be infeasible under the candidate. This statement applies to any convex safety layer with a shared proposal objective and shared non-target constraints, not only to voltage shielding. A service-relaxed fallback changes C; solver success and strict voltage slack alone do not prove optimality.

4 Frozen Evaluation Protocol

Development. IEEE 34, IEEE 37, and two previously inspected SMART-DS feeders were used to compare E/R/U/I, freeze q, and define a mechanism diagnostic. On all four, I covered 20/20 interior matrices in leave-one-feeder-out development. A boundary case on IEEE 34 produced a large E–R service gap and motivated the action-aligned diagnostic described below. These results are development evidence and cannot establish external replication.

Confirmation feeders and static endpoint. Before target method outcomes, we selected the first lexicographic non-analysis feeder under each of the next three unused SMART-DS Austin substations: T1 (`p1rhs2--p1rdt1528`), T2 (`p1rhs3--p1rdt2378`), and T3 (`p1rhs4--p1rdt2300`). Their source-connected components contain 863, 1,357, and 847 buses. Eight stations are chosen by deterministic topology-only farthest-point sampling. SMART-DS is synthetic (Palmintier et al., 2020); the three feeders are distinct deterministic instances from one city and data-generation system, not random field samples.

Each target supplies only the three probe matrices. We evaluate ten strict interior points per adjacent interval, for 20 matrices per target. The primary endpoint is simultaneous upper coverage of every phase–station entry in a matrix. Passing was fixed as 20/20 per target; confirmation data could not be used to increase q.

Dates, pressures, and replay. An exposure audit found 2,431 eligible dates absent from prior result files. Using only session inputs, we selected two dates per month near the one-third and two-thirds ranks of a demand/port pressure score, for 24 dates. The three main load scales are $\{0.25, 0.50, 0.75\}$; strict-interior audit scales are $\{0.375, 0.625\}$. Every method, target, date, and pressure is replayed for 96 slots in phase-resolved OpenDSS, yielding 2160 trajectories and 207360 retained slot rows. Terminal power is read back for all eight stations.

Demand, policy, and metric. Public Palo Alto ChargePoint sessions (City of Palo Alto, 2021) are demand templates: each selected day's filtered sessions are copied four times, then six 11.5-kWh sessions per station are added from 17:00 to 19:00. All methods use a max-charge proposal with EDF disaggregation, 7-kW ports, 42-kW station caps, and a 260-kW aggregate EV cap. Service is delivered energy divided by total constructed demand; historical delivered energy is a demand proxy. Background load is fixed within a day. Control updates session energy online; the complete action trace is then replayed in AC power flow, without grid-state feedback. Appendix C details preprocessing and this workload.

Primary comparisons and mechanism falsification. Date is the demand statistical unit. At the three main pressures, each date is averaged first within target and then equally across the three targets. The two predeclared paired comparisons are I–E and E–R. We report 10,000- replicate paired date-bootstrap intervals and Holm-adjusted Wilcoxon tests; phases, slots, and target repetitions do not enlarge $n = 24$.

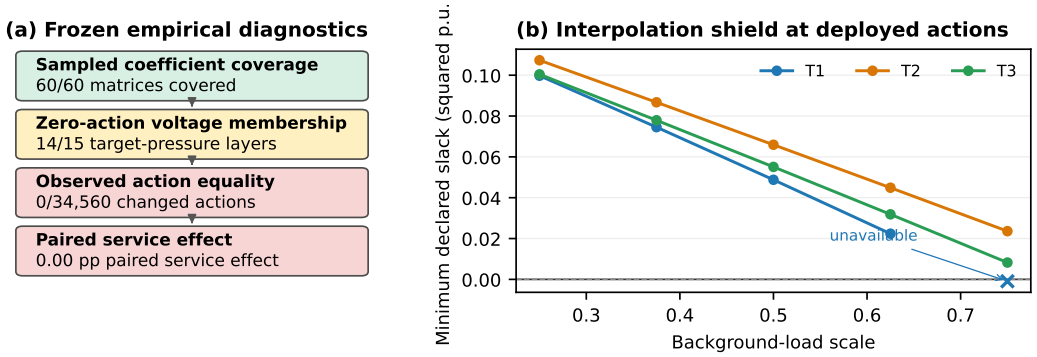


Figure 1: From coefficient coverage to control value. Left: the frozen audit passes sampled coefficient coverage but yields no changed query-matched action or service effect. The full error premise and optimizer-level certificate remain unverified. Right: minimum interpolation-shield slack over all dates and slots in each target–pressure layer. The cross is the base-feasible T1/0.75 layer where the fixed reserve makes even zero action unavailable. All other deployed solutions remain strictly inside the declared voltage set.

On I's actions P_t^I, development defined

$$M_{\text{day}} = \frac{1}{96} \sum_t \max_j \left[(C^R - C^E) P_t^I - s_t^E \right]_+ , \tag{10}$$

where s_t^E is E's declared phase-row slack. The frozen prediction is $2301.176588 M_{\text{day}}$ service percentage points for E–R. Confirmation reports its error without refitting.

5 Results

Figure 1 separates the observed diagnostics. Sampled coefficient coverage passes; under the fixed reserve, 14/15 layers retain zero-action voltage membership. Stored query-matched actions and paired service effects are identical. These observations do not establish the full residual premise or optimizer-level invariance: shared-problem and optimality evidence are missing. The denominators differ; this is not a single attrition rate.

5.1 Static confirmation: update the center first

Table 1 gives the frozen static result. InterpBox covers all 60/60 unqueried matrices. E and R also cover every matrix, so the comparison is not coverage versus safety. It is the amount of upper coefficient overstatement under the same three target queries. A post hoc descriptive reanalysis computes $\text{mean}[U - A]_+$ over all phase–station entries and 20 interior matrices per target. This center-independent metric gives a 38.74–61.96% reduction versus E across T1–T3. Additive margins around the different centers instead suggest 97–99% reductions; they are not comparable as final-model errors (Appendix G). Coverage is the prespecified endpoint; overstatement is descriptive, not a voltage or control-performance metric.

This changes the baseline hierarchy. The topology-centered element box is a valid but unnecessarily indirect use of complete target matrices. Directly interpolating those matrices reduces final overstatement. The conclusion is empirical and in-range: it does not prove continuous coverage between the 20 checked points.

5.2 Closed loop: the tighter model changes no query-matched action

The independent closed loop does not reproduce the development service gap. Across all 34,560 I slots, I and E choose the same eight-dimensional station action; I and R do as well. Consequently, the prespecified main effects are I–E $= 0.0000$ and E–R $= 0.0000$ percentage points, both with

Table 1: Static confirmation and post hoc final overstatement $\mathrm{mean}[U - A]_+$ (squared p.u./kW). "Matrices" requires coverage of every phase–station coefficient at each of 20 strict-interior scales.

Target	Upper model	Matrices	Final overstatement
T1	Interp.	20/20	3.04e-07
	Nominal+element	20/20	8.00e-07
	Nominal+row	20/20	1.36e-05
T2	Interp.	20/20	2.27e-07
	Nominal+element	20/20	3.71e-07
	Nominal+row	20/20	8.35e-06
T3	Interp.	20/20	5.57e-07
	Nominal+element	20/20	1.07e-06
	Nominal+row	20/20	1.16e-05

Table 2: Frozen confirmation for InterpBox. Service shows the three main pressures; all five layers are used for voltage membership and action equality. "Voltage layers" requires all deployed actions to satisfy the declared rows, not the full error premise or service constraints. Each layer has 24 dates.

Target	Service .25/.50/.75 (%)	Min. AC	Viol.	Voltage layers	I/E/R equal
T1	94.75/94.75/0.00	0.95214	0/120	4/5	yes
T2	94.75/94.75/94.75	0.96498	0/120	5/5	yes
T3	94.75/94.75/94.75	0.95701	0/120	5/5	yes

date-bootstrap interval $[0, 0]$ and Holm $p = 1$. This is a null result, not a population equivalence guarantee. Each query-matched method has 32,256 voltage-member slots, all with strict voltage slack. The smallest slack across I/E/R is 0.003124 squared p.u., more than 300 times the 10^{-5} active-constraint tolerance. The remaining 2,304 slots belong to the unavailable T1/0.75 layer. However, the available layers include 3,570 grid-safe fallbacks that remove service lower bounds. The ledger lacks per-slot branch labels, session actions, proposals, and optimality witnesses. Thus Proposition 1 cannot certify all stored actions as optima of one shared non-voltage problem. The 161,280 E/R/U/F/N pairs establish retrospective station-action equality, not independent forward screening or saved solves. On positive-power voltage-member slots, slack divided by station-power norm gives a coefficient-row *feasibility* radius. Its frozen minimum across I/E/R is 1.20×10^{-5} squared p.u./kW; the pooled fifth percentile and median are 1.24×10^{-4} and 7.80×10^{-4}. Optimizer invariance additionally requires the theorem's unverified premises. The higher-information current-matrix reference F also differs from I on only 0/34,560 actions under the same reserve. N differs from I on 1,650 actions, all in the unavailable T1/0.75 layer (0 elsewhere). This post hoc triangulation localizes the observed action separation to the availability failure, not the fourteen available layers where coefficient fidelity differs. These observations do not isolate a causal reserve effect or prove that service-relaxed and original problems have identical optimizers.

InterpBox has 0 AC-violating trajectories among its 360 replays, with minimum observed voltage 0.95214 p.u. These deterministic zeros do not estimate a small population failure probability. They also do not rescue the mechanism claim: all confirmation E–R gaps are zero, while the frozen predictor is nonzero on 24 rows and has main-pressure MAE 0.2376 points; rank correlation is undefined because the observed effect is constant. The predictor therefore fails its external gate.

5.3 A FIXED RESERVE CAN ERASE CERTIFICATE AVAILABILITY

T1 at load scale 0.75 has no-EV minimum voltage 0.952140 p.u., so it is base- feasible. Its weakest squared-voltage headroom is nevertheless below 0.005, violating Eq. (9). I/E/R/U and even the current-matrix reference F report infeasible contexts for every slot and serve 0.00%. Their AC replay stays at the safe no-EV base because no charging is deployed. N, which removes both coefficient and nonlinear reserve, serves 94.75% and reaches 0.950185 p.u. without an observed violation. N is informative but uncertified: one replay set does not prove that the omitted residual is harmless.

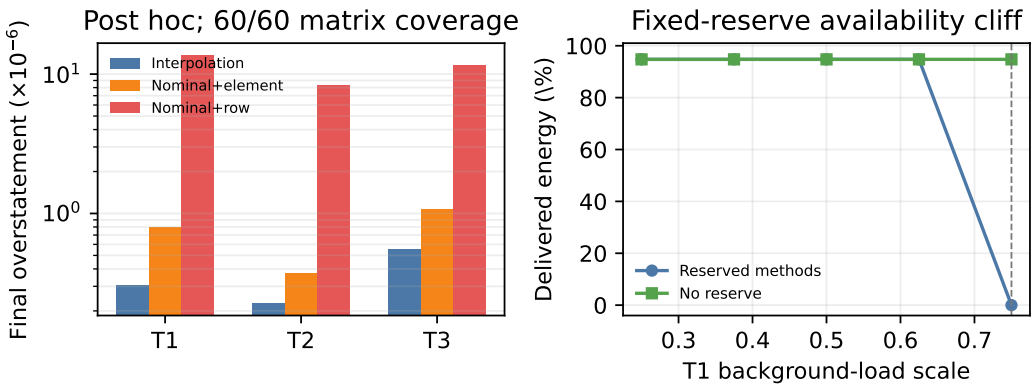


Figure 2: Left: post hoc final overstatement $\mathrm{mean}[U - A]_+$ in 10^{-6} squared p.u./kW; all three models cover 60/60 interior matrices. Right: the reserved methods have the same service on T1; the fixed reserve creates an availability cliff at 0.75, while the no-reserve reference continues charging without the same certificate.

Table 3: Development-only IEEE 34 boundary stress test at load scale 0.4773, averaged over six reused dates. It explains a possible mechanism but is not an external effect estimate.

Method	Service (%)	Min. AC	Viol. traj.
Interpolation, reserve	92.99	0.9529	0/6
Nominal+element, reserve	92.98	0.9530	0/6
Nominal+row, reserve	64.37	0.9602	0/6
Current matrix, reserve	92.92	0.9525	0/6
Nominal, no reserve	93.07	0.9491	1/6

The distinction is operationally important. "No AC violation" can mean either that a useful action survived replay or that the controller refused all service. A useful safety report must include certificate availability, service, and base headroom alongside violation counts.

5.4 A SIMPLE STATE RESERVE FAILS ITS DEVELOPMENT TRANSFER GATE

After the frozen study, we implemented Eq. (6) without changing the primary outputs. On four already-seen feeders, the development ledger contains 13,824 slot–state rows under an ideal online base snapshot and a query-interpolated base estimate. All 12/12 tested upper matrices cover their measured development matrices, so the residual is computed relative to an independently measured A rather than absorbing coefficient error again. Within each model source, a global constant and a single prespecified load-binned constant are fitted on one feeder and audited on the other.

The state-binned rule improves coverage in 0/8 held-out feeder–information comparisons. The global constant is itself incomplete in 5/8 comparisons. Interpolated base voltage additionally fails across discrete regulator/control changes on the IEEE feeders. Appendix I gives the complete counts. This failed gate stops confirmation. Later action-dependent development also fails its transfer and availability gate (Appendix J).

5.5 WHY THE DEVELOPMENT MECHANISM REMAINS ONLY A STRESS TEST

On the already inspected IEEE 34 interior boundary, R excludes 198 of I's 576 slot actions. I serves 92.99%, E 92.98%, and R 64.37%, all without AC violations; F serves 92.92%. The no-reserve nominal reference violates the AC limit on 1/6 trajectories. Table 3 confirms that row compression *can* matter when its discarded station structure aligns with a binding phase constraint. The new confirmation shows that this condition is not common enough in the selected targets to support a general service claim.

6 Discussion and Limitations

What is supported. Using already-purchased target matrices to update the center reduces coefficient slack in the frozen three-topology test, without improving control. Proposition 1 and Corollary 1 provide conditional invariance tests, not a new projection algorithm. A separate forward implementation certifies 25/160 development candidate pairs using an analytic box-optimality witness, with zero mismatches, but 135 abstentions. Its measured total runtime is 1.01 times full solving and 1.00 times a shared-box baseline in this run; we claim no speed advantage (Appendix A.4).

An earlier frozen protocol on IEEE 37 and SMART-DS likewise covered 40/40 matrices while all 5 variants tied on 24/24 service rows, with 0/24 primary AC violations. Its different upper model and 6 reused dates make this a replication only of the coverage–value separation, not an independent confirmation of the v0.6 interpolation method (Appendix F.2).

What is falsified. Narrower coefficient sets do not automatically improve closed-loop service. The action-aligned prediction failed because the held-out conditions never separated I, E, and R. We claim neither a general service advantage nor a transferable mechanism; more similar loose feeders cannot repair this result.

Decision-relevance gate. On four seen feeders, our input-only pressure-selection gate finds eligible interior scales on only two and the planned four on none. Its continuation rule therefore fails and no new confirmation is launched (Appendix H); no outcome-selected pressures are substituted.

Reserve design. The inherited 0.005 residual buffer is not an analytic global bound and can consume all base headroom. Equation (6) instead separates coefficient, base-voltage, and nonlinear/discrete-control errors. Both the state-binned and later action-normalized candidates fail development transfer gates. A replacement must explain transfer variation, beat a calibrated global constant under identical information, and retain an availability gate; merely reducing 0.005 is not a contribution.

Information and data limits. "Three queries" means three full phase-by-station matrices, not three scalar measurements. The frozen controller also assumes a trustworthy current no-EV voltage at every protected phase. The reserve development audit shows that query-interpolated base voltage can fail at discrete control changes; it does not yet test sensor noise, state-estimation delay, communication faults, or hardware dynamics. The primary study uses one synthetic SMART-DS setting and 24 transformed demand dates. A prospectively frozen unbalanced IEEE 8500 stress test stops at the model-premise gate (I: 3/20 matrices; Appendix F.1); it does not confirm interpolation across sources. Earlier IEEE 37 results broaden only the coverage–value pattern.

Direct-query baseline audit. After the frozen confirmation, development-only leave-one-feeder-out tests added elementwise-maximum and nearest-query rules with equally calibrated one-sided margins. They cover 62/80 and 77/80 matrices versus 80/80 for interpolation; uncalibrated interpolation covers 40/80. Neither simple rule passes the gate, but this post hoc result does not prove future dominance; a new confirmation must freeze a safe simple baseline.

Operational and social limits. Unavailable certificates can deny charging to users with inflexible departures. Aggregation is not privacy; deployment requires utility validation, fairness, monitoring, and human override.

7 Conclusion

Model coverage alone does not establish control value. On three frozen topologies, tighter interpolation changes no action; the external IEEE 8500 test instead fails coefficient coverage. The four-stage audit separates error validity, availability, decision relevance and value. Its per-decision invariance screen requires verified shared-problem optimality; missing error or optimizer evidence requires abstention. It turns null effects into falsifiable diagnoses without claiming arbitrary-topology or nonlinear-AC safety.

Reproducibility statement

Appendices A–N give the complete proof, information ledger, target/date selection, protocol evolution, precheck repairs, and claim boundaries. The experiments use named public feeder and charging-session data. The local research record retains static points, trajectory and slot rows, command/readback vectors, frozen manifests, and source hashes; paper tables, macros, and figures are regenerated from those records. Full external replication also requires the implementation.

Ethics statement

The study uses public charging-session records after station-level processing and does not attempt to identify drivers. Aggregation is not a privacy guarantee. A conservative but unavailable certificate can deny service to drivers with inflexible departures, so field use would require utility review, fairness and override policies, monitored state estimation, and validation on the deployed feeder. This simulator study does not authorize deployment.

AI use statement

Generative AI tools were used only to polish the language of the manuscript. The authors reviewed the edits and take responsibility for the final text.

References

Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J. Zico Kolter. Differentiable convex optimization layers. In *Advances in Neural Information Processing Systems*, 2019.

Brandon Amos and J. Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, 2017.

Mohammad Bajelani and Klaske van Heusden. From raw data to safety: Reducing conservatism by set expansion. In *Proceedings of the 6th Annual Learning for Dynamics and Control Conference*, volume 242 of *Proceedings of Machine Learning Research*, pp. 1305–1317. PMLR, 2024. URL https://proceedings.mlr.press/v242/bajelani24a.html.

City of Palo Alto. Electric vehicle charging station usage (july 2011–dec 2020). Open Data Portal, 2021. URL https://data.cityofpaloalto.org/datasets/194693/electric-vehicle-charging-station-usage-july-2011-dec-2020/.

Priya L. Donti, David Rolnick, and J. Zico Kolter. DC3: A learning method for optimization with hard constraints. In *International Conference on Learning Representations*, 2021.

Adam N. Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1): 9–26, 2022. doi: 10.1287/mnsc.2020.3922.

Yuxiang Guan, Jin Zhang, Wenhao Ma, and Liang Che. Rule-based shields embedded safe reinforcement learning approach for electric vehicle charging control. *International Journal of Electrical Power & Energy Systems*, 158:109863, 2024. doi: 10.1016/j.ijepes.2024.109863.

Jiayu Han, Wei Wang, Chao Yang, Mengyang Niu, Cheng Yang, Lei Yan, and Zuyi Li. FRMNet: A feasibility restoration mapping deep neural network for AC optimal power flow. *IEEE Transactions on Power Systems*, 2024. doi: 10.1109/TPWRS.2024.3354733.

Sacha Huriot, Ihab Tabbara, and Hussein Sibai. Safe control using learned safety filters and adaptive conformal inference. In *Proceedings of the 8th Annual Learning for Dynamics and Control Conference*, volume 331 of *Proceedings of Machine Learning Research*, pp. 833–847. PMLR, 2026. URL https://proceedings.mlr.press/v331/huriot26a.html.

Hepeng Li, Zhiqiang Wan, and Haibo He. Constrained EV charging scheduling based on safe deep reinforcement learning. *IEEE Transactions on Smart Grid*, 11(3), 2020. doi: 10.1109/TSG.2019.2955437.

Shaohui Liu, Chengyang Wu, and Hao Zhu. Topology-aware graph neural networks for learning feasible and adaptive AC-OPF solutions. *IEEE Transactions on Power Systems*, 38(6), 2023. doi: 10.1109/TPWRS.2022.3230555.

Aoxiang Ma, Salah Ghamizi, Aleksandar Bojchevski, Pei Zhang, and Jun Cao. SafePowerGraph: A safety-aware benchmarking framework for scalable graph neural networks in power flow and optimal power flow. *Energy and AI*, 25:100819, 2026. doi: 10.1016/j.egyai.2026.100819.

Pierre-François Massiani, Steve Heim, and Sebastian Trimpe. On exploration requirements for learning safety constraints. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pp. 905–916. PMLR, 2021. URL https://proceedings.mlr.press/v144/massiani21a.html.

Sidhant Misra, Line Roald, and Yeesian Ng. Learning for constrained optimization: Identifying optimal active constraint sets. *INFORMS Journal on Computing*, 34(1):463–480, 2022. doi: 10.1287/ijoc.2020.1037.

Bryan Palmintier, Carlos Mateo Domingo, Fernando Emilio Postigo Marcos, Tomas Gomez San Roman, Fernando de Cuadra, Nicolas Gensollen, Tarek Elgindy, and Pablo Duenas. SMART-DS synthetic electrical network data OpenDSS models for SFO, GSO, and AUS. Open Energy Data Initiative, National Renewable Energy Laboratory, 2020. URL https://data.openei.org/submissions/2981.

Kaizer Rahaman, Jyotirmoy V. Deshmukh, Ashish R. Hota, and Lars Lindemann. When environments shift: Safe planning with generative priors and robust conformal prediction. In *Proceedings of the 8th Annual Learning for Dynamics and Control Conference*, volume 331 of *Proceedings of Machine Learning Research*, pp. 601–623. PMLR, 2026. URL https://proceedings.mlr.press/v331/rahaman26a.html.

Akylas Stratigakos, Salvador Pineda, Juan Miguel Morales, and Georges Kariniotakis. Interpretable machine learning for DC optimal power flow with feasibility guarantees. *IEEE Transactions on Power Systems*, 2023. doi: 10.1109/TPWRS.2023.3333165.

Kegan J. Strawn, Nora Ayanian, and Lars Lindemann. Conformal predictive safety filter for RL controllers in dynamic environments. *IEEE Robotics and Automation Letters*, 2023. doi: 10.1109/LRA.2023.3322644. URL https://arxiv.org/abs/2306.02551.

Kim Peter Wabersich and Melanie N. Zeilinger. A predictive safety filter for learning-based control of constrained nonlinear dynamical systems. *Automatica*, 129:109597, 2021. doi: 10.1016/j.automatica.2021.109597.

Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 1658–1665, 2019. doi: 10.1609/aaai.v33i01.33011658.

Mei Yang, Gao Qiu, Junyong Liu, Youbo Liu, Tingjian Liu, Zhiyuan Tang, Lijie Ding, Yue Shui, and Kai Liu. Topology-transferable physics-guided graph neural network for real-time optimal power flow. *IEEE Transactions on Industrial Informatics*, 2024. doi: 10.1109/TII.2024.3398058.

Jin Zhang, Yuxiang Guan, Liang Che, and Mohammad Shahidehpour. EV charging command fast allocation approach based on deep reinforcement learning with safety modules. *IEEE Transactions on Smart Grid*, 15(1):757–769, 2024. doi: 10.1109/TSG.2023.3281782.

Mengfan Zhang, Guodong Guo, Tianyang Zhao, and Qianwen Xu. DNN assisted projection based deep reinforcement learning for safe control of distribution grids. *IEEE Transactions on Power Systems*, 2023. doi: 10.1109/TPWRS.2023.3336614.

Yuntian Zhang, Tiance Zhang, Gengyin Li, Ming Zhou, and Qinyong Zhou. Adaptive scheduling of distribution networks with massive DERs and EVs: A constraint-guided deep reinforcement learning approach. *International Journal of Electrical Power & Energy Systems*, 178:111815, 2026. doi: 10.1016/j.ijepes.2026.111815.

Min Zhou, Minghua Chen, and Steven H. Low. DeepOPF-FT: One deep neural network for multiple AC-OPF problems with flexible topology. *IEEE Transactions on Power Systems*, 38(1), 2023. doi: 10.1109/TPWRS.2022.3217407.

A PROOFS

A.1 CONDITIONAL ONE-SIDED PHASE-VOLTAGE FEASIBILITY

Restatement of Theorem 1. Let $P \geq 0$ and suppose the physical squared voltage is $u = u_0 - AP - r$. If $A \leq U$ elementwise, $\widehat{u}_0 - u_0 \leq \eta_0$, $r(P, s) \leq \rho(P, s)$, and

$$UP + \eta_0 + \rho(P, s) \leq \widehat{u}_0 - v_{\min}^2 \mathbf{1},$$

then $u \geq v_{\min}^2 \mathbf{1}$.

Proof. For every phase row j, elementwise upper coverage and nonnegative station power imply

$$(AP)_j = \sum_m A_{jm} P_m \leq \sum_m U_{jm} P_m = (UP)_j.$$

Therefore

$$\begin{aligned} u_j &= u_{0,j} - (AP)_j - r_j \\ &\geq \widehat{u}_{0,j} - \eta_{0,j} - (UP)_j - \rho_j(P, s) \\ &\geq v_{\min}^2. \end{aligned}$$

□

Only the upper coefficient residual matters because $P \geq 0$; a symmetric box is sufficient but can be wider than necessary. The theorem is independent of how the proposal was produced, but it proves none of its three error premises. Finite-point static confirmation audits $A \leq U$, while base and residual bounds require separate evidence; AC replay probes only their combined consequence on named trajectories.

If a solver-side post-check accepts the implemented shield inequality with violation at most $\tau_{\text{num}} \geq 0$, the same argument gives only $u_j \geq v_{\min}^2 - \tau_{\text{num}}$. Subtracting τ_{num} from the implemented right-hand side restores the exact mathematical endpoint; a tolerance-accepted result must not silently be reported as exact feasibility.

A.2 ROW-COMPRESSION SET INCLUSION

Proposition 2. *Let the nominal-centered element model be $H + C^E$ and let $C_{jm}^R = \max_\ell C_{j\ell}^E$. With identical right-hand side and non-voltage constraints, the feasible set using $H + C^R$ is a subset of the one using $H + C^E$ for $P \geq 0$.*

Proof. $C_{jm}^E \leq C_{jm}^R$ for every j, m. Multiplication by $P_m \geq 0$ and summation give $((H+C^E)P)_j \leq ((H + C^R)P)_j$. □

This proposition does not compare either nominal-centered set with InterpBox, whose center is the interpolated target matrix. It also does not guarantee a service difference: the removed constraints must become active for the observed proposal sequence.

A.3 DIAGNOSTICS FOR IDENTICAL ACTIONS

Proposition 3. *Let C be the common non-voltage feasible set and let f be strictly convex on C, with unique session-space minimizer p^0. For two voltage-feasible sets V_1 and V_2, if $p^0 \in V_1 \cap V_2$, then the unique minimizer of f over $C \cap V_i$ is p^0 for both $i \in \{1, 2\}$.*

Proof. Because $C \cap V_i \subseteq C$, no point in $C \cap V_i$ has objective below $f(p^0)$. Since p^0 belongs to each restricted set, it remains optimal; strict convexity makes it unique. □

Thus different voltage sets can produce identical actions without being equivalent. The diagnostic requires the counterfactual optimizer with voltage constraints removed. That action was not retained in the frozen v0.6 slot ledger, so this first proposition alone cannot assign a post hoc cause to all 34,560 equal-action slots.

Restatement of Proposition 1. Let C be convex, let f be strictly convex, and let $V_0 = \{p : B_0 p \leq d_0\}$. If $p^\star$ minimizes f over $C \cap V_0$ and $B_0 p^\star < d_0$ componentwise, then $p^\star$ is the unique minimizer of f over C. Hence it is also the unique minimizer over every $C \cap V_k$ for which $p^\star \in V_k$.

Proof. Suppose instead that some $q \in C$ has $f(q) < f(p^\star)$. Convexity of C puts $p_t = (1-t)p^\star + tq$ in C for $t \in [0,1]$. Because the finite affine rows are strictly satisfied at $p^\star$, continuity gives a sufficiently small $t > 0$ for which $B_0 p_t \leq d_0$, so $p_t \in C \cap V_0$. Convexity of f then gives

$$f(p_t) \leq (1-t)f(p^\star) + tf(q) < f(p^\star),$$

contradicting optimality on $C \cap V_0$. Strict convexity makes the optimizer over C unique. For any candidate V_k containing $p^\star$, restricting C to $C \cap V_k$ cannot introduce a lower-objective point, while $p^\star$ remains feasible; it is therefore the same unique optimizer. Candidate rows need not be strict, and no candidate optimization is required. □

Executable first-stop audit. Given a declared error contract, a reference shield, and candidate voltage models, the audit is: (1) return `premise_failed` when a required contract inequality fails on the declared audit set, or `unverified` when required error-bound evidence is missing; (2) solve the reference shield once and return `unavailable` for proven voltage infeasibility, or `unverified` when optimality or shared-problem evidence is missing; (3) if a verified reference optimum has strict voltage rows, evaluate every candidate row at the reference action and return `decision_invariant` when all candidates contain it; otherwise return `effect_required` and run the prespecified paired control comparison. These stopping outputs stop downstream attribution; they do not assert that an unevaluated downstream quantity is physically false.

This is a per-decision certificate. For a trajectory claim, start from the same session state and exogenous inputs. At every visited state, verify the shared proposal, objective, constraints, branch, reference optimum and candidate membership. Identical session actions, together with identical deterministic state updates, then give identical next states; induction proves trajectory invariance. A certificate at one time cannot replace these checks at later times. Zero-action voltage membership also does not establish feasibility of positive service lower bounds. Without the trajectory premises, report a local certificate or observed equality, not a certified closed-loop null.

Proof of Corollary 1. For a perturbed row, Hölder's inequality gives

$$\begin{aligned}(b_{0j} + \Delta b_j)^\top P^\star - (d_{0j} + \Delta d_j) &= -s_j + \Delta b_j^\top P^\star - \Delta d_j \\ &\leq -s_j + \|\Delta b_j\|_\infty \|P^\star\|_1 + |\Delta d_j| < 0.\end{aligned}$$

Here $P^\star = Sp^\star$; station rows are lifted by the same aggregation S. Thus $p^\star$ remains feasible for every perturbed voltage row. Proposition 1 already makes $p^\star$ the unique optimizer over C, so it is also the unique optimizer over the perturbed restricted set. With fixed right-hand sides and $P^\star \neq 0$, the common coefficient-only radius is $\min_j s_j / \|P^\star\|_1$; when $P^\star = 0$, coefficient changes do not alter the row value at the current decision.

Because the proposition proves optimality over all of C, the same $p^\star$ remains the unique optimizer over $C \cap \widetilde{V}$ for every relaxation $\widetilde{V} \supseteq V_0$. In particular, reducing a constant reserve while holding the coefficient rows fixed cannot change a reference optimizer with strict target-row slack, on the same branch and shared problem.

Applying the proposition separately to two shields with common C and f shows that strict voltage slack at both verified optimizers implies identical session actions. Station-action equality alone does not establish equal session allocations, next states, proposals or branches.

Frozen-ledger limitation. All 32,256 voltage-member slots per query-matched method have strict voltage slack, with minimum 0.003124 squared p.u. However, I has 3,570 service-relaxed grid fallbacks within those layers. The frozen ledger retains station powers and voltage-row summaries, not per-slot branch labels, session powers, raw session proposals, or optimality witnesses. The solver's `certified` status means numerical feasibility and solver success, not an exact optimality certificate. Thus the theorem cannot be applied to all ledger rows. Independent E/R diagnostics at I's actions do establish voltage membership on 32,256 slots. The five-method comparison of 161,280 stored pairs instead reads candidate answers; it has 0 station-action mismatches, but establishes

neither a forward screen nor saved solves. The earlier saved-solve interpretation is withdrawn. The reported slack-to-$\|P_i\|_1$ ratios are voltage-feasibility radii only; optimizer invariance requires verified shared- problem optimality. No missing session or branch fields were reconstructed by rerunning frozen experiments.

A.4 Independent forward-screen implementation and timing

The v0.12 development implementation is separate from the frozen controller. Its interface receives the shared session problem (proposal, positive diagonal weights, bounds, shared rows, state identifier and requested/fallback branch), reference action/status, and candidate rows. It accepts no candidate action, objective or solver status. It verifies the restricted analytic witness $p^\star = \mathrm{clip}(z, \ell, u)$ for the weighted quadratic box problem; if this point also satisfies the shared rows, it uniquely minimizes over C. The implementation abstains for general shared-row optima without this witness, even when a numerical solver reports success. This is a standard common- problem baseline, not an implementation of a general-purpose KKT verifier.

All objective, bound and shared-row arrays, the state identifier and branch must match. Reference and candidate rows are evaluated only at the reference point. Each multiply and accumulation is rounded outwards with `nextafter`; supplied absolute row-error allowances are subtracted from the slack bound and a 10^{-9} guard is required. This conclusion concerns represented IEEE-754 inputs without overflow, not unmodeled physical error. Near-boundary rows abstain, although the exact theorem permits candidate equality. Tests include feasible nonoptima, active rows, changed branches and objectives, missing evidence, zero power, overflow, and exact rational checks of outward row bounds.

Before evaluation we froze the implementation, protocol, and all synthetic inputs: seed 20260922; $(n, J) = (8, 32), (32, 256)$; four pressure factors $0.1, 0.7, 1.3, 2.3$; four instances per size/pressure; and six models including the reference. Proposals, boxes, weights, target matrices and right-hand sides are sampled without consulting any optimizer. Shared capacity bounds alternate between restrictive and loose values. These 32 problem families are mechanism tests, not independent application evidence or new EV confirmation. We save all 160 candidate predictions before any candidate solve, then compare with the unchanged projection solver at action/objective tolerances 10^{-7}. There are 25 sufficient certificates, zero false certificates, and zero action or objective discrepancy on the accepted pairs. Of 135 abstentions, 125 lack the analytic witness and ten have an unavailable reference.

Three pipelines use identical problems and solver settings: all six full projections; one reference plus checks and unscreened projections; and direct shared-box optimization plus row checks, without a reference requirement. Method order is seeded and randomized. First measured traversal and three repetitions are recorded separately; none uses solver-state warm starts or measures machine cold-start latency. Input generation/serialization and verification are reported separately. The table includes all 32 families, including nine infeasible model problems per traversal and no solver failures. Repetitions are systems measurements, not independent scientific samples.

Table 4: Development timing over three repeated traversals. Median/P95 are per-family times; calls and totals sum all repetitions within a size. Calls count `project`, including its existing analytic fast path, not only SLSQP runs.

(n, J)	Pipeline	Median/P95 (ms)	Calls	Total (ms)
(8,32)	Full solve	6.44/8.77	288	247.50
(8,32)	Reference + screen	6.50/9.59	243	269.13
(8,32)	Shared box	6.51/8.93	228	256.23
(32,256)	Full solve	91.13/120.47	288	3624.00
(32,256)	Reference + screen	90.65/118.02	258	3635.29
(32,256)	Shared box	88.36/123.59	237	3656.97

Per traversal, the reference screen calls projection 167 times versus 192 for full solving and 155 for the shared-box baseline. Across three repetitions, total times are 3904.42, 3871.50, and 3913.20 ms, respectively: a screened/full ratio of 1.0085 and screened/baseline ratio of 0.9978. Small single-machine differences do not establish a speed advantage. The full solver already has a box-feasible

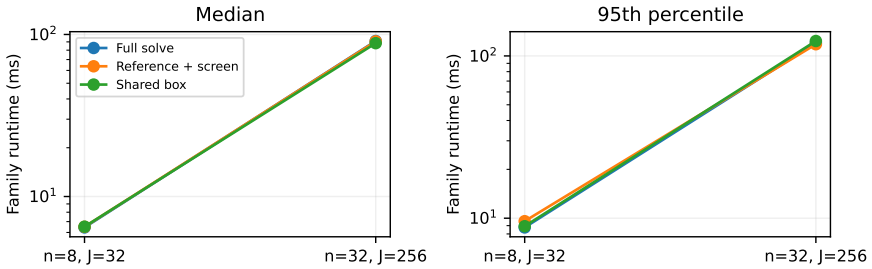


Figure 3: Two tested scales only; connected points do not establish asymptotic runtime. The restricted screen provides no demonstrated acceleration.

fast path, so fewer projection calls need not reduce time. We retain this negative result and do not tune thresholds on these inputs.

A.5 Zero-action reserve availability

Proposition 4. *The zero station-power action satisfies the voltage part of the reserved certificate if and only if* $\widehat{u}_0 \geq v_{\min}^2 \mathbf{1} + \eta_0 + \rho(0, s)$.

Proof. Set $P = 0$ in Eq. (7). The linear term UP vanishes, leaving $\eta_0 + \rho(0, s) \leq \widehat{u}_0 - v_{\min}^2 \mathbf{1}$, which is exactly the stated componentwise condition. □

This is stronger than nominal base feasibility $u_0 \geq v_{\min}^2 \mathbf{1}$. The v0.6 special case has $\widehat{u}_0 = u_0$, $\eta_0 = 0$, and $\rho = 0.005\mathbf{1}$. Failure of the condition means zero is not certified; without an additional sign premise on every upper-model coefficient, it does not by itself prove that every nonzero action is infeasible. In the T1/0.75 confirmation layer, the optimization reported infeasible contexts for all reserved methods and every slot. A scalar counterexample to "base feasible implies certificate available" is any $u_0 \in [v_{\min}^2, v_{\min}^2 + \rho(0, s))$ with positive reserve and $\eta_0 = 0$: the physical base meets the endpoint, while even $P = 0$ fails the certificate inequality.

A.6 One-step session viability

Lemma 1. *For remaining energy* $E \geq 0$, $n \geq 1$ *available slots including the current slot, duration* $\Delta > 0$, *and port cap* $\bar{p} \geq 0$, *current power* p *leaves a port-feasible completion schedule if and only if*

$$\frac{\max\{0, E - \Delta\bar{p}(n-1)\}}{\Delta} \leq p \leq \min\{\bar{p}, E/\Delta\}.$$

Proof. After the current slot, $E' = E - \Delta p$. Completion over the remaining $n - 1$ slots is possible exactly when $0 \leq E' \leq \Delta\bar{p}(n-1)$. Rearranging and intersecting with the current port cap yields the interval. □

Recursive completion additionally requires the joint grid/service set to stay nonempty at future slots. The implementation reports `infeasible_context` rather than claiming that the one-step interval proves future grid feasibility.

B Information and Query Ledger

Table 5: Information used by each layer. F is an evaluation reference rather than a query-matched deployable method.

Layer	Topology H	Three target matrices	Live base phases	Current true A
I: interpolation	optional	yes	yes	no
E/U: element baselines	yes	yes	yes	no
R: row baseline	yes	yes	yes	no
F: current matrix	no	no	yes	yes
N: nominal/no reserve	yes	no	yes	no
AC replay	full simulator	n/a	recomputed	full nonlinear solve

Each target matrix uses one no-EV base solve and eight single-station finite differences. Three matrices therefore cost 27 solves, independent of the number of protected phase rows. T1, T2, and T3 contain 1,809, 2,624, and 1,600 protected phase nodes, respectively, crossed with eight stations. The current no-EV phase-voltage snapshot is a separate online input. The candidate does not select actions using interior evaluation matrices, AC replay, or future sessions.

The source files retain original ratings, connections, phases, and kW/kvar values. The yearly loadshape attribute is removed mechanically for static load-multiplier experiments. Topology features and station selection use only the source-connected tree; OpenDSS replay still compiles the complete feeder.

For completeness, the nominal bus matrix is the radial LinDistFlow common-path resistance model. If $\mathcal{P}(b)$ is the root-to-bus path, b_m is station m's bus and S_{base} is in MVA, then

$$H_{(b,\phi),m} = \frac{2}{1000 S_{\text{base}}} \sum_{e \in \mathcal{P}(b) \cap \mathcal{P}(b_m)} r_e^{\text{pu}}.$$

Each protected phase ϕ at bus b receives the same nominal bus row; measured target matrices are phase-resolved. Thus H does not encode phase imbalance, although target queries and AC replay do. Its units, like A and U, are squared p.u./kW.

C Frozen Target, Date, and Workload Construction

Public demand templates. We use the City of Palo Alto ChargePoint charging-session dataset (City of Palo Alto, 2021), with station name, start/end timestamps and delivered energy. The input contains 259,415 records. We remove invalid timestamps/energies, nonpositive durations and energies at most 0.05 kWh; energy is clipped to 7 kW times duration rounded up to 15-minute slots. The eight station names with most retained records are selected (ties broken lexicographically), and days with at least 18 retained sessions are eligible. There are 111,533 retained records across 2,804 eligible dates. Historical delivered energy is a proxy for desired energy, not an observation of unmet demand. Station indices map deterministically to the eight topology-selected feeder buses; this is a synthetic coupling, not Palo Alto grid reconstruction.

For each date, arrivals are rounded down and departures up to 15-minute slots, with departure clipped at midnight (96 slots). Demand is again clipped to 7 kW times the remaining modeled dwell time. Each resulting session is replicated four times. At every station we add six 11.5-kWh sessions arriving at slot 68 and departing at slot 76 (17:00–19:00), giving 48 added sessions. This fixed stress transformation is applied identically to every method; the experiment is not an unmodified replay of historical demand.

Proposal, feedback, and service. At each slot, the common max-charge proposal at station m is $z_m = \min\{42, \sum_{i \in \mathcal{A}_m} \min(7, E_i/0.25)\}$ kW, where $\mathcal{A}_m$ contains active sessions and E_i is remaining kWh. EDF disaggregation gives the session-space proposal. The shield enforces 7-kW port, 42-kW station and 260-kW aggregate EV caps, along with the voltage and applicable service rows. Session energy is updated after each action. For each day, service is $100 \sum_{t,i} 0.25 p_{it} / \sum_i E_i^{\text{initial}}$: it

is an energy fraction, not the fraction of sessions fully charged. Background load is constant at the chosen scale throughout each 96-slot trajectory. After generating all actions, we replay that trace in OpenDSS. The feedback loop concerns session energy, not measured AC grid-state updates.

The first two SMART-DS substations used in earlier versions were classified as development data. T1–T3 were selected as the first lexicographic non-analysis feeder in each of the next three unused substations. Selection preceded target method results and used input-only compile, tree, station-count, and base checks.

Table 6: Frozen target inputs.

Key	Feeder suffix	Source buses	Protected phases
T1	`p1rhs2--p1rdt1528`	863	1,809
T2	`p1rhs3--p1rdt2378`	1,357	2,624
T3	`p1rhs4--p1rdt2300`	847	1,600

An exposure audit searched prior JSON and CSV outputs. Of 2,804 dates passing the common session filters, 373 appeared in prior result artifacts. The remaining 2,431 were eligible input candidates. Within each month, dates were ranked by the average percentile of total requested energy and peak active ports; the closest dates to one-third and two-thirds ranks were selected. This produced 24 dates spanning all months. Absence from result artifacts is a conservative exposure screen, not proof that no human ever inspected a date. The 24 selected dates also occupy 24 distinct ISO year–week blocks, so the predeclared date and week-block bootstrap resampling units coincide here.

D Base Cases and Reserve Availability

Table 7: No-EV minimum phase voltage on the five frozen closed-loop layers.

Target	0.25	0.375	0.50	0.625	0.75
T1	1.00525	0.99270	0.97969	0.96619	0.95214
T2	1.00932	0.99910	0.98866	0.97799	0.96707
T3	1.00824	0.99710	0.98566	0.97389	0.96178

All fifteen states are above the 0.95 p.u. operational endpoint. With $\varepsilon = 0.005$ in squared voltage, however, zero-action availability requires $v \geq \sqrt{0.95^2 + 0.005} \approx 0.95263$ p.u. at every protected phase. T1/0.75 fails this stronger test. This threshold was not used to remove or replace the frozen pressure layer.

E Static and Closed-Loop Detail

Static coverage. I, E, R, and the one-sided nominal baseline U each cover all 20 matrices on every target. I's worst positive residual-to-bound ratios are 0.395 on T1 and 0 on T2/T3; the latter means the interpolated center already lies above all checked entries. E and R have worst ratios near 0.975 because their 1.025 inflation is active. The zero-bound nominal model covers 0/20 matrices on every target.

Action equality. The final slot audit parses the stored eight-dimensional command vectors. I–E, I–R, E–R, and E–U each have 34,560/34,560 identical slots at tolerance 10^{-9} kW. The largest observed difference is exactly zero in the stored double-precision outputs. This justifies the descriptive null result but not a population equivalence claim. The same read-only comparison gives I–F equality on 34,560/34,560 slots even though F uses the current evaluation matrix. I and the no-reserve reference N differ on 1,650 slots, all inside T1/0.75; they differ on 0 slots outside that unavailable layer. Because N changes both the coefficient treatment and reserve, this localizes the difference but does not by itself identify a single causal component.

Table 8: Empirical diagnostics from retained primary outputs. Denominators differ; sampled coefficient coverage and voltage membership do not establish the full error premise. The optimizer-level certificate remains unverified.

Evidence	Frozen diagnostic	Outcome
Sampled coefficients	Complete-matrix upper coverage, not the full error premise	60/60 for I, E, and R
Voltage membership	Zero action satisfies the declared reserved voltage rows	14/15 target–pressure layers
Observed equality	Stored station actions; optimizer certificate unverified	I–E/R/F: 0/34,560 actions differ; I–N: 1,650, all in the unavailable layer
Service effect	Date-paired delivered-energy fraction	I–E $= 0.00$ pp; E–R $= 0.00$ pp

Safety and readback. I has zero AC-violating trajectories and zero true-linear violating slots in 360 trajectories. Its minimum replay voltage is 0.952140 p.u., attained when T1/0.75 deploys no charging. Across all methods, the maximum station-power readback error is 2.81×10^{-9} kW. Solver-failure slots are zero.

Primary inference. For each of 24 dates, service is averaged across the three main pressures within target and then equally across T1–T3. Both paired differences are identically zero. Their empirical means, medians, and 10,000-replicate paired date-bootstrap intervals are therefore zero; the Holm-adjusted Wilcoxon p-values are one. Slots, phase rows, pressure layers, and targets are not treated as extra demand samples.

Mechanism falsification. The frozen action-aligned model predicts nonzero E–R service gaps in 24 of 216 main-pressure target-date rows, but every observed gap is zero. Its mean absolute error is 0.2376 percentage points. Spearman correlation is undefined because the outcome has no variation. The mechanism therefore does not beat the predeclared simple controls and is not presented as a transferable law.

F Protocol Evolution and Evidence Roles

The retained history explains why the final method is simpler than the original learned proposal.

v0.3: learned row envelopes. On 33-, 69-, and IEEE 123-bus models, a learned row scale and a direct row baseline were compared with equal target-query counts. Corrected replay was AC feasible for every method, while the learned robust shield lost service. This motivated a decomposition of conservatism rather than further fitting.

v0.4–v0.5: phase and station structure. Station-specific endpoint boxes exposed row-max slack; phase-resolved voltage checks exposed failures hidden by bus averages. A reserved phase-by-station box covered unseen IEEE 37 and SMART-DS-2 matrices, but their closed-loop conditions were non-discriminative. IEEE 34 remained a development-only boundary case with a large row-compression service loss.

v0.6: query-matched center and independent mechanism test. Four seen feeders selected interpolation, the scale floor, $q = 0.0203262$, and the frozen mechanism mapping. Three new feeders, 24 input-selected dates, and five pressures were frozen before method outcomes. Static confirmation passed; closed-loop service and mechanism effects did not. The paper retains both.

v0.11: prospectively frozen external-source failure. After the primary audit was complete, the same three-query rule, q, 20 interior scales, and stop criterion were frozen for the unbalanced EPRI IEEE 8500-node feeder. Its 4,876-bus source component, 8,531 energized phase rows, eight topology-selected 12.47 kV stations, and 24 fresh demand dates were frozen before sensitivities. All 21

input-only base checks were feasible (minimum 0.95090 p.u.), but I covered only 3/20 matrices and the protocol stopped before closed loop. This is a failure at the first funnel stage, not an additional confirmation or a reason to retune q on the target.

F.1 External-model stress test details

The source was the official OpenDSS `Version8/Distrib/IEEETestCases/8500-Node` directory (repository page revision 4176), with the BSD-style EPRI license and every used file hash retained. The unbalanced master keeps regulator and capacitor controls. Each scale settles those controls, freezes them, then measures an $8{,}531 \times 8$ phase–station matrix with constant-power 3 kW perturbations. Table 9 reports every prespecified method; no row is selected post hoc.

Table 9: Prospectively frozen IEEE 8500 external-source stress test. The frozen I criterion was 20/20 complete upper-matrix coverage; it failed, and no closed loop was run.

Method	Matrices	Entry coverage	Worst residual/bound
I: interpolated center	3/20	89.305%	5.708
E: element box	8/20	97.294%	1.043
R: row-max box	15/20	99.877%	1.023
U: one-sided element	8/20	97.294%	1.043
N: nominal, no bound	0/20	0.000%	–

I's worst positive residual was 5.708 times its frozen bound. A post-confirmation diagnostic then held the left endpoint's tap/capacitor mode fixed throughout each interval. Coverage improved from 3/20 to 7/20, but the worst residual remained 2.365 times the bound. Thus discrete mode changes contribute, but do not by themselves explain the cross-source failure. This counterfactual is development-only and cannot revise the frozen FAIL.

F.2 Cross-source synthesis of frozen protocols

Table 10 is a post hoc synthesis of two retained frozen studies, not a pooled statistical analysis. The v0.5 method bounds endpoint error around a topology-derived nominal center; it is therefore not the v0.6 interpolation method. Its six demand dates were reused from an older protocol, and only service aggregates—not station action vectors—were retained. We may conclude that the empirical pattern "complete static coverage without downstream discrimination" also occurred on an IEEE test feeder, but not that v0.6 interpolation was independently confirmed across model sources or that the two protocols estimate one common effect.

Table 10: Read-only synthesis of independently frozen protocols. Denominators retain their original meanings and are not pooled as independent samples.

Study	target sources	frozen upper model	static coverage	control result
v0.5	IEEE 37 + SMART-DS	nominal-centered phase-element endpoint envelope	40/40	0/24 service-separated rows (5 variants)
v0.6	SMART-DS (three feeders)	query-interpolated one-sided upper model	60/60	0/34560 action-separated slots (I–E; I–R)

G Post hoc Metric and Strong-Baseline Development

The frozen v0.6 summary reports the mean additive margin around each method's own center. Since those centers differ, the percentage reduction is useful for diagnosing calibration residuals

but is not invariant to the choice of center. We therefore added a descriptive final-model quantity, $\mathrm{mean}[U - A]_+$, using only the saved v0.6 static summaries:

Table 11: Post hoc re-expression of the frozen static result. Both I and E cover 20/20 matrices on every target.

Target	Added-margin reduction	Final-overstatement reduction
T1	97.35%	61.96%
T2	97.52%	38.74%
T3	97.54%	47.79%

No action-weighted voltage-budget value is reported because the saved static aggregate does not retain the complete matrices and a common reference action needed to reconstruct it. We do not rerun or alter the frozen confirmation to fill that gap.

Separately, four already-seen feeders were used for an equal-budget baseline audit. I, Qmax (the elementwise maximum of adjacent queries), and Qnear (the nearest query) each use the same three complete matrices, endpoint-derived normalization, 1.025 inflation, and a method-specific margin fitted on the other three feeders. Their held-out complete-matrix counts are 80/80, 62/80, and 77/80; zero-margin I0 covers 40/80. Because neither simple baseline covers all held-out matrices, no strong safety baseline is declared frozen for a new confirmation. These data are development evidence only and were obtained after the v0.6 confirmation was known.

H Development Constraint-Activation Gate

We implemented the proposed input-side gate on the four already-seen feeders. For each load scale from 0.25 to 0.75 in increments of 0.0125, the gate uses only the no-EV base voltage and a common capacity-only EDF trajectory. A strict-interior scale qualifies when the base case exceeds 0.95 p.u., the zero action remains voltage-certifiable with the fixed 0.005 squared-voltage reserve, and nominal voltage-budget occupancy reaches 90% on at least 20% of reference demand slots. Candidate shield actions, service differences, interior coefficient errors, and AC replay outcomes are not gate inputs.

Table 12: Development-only activation map. Base and zero-certificate columns give passing scales out of 41; Gate counts qualifying strict interiors; Selected takes at most two per query bracket.

Seen feeder	Base	Zero-cert.	Gate	Selected
IEEE 34	37/41	35/41	3	2
IEEE 37	34/41	32/41	1	1
SMART-DS-1	41/41	41/41	0	0
SMART-DS-2	41/41	41/41	0	0

No feeder provides the planned four selected pressures, and both seen SMART-DS feeders provide none. Thus the prespecified continuation condition for a method-superiority confirmation is not met. The result diagnoses the current experimental design; it neither establishes a population frequency nor uses target outcomes to manufacture a favorable stress case. The development implementation obtains the base-voltage field through the existing sensitivity routine, although the gate reads only that field. Any future unseen-target selector must instead use a base-only solve so the scan cannot expose interior matrices before selection is frozen.

I Availability-Reserve Development Gate

After v0.6 and v0.7 were known, we implemented the separated contract in Eq. (6) as a new development-only branch. The residual sample is

$$r(P, s) = u_0(s) - A(s)P - u(P, s),$$

where $A(s)$ is independently measured at the development state. Thus the coefficient error handled by $U - A$ is not counted again in r. We retain the safety-relevant positive side of r and the positive base overestimate $[\widehat{u}_0 - u_0]_+$. Empirical maxima are multiplied by 1.025; they are finite development bounds, not confidence or conformal bounds.

The ledger uses four already-seen feeders, load scales 0.40, 0.60, and 0.675, six reused dates, 96 slots, and two information settings, for 13,824 rows. "Online base" sets $\widehat{u}_0 = u_0$; "interpolated base" obtains $\widehat{u}_0$ from the three queried base states. Within IEEE and SMART-DS separately, each target is audited using a model fitted only on the other same-source feeder. The global baseline is one constant. The state model has one break fixed at load scale 0.55. All 12/12 development upper matrices cover their measured counterparts.

Table 13: Development-only joint coverage of base and nonlinear-residual error. Counts are covered slot–states out of 1,728; they are not independent demand samples.

	Online base		Interpolated base	
Feeder	Global	Piecewise	Global	Piecewise
IEEE 34	1419/1728	1412/1728	1419/1728	942/1728
IEEE 37	1728/1728	1728/1728	1152/1728	576/1728
SMART-DS-1	1728/1728	1728/1728	1728/1728	1728/1728
SMART-DS-2	1527/1728	1521/1728	1527/1728	1521/1728

The piecewise rule improves coverage in 0/8 comparisons, while the global constant remains incomplete in 5/8. Interpolated base voltage is particularly brittle across IEEE regulator/control changes. The prespecified state variable therefore fails to explain transferable reserve variation. We stop before closed-loop method comparison and do not create a v0.8 confirmation set. A subsequent action-dependent investigation uses an explicit new development protocol (Appendix J); it does not change this stopping decision or convert these seen feeders into confirmation.

J Action-Dependent Reserve Development

After inspecting the v0.8 failures, we froze three further reserve forms and their comparison rule before collecting new full-vector replays. This is post hoc development on the same four seen feeders and six dates, not a new independent test. The 6,912 physical slot–state samples each have a saved command and actual eight-station action, base and action voltage for every protected phase, and an independently measured sensitivity matrix. Fresh replay matches the previous ideal-base scalar residuals within 3.1×10^{-16} squared p.u. and the action sums within 5.7×10^{-14} kW. All 12 measured development sensitivity matrices are upper-covered. We use the command residual $[\max_j\{u_{0j} - (AP_{\text{command}})_j - u_j(P_{\text{actual}})\}]_+$, which includes the recorded command/readback discrepancy; its maximum absolute station-power error is below 3.7×10^{-9} kW. Evaluation matrices and action voltages do not enter the predictor.

Let $d(P) = \max_j[(U_+)P]_j$, $b = \min_j u_{0j}$, and $U_+ = \max(U, 0)$ elementwise. Besides the old global and load-binned constants, we fit $\rho(P) = \epsilon + \alpha f(P)$ for $f \in \{\|P\|_1^2, d(P)^2, d(P)^2/(b - d(P))\}$, the last only for $d(P) < b$. With nonnegative P and fixed queried U and base state, each f is convex on its declared domain: d is a maximum of nonnegative linear forms, and t^2 and $t^2/(b - t)$ are convex and nondecreasing for $0 \le t < b$. The intercept is 1.025 times the larger of the training zero-action residual and 10^{-10}; α is 1.025 times the largest remaining training residual-to-feature ratio. No ratio is fitted on the target feeder. Invalid domains, zero-feature failures, and all uncovered samples remain failures.

Table 14: Exposed-data, reserve-only leave-one-feeder-out residual coverage with the other same-source feeder used for calibration; ideal online base. Each cell is covered slots out of 1,728, with dates and slots shared across models. SMART-DS-1 and SMART-DS-2 denote p1rdt5663 and p1rdt2397.

Feeder	Global	Binned	$\|P\|_1^2$	d^2	$d^2/(b-d)$
IEEE 34	1419/1728	1412/1728	1126/1728	1728/1728	1728/1728
IEEE 37	1728/1728	1728/1728	1728/1728	390/1728	390/1728
SMART-DS-1	1728/1728	1728/1728	1728/1728	1728/1728	1728/1728
SMART-DS-2	1527/1728	1521/1728	1639/1728	1728/1728	1728/1728

Under three fitting rules (other same-source feeder, all three other feeders, and two other-source feeders), the global and load-binned constants cover all samples in 7/12 folds each; action power squared also covers 7/12, and both voltage-drop forms cover 9/12. Folds reuse the same 6,912 physical samples and are not independent replicates. In the same-source comparison, drop-normalized forms cover IEEE 34 and SMART-DS-2 completely but only 390/1,728 IEEE 37 samples. The maximum positive residual divided by d^2 is 5.31 on IEEE 37 versus 0.397 on IEEE 34, explaining why calibration on IEEE 34 misses IEEE 37. The improved coverage does not translate into greater common-action availability: on IEEE 34 the drop forms certify 1,355/1,728 observed actions against 1,380/1,728 for the global constant; no fold gains a certifiable observed action. These are offline voltage checks, not optimized closed-loop service. All three candidates fail the fixed 12-fold coverage and availability gate, so no further confirmation or method-benefit claim follows. The unchanged zero-action state counts also show that the available headroom in these development states did not recreate the fixed-reserve availability cliff of the frozen T1/0.75 case.

K Precheck Repairs and Failure Accounting

The first v0.6 closed-loop attempt stopped before T1/0.75 produced a trajectory. The harness asserted that every deployed action must lie in the reserve-shrunk set, even when the controller had explicitly labeled the state uncertified. Because zero itself fails Eq. (9), the assertion confused a scientific availability boundary with an execution failure. Partial atomic checkpoints were archived, the assertion was restricted to actions labeled certified, and the active result directory was restarted empty.

A second pre-analysis schema review stopped the processes during the first pressure layer, before any checkpoint was written. The slot schema had total power and maximum readback error but lacked the full commanded and actual eight-station vectors needed to audit exact action equality. The vectors were added and all three targets restarted from empty fragments. Neither repair changed targets, dates, pressures, methods, parameters, or statistical rules. Script hashes and archived fragments are retained in the precheck failure log.

The final failure ledger records every trajectory with fallback, infeasible, or readback diagnostics. Some fallback/infeasible session states are shared capacity or deadline behavior rather than voltage-certificate failure; the T1/0.75 layer is separated by its zero grid-certificate rate and zero service.

L Legacy Implementation Contracts

The research implementation retains earlier shield components covered by executable tests even though they are not contributions of the v0.6 method. We state their contracts here to keep the proof–code audit explicit.

Feasibility and service contracts. *Conditional robust row feasibility.* For the earlier symmetric row model, the deployed left-hand side is $\hat{a}_j^\top p + \rho_j \|p\|_1$. The proof is the ℓ_∞/ℓ_1 dual inequality. The v0.6 result instead exploits nonnegative charging and the one-sided upper matrix in the main paper. *Exact one-step service viability.* In the indexed implementation, the lower action is $\ell_i = \frac{\max\{0, E_i - \Delta u_i(n_i-1)\}}{\Delta}$; together with the upper port/energy limit, this interval is necessary and sufficient for a port-feasible remaining schedule. *Conditional recursive completion.* Induction additionally assumes that the combined robust grid/service set is nonempty at every future slot. Present

feasibility alone does not imply this future nonemptiness. *Learned row-wise versus learned global inflation.* Under their shared nominal rows and right-hand sides, when the global radius dominates every row radius, the globally inflated robust feasible set is a subset of the row-wise set. Set inclusion does not order projection service or objectives across different proposals.

Status and scope. *Certificate status.* The controller has three states: `certified`, `infeasible_context`, and `solver_failure`. Either failure state deploys zero without relabeling the outcome as safe. *Target-envelope boundary.* The earlier endpoint envelopes are interval rules, not proofs of between-point dominance; in particular, they do not establish dominance between probes. A changed topology signature invalidates target adaptation; the v0.6 interpolation claim is likewise restricted to the frozen topology and the tested in-range points.

M OpenDSS Power and Connection Semantics

Background demand is varied with `LoadMult`. Added EV loads use `Status=Fixed`, and `kvar=0` is reset after each kW setter because OpenDSS can otherwise retain a default power factor when a zero-kW load is activated. Finite differences divide by measured terminal kW. Replay reads all eight station powers after every solve.

The first v0.8 residual run exposed another simulator semantic: an added constant-power load without an explicit `Vminpu` inherits a 0.95 p.u. threshold and is silently voltage-de-rated below that level. One 41.824 kW command read back as 41.588 kW. The run stopped, no partial result was promoted, and v0.8 residual instrumentation now sets `Vminpu=0.5` on added EV loads while retaining strict terminal-power readback. This prevents simulator load shedding from being counted as safe control; it is a development instrumentation change, not a rewrite of frozen v0.6 trajectories.

The v0.11 preflight exposed the symmetric high-voltage case. Two added `Model=1` loads at base voltages above OpenDSS's default `Vmaxpu=1.05` entered the constant-impedance region, so a 3 kW command read back as 3.02246 kW and the run stopped before any sensitivity matrix or method result. The target-specific test loads were then frozen with `Vminpu=0,Vmaxpu=2`; the original failure and code hashes remain in the artifact. This enforces the declared constant-power perturbation rather than using voltage-dependent de-rating to improve coverage.

The v0.6 SMART-DS feeders use energized phase-to-neutral magnitudes and balanced wye charging loads. Earlier IEEE 37 experiments used line-to-line magnitudes and delta charging; mixing those definitions would produce an apparently phase-resolved certificate for the wrong physical quantity. OpenDSSDirect holds process-global circuit state, so each requested operating point is compiled after topology loading before voltages are read.

N Reproducibility and Claim Boundary

The frozen protocol, inputs, dates, source, station order, topology, code, static/trajectory/slot rows, tables, figures, and claim macros carry SHA-256 records. Static confirmation retains all 20 points per target; closed-loop outputs retain every trajectory and slot with command/readback vectors, service, AC voltage, violations, certificate status, intervention, and runtime.

These boundaries are part of the result. Coverage is a deterministic in-range count, not a probability or arbitrary-topology theorem; feasibility is conditional on an unproved-global residual premise; and three probes are full matrices plus live no-EV voltages. The independent closed loop has zero method contrast; zero violations with zero service is not useful; and base feasibility differs from fixed-reserve availability. The state-binned reserve fails its transfer gate and has no confirmation claim. There is no field hardware, privacy guarantee, or deployment authorization.