Certified Single-Level Reformulations for Tri-Level Cyber-Physical Grid Security

Debraj Banerjee
Department of Electrical Engineering
Indian Institute of Science
Bengaluru, India
debrajb@iisc.ac.in

Gurunath Gurrala
Department of Electrical Engineering
Indian Institute of Science
Bengaluru, India
gurunath@iisc.ac.in

Kunal N. Chaudhury
Department of Electrical Engineering
Indian Institute of Science
Bengaluru, India
kunal@iisc.ac.in

Abstract

Modern power grids are increasingly exposed to coordinated cyber-physical disruptions, where the outage of a small number of transmission components can trigger severe post-contingency load shedding. Designing proactive *security defense* against such adversarial contingencies is naturally modeled as a tri-level Defender-Attacker-Operator (DAO) problem: the defender allocates limited hardening resources, the attacker selects worst-case line outages, and the operator solves a post-attack DC/AC-optimal power flow (OPF). However, the resulting tri-level program is computationally prohibitive even for moderate network sizes.

We develop a certified four-step reformulation pipeline that converts this intractable tri-level DAO problem into a tractable single-level optimization, yielding a Single-Level Mixed-Integer Linear Program (SL-MILP) under DC-OPF and a Single-Level Mixed-Integer Conic Program (SL-MICP) under a convex SOCP relaxation of AC-OPF. The construction is exact except for one controlled relaxation of the attacker's binary decisions. It provides two complementary guarantees: the single-level value S_{SL} is an a-priori conservative bound, while one post-hoc binary attacker-operator solve at the returned defense gives the exact realized certificate $S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \leqslant S_{\mathrm{SL}}$. Experiments on IEEE power systems up to 118 buses for DC and 57 buses for AC quantify both certificate gaps and runtime while demonstrating substantial computational gains.

1 Introduction

Power transmission grids constitute critical cyber-physical infrastructure, where localized failures or coordinated malicious disruptions can propagate rapidly into large-scale service interruptions and substantial emergency load shedding. Hence, reliable operation of electric power grids is of paramount societal importance. Transmission lines are the most physically exposed assets: they span large geographic areas, can be damaged with relatively modest effort, and their failure can cascade into widespread outages [1, 2]. Protecting a grid under a limited budget, therefore requires deciding *which* lines to harden (protect) so that, even after an intelligent adversary attacks the remaining lines, the grid operator can maintain service. This hierarchical decision problem is naturally captured by a *Tri-Level Defender-Attacker-Operator (DAO)* model: the defender moves first (strategic, long-term

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

protection), the attacker responds (worst-case interdiction), and the operator re-dispatches generation to minimize load shedding [3].

Prior work on tri-level DC-OPF models. The bilevel interdiction problem for power grids was pioneered by Salmerón, Wood, and Baldick [1], who formulated a max-min attacker problem under DC power flow. Motto, Arroyo, and Galiana [4] analysed the same setting with a mixed-integer LP. The introduction of a third (defender) level, yielding the min-max-min DAO structure, was proposed by Brown *et al.* [3] and Alguacil *et al.* [5]; an implicit enumeration algorithm that gives the exact tri-level solution under DC power flow was established in [6, 3]. Iterative decomposition methods based on Benders cuts and Column-and-Constraint Generation (C&CG) [7, 8, 9, 10] achieve better practical scalability at the cost of approximate solutions or multiple MILP iterations. Extensions to cyber-physical attacks [11, 12, 13] and resilience-oriented models [14, 15] also operate predominantly under the DC power flow approximation.

The gap: tri-level AC-OPF models. Despite the widespread use of DC approximations, the *AC power flow* equations govern physical reality. AC-OPF captures reactive power, voltage magnitudes, and ohmic losses, all of which matter for realistic load-shedding estimation, especially under stressed post-attack conditions. Solving exact nonconvex AC-OPF is NP-hard in general [16, 17]; the standard engineering practice is to use a *conic (SOCP) relaxation* that is tight for many practical instances [18, 19, 20]. *To our knowledge, no prior work has derived a tractable single-level reformulation of the tri-level DAO problem with an AC-OPF (or its conic relaxation) as the inner problem, nor provided accompanying mathematical guarantees.* This gap is the central motivation of the present paper. The main contributions are listed below.

1. **Single-level reformulation.** We develop a four-step reformulation pipeline that converts the tri-level Defender-Attacker-Operator problem into a tractable single-level optimization, yielding an SL-MILP under DC-OPF and an SL-MICP under AC-OPF. A visual illustration of the pipeline is depicted in Figure 3.
2. **Certified robustness guarantees.** We prove that the single-level objective provides an a-priori conservative upper bound (Theorem 5.2). For the returned defense, one exact binary attacker-operator solve provides the tighter post-hoc value $S_{\mathrm{true}}(d^*_{\mathrm{SL}})$ and decomposes the total gap into conservatism and defense-optimality components. No feasible binary attack can exceed either applicable certificate (Theorem 5.4).
3. **Unified theory and validation.** Our framework applies to LP-based DC and SOCP-relaxed AC operator models. The guarantee is unconditional for DC and model-relative for AC, i.e., certified within the SOCP relaxation rather than the original nonconvex AC equations. Experiments on IEEE benchmark grids quantify the relaxation gaps and demonstrate substantial runtime gains with strong worst-case protection.

2 Optimal Power Flow: DC and AC

2.1 DC Optimal Power Flow

Consider a power grid graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $n = |\mathcal{V}|$ buses and $m = |\mathcal{E}|$ transmission lines. The DC-OPF model linearizes the full AC equations by assuming flat voltage profiles ($|v_i| \approx 1$ p.u., small angle differences) and lossless lines. Let $x := (g, \theta, s) \in \mathbb{R}^{3n}$ collect all operator variables, where $g_i \in [0, g_i^{\max}]$ is generation, $\theta_i \in [\theta_i^{\min}, \theta_i^{\max}]$ is phase angle, and $s_i \in [0, c_i]$ is load shedding at bus $i \in \mathcal{V}$, with c_i the power demand. We write the box constraints compactly as $x^{\min} \leqslant x \leqslant x^{\max}$, where $x^{\min} := (0, \theta^{\min}, 0)$ and $x^{\max} := (g^{\max}, \theta^{\max}, c)$. The nodal-balance matrix $B(z) \in \mathbb{R}^{n \times 3n}$ is defined by

$$[B(z)\, x]_i \;:=\; g_i + s_i - \sum_{j \in \mathcal{V}} Y_{ij} z_{ij} (\theta_i - \theta_j), \quad \forall\, i \in \mathcal{V}, \tag{1}$$

so (applying KCL [16]) that nodal balance reads $B(z)x = c$. The line-flow matrix $F(z) \in \mathbb{R}^{m \times 3n}$ is defined by $[F(z)\, x]_{ij} := Y_{ij} z_{ij} (\theta_i - \theta_j)$, so that the flow bound constraints read $-f \leqslant F(z)x \leqslant f$. The vector $e_s \in \mathbb{R}^{3n}$ selects the shedding components of x, i.e. $(e_s)_k = 1$ if component k corresponds to some s_i, and zero otherwise. Figure 1 illustrates the generic layout and the graphical model of the WSCC 9-bus power system [21].

The operator minimizes total shedding $S = e_s^\top x$, subject to nodal power balance (2b) and line-flow (2c) limits:

$$\mathcal{P}^*_{\mathrm{DC}}(z) := \min_x \; e_s^\top x \tag{DC-OPF}$$

$$\text{s.t. } x^{\min} \leqslant x \leqslant x^{\max}, \tag{2a}$$

$$B(z)\, x = c, \tag{2b}$$

$$-f \leqslant F(z)\, x \leqslant f. \tag{2c}$$

Here $z_{ij} \in \{0,1\}$ encodes line availability. We assume $c, g^{\max}, f \geqslant 0$ and $\theta^{\min} \leqslant 0 \leqslant \theta^{\max}$. Each undirected line is assigned one orientation in $\mathcal{E}$; nodal sums include all incident lines with the corresponding incidence signs. For fixed z, (DC-OPF) is a *Linear Program* (LP), solvable efficiently at scale by interior-point or simplex methods [22]. Load-shedding slack variables $s_i \geqslant 0$ ensure feasibility and boundedness, and we get the following result.

Lemma 2.1. *Strong duality holds for* (DC-OPF)

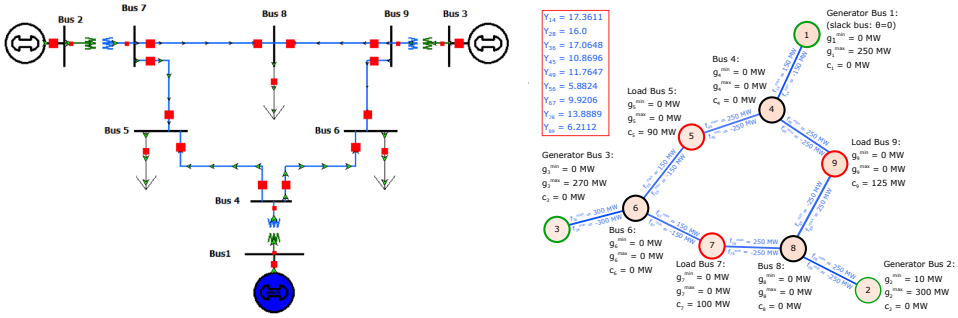


Figure 1: Illustration of the layout (left) and the graphical model (right) for the WSCC 9-bus power grid with DC parameters. Different buses are colour-coded in red (load bus), green (generator bus) and black (connection bus).

2.2 AC Optimal Power Flow and Conic Relaxation

AC-OPF introduces voltage magnitudes $|v_i|^2 = u_i$, and the phasor products $\Phi_{ij} = |v_i||v_j|\cos(\theta_i - \theta_j)$ and $\Psi_{ij} = |v_i||v_j|\sin(\theta_i - \theta_j)$. Active and reactive nodal balance read

$$g_i^P + s_i^P - \textstyle\sum_j z_{ij}\big[G_{ij}(u_i - \Phi_{ij}) - B_{ij}\Psi_{ij}\big] = p_i^d, \tag{3a}$$

$$g_i^Q + s_i^Q - \textstyle\sum_j z_{ij}\big[-B_{ij}(u_i - \Phi_{ij}) - G_{ij}\Psi_{ij}\big] = q_i^d, \tag{3b}$$

and the active, reactive power-flow limits

$$-|P_{ij}| \leqslant z_{ij}\big[G_{ij}(u_i - \Phi_{ij}) - B_{ij}\Psi_{ij}\big] \leqslant |P_{ij}|, \tag{4a}$$

$$-|Q_{ij}| \leqslant z_{ij}\big[-B_{ij}(u_i - \Phi_{ij}) - G_{ij}\Psi_{ij}\big] \leqslant |Q_{ij}|. \tag{4b}$$

where (G_{ij}, B_{ij}) are the series conductance/susceptance, with $\Phi_{ji} = \Phi_{ij}$ and $\Psi_{ji} = -\Psi_{ij}$; the flow limits apply at both line ends. Here (P_{ij}, Q_{ij}) are max active/reactive power-flow limits of line (ij) and $(g_i^P, g_i^Q), (s_i^P, s_i^Q), (p_i^d, q_i^d)$ are active/reactive generation/shedding/demands respectively. A necessary exact power-flow constraint is $\Phi_{ij}^2 + \Psi_{ij}^2 = u_i u_j$, which is *nonconvex*: the feasible set is a non-convex quadric. Thus solving AC-OPF exactly is NP-hard in general [16, 23], and even approximating it remains computationally challenging.

Relaxation landscape. Three families of relaxations are prevalent in the literature [17, 19]: (i) *piecewise-linear* approximations of trigonometric terms, which yield MILPs but introduce discretization error; (ii) *semidefinite (SDP)* relaxations based on a rank-1 matrix lift, which are tight in many instances [18] but scale poorly; (iii) *second-order cone (SOCP)* relaxations that drop the equality $\Phi_{ij}^2 + \Psi_{ij}^2 = u_i u_j$ to the inequality $\Phi_{ij}^2 + \Psi_{ij}^2 \leqslant u_i u_j$, yielding a convex conic program solvable in polynomial time [20, 24].

Our conic (SOCP) relaxation. We adopt the SOCP relaxation throughout the AC development. Accordingly, *AC* below refers to this convex operator model unless the original nonconvex AC-OPF

is named explicitly. Let $x = (g^P, g^Q, s^P, s^Q, u, \Phi, \Psi)$ collect all operator variables and let t be an epigraph variable for generation cost. The inner operator problem for fixed line availability z is:

$$\mathcal{P}^*_{AC}(z) := \min_{x,t} \; c^\top x + t \tag{AC-OPF}$$

$$\text{s.t. } x^{\min} \leqslant x \leqslant x^{\max}, \tag{5a}$$

$$A(z)x = b, \tag{5b}$$

$$-f \leqslant H(z)x \leqslant f, \tag{5c}$$

$$E_{ij}x \in \mathcal{K}_4 \; \forall (ij) \in \mathcal{E}, \tag{5d}$$

$$[2J^{1/2}x \quad t-1 \quad t+1]^\top \in \mathcal{K}_{N+2}, \tag{5e}$$

where $\mathcal{K}_n$ is the Lorentz (second-order) cone of dimension $n \geqslant 2$ defined as

$$\mathcal{K}_n := \{v \in \mathbb{R}^n : \|(v_1, \ldots, v_{n-1})\| \leqslant v_n\}. \tag{6}$$

$A(z)$ encodes the nodal balances (3a)–(3b), $H(z)$ encodes line-flow limits (4a)–(4b). Constraints (5d) enforces the rotated SOC constraint $\Phi_{ij}^2 + \Psi_{ij}^2 \leqslant u_i u_j$, with $E_{ij}x = (2\Phi_{ij}, 2\Psi_{ij}, u_i - u_j, u_i + u_j)^\top$ and $u_i, u_j \geqslant 0$. Constraint (5e) encodes the quadratic generation cost $x^\top Jx$ via a second-order cone epigraph, with $J = J^\top \succeq 0$. This is a *conic program* (SOCP) for each fixed z. Similar to Lemma 2.1, under Slater's condition we have

Lemma 2.2. *Strong duality holds for* (AC-OPF).

Significance of the objective. The DC objective is pure load shedding, $e_s^\top x$. In contrast, the AC operator objective (AC-OPF) minimizes a weighted combination of load shedding and generation cost (we use suitable scaling to control their relative priority). Thus, the AC certificate applies to this combined objective; a pure load-shedding certificate is obtained by setting the generation-cost coefficients to zero. The generation cost-epigraph formulation allows the quadratic generation cost $x^\top Jx = {g^P}^\top C^{(2)} g^P$ (with non-negative diagonal matrix $C^{(2)}$ for quadratic generation cost coefficients) to be handled exactly within a conic framework, without piecewise-linear approximation. At optimality, $t = x^\top Jx$ recovers the true generation cost.

3 The Tri-Level DAO Model

3.1 Interdiction Variables

Table 1 summarises the interdiction variables. For each line $(ij) \in \mathcal{E}$, $d_{ij} \in \{0, 1\}$ is the *defender*'s binary protection decision and $a_{ij} \in \{0, 1\}$ is the *attacker*'s binary attack decision. Post-event line availability is

$$z_{ij} = 1 - a_{ij}(1 - d_{ij}) \in \{0, 1\}, \tag{7}$$

so a protected line ($d_{ij} = 1$) cannot be disabled.

Tri-level DAO problem. Combining the three agents' objectives yields the min-max-min program:

$$S^* := \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \; \max_{\substack{a \in \{0,1\}^m \\ \mathbf{1}^\top a \leqslant A}} \; \mathcal{P}^*(z(a, d)), \tag{TL}$$

where $\mathcal{P}^*(z(a, d))$ is the inner operator's optimal value (DC or AC-OPF). This is a sequential Stackelberg game [25, 26]: the defender commits first (long-term protection), the attacker observes the defense and responds, and the operator minimizes damage given the resulting grid topology.

Table 1: Interdiction variables.

Symbol	Domain	Interpretation
d_{ij}	$\{0, 1\}$	Defender: protect line ij (1=protected)
a_{ij}	$\{0, 1\}$	Attacker: disable line ij (1=attacked)
z_{ij}	$\{0, 1\}$	Availability: $1 - a_{ij}(1 - d_{ij})$
$D,\ A$	$\mathbb{Z}_{\geqslant 0}$	Defense and attack budgets

The Stackelberg formulation is necessary, as a simultaneous Nash equilibrium would require the defender to know the attack strategy beforehand, which is physically unrealistic. Moreover, the attacker's feasible region depends on the defender's choice (protected lines cannot be attacked), so simultaneous best-response functions are not well-defined [2].

Computational difficulty. Brute-force enumeration requires $\mathcal{O}(m^{A+D})$ operator solves. For even moderate budgets ($A = D = 3$, $m = 20$), this is intractable. Implicit enumeration [6] reduces to

$\mathcal{O}(A^{D-1})$ MILP solves [3], but remains exponential in the budgets. For the DC-OPF case, iterative Benders/C&CG methods [2, 7, 8] achieve linear iteration growth at the cost of approximate solutions. No comparable single-level solver exists for the AC-OPF (or SOCP) inner problem.

Existing DC-OPF solvers. The *Implicit Enumeration* algorithm of [3, 6] constructs a branch-and-bound tree where each node solves the attacker MILP with the current defense fixed; it is exact but exponentially expensive for large budgets. The *Benders/C&CG* approach [7, 27] alternates between a defender master problem (providing a lower bound) and an attacker-operator subproblem (providing an upper bound), iterating until convergence. Each iteration requires two MILP solves, and the number of iterations grows with problem difficulty. Both approaches exploit the LP structure of the DC inner problem; extending them to the conic AC inner problem introduces significant additional complexity.

4 Single-Level Reformulations

Throughout the reformulations, the attacker envelope constants are assumed to retain at least one optimal inner-dual solution for every binary topology. The AC statements additionally use the regularity hypothesis for Lemma 2.2.

4.1 SL-MILP: DC-OPF Case

We reduce (TL) with inner (DC-OPF) to a single-level MILP in four steps. The detailed derivation is provided in Appendix C.

Step 1: Dual of the inner LP. Attach dual variables $\lambda \in \mathbb{R}^n$ to the balance equalities (2b), $\mu^- \in \mathbb{R}^{3n}_{\geqslant 0}$ to the lower bounds $x \geqslant x^{\min}$ (2a), $\mu^+ \in \mathbb{R}^{3n}_{\geqslant 0}$ to the upper bounds $x \leqslant x^{\max}$ (2a), $\phi^- \in \mathbb{R}^m_{\geqslant 0}$ to the lower flow bounds $F(z)x \geqslant -f$ (2c), and $\phi^+ \in \mathbb{R}^m_{\geqslant 0}$ to the upper flow bounds $F(z)x \leqslant f$ (2c). By Lemma 2.1, $\mathcal{P}^*_{\mathrm{DC}}(z) = \mathcal{D}^*_{\mathrm{DC}}(z)$, where

$$\mathcal{D}^*_{\mathrm{DC}}(z) := \max_{\lambda,\mu^-,\mu^+,\phi^-,\phi^+} \; c^\top\lambda + (x^{\min})^\top\mu^- - (x^{\max})^\top\mu^+ - f^\top(\phi^- + \phi^+) \tag{8a}$$

$$\text{s.t. } B(z)^\top\lambda + \mu^- - \mu^+ + F(z)^\top(\phi^- - \phi^+) = e_s, \tag{8b}$$

$$\mu^-,\ \mu^+,\ \phi^-,\ \phi^+ \geqslant 0, \quad \lambda : \text{free}. \tag{8c}$$

Step 2: Merge attacker and inner dual then linearize. Substituting $z_{ij} = 1 - a_{ij}(1 - d_{ij})$ into the stationarity constraint (8b) and expanding introduces two families of bilinear terms. Defining $w_{ij} := a_{ij}\Delta\lambda_{ij}$ and $v_{ij} := a_{ij}\Delta\phi_{ij}$, where $\Delta\lambda_{ij} := \lambda_i - \lambda_j$ and $\Delta\phi_{ij} := \phi^-_{ij} - \phi^+_{ij}$, and applying the Big-$M$ method [28] gives the exact linearized constraints for $a_{ij} \in \{0,1\}$. The merged attacker-operator dual (over $a_{ij} \in \{0,1\}$, d fixed, jointly in $a, w, v, \lambda, \mu^-, \mu^+, \phi^-, \phi^+$) becomes a MILP $\widetilde{\mathcal{D}}^*_{\mathrm{DC}}(d)$. The defender envelope uses the angle boxes in (2a).

Step 3: Relax attacker variables and dualize. We relax $a_{ij} \in \mathcal{A}_{\mathrm{bin}}$ to $a_{ij} \in \mathcal{A}_{\mathrm{rel}}$, where

$$\mathcal{A}_{\mathrm{bin}} := \{a \in \{0,1\}^m : \mathbf{1}^\top a \leqslant A\}, \quad \mathcal{A}_{\mathrm{rel}} := \{a \in [0,1]^m : \mathbf{1}^\top a \leqslant A\} \tag{9}$$

Thus the MILP $\widetilde{\mathcal{D}}^*_{\mathrm{DC}}(d)$ becomes the relaxed LP $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$. Applying strong LP duality gives the following lemma.

Lemma 4.1. *Strong duality holds for the relaxed LP: $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$.*

Let the dual be $\bar{\mathcal{P}}^*_{DC}(d)$, then $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d) = \bar{\mathcal{P}}^*_{DC}(d)$. We introduce dual variables be: $x \in \mathbb{R}^{3n}$ for the stationarity constraint (8b) (which *recovers the primal operator variable* $x = (g, \theta, s)$ by construction), $\alpha \geqslant 0$ for the attack budget $\mathbf{1}^\top a \leqslant A$, $\rho^{1:4}_{ij} \geqslant 0$ for the Big-M constraints on w_{ij}, $\sigma^{1:4}_{ij} \geqslant 0$ for the Big-M constraints on v_{ij}, and $\kappa^{0,1}_{ij} \geqslant 0$ for the box $a_{ij} \in [0,1]$. Then $\bar{\mathcal{P}}^*_{DC}(d)$ is a *minimization* over $(x, \alpha, \rho, \sigma, \kappa)$ for fixed d, with objective

$$e_s^\top x + \alpha A + \sum_{(ij)\in\mathcal{E}} \Big[M_\lambda(\rho^3_{ij} + \rho^4_{ij}) + M_\phi(\sigma^3_{ij} + \sigma^4_{ij}) + \kappa^1_{ij}\Big]. \tag{10}$$

Step 4: Linearize defender-dual product and merge with outer defender. Stationarity of the relaxed dual w.r.t. w_{ij} and v_{ij} yields nonlinear constraints, which in the re-parameterized dual read $d_{ij}(x_{\theta_i} - x_{\theta_j})$. We define $y_{ij} := d_{ij}(x_{\theta_i} - x_{\theta_j})$ and apply Big-M linearization (with finite bound $M_y \geqslant \sup|x_{\theta_i} - x_{\theta_j}|$, guaranteed by the box constraint (2a)). Substituting y_{ij} into the stationarity conditions linearizes all remaining nonlinearities. Merging with the outer defender's minimization yields (SL-MILP).

Definition 4.2 (SL-MILP)**.** The *Single-Level MILP*, obtained from (TL) for (DC-OPF) is:

$$S_{\mathrm{SL}} := \min_{x,d,y,\rho,\sigma,\kappa,\alpha} e_s^\top x + \alpha A + \sum_{(ij)\in\mathcal{E}} \left[M_\lambda(\rho^3_{ij} + \rho^4_{ij}) + M_\phi(\sigma^3_{ij} + \sigma^4_{ij}) + \kappa^1_{ij}\right] \tag{SL-MILP}$$

$$\text{s.t. } x^{\min} \leqslant x \leqslant x^{\max}, \tag{11a}$$

$$-f \leqslant F_0 x - \sigma^3 + \sigma^4 \leqslant f, \tag{11b}$$

$$c_i - (B_0 x)_i + \sum_{j:(ij)\in\mathcal{E}} (\rho^3_{ij} - \rho^4_{ij}) - \sum_{j:(ji)\in\mathcal{E}} (\rho^3_{ji} - \rho^4_{ji}) = 0, \quad \forall i \in \mathcal{V}, \tag{11c}$$

$$\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} - \rho^4_{ij} = -Y_{ij}\left[(x_{\theta_i} - x_{\theta_j}) - y_{ij}\right], \quad \forall (ij) \in \mathcal{E}, \tag{11d}$$

$$\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} - \sigma^4_{ij} = Y_{ij}\left[(x_{\theta_i} - x_{\theta_j}) - y_{ij}\right], \quad \forall (ij) \in \mathcal{E}, \tag{11e}$$

$$\alpha + (\kappa^1_{ij} - \kappa^0_{ij}) + M_\lambda(-\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} + \rho^4_{ij}) + M_\phi(-\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} + \sigma^4_{ij}) = 0, \quad \forall (ij) \in \mathcal{E}, \tag{11f}$$

$$-M_y d_{ij} \leqslant y_{ij} \leqslant M_y d_{ij}, \quad \forall (ij) \in \mathcal{E}, \tag{11g}$$

$$(x_{\theta_i} - x_{\theta_j}) - M_y(1 - d_{ij}) \leqslant y_{ij} \leqslant (x_{\theta_i} - x_{\theta_j}) + M_y(1 - d_{ij}), \quad \forall (ij) \in \mathcal{E}, \tag{11h}$$

$$\rho^{1:4}_{ij},\ \sigma^{1:4}_{ij},\ \kappa^{0,1}_{ij},\ \alpha \geqslant 0, \tag{11i}$$

$$d_{ij} \in \{0,1\}, \quad \forall (ij) \in \mathcal{E}, \qquad \mathbf{1}^\top d \leqslant D. \tag{11j}$$

where $B_0 = B(\mathbf{1})$, $F_0 = F(\mathbf{1})$ (no-interdiction case). The *only integer variables* are $d_{ij} \in \{0,1\}$, a vector of length m, independent of budgets A and D.

4.2 SL-MICP: SOCP-Relaxed AC-OPF Case

We now extend the derivation to the AC-OPF inner problem (AC-OPF). The steps are similar to the DC-OPF case. Let $x \in \mathbb{R}^N$ denote all operator variables in compact form, t the epigraph variable, and $\mathcal{K}_n$ the Lorentz (second-order) cone of dimension $n \geqslant 2$ defined in (6). A detailed derivation is provided in Appendix D.

Step 1: Conic dual of the inner SOCP. For any fixed z, from Lemma 2.2 the inner problem (AC-OPF) satisfies $\mathcal{P}^*_{AC}(z) = \mathcal{D}^*_{AC}(z)$. Attach dual variables $\nu \in \mathbb{R}^{2n}$ to the equalities (5b), $\phi^-, \phi^+ \geqslant 0$ to the lower and upper inequalities (5c), $\mu^+, \mu^- \geqslant 0$ to the box (5a), $s_{ij} \in \mathcal{K}_4$ to the SOC constraints (5d), and $r = (w, \alpha_c, \beta_c) \in \mathcal{K}_{N+2}$ to the cost epigraph (5e). Stationarity of the Lagrangian with respect to t gives $\alpha_c + \beta_c = 1$ and yields the contribution $-\|w\|^2$ to the dual objective. Stationarity with respect to x gives

$$A(z)^\top \nu + H(z)^\top(\phi^- - \phi^+) + \textstyle\sum_{(ij)\in\mathcal{E}} E_{ij}^\top s_{ij} + 2J^{1/2} w - \mu^+ + \mu^- = \tilde{c}, \tag{12}$$

where $\tilde{c}$ is the cost vector restricted to x. The dual problem is

$$\mathcal{D}^*_{\mathrm{AC}}(z) := \max_{\nu,\phi^-,\phi^+,\mu^+,\mu^-,s,w} b^\top \nu - f^\top(\phi^- + \phi^+) + (x^{\min})^\top \mu^- - (x^{\max})^\top \mu^+ - \|w\|^2 \tag{13a}$$

$$\text{s.t. } A(z)^\top \nu + H(z)^\top(\phi^- - \phi^+) + \textstyle\sum_{(ij)} E_{ij}^\top s_{ij} + 2J^{1/2} w - \mu^+ + \mu^- = \tilde{c}, \tag{13b}$$

$$s_{ij} \in \mathcal{K}_4\ \forall (ij) \in \mathcal{E}, \quad \mu^+, \mu^-, \phi^-, \phi^+ \geqslant 0. \tag{13c}$$

Step 2: Merge attacker and inner dual then linearize. Substituting $z_{ij} = 1 - a_{ij}(1 - d_{ij})$ into $A(z)$ and $H(z)$ in (13b) decomposes them as

$$A(z)^\top \nu = \widetilde{A}^\top \nu - \textstyle\sum_{(ij)} (1 - d_{ij}) \widehat{A}_{ij}^\top (a_{ij} \nu), \tag{14}$$

$$H(z)^\top(\phi^- - \phi^+) = \widetilde{H}^\top(\phi^- - \phi^+) - \textstyle\sum_{(ij)} (1 - d_{ij}) \widehat{H}_{ij}^\top (a_{ij}(\phi^-_{ij} - \phi^+_{ij})), \tag{15}$$

where $\widetilde{A} := A(\mathbf{1})$ and $A(z) = \widetilde{A} - \sum_{(ij)\in\mathcal{E}}(1 - z_{ij})\widehat{A}_{ij}$, with $\widehat{A}_{ij} \in \mathbb{R}^{2n\times N}$ the contribution of line (ij). The matrices $\widetilde{H}_{ij}$ and $\widehat{H}_{ij}$ are the corresponding row blocks of $\widetilde{H} = H(\mathbf{1})$ and its line contributions; $\phi^{\pm}_{ij}$ and v_{ij} collect all flow rows for that line. Thus $w^A_{ij}, \rho^{1:4}_{ij} \in \mathbb{R}^N$; $v_{ij}, \sigma^{1:4}_{ij}$ have the dimension of the line's flow block. All envelope inequalities are component-wise. This introduces the line-wise products $w^A_{ij} := a_{ij}\widehat{A}^{\top}_{ij}\nu$ and $v_{ij} := a_{ij}(\phi^-_{ij} - \phi^+_{ij})$.

Applying the exact four-inequality envelopes gives

$$\begin{aligned} |w^A_{ij}| &\leqslant M_\nu a_{ij}, \qquad & |w^A_{ij} - \widehat{A}^{\top}_{ij}\nu| &\leqslant M_\nu(1 - a_{ij}),\\ |v_{ij}| &\leqslant M_\phi a_{ij}, \qquad & |v_{ij} - (\phi^-_{ij} - \phi^+_{ij})| &\leqslant M_\phi(1 - a_{ij}). \end{aligned}$$

The stationarity constraint (13b) becomes linear in $(a, w^A, v, \nu, \phi^{\pm}, \mu^{\pm}, s, w)$ for fixed d, and the merged attacker-operator dual maximization over $a \in \mathcal{A}_{\text{bin}}$ becomes a mixed-integer conic program (MICP) $\widetilde{\mathcal{D}}^*_{\text{AC}}(d)$. **Step 3: Relax attacker variables and dualize.** We relax $a_{ij} \in [0, 1]$ and obtain the relaxed conic program (CP) $\bar{\mathcal{D}}^*_{\text{AC}}(d)$ with primal variables $(a, w^A, v, \nu, \phi^+, \phi^-, \mu^+, \mu^-, s, w)$. Verifying Slater's condition gives

Lemma 4.3. *Strong duality holds for the relaxed CP: $\bar{\mathcal{D}}^*_{\text{AC}}(d)$.*

So, taking the conic dual of $\bar{\mathcal{D}}^*_{\text{AC}}(d)$ as $\bar{\mathcal{P}}^*_{\text{AC}}(d)$, we get $\bar{\mathcal{D}}^*_{\text{AC}}(d) = \bar{\mathcal{P}}^*_{\text{AC}}(d)$. We use the dual variables: $x \in \mathbb{R}^N$ for the stationarity equality, $\xi^+ \geqslant 0$ for $\mu^+ \geqslant 0$, $\xi^- \geqslant 0$ for $\mu^- \geqslant 0$, $\psi^{\pm} \geqslant 0$ for $\phi^{\pm} \geqslant 0$, $\rho^{1:4}_{ij} \geqslant 0$ for the Big-M constraints on w^A_{ij}, $\sigma^{1:4}_{ij} \geqslant 0$ for the Big-M constraints on v_{ij}, $\kappa^{0,1}_{ij} \geqslant 0$ for $0 \leqslant a_{ij} \leqslant 1$, $\alpha \geqslant 0$ for the budget $\mathbf{1}^{\top}a \leqslant A$, and $k_{ij} \in \mathcal{K}_4$ dual to $s_{ij} \in \mathcal{K}_4$.

Step 4: Linearize defender-dual product and merge with outer defender. Stationarity of the dual Lagrangian with respect to w^A_{ij} introduces the product $(1 - d_{ij})x$, and with respect to v_{ij} introduces $(1 - d_{ij})\widehat{H}_{ij}x$. Setting $y_{ij} := d_{ij}x$ and applying Big-M linearization with $M_k := \max(|x^{\min}_k|, |x^{\max}_k|)$ component-wise:

$$y_{ij} \leqslant x^{\max}d_{ij}, \quad y_{ij} \geqslant x^{\min}d_{ij}, \quad y_{ij} \leqslant x - x^{\min}(1 - d_{ij}), \quad y_{ij} \geqslant x - x^{\max}(1 - d_{ij}), \quad (16)$$

makes all remaining constraints linear or conic. Merging with the outer defender's minimization yields:

Definition 4.4 (SL-MICP)**.** The *Single-Level MICP*, obtained from (TL) for (AC-OPF) is:

$$\begin{aligned} S_{\text{SL}} := \min_{x,t,d,y,\rho,\sigma,\kappa,\alpha}\ & \tilde{c}^{\top}x + t + \alpha A \\ & + \sum_{(ij)\in\mathcal{E}} \left[M_\nu \mathbf{1}^{\top}(\rho^3_{ij} + \rho^4_{ij}) + M_\phi \mathbf{1}^{\top}(\sigma^3_{ij} + \sigma^4_{ij}) + \kappa^1_{ij}\right] && \text{(SL-MICP)}\\ \text{s.t. } & x^{\min} \leqslant x \leqslant x^{\max}, && (17a)\\ & E_{ij}x \in \mathcal{K}_4, \quad \forall (ij) \in \mathcal{E}, && (17b)\\ & [2J^{1/2}x;\ t-1;\ t+1] \in \mathcal{K}_{N+2}, && (17c)\\ & \widetilde{A}x - \sum_{(ij)\in\mathcal{E}} \widehat{A}_{ij}(\rho^3_{ij} - \rho^4_{ij}) = b, && (17d)\\ & -f_{ij} \leqslant \widetilde{H}_{ij}x - \sigma^3_{ij} + \sigma^4_{ij} \leqslant f_{ij}, \quad \forall (ij) \in \mathcal{E}, && (17e)\\ & \rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} - \rho^4_{ij} = (x - y_{ij}), \quad \forall (ij) \in \mathcal{E}, && (17f)\\ & \sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} - \sigma^4_{ij} = \widehat{H}_{ij}(x - y_{ij}), \quad \forall (ij) \in \mathcal{E}, && (17g)\\ & \alpha + (\kappa^1_{ij} - \kappa^0_{ij}) + M_\nu \mathbf{1}^{\top}(-\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} + \rho^4_{ij}) \\ & \qquad + M_\phi \mathbf{1}^{\top}(-\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} + \sigma^4_{ij}) = 0, \quad \forall (ij) \in \mathcal{E}, && (17h)\\ & -M_x d_{ij} \leqslant y_{ij} \leqslant M_x d_{ij}, \\ & x - M_x(1 - d_{ij}) \leqslant y_{ij} \leqslant x + M_x(1 - d_{ij}), \quad \forall (ij) \in \mathcal{E}, && (17i)\\ & \rho^{1:4}_{ij},\ \sigma^{1:4}_{ij},\ \kappa^{0,1}_{ij},\ \alpha \geqslant 0, && (17j)\\ & d_{ij} \in \{0, 1\}, \quad \forall (ij) \in \mathcal{E}, \qquad \mathbf{1}^{\top}d \leqslant D. && (17k) \end{aligned}$$

The *only integer variables* are $d_{ij} \in \{0, 1\}$, a vector of length m, independent of budgets A and D.

5 Theoretical Analysis

We analyze the main guarantees for both formulations. Full proofs are given in Appendix E. For DC, $\mathcal{P}^*$ denotes the original DC-OPF shedding objective. For AC, it denotes the convex SOCP operator

objective in (AC-OPF); it does not denote the globally optimal value of the original nonconvex AC-OPF.

Theorem 5.1. *Let S^* be the optimal value of the tri-level problem* (TL) *for the applicable convex operator model and S_{SL} the optimal value of* (SL-MILP) *(or* (SL-MICP)*). Then $S_{\mathrm{SL}} \geqslant S^*$.*

Proof sketch. Steps 1 and 2 are exact: the inner operator problem is replaced by its exact dual via strong duality, and the attacker-dual bilinearities are removed by exact Big-M linearization. In Step 3, relaxing $a_{ij} \in \{0,1\}$ to $a_{ij} \in [0,1]$ enlarges the attacker's feasible set, hence $\bar{\mathcal{D}}^*(d) \geqslant \widetilde{\mathcal{D}}^*(d)$ for every fixed defense d. Strong duality then gives $\bar{\mathcal{D}}^*(d) = \bar{\mathcal{P}}^*(d)$, and Step 4 again applies exact Big-M linearization to obtain the final single-level problem. Therefore

$$S_{\mathrm{SL}} = \min_d \bar{\mathcal{P}}^*(d) = \min_d \bar{\mathcal{D}}^*(d) \geqslant \min_d \widetilde{\mathcal{D}}^*(d) = S^*.$$

Theorem 5.2. *Let d^*_{SL} be an optimal defense from the SL-MILP (or SL-MICP), and define the* realized worst-case value

$$S_{\mathrm{true}}(d^*_{\mathrm{SL}}) := \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a, d^*_{\mathrm{SL}})). \tag{18}$$

Then $S_{\mathrm{SL}} \geqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \geqslant S^*$.

Proof sketch. The proof of Theorem 5.1 actually gives the pointwise bound

$$\bar{\mathcal{P}}^*(d) \geqslant \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a,d)) \qquad \forall d.$$

Evaluating at the single-level optimizer d^*_{SL} yields $S_{\mathrm{SL}} \geqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}})$. Since S^* is the minimum of the same worst-case binary-attacker objective over all feasible defenses, the realized value at the particular defense d^*_{SL} must satisfy $S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \geqslant S^*$.

Remark 5.3. The two gaps $\Delta_C := S_{\mathrm{SL}} - S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \geqslant 0$ (*conservative gap*) and $\Delta_0 := S_{\mathrm{true}}(d^*_{\mathrm{SL}}) - S^* \geqslant 0$ (*optimality gap*) quantify the two sources of sub-optimality. The first is computed by one additional exact binary attacker-operator solve with $d = d^*_{\mathrm{SL}}$ fixed; when exact tri-level enumeration is tractable, the same recomputation also gives Δ_0. Thus $S_{\mathrm{true}}(d^*_{\mathrm{SL}})$ is a tighter post-hoc certificate for the deployed defense, while S_{SL} remains the a-priori certificate.

Theorem 5.4. *Under the defence d^*_{SL}, no binary attack $a \in \mathcal{A}_{\mathrm{bin}}$ can cause an applicable operator-objective value exceeding S_{SL}: $\mathcal{P}^*(z(a, d^*_{\mathrm{SL}})) \leqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \leqslant S_{\mathrm{SL}}$.*

Proof sketch. The first inequality is immediate from the definition of $S_{\mathrm{true}}(d^*_{\mathrm{SL}})$ as the maximum over all binary attacks, and the second follows directly from Theorem 5.2.

Remark 5.5. S_{SL} is the worst-case damage achievable by a "fractional super-attacker" who can partially disable lines. Since real attackers are binary-constrained, the defence d^*_{SL} works as a conservative safety certificate.

Complexity. Table 2 compares the four approaches. The SL-MILP/MICP requires exactly one MILP/MICP solve with m binary variables, independent of A and D.

Table 2: Computational comparison.

Method	Exact?	OPF	# Solves	Scales with A, D / iter?
Brute force	Yes	DC	$\mathcal{O}(m^{A+D})$	Yes (exponential)
Implicit enumeration	Yes	DC	$\mathcal{O}(A^{D-1})$ MILP	Yes (exponential)
Benders/C&CG	Approx.	DC	$\mathcal{O}(\text{iter}) \times 2$ MILP	Yes (linear)
SL-MILP/MICP (ours)	**Upper bd.**	**DC/AC**	**1 MILP/MICP**	**No (1 iter)**

6 Experiments

We implement all models in Python using the Gurobipy library [29] and evaluate the tri-level DC/AC-OPF problem on standard IEEE benchmark systems from MATPOWER [30]. All experiments are conducted on a machine with 32 GB RAM and an AMD Ryzen 7 9700X 8-core 64-bit processor. We compare against (i) Implicit enumeration methods [3, 6] and (ii) Benders cut-based bi-level solvers [8, 27], where applicable.

Table 3: **Optimal shedding and runtime comparison on IEEE 14-bus system for tri-level DC-OPF.** We report optimal load shedding (MW) and runtime (seconds) across attacker budget A and defender budget D.Bold values indicate optimal shedding or fastest runtime; underlined values indicate second-best values.SL-MILP matches or upper-bounds the exact optimum with significantly lower computational cost.

		Implicit Enumeration					Bi-level Solver					SL-MILP (ours)				
	$A\backslash D$	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Shed.(MW)	1	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
	2	**14.9**	**12.5**	**6.1**	**3.5**	**0.0**	**14.9**	14.9	12.9	8.7	5.7	**14.9**	**12.5**	12.5	6.2	4.1
	3	**87.7**	**34.5**	**14.9**	**12.5**	**6.1**	**87.7**	87.7	56.9	34.5	28.4	**87.7**	**34.5**	20.3	**12.5**	6.3
Runtime (s)	1	0.023	0.023	0.123	0.723	3.241	0.016	0.045	0.182	0.910	3.834	**0.007**	**0.019**	**0.042**	**0.046**	**0.016**
	2	0.118	2.326	20.801	130.186	577.160	0.137	0.264	0.475	1.610	4.172	**0.007**	**0.018**	**0.042**	**0.048**	**0.015**
	3	0.744	12.771	112.187	374.238	1094.449	0.846	1.479	2.649	2.969	6.184	**0.008**	**0.018**	**0.028**	**0.045**	**0.015**

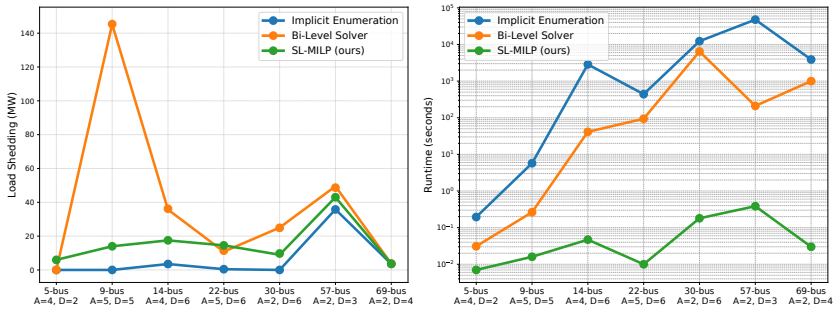


Figure 2: **Shedding and runtime comparison across IEEE systems for tri-level DC-OPF.** Solution obtained by SL-MILP achieves tight upper bounds on load shedding, near-constant runtimes across (A, D) combinations, and is consistently 1–4 orders of magnitude speed-up over competing methods.

Table 3 reports optimal load shedding and runtime on the IEEE 14-bus system across attacker and defender budgets (A, D). The SL-MILP consistently produces values greater than or equal to the exact values obtained by implicit enumeration, confirming the upper-bound guarantee in Theorem 5.1. The total difference is decomposed by recomputing the exact binary worst case at d^*_{SL}: for example, at $(A, D) = (2, 2)$, $S_{\mathrm{SL}} = 12.5$, $S_{\mathrm{true}}(d^*_{\mathrm{SL}}) = 9.0$, and $S^* = 6.1$ MW, giving $(\Delta_C, \Delta_0) = (3.5, 2.9)$ MW. In terms of runtime, SL-MILP is significantly faster, with nearly constant solve times (≈10–50 ms on average) across all (A, D) combinations, whereas implicit enumeration and bi-level methods scale from milliseconds to seconds (close to an hour for exact enumeration). Figure 2 further shows one to four orders-of-magnitude speed-ups across seven IEEE systems. Appendix F reports the gap decomposition and an IEEE 118-bus extension, where $\Delta_C, \Delta_0 \geqslant 0$ for every reported defense budget.

Table 4 reports optimal defenses, worst-case attacks, load shedding, and generation cost on the WSCC 9-bus system (Figure 1). Consistent with (AC-OPF), the model jointly minimizes load shedding and quadratic generation cost. The generation-cost column need not be monotone in D: the worst-case topology changes between rows, and the shedding-dominant objective can deliberately use a more expensive dispatch to reduce shedding. At $D = 3$, for example, shedding drops from 97.3 to 66.1 MW while the surviving topology forces a costlier dispatch; additional protected delivery paths reduce cost

Table 4: **Tri-level security defense AC-OPF using SL-MICP (WSCC 9-bus 1).** Optimal defended lines, worst-case attacks, load shedding, and generation cost.

D	Protected lines	Attacked lines	Load shedding (MW)	Generation cost ($/h)
0	—	8–9, 9–4	125.0	2947.96
1	9–4	6–7, 7–8	101.3	3252.21
2	6–7, 4–9	4–5, 5–6	97.3	3533.50
3	4–5, 6–7, 4–9	3–6, 8–2	66.1	8991.25
4	1–4, 4–5, 6–7, 4–9	3–6, 2–8	65.7	6697.50
5	1–4, 4–5, 6–7, 8–2, 9–4	—	0.00	6505.94

again at $D = 4, 5$. Appendix F gives the realized binary-attack gaps and results for IEEE 14, 30, and 57-bus systems. These are certificates for the SOCP model.

7 Conclusion

We develop certified single-level reformulations of the tri-level cyber-physical grid security problem under DC-OPF (SL-MILP) and SOCP-relaxed AC-OPF (SL-MICP). The approach follows a four-step exact-relax-exact pipeline combining duality and Big-M linearization. We prove a conservative a-priori bound and obtain a tighter exact post-hoc certificate for the returned defense through one binary attacker-operator solve. Experiments on IEEE benchmarks show substantial computational gains and quantify both components of the certificate gap.

Impact Statement. Our work enables tractable, certified defense planning for power grids under adversarial contingencies, improving resilience and operational reliability. As adversarial modeling can also reveal vulnerabilities, deployment should follow responsible practices, including safeguards, controlled disclosure, and domain-specific oversight.

Future directions. (i) *Tightness analysis:* sharper bounds on conservativeness from the attacker relaxation, especially for structured networks; (ii) *Improved AC fidelity:* tighter relaxations or nonlinear AC-OPF to reduce approximation gaps; (iii) *Robust extensions:* stochastic/robust formulations with uncertain demand and generation for security-aware planning.

Limitations. The AC certificate is with respect to the SOCP-relaxed operator and does not automatically transfer to a globally solved nonconvex AC-OPF. The continuous attacker relaxation can also be conservative, as the reported Δ_C values make explicit, and mixed-integer runtimes can depend on numerical scaling even when valid Big-M thresholds leave the optimum unchanged.

Acknowledgment

This work was supported by the POWERGRID Center of Excellence in Cyber Security for Transmission and Grid Operations (PGCoE) at the Indian Institute of Science, under the project *Multilevel Optimization for POWERGRID Security and Defense*.

Ethics statement

This work is purely theoretical and algorithmic, uses only publicly available data, and involves no human subjects, crowdsourcing, or personal/sensitive data.

References

[1] Salmeron, J., K. Wood, R. Baldick. Analysis of electric grid security under terrorist threat. *IEEE Transactions on Power Systems*, 19(2):905–912, 2004.

[2] Brown, G., M. Carlyle, J. Salmerón, et al. Defending critical infrastructure. *Interfaces*, 36(6):530–544, 2006.

[3] Wu, X., A. J. Conejo. An efficient tri-level optimization model for electric grid defense planning. *IEEE Transactions on Power Systems*, 32(4):2984–2994, 2017.

[4] Motto, A. L., J. M. Arroyo, F. D. Galiana. A mixed-integer lp procedure for the analysis of electric grid security under disruptive threat. *IEEE Transactions on Power Systems*, 20(3):1357–1365, 2005.

[5] Alguacil, N., A. Delgadillo, J. M. Arroyo. A trilevel programming approach for electric grid defense planning. *Computers & Operations Research*, 41:282–290, 2014.

[6] Scaparra, M. P., R. L. Church. A bilevel mixed-integer program for critical infrastructure protection planning. *Computers & Operations Research*, 35(6):1905–1923, 2008.

[7] Zeng, B., L. Zhao. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters*, 41(5):457–461, 2013.

[8] Davarikia, H., M. Barati. A tri-level programming model for attack-resilient control of power grids. *Journal of Modern Power Systems and Clean Energy*, 6(5):918–929, 2018.

[9] Lai, K., Y. Wang, D. Shi, et al. A resilient power system operation strategy considering transmission line attacks. *IEEE Access*, 6:70633–70643, 2018.

[10] Fang, Y., G. Sansavini. Optimizing power system investments and resilience against attacks. *Reliability Engineering & System Safety*, 159:161–173, 2017.

[11] Liu, X., Z. Li. Trilevel modeling of cyber attacks on transmission lines. *IEEE Transactions on Smart Grid*, 8(2):720–729, 2017.

[12] Yuan, Y., Z. Li, K. Ren. Modeling load redistribution attacks in power systems. *IEEE Trans. Smart Grid*, 2(2):382–390, 2011.

[13] Liu, Y., P. Ning, M. K. Reiter. False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur.*, 14(1):1–33, 2011.

[14] He, H., S. Huang, Y. Liu, et al. A tri-level optimization model for power grid defense with the consideration of post-allocated dgs against coordinated cyber-physical attacks. *International Journal of Electrical Power & Energy Systems*, 130:106903, 2021.

[15] Zhang, Z., X. Huang, T. Zhang. A tri-level distributionally robust defender–attacker–defender model for grid resilience enhancement under repair time uncertainty. *Applied System Innovation*, 8(4):115, 2025.

[16] Frank, S., I. Steponavice, S. Rebennack. Optimal power flow: a bibliographic survey I. *Energy Syst.*, 3(3):221–258, 2012.

[17] Molzahn, D. K., I. A. Hiskens. A survey of relaxations and approximations of the power flow equations. *Found. Trends Electr. Energy Syst.*, 4(1–2):1–221, 2019.

[18] Lavaei, J., S. H. Low. Zero duality gap in optimal power flow problem. *IEEE Trans. Power Syst.*, 27(1):92–107, 2012.

[19] Low, S. H. Convex relaxation of optimal power flow—Part I: Formulations and equivalence. *IEEE Trans. Control Netw. Syst.*, 1(1):15–27, 2014.

[20] Jabr, R. A. Radial distribution load flow using conic programming. *IEEE Trans. Power Syst.*, 21(3):1458–1459, 2006.

[21] Illinois Center for a Smarter Electric Grid (ICSEG). WSCC 9-bus system. `https://icseg.iti.illinois.edu/wscc-9-bus-system/`. Accessed: 2026-05-05.

[22] Bertsimas, D., J. N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, Belmont, MA, 1997.

[23] Capitanescu, F., J. L. Martinez Ramos, P. Panciatici, et al. State-of-the-art, challenges, and future trends in security constrained optimal power flow. *Electr. Power Syst. Res.*, 81(8):1731–1741, 2011.

[24] Lobo, M. S., L. Vandenberghe, S. Boyd, et al. Applications of second-order cone programming. *Linear Algebra Appl.*, 284:193–228, 1998.

[25] von Stackelberg, H. *The Theory of the Market Economy*. William Hodge & Co., London, UK, 1952. Translated from German.

[26] Owen, G. *Game Theory*. Academic Press, San Diego, CA, 3 edn., 2001.

[27] Guo, Y., C. Guo, J. Yang. A tri-level optimization model for power systems defense considering cyber-physical interdependence. *IET Generation, Transmission & Distribution*, 17(7):1477–1490, 2023.

[28] Cococcioni, M., L. Fiaschi. The big-m method with the numerical infinite m. *Optimization Letters*, 15:2455–2468, 2021.

[29] Gurobi Optimization, LLC, Houston, Texas. *Gurobi Optimizer Reference Manual*, 2025.

[30] Zimmerman, R. D., C. E. Murillo-Sánchez, R. J. Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, 2011.

[31] Nesterov, Y., A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, 1994.

[32] Vielma, J. P. Mixed integer linear programming formulation techniques. *SIAM Review*, 57(1):3–57, 2015.

Appendix

This appendix provides a complete and self-contained derivation of the proposed tri-level-to-single-level reformulations, together with supporting theory and experiments. Section A introduces conic duality, Slater's condition, and Big-M linearization. Section B gives computable, physics-based constants and the verification procedure used in our implementation. Sections C and D develop the DC and SOCP-relaxed AC reformulations, respectively, and Section E proves the theoretical results. Section F reports certificate-gap decompositions, larger-system results, and the empirical SOCP comparison.

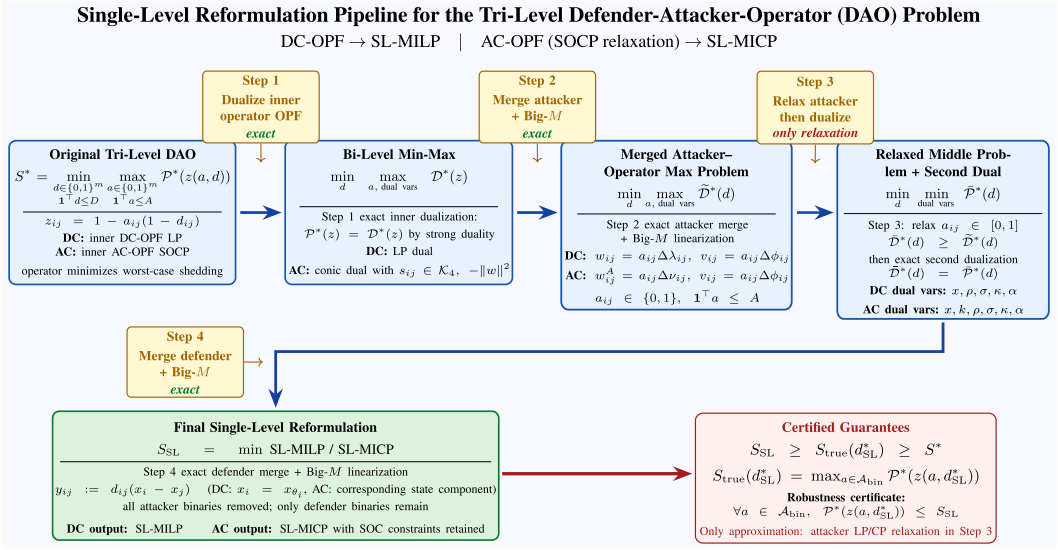


Figure 3: Illustration of the 4-step pipeline to convert the tri-level DC/AC-OPF model into single-level mixed-integer linear/conic programming.

A Mathematical Preliminaries

Definition A.1 (Second-order cone)**.** The Lorentz cone of dimension $n \geqslant 2$ is $\mathcal{K}_n := \{v \in \mathbb{R}^n : \|(v_1, \ldots, v_{n-1})\| \leqslant v_n\}$. Its dual is $\mathcal{K}_n^* = \mathcal{K}_n$ (self-dual). A rotated second-order cone is $\mathcal{K}_n^r := \{(x_0, x_1, y) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n-2} : \|y\|^2 \leqslant 2x_0x_1,\ x_0, x_1 \geqslant 0\}$.

Definition A.2 (Conic program and dual)**.** A primal conic program in standard form is $\mathcal{P}^* := \inf\{c^\top x : Ax = b,\ x \in K\}$, with Lagrangian dual $\mathcal{D}^* := \sup\{b^\top y : A^\top y + s = c,\ s \in K^*\}$. Weak duality $\mathcal{D}^* \leqslant \mathcal{P}^*$ always holds; strong duality $\mathcal{D}^* = \mathcal{P}^*$ holds under Slater's condition [31].

Definition A.3 (Slater's condition)**.** Problem $\{\mathcal{P}^* : Ax = b, Dx \leqslant c, x \in K\}$ satisfies Slater's condition if there exists $\hat{x}$ with $A\hat{x} = b, D\hat{x} < c$ and $\hat{x} \in \mathrm{ri}(K)$, where $\mathrm{ri}()$ denotes the relative interior.

Definition A.4 (Big-M linearization)**.** Given $\delta \in \{0, 1\}$ and $v \in [\underline{v}, \bar{v}]$, the product $w = \delta v$ is exactly represented by: $w \leqslant \bar{v}\delta$, $w \geqslant \underline{v}\delta$, $w \leqslant v - \underline{v}(1 - \delta)$, $w \geqslant v - \bar{v}(1 - \delta)$. These are exact whenever $\delta \in \{0, 1\}$ and $M \geqslant \max(|\underline{v}|, |\bar{v}|)$.

B Valid Attacker-Envelope Constants and Defender Bounds

The two attacker envelopes require bounds on dual-variable differences, whereas the defender envelopes use primal box bounds. We use the following construction rather than an ad-hoc common constant.

Lemma B.1 (Dual bounds for DC)**.** *For every binary topology z, select an optimal dual solution $(\lambda^*(z), \mu^{-*}(z), \mu^{+*}(z), \phi^{-*}(z), \phi^{+*}(z))$. Valid finite envelope constants satisfy*

$$M_\lambda \geqslant \max_{z\in\{0,1\}^m} \max_{(ij)\in\mathcal{E}} |\lambda_i^*(z) - \lambda_j^*(z)|, \qquad M_\phi \geqslant \max_{z\in\{0,1\}^m} \max_{(ij)\in\mathcal{E}} |\phi_{ij}^{-*}(z) - \phi_{ij}^{+*}(z)|.$$

Proof. The all-shed point is feasible under the stated box assumptions and the objective is bounded, so each binary-topology LP has an attained dual optimum. There are finitely many binary topologies. Selecting one finite optimal dual solution for each topology and taking the displayed maxima therefore gives finite valid bounds. The envelopes retain an optimal solution for every binary attack. □

Defender envelope for DC. The angle boxes in (2a) give the exact line-wise bound

$$M_{y,ij} := \max\left\{\left|\theta_i^{\min} - \theta_j^{\max}\right|, \left|\theta_i^{\max} - \theta_j^{\min}\right|\right\},$$

and one may use $M_y = \max_{(ij)\in\mathcal{E}} M_{y,ij}$ when a scalar constant is convenient.

Uniform AC bounds. Under the hypothesis of Lemma 2.2, select an attained optimal dual solution for every binary topology. The finite constants satisfy

$$M_\nu \geqslant \max_{z\in\{0,1\}^m} \max_{(ij)\in\mathcal{E}} \|\widehat{A}_{ij}^\top \nu^*(z)\|_\infty, \qquad M_\phi \geqslant \max_{z\in\{0,1\}^m} \max_{(ij)\in\mathcal{E}} \|\phi_{ij}^{-*}(z) - \phi_{ij}^{+*}(z)\|_\infty.$$

The defender products use the component-wise primal bounds

$$(M_x)_k = \max\{|x_k^{\min}|, |x_k^{\max}|\}.$$

Implementation. After solving SL-MILP or SL-MICP, we inspect every attacker and defender envelope. If a returned solution touches an artificial M boundary, we double the corresponding constant and resolve. With the analytic bounds above, exactness already follows from Definition A.4; the activity test is an additional diagnostic against scaling or implementation errors. The reported systems used the data-derived M_y, M_λ, M_ϕ values for DC, and the uniform Slater/primal-box constants for AC.

C Derivation of the SL-MILP (DC-OPF)

Step 1: Dual of the Inner DC-OPF LP

Primal. For fixed $z \in \{0,1\}^m$, the inner problem is the LP (DC-OPF) with $x = (g, \theta, s) \in \mathbb{R}^{3n}$, and Lemma 2.1 gives $\mathcal{P}_{\mathrm{DC}}^*(z) = \mathcal{D}_{\mathrm{DC}}^*(z)$.

Lagrangian. Form the Lagrangian with $\lambda \in \mathbb{R}^n$ for (2b), $\mu^- \in \mathbb{R}^{3n}_{\geqslant 0}$ for $x \geqslant x^{\min}$, $\mu^+ \in \mathbb{R}^{3n}_{\geqslant 0}$ for $x \leqslant x^{\max}$, $\phi^- \in \mathbb{R}^m_{\geqslant 0}$ for $F(z)x \geqslant -f$, and $\phi^+ \in \mathbb{R}^m_{\geqslant 0}$ for $F(z)x \leqslant f$:

$$\begin{aligned}\mathcal{L}_{DC} \;&= e_s^\top x + \lambda^\top\big(c - B(z)x\big) - (\mu^-)^\top(x - x^{\min}) + (\mu^+)^\top(x - x^{\max}) \\ &\quad - (\phi^-)^\top\big(F(z)x + f\big) + (\phi^+)^\top\big(F(z)x - f\big).\end{aligned} \tag{19}$$

Stationarity w.r.t. x. Setting $\nabla_x \mathcal{L}_{DC} = 0$ gives the single compact condition

$$B(z)^\top\lambda + \mu^- - \mu^+ + F(z)^\top(\phi^- - \phi^+) = e_s, \tag{8b}$$

which decomposes component-by-component as follows. *w.r.t. g_i*: since $[B(z)]_{i,g_i} = 1$ and $[F(z)]_{\cdot,g_i} = 0$,

$$\lambda_i + \mu_{g_i}^- - \mu_{g_i}^+ = 0 \implies \lambda_i = \mu_{g_i}^+ - \mu_{g_i}^-. \tag{20}$$

w.r.t. s_i: since $[B(z)]_{i,s_i} = 1$ and $(e_s)_{s_i} = 1$,

$$\lambda_i + \mu_{s_i}^- - \mu_{s_i}^+ = 1 \implies \lambda_i = 1 - \mu_{s_i}^- + \mu_{s_i}^+. \tag{21}$$

w.r.t. θ_i: since $[B(z)]_{k,\theta_i}$ contributes $-Y_{ij}z_{ij}$ for $k = i$ and $+Y_{ij}z_{ij}$ for $k = j$, and $[F(z)]_{ij,\theta_i} = Y_{ij}z_{ij}$ (and $-Y_{ij}z_{ij}$ for θ_j),

$$\sum_{(k\ell)\in\mathcal{E}} [F(z)]_{k\ell,\theta_i}\big(\lambda_k - \lambda_\ell - \phi_{k\ell}^- + \phi_{k\ell}^+\big) - \mu_{\theta_i}^- + \mu_{\theta_i}^+ = 0. \tag{22}$$

Dual objective. Taking $\inf_x \mathcal{L}_{DC}$ subject to stationarity and collecting constant terms yields the dual objective (8a):

$$\mathcal{D}^*_{\mathrm{DC}}(z) = \max\Big\{ c^\top \lambda + (x^{\min})^\top \mu^- - (x^{\max})^\top \mu^+ - f^\top(\phi^- + \phi^+) \Big\}. \tag{23}$$

Expanding with $x^{\min} = (0, \theta^{\min}, 0)$, $x^{\max} = (g^{\max}, \theta^{\max}, c)$:

$$\mathcal{D}^*_{\mathrm{DC}}(z) = \max\Big\{ \sum_{i\in\mathcal{V}} (c_i\lambda_i - g_i^{\max}\mu_{g_i}^+ + (\theta_i^{\min})\mu_{\theta_i}^- - \theta_i^{\max}\mu_{\theta_i}^+ - c_i\mu_{s_i}^+) - \sum_{(ij)\in\mathcal{E}} f_{ij}(\phi_{ij}^- + \phi_{ij}^+) \Big\}, \tag{24}$$

subject to (8b) and $\mu^-, \mu^+, \phi^-, \phi^+ \geqslant 0$, λ : free.

Step 2: Merging and Attacker Linearization

Substituting $z_{ij} = 1 - a_{ij}(1 - d_{ij})$ into the stationarity constraint (8b) (equivalently into (22)) introduces two families of bilinear terms involving the attacker decision a_{ij}:

$$a_{ij}\,\Delta\lambda_{ij} \quad \text{where} \quad \Delta\lambda_{ij} := \lambda_i - \lambda_j, \qquad \text{and} \qquad a_{ij}\,\Delta\phi_{ij} \quad \text{where} \quad \Delta\phi_{ij} := \phi_{ij}^- - \phi_{ij}^+. \tag{25}$$

We introduce auxiliary variables $w_{ij} := a_{ij}\Delta\lambda_{ij}$ and $v_{ij} := a_{ij}\Delta\phi_{ij}$, and apply exact Big-M linearization [28, 32] (Definition A.4) with sufficiently large M-values ($M_\lambda \geqslant \sup|\Delta\lambda_{ij}|$ and $M_\phi \geqslant \sup|\Delta\phi_{ij}|$ over the selected optimal dual solutions):

$$|w_{ij}| \leqslant M_\lambda\, a_{ij}, \qquad w_{ij} - \Delta\lambda_{ij} \in [-M_\lambda(1 - a_{ij}),\ M_\lambda(1 - a_{ij})], \tag{26a}$$

$$|v_{ij}| \leqslant M_\phi\, a_{ij}, \qquad v_{ij} - \Delta\phi_{ij} \in [-M_\phi(1 - a_{ij}),\ M_\phi(1 - a_{ij})]. \tag{26b}$$

So the equality constraint in (22) becomes:

$$\sum_{(k\ell)\in\mathcal{E}} (F_0)_{k\ell,\theta_i}\big[\Delta\lambda_{k\ell} - \Delta\phi_{k\ell} - (1 - d_{k\ell})(w_{k\ell} - v_{k\ell})\big] - \mu_{\theta_i}^- + \mu_{\theta_i}^+ = 0. \tag{27}$$

Finally with $a_{ij} \in \{0, 1\}$ and $\mathbf{1}^\top a \leqslant A$, the merged attacker-operator MILP problem becomes

$$\begin{aligned}
\widetilde{\mathcal{D}}^*_{\mathrm{DC}}(d) := \max_{a,\lambda,\mu^-,\mu^+,\phi^-,\phi^+,w,v}\ & c^\top\lambda + (x^{\min})^\top\mu^- - (x^{\max})^\top\mu^+ - f^\top(\phi^- + \phi^+) && \text{(28a)}\\
\text{s.t. } & \lambda_i = \mu_{g_i}^+ - \mu_{g_i}^-, \quad \forall i \in \mathcal{V}, && \text{(28b)}\\
& \lambda_i = 1 - \mu_{s_i}^- + \mu_{s_i}^+, \quad \forall i \in \mathcal{V}, && \text{(28c)}\\
& \sum_{(k\ell)\in\mathcal{E}} (F_0)_{k\ell,\theta_i}\Big[\Delta\lambda_{k\ell} - \Delta\phi_{k\ell} - (1 - d_{k\ell})(w_{k\ell} - v_{k\ell})\Big]\\
& \qquad - \mu_{\theta_i}^- + \mu_{\theta_i}^+ = 0, \quad \forall i \in \mathcal{V}, && \text{(28d)}\\
& \text{big-M constraints (26a)–(26b)},\\
& \mu^-,\ \mu^+,\ \phi^-,\ \phi^+ \geqslant 0, \qquad \lambda \text{ : free}, && \text{(28e)}\\
& a_{ij} \in \{0, 1\}, \quad \forall (ij) \in \mathcal{E}, \quad \mathbf{1}^\top a \leqslant A. && \text{(28f)}
\end{aligned}$$

Step 3: LP Relaxation and Dual of the Merged Problem

Relaxing $a_{ij} \in [0, 1]$ we get the relaxed LP $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$:

$$\begin{aligned}
\bar{\mathcal{D}}^*_{\mathrm{DC}}(d) := \max_{a,\lambda,\mu^-,\mu^+,\phi^-,\phi^+,w,v}\ & c^\top\lambda + (x^{\min})^\top\mu^- - (x^{\max})^\top\mu^+ - f^\top(\phi^- + \phi^+) && \text{(29a)}\\
\text{s.t. } & \text{constraints (28b)–(28d)},\\
& \text{Big-M constraints (26a)–(26b)},\\
& \mu^-,\ \mu^+,\ \phi^-,\ \phi^+ \geqslant 0, \qquad \lambda \text{ : free}, && \text{(29b)}\\
& a_{ij} \in [0, 1], \quad \forall (ij) \in \mathcal{E}, \qquad \mathbf{1}^\top a \leqslant A. && \text{(29c)}
\end{aligned}$$

For the dual $\bar{\mathcal{P}}^*_{\mathrm{DC}}(d)$ we introduce one dual variable per constraint of $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$:

Primal constraint of $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$	Sign	Dual variable
Stationarity (28b): $\lambda_i = \mu^+_{g_i} - \mu^-_{g_i}$	eq.	$x_{g_i} \in \mathbb{R}$
Stationarity (28c): $\lambda_i = 1 - \mu^-_{s_i} + \mu^+_{s_i}$	eq.	$x_{s_i} \in \mathbb{R}$
Stationarity (28d) (flow balance)	eq.	$x_{\theta_i} \in \mathbb{R}$
$\mu^- \geqslant 0$	$\geqslant 0$	$\xi^- \geqslant 0$
$\mu^+ \geqslant 0$	$\geqslant 0$	$\xi^+ \geqslant 0$
$\phi^- \geqslant 0$	$\geqslant 0$	$\psi^- \geqslant 0$
$\phi^+ \geqslant 0$	$\geqslant 0$	$\psi^+ \geqslant 0$
Big-M on w_{ij} (26a) ($\lvert w_{ij}\rvert \leqslant M_\lambda a_{ij}$)	$\geqslant 0$	$\rho^1_{ij}, \rho^2_{ij} \geqslant 0$
Big-M on w_{ij} (26a) ($\lvert w_{ij} - \Delta\lambda_{ij}\rvert \leqslant M_\lambda(1 - a_{ij})$)	$\geqslant 0$	$\rho^3_{ij}, \rho^4_{ij} \geqslant 0$
Big-M on v_{ij} (26b) ($\lvert v_{ij}\rvert \leqslant M_\phi a_{ij}$)	$\geqslant 0$	$\sigma^1_{ij}, \sigma^2_{ij} \geqslant 0$
Big-M on v_{ij} (26b) ($\lvert v_{ij} - \Delta\phi_{ij}\rvert \leqslant M_\phi(1 - a_{ij})$)	$\geqslant 0$	$\sigma^3_{ij}, \sigma^4_{ij} \geqslant 0$
$0 \leqslant a_{ij} \leqslant 1$	$\geqslant 0$	$\kappa^0_{ij}, \kappa^1_{ij} \geqslant 0$
$\mathbf{1}^\top a \leqslant A$	$\geqslant 0$	$\alpha \geqslant 0$

The multipliers $\rho^{1:4}_{ij}$ are assigned, in order, to the nonnegative residuals $M_\lambda a_{ij} - w_{ij}$, $M_\lambda a_{ij} + w_{ij}$, $M_\lambda(1 - a_{ij}) - w_{ij} + \Delta\lambda_{ij}$, and $M_\lambda(1 - a_{ij}) + w_{ij} - \Delta\lambda_{ij}$. The same ordering applies to $\sigma^{1:4}_{ij}$ with $(M_\phi, v, \Delta\phi)$. The upper Lagrangian uses $x^\top(e_s - \text{stationarity left-hand side})$, with the compact sign convention of (8b); equivalently, the theta rows displayed in (28d) enter with a positive x_{θ_i} multiplier.

Stationarity of the Lagrangian of $\bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$ with respect to each primal variable:

w.r.t. λ_i (appears in objective via $c_i\lambda_i$, in equalities (28b)–(28d); λ is free):

$$c_i - (B_0 x)_i + \sum_{j:(ij)\in\mathcal{E}} (\rho^3_{ij} - \rho^4_{ij}) - \sum_{j:(ji)\in\mathcal{E}} (\rho^3_{ji} - \rho^4_{ji}) = 0. \tag{30}$$

w.r.t. μ^-:

$$x^{\min} - x + \xi^- = 0, \quad \xi^- \geqslant 0 \;\Rightarrow\; x \geqslant x^{\min}. \tag{31}$$

w.r.t. μ^+:

$$-x^{\max} + x + \xi^+ = 0, \quad \xi^+ \geqslant 0 \;\Rightarrow\; x \leqslant x^{\max}. \tag{32}$$

Conditions (31)–(32) together recover the box constraint $x^{\min} \leqslant x \leqslant x^{\max}$ (2a) ; the balance equation is (30), with $B_0 = B(\mathbf{1})$.

w.r.t. ϕ^-_{ij}:

$$-f_{ij} - [F_0 x]_{ij} + \sigma^3_{ij} - \sigma^4_{ij} + \psi^-_{ij} = 0, \quad \psi^-_{ij} \geqslant 0 \;\Rightarrow\; [F_0 x]_{ij} - \sigma^3_{ij} + \sigma^4_{ij} \geqslant -f_{ij}. \tag{33}$$

w.r.t. ϕ^+_{ij}:

$$-f_{ij} + [F_0 x]_{ij} - \sigma^3_{ij} + \sigma^4_{ij} + \psi^+_{ij} = 0, \quad \psi^+_{ij} \geqslant 0 \;\Rightarrow\; [F_0 x]_{ij} - \sigma^3_{ij} + \sigma^4_{ij} \leqslant f_{ij}. \tag{34}$$

Conditions (33)–(34) together recover the line-flow constraint $-f \leqslant F_0 x - \sigma^3 + \sigma^4 \leqslant f$, with $F_0 = F(\mathbf{1})$.

w.r.t. w_{ij} (appears in (28d) via $-(1 - d_{ij})w_{ij}$ and in Big-M constraints (26a)):

$$\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} - \rho^4_{ij} = -Y_{ij}(1 - d_{ij})(x_{\theta_i} - x_{\theta_j}). \tag{35}$$

w.r.t. v_{ij} (appears in (28d) via $+(1 - d_{ij})v_{ij}$ and in Big-M constraints (26b)):

$$\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} - \sigma^4_{ij} = Y_{ij}(1 - d_{ij})(x_{\theta_i} - x_{\theta_j}). \tag{36}$$

w.r.t. a_{ij} (appears in Big-M constraints on w_{ij} and v_{ij} and in the box-and-budget constraints):

$$\alpha + (\kappa^1_{ij} - \kappa^0_{ij}) + M_\lambda(-\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} + \rho^4_{ij}) + M_\phi(-\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} + \sigma^4_{ij}) = 0. \tag{37}$$

The dual objective (in minimization form) is:

$$\bar{\mathcal{P}}^*_{\mathrm{DC}}(d) = \min_{x,\rho,\sigma,\kappa,\alpha} e_s^\top x + \alpha A + \sum_{(ij)\in\mathcal{E}} \big[M_\lambda(\rho^3_{ij} + \rho^4_{ij}) + M_\phi(\sigma^3_{ij} + \sigma^4_{ij}) + \kappa^1_{ij}\big]. \tag{10}$$

By Lemma 4.1, $\bar{\mathcal{P}}^*_{\mathrm{DC}}(d) = \bar{\mathcal{D}}^*_{\mathrm{DC}}(d)$. Note that (31)–(34) recover the box bounds and the modified flow constraints; the modified balance is (30). The variables $\xi^\pm$ and $\psi^\pm$ are eliminated algebraically.

Step 4: Defender Linearization

Equations (35)–(36) contain the product $d_{ij}(x_{\theta_i} - x_{\theta_j})$. Define $y_{ij} := d_{ij}(x_{\theta_i} - x_{\theta_j})$. Since $x_{\theta_i} - x_{\theta_j} \in [\theta_i^{\min} - \theta_j^{\max},\, \theta_i^{\max} - \theta_j^{\min}]$ from (2a), setting $M_y := \max_{(ij)\in\mathcal{E}} \max\{|\theta_i^{\min} - \theta_j^{\max}|, |\theta_i^{\max} - \theta_j^{\min}|\}$ makes Big-M linearization exact:

$$-M_y\, d_{ij} \leqslant y_{ij} \leqslant M_y\, d_{ij}, \quad (x_{\theta_i} - x_{\theta_j}) - M_y(1 - d_{ij}) \leqslant y_{ij} \leqslant (x_{\theta_i} - x_{\theta_j}) + M_y(1 - d_{ij}). \tag{38}$$

Substituting y_{ij} into (35)–(36) and merging with the outer defender's minimization over $d \in \{0,1\}^m$, $\mathbf{1}^\top d \leqslant D$ recovers the SL-MILP from $\widehat{\mathcal{P}}^*_{DC}(d)$ (Definition 4.2).

Remark C.1. Constraints (11a)–(11b) are recovered exactly from the stationarity conditions (31)–(34) of the second dual, with the auxiliary variables $\xi^{\pm}$ and $\psi^{\pm}$ eliminated algebraically. Constraint (11c) is the nodal power-balance stationarity (30). Constraints (11d)–(11e) are obtained from (35)–(36) after substituting $y_{ij} := d_{ij}(x_{\theta_i} - x_{\theta_j})$, so that $(1 - d_{ij})(x_{\theta_i} - x_{\theta_j}) = (x_{\theta_i} - x_{\theta_j}) - y_{ij}$. Note that (11d) and (11e) have opposite right-hand sides; the two sets of multipliers $\rho_{ij}^{1:4}$ and $\sigma_{ij}^{1:4}$ are nevertheless distinct, as they dual the Big-M constraints on w_{ij} and v_{ij} respectively.

D Derivation of the SL-MICP (AC-OPF)

Step 1: Conic Dual of the Inner SOCP

Primal. For fixed $z \in \{0,1\}^m$, the inner problem is the SOCP (AC-OPF)–(5e) with $x = (g^P, g^Q, s^P, s^Q, u, \Phi, \Psi) \in \mathbb{R}^N$ and epigraph variable $t \in \mathbb{R}$, and Lemma 2.2 gives $\mathcal{P}^*_{\mathrm{AC}}(z) = \mathcal{D}^*_{\mathrm{AC}}(z)$.

Lagrangian. Form the Lagrangian with $\nu \in \mathbb{R}^{2n}$ for the equalities (5b), $\mu^- \in \mathbb{R}^N_{\geqslant 0}$ for $x \geqslant x^{\min}$, $\mu^+ \in \mathbb{R}^N_{\geqslant 0}$ for $x \leqslant x^{\max}$, $\phi^- \in \mathbb{R}^{m_H}_{\geqslant 0}$ for $H(z)x \geqslant -f$, $\phi^+ \in \mathbb{R}^{m_H}_{\geqslant 0}$ for $H(z)x \leqslant f$, $s_{ij} \in \mathcal{K}_4$ for the SOC constraints (5d), and $r = (w, \alpha_c, \beta_c) \in \mathcal{K}_{N+2}$ for the cost epigraph (5e):

$$\begin{aligned}\mathcal{L}_{\mathrm{AC}} \;&= \tilde{c}^\top x + t + \nu^\top\big(b - A(z)x\big) - (\mu^-)^\top(x - x^{\min}) + (\mu^+)^\top(x - x^{\max}) \\ &\quad - (\phi^-)^\top\big(H(z)x + f\big) + (\phi^+)^\top\big(H(z)x - f\big) - \textstyle\sum_{(ij)\in\mathcal{E}} s_{ij}^\top E_{ij} x \\ &\quad - \big[w^\top(2J^{1/2}x) + \alpha_c(t-1) + \beta_c(t+1)\big].\end{aligned} \tag{39}$$

Stationarity w.r.t. t.

$$1 - \alpha_c - \beta_c = 0 \;\Rightarrow\; \alpha_c + \beta_c = 1. \tag{40}$$

The dual cone constraint on $r \in \mathcal{K}_{N+2}$ requires $\|w\|^2 + \alpha_c^2 \leqslant \beta_c^2$. Substituting $\beta_c = 1 - \alpha_c$ gives $\|w\|^2 \leqslant 1 - 2\alpha_c$, and maximizing $(\alpha_c - \beta_c) = 2\alpha_c - 1$ subject to this constraint yields the contribution $-\|w\|^2$ to the dual objective.

Stationarity w.r.t. x. Setting $\nabla_x \mathcal{L}_{\mathrm{AC}} = 0$ gives the single compact condition

$$A(z)^\top\nu + H(z)^\top(\phi^- - \phi^+) + \textstyle\sum_{(ij)\in\mathcal{E}} E_{ij}^\top s_{ij} + 2J^{1/2}w - \mu^+ + \mu^- = \tilde{c}. \tag{12}$$

Dual objective. Taking $\inf_{x,t} \mathcal{L}_{\mathrm{AC}}$ subject to stationarity and collecting constant terms yields

$$\mathcal{D}^*_{\mathrm{AC}}(z) = \max\Big\{ b^\top\nu + (x^{\min})^\top\mu^- - (x^{\max})^\top\mu^+ - f^\top(\phi^- + \phi^+) - \|w\|^2 \Big\}, \tag{41}$$

subject to (12), $s_{ij} \in \mathcal{K}_4$, $\mu^-, \mu^+, \phi^-, \phi^+ \geqslant 0$, ν free.

Step 2: Merging and Attacker Linearization

Substituting $z_{ij} = 1 - a_{ij}(1 - d_{ij})$ into (12) via $A(z)^\top\nu = \widetilde{A}^\top\nu - \sum_{(ij)}(1 - d_{ij})\widehat{A}_{ij}^\top(a_{ij}\nu)$ and $H(z)^\top(\phi^- - \phi^+) = \widetilde{H}^\top(\phi^- - \phi^+) - \sum_{(ij)}(1 - d_{ij})\widehat{H}_{ij}^\top(a_{ij}(\phi_{ij}^- - \phi_{ij}^+))$ introduces two families of bilinear terms involving the attacker decision a_{ij}:

$$a_{ij}\,\Delta\nu_{ij} \quad \text{where} \quad \Delta\nu_{ij} := \widehat{A}_{ij}^\top\nu \in \mathbb{R}^N, \qquad \text{and} \qquad a_{ij}\,\Delta\phi_{ij} \quad \text{where} \quad \Delta\phi_{ij} := \phi_{ij}^- - \phi_{ij}^+. \tag{42}$$

We introduce auxiliary variables $w_{ij}^A := a_{ij}\Delta\nu_{ij}$ and $v_{ij} := a_{ij}\Delta\phi_{ij}$, and apply exact Big-M linearization (Definition A.4) with sufficiently large M-values ($M_\nu \geqslant \sup\|\Delta\nu_{ij}\|_\infty$ and $M_\phi \geqslant \sup\|\Delta\phi_{ij}\|_\infty$ over the selected optimal dual solutions):

$$|w_{ij}^A| \leqslant M_\nu\, a_{ij}, \qquad w_{ij}^A - \Delta\nu_{ij} \in [-M_\nu(1-a_{ij}),\ M_\nu(1-a_{ij})], \tag{43a}$$

$$|v_{ij}| \leqslant M_\phi\, a_{ij}, \qquad v_{ij} - \Delta\phi_{ij} \in [-M_\phi(1-a_{ij}),\ M_\phi(1-a_{ij})]. \tag{43b}$$

The stationarity constraint (12) becomes

$$\begin{aligned}\widetilde{A}^\top\nu - \textstyle\sum_{(ij)}(1-d_{ij})w_{ij}^A + \widetilde{H}^\top(\phi^- - \phi^+) \\ - \textstyle\sum_{(ij)}(1-d_{ij})\widehat{H}_{ij}^\top v_{ij} + \sum_{(ij)} E_{ij}^\top s_{ij} \\ + 2J^{1/2}w - \mu^+ + \mu^- = \tilde{c}.\end{aligned} \tag{44}$$

With $a_{ij} \in \{0,1\}$ and $\mathbf{1}^\top a \leqslant A$, the merged attacker-operator problem is

$$\widetilde{\mathcal{D}}_{\mathrm{AC}}^*(d) := \max_{a,\nu,\mu^-,\mu^+,\phi^-,\phi^+,w^A,v,s,w} b^\top\nu + (x^{\min})^\top\mu^- - (x^{\max})^\top\mu^+ - f^\top(\phi^- + \phi^+) - \|w\|^2 \tag{45a}$$

$$\text{s.t. constraint (44)},$$

$$\text{Big-M constraints (43a)–(43b)},$$

$$s_{ij} \in \mathcal{K}_4, \quad \forall(ij) \in \mathcal{E}, \tag{45b}$$

$$\mu^-,\ \mu^+,\ \phi^-,\ \phi^+ \geqslant 0, \qquad \nu \text{ : free}, \tag{45c}$$

$$a_{ij} \in \{0,1\}, \quad \forall(ij) \in \mathcal{E}, \qquad \mathbf{1}^\top a \leqslant A. \tag{45d}$$

Step 3: Conic Relaxation and Dual of the Merged Problem

Relaxing $a_{ij} \in [0,1]$ we get the relaxed conic program $\bar{\mathcal{D}}_{\mathrm{AC}}^*(d)$:

$$\bar{\mathcal{D}}_{\mathrm{AC}}^*(d) := \max_{a,\nu,\mu^-,\mu^+,\phi^-,\phi^+,w^A,v,s,w} b^\top\nu + (x^{\min})^\top\mu^- - (x^{\max})^\top\mu^+ - f^\top(\phi^- + \phi^+) - \|w\|^2 \tag{46a}$$

$$\text{s.t. constraint (44)},$$

$$\text{Big-M constraints (43a)–(43b)},$$

$$\text{constraints (45b)–(45c)},$$

$$a_{ij} \in [0,1], \quad \forall(ij) \in \mathcal{E}, \qquad \mathbf{1}^\top a \leqslant A. \tag{46b}$$

Since $\mathcal{A}_{\mathrm{bin}} \subset \mathcal{A}_{\mathrm{rel}}$, we have $\bar{\mathcal{D}}_{\mathrm{AC}}^*(d) \geqslant \widetilde{\mathcal{D}}_{\mathrm{AC}}^*(d)$ for every feasible d. By Lemma 4.3, $\bar{\mathcal{P}}_{\mathrm{AC}}^*(d) = \bar{\mathcal{D}}_{\mathrm{AC}}^*(d)$, and we now form the dual $\bar{\mathcal{P}}_{\mathrm{AC}}^*(d)$ by introducing one dual variable per constraint of $\bar{\mathcal{D}}_{\mathrm{AC}}^*(d)$:

Primal constraint of $\bar{\mathcal{D}}_{\mathrm{AC}}^*(d)$	Sign	Dual variable
Equality (44)	eq.	$x \in \mathbb{R}^N$
$\mu^- \geqslant 0$	$\geqslant 0$	$\xi^- \geqslant 0$
$\mu^+ \geqslant 0$	$\geqslant 0$	$\xi^+ \geqslant 0$
$\phi^- \geqslant 0$	$\geqslant 0$	$\psi^- \geqslant 0$
$\phi^+ \geqslant 0$	$\geqslant 0$	$\psi^+ \geqslant 0$
$s_{ij} \in \mathcal{K}_4$ (45b)	cone	$k_{ij} \in \mathcal{K}_4$
Big-M on w_{ij}^A (43a) ($\lvert w_{ij}^A\rvert \leqslant M_\nu a_{ij}$)	$\geqslant 0$	$\rho_{ij}^1, \rho_{ij}^2 \geqslant 0$
Big-M on w_{ij}^A (43a) ($\lvert w_{ij}^A - \Delta\nu_{ij}\rvert \leqslant M_\nu(1-a_{ij})$)	$\geqslant 0$	$\rho_{ij}^3, \rho_{ij}^4 \geqslant 0$
Big-M on v_{ij} (43b) ($\lvert v_{ij}\rvert \leqslant M_\phi a_{ij}$)	$\geqslant 0$	$\sigma_{ij}^1, \sigma_{ij}^2 \geqslant 0$
Big-M on v_{ij} (43b) ($\lvert v_{ij} - \Delta\phi_{ij}\rvert \leqslant M_\phi(1-a_{ij})$)	$\geqslant 0$	$\sigma_{ij}^3, \sigma_{ij}^4 \geqslant 0$
$0 \leqslant a_{ij} \leqslant 1$ (46b)	$\geqslant 0$	$\kappa_{ij}^0, \kappa_{ij}^1 \geqslant 0$
$\mathbf{1}^\top a \leqslant A$ (46b)	$\geqslant 0$	$\alpha \geqslant 0$

The envelope multipliers use the component-wise residual ordering specified in the DC derivation, with $(M_\nu, w^A, \Delta\nu)$ and $(M_\phi, v, \Delta\phi)$ respectively. The upper Lagrangian uses $x^\top(\tilde{c} -$ stationarity left-hand side$)$.

Stationarity of the Lagrangian of $\bar{\mathcal{D}}^*_{\mathrm{AC}}(d)$ with respect to each primal variable:

w.r.t. ν (appears in objective via $b^\top\nu$, in equality (44) via $\widetilde{A}^\top\nu$; ν is free):

$$b - \widetilde{A}x + \sum_{(ij)\in\mathcal{E}} \widehat{A}_{ij}(\rho^3_{ij} - \rho^4_{ij}) = 0 \;\Rightarrow\; \widetilde{A}x - \sum_{(ij)\in\mathcal{E}} \widehat{A}_{ij}(\rho^3_{ij} - \rho^4_{ij}) = b. \tag{47}$$

w.r.t. μ^- (appears in objective via $+(x^{\min})^\top\mu^-$, in equality (44) via $+\mu^-$, and $\mu^- \geqslant 0$ dual $\xi^- \geqslant 0$):

$$x^{\min} - x + \xi^- = 0, \quad \xi^- \geqslant 0 \;\Rightarrow\; x \geqslant x^{\min}. \tag{48}$$

w.r.t. μ^+ (appears in objective via $-(x^{\max})^\top\mu^+$, in equality (44) via $-\mu^+$, and $\mu^+ \geqslant 0$ dual $\xi^+ \geqslant 0$):

$$-x^{\max} + x + \xi^+ = 0, \quad \xi^+ \geqslant 0 \;\Rightarrow\; x \leqslant x^{\max}. \tag{49}$$

Conditions (48)–(49) together recover the box constraint $x^{\min} \leqslant x \leqslant x^{\max}$ (5a); $\xi^\pm$ are eliminated.

w.r.t. ϕ^-_{ij}:

$$-f_{ij} - \widetilde{H}_{ij}x + \sigma^3_{ij} - \sigma^4_{ij} + \psi^-_{ij} = 0, \quad \psi^-_{ij} \geqslant 0 \;\Rightarrow\; \widetilde{H}_{ij}x - \sigma^3_{ij} + \sigma^4_{ij} \geqslant -f_{ij}. \tag{50}$$

w.r.t. ϕ^+_{ij}:

$$-f_{ij} + \widetilde{H}_{ij}x - \sigma^3_{ij} + \sigma^4_{ij} + \psi^+_{ij} = 0, \quad \psi^+_{ij} \geqslant 0 \;\Rightarrow\; \widetilde{H}_{ij}x - \sigma^3_{ij} + \sigma^4_{ij} \leqslant f_{ij}. \tag{51}$$

Conditions (50)–(51) together recover the modified flow constraint $-f_{ij} \leqslant \widetilde{H}_{ij}x - \sigma^3_{ij} + \sigma^4_{ij} \leqslant f_{ij}$; $\psi^\pm$ are eliminated.

w.r.t. s_{ij} (appears in equality (44) via $+E^\top_{ij}s_{ij}$ and in cone constraint $s_{ij} \in \mathcal{K}_4$ dual $k_{ij} \in \mathcal{K}_4$):

$$E_{ij}x = k_{ij} \in \mathcal{K}_4, \quad \forall (ij) \in \mathcal{E}. \tag{52}$$

This recovers the SOCP relaxation constraint (5d).

w.r.t. w (appears in equality (44) via $+2J^{1/2}w$ and in the cost-epigraph cone):

$$-2J^{1/2}x - 2w = 0 \;\Rightarrow\; w = -J^{1/2}x, \text{ so } \|w\|^2 = x^\top Jx, \text{ and } [2J^{1/2}x;\ t-1;\ t+1] \in \mathcal{K}_{N+2}. \tag{53}$$

This recovers the cost-epigraph constraint (5e).

w.r.t. w^A_{ij} (appears in equality (44) via $-(1-d_{ij})w^A_{ij}$ and in Big-M constraints (43a)):

$$\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} - \rho^4_{ij} = (1-d_{ij})x. \tag{54}$$

w.r.t. v_{ij} (appears in equality (44) via $-(1-d_{ij})\widehat{H}^\top_{ij}v_{ij}$ and in Big-M constraints (43b)):

$$\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} - \sigma^4_{ij} = (1-d_{ij})\widehat{H}_{ij}x. \tag{55}$$

w.r.t. a_{ij} (appears in Big-M constraints on w^A_{ij} and v_{ij} and in the box-and-budget constraints):

$$\alpha + (\kappa^1_{ij} - \kappa^0_{ij}) + M_\nu \mathbf{1}^\top(-\rho^1_{ij} - \rho^2_{ij} + \rho^3_{ij} + \rho^4_{ij}) + M_\phi \mathbf{1}^\top(-\sigma^1_{ij} - \sigma^2_{ij} + \sigma^3_{ij} + \sigma^4_{ij}) = 0. \tag{56}$$

The dual objective (in minimization form) is:

$$\bar{\mathcal{P}}^*_{\mathrm{AC}}(d) = \min_{x,t,\rho,\sigma,\kappa,\alpha} \tilde{c}^\top x + t + \alpha A + \sum_{(ij)\in\mathcal{E}} \Big[M_\nu \mathbf{1}^\top(\rho^3_{ij} + \rho^4_{ij}) + M_\phi \mathbf{1}^\top(\sigma^3_{ij} + \sigma^4_{ij}) + \kappa^1_{ij} \Big]. \tag{57}$$

By Lemma 4.3, $\bar{\mathcal{P}}^*_{\mathrm{AC}}(d) = \bar{\mathcal{D}}^*_{\mathrm{AC}}(d)$. Note that conditions (48)–(53) recover the box and cone constraints and the modified flow constraints. The modified balance is (47); no attacker variable remains in this second dual. The variables $\xi^\pm$ and $\psi^\pm$ are eliminated algebraically.

Step 4: Defender Linearization

Equations (54)–(55) contain the product $d_{ij}x$, since $(1-d_{ij})x = x - d_{ij}x$. Define $y_{ij} := d_{ij}x \in \mathbb{R}^N$. Since $d_{ij} \in \{0,1\}$ and $x_k \in [x_k^{\min}, x_k^{\max}]$ component-wise from (48)–(49), setting $M_{x_k} := \max(|x_k^{\min}|, |x_k^{\max}|)$ makes Big-M linearization exact (Definition A.4, applied component-wise):

$$-M_x d_{ij} \leqslant y_{ij} \leqslant M_x d_{ij}, \quad x - M_x(1-d_{ij}) \leqslant y_{ij} \leqslant x + M_x(1-d_{ij}), \quad \forall (ij) \in \mathcal{E}, \tag{58}$$

where $M_x \in \mathbb{R}^N$ is the component-wise bound vector. Substituting $d_{ij}x = y_{ij}$ and $(1-d_{ij})x = x - y_{ij}$ into (47), (54), (55) makes all constraints linear, while (52) and (53) remain conic. Merging the minimization (57) subject to (47)–(58) with the outer defender's minimization over $d \in \{0,1\}^m$, $\mathbf{1}^\top d \leqslant D$ recovers the SL-MICP (Definition 4.4).

Remark D.1. Constraints (17a)–(17e) are recovered exactly from the stationarity conditions (48)–(51) of the second dual, with $\xi^{\pm}$ and $\psi^{\pm}$ eliminated algebraically. The SOC constraint (17b) and cost-epigraph constraint (17c) are recovered from the cone stationarity conditions (52)–(53). Constraints (17f)–(17g) are obtained from (54)–(55) after substituting $(1-d_{ij})x = x - y_{ij}$. The two sets of multipliers $\rho_{ij}^{1:4}$ and $\sigma_{ij}^{1:4}$ have the respective right-hand sides $x - y_{ij}$ and $\widehat{H}_{ij}(x - y_{ij})$ and remain distinct, dualizing the Big-M constraints on w_{ij}^A and v_{ij} respectively.

E Mathematical Proofs

E.1 Proof of Lemma 2.1

Re-statement: Strong duality holds for (DC-OPF).

Proof. Since $s_i \in [0, c_i]$, the objective in (DC-OPF) is bounded. And setting $s_i = c_i, g_i = \theta_i = 0$ satisfies all the constraints. Thus, the LP is bounded and feasible, strong duality holds, and we get $\mathcal{P}_{DC}^*(z) = \mathcal{D}_{DC}^*(z)$. □

E.2 Proof of Lemma 2.2

Re-statement: Strong duality holds for (AC-OPF).

Proof. Under the stated hypothesis, a feasible point satisfies the affine constraints and lies in the relative interior of the conic domain after fixed variables are eliminated. The operator objective is bounded below because x is in a finite box and $J \succeq 0$; the cost epigraph can be chosen strictly above $x^\top Jx$. Slater's theorem, therefore, gives zero duality gap and an attained dual optimum. □

E.3 Proof of Lemma 4.1

Re-statement: Strong duality holds for the relaxed LP: $\bar{\mathcal{D}}_{\mathrm{DC}}^*(d)$ (29a)–(29c).

Proof. Take $a = w = v = \lambda = \phi^- = \phi^+ = 0$, $\mu^- = e_s$, and $\mu^+ = 0$. These values satisfy all constraints for every $A \geqslant 0$. Moreover, $\lambda_i + \mu_{s_i}^- - \mu_{s_i}^+ = 1$ gives $c_i\lambda_i - c_i\mu_{s_i}^+ = c_i(1 - \mu_{s_i}^-) \leqslant c_i$. All remaining terms of the corrected dual objective are nonpositive under $\theta^{\min} \leqslant 0 \leqslant \theta^{\max}$, $g^{\max}, f \geqslant 0$. Thus the relaxed LP is feasible and bounded above by $\sum_i c_i$, and strong LP duality yields $\bar{\mathcal{D}}_{\mathrm{DC}}^*(d) = \bar{\mathcal{P}}_{DC}^*(d)$. □

E.4 Proof of Lemma 4.3

Re-statement: Strong duality holds for the relaxed CP: $\bar{\mathcal{D}}_{\mathrm{AC}}^*(d)$ (46a)–(46b).

Proof. For $m > 0$ and some $\epsilon > 0$, take

$$\begin{aligned} &a_{ij} = \min\{1/2, A/(2m)\} \in (0,1) \quad (A > 0), \\ &\phi_{ij}^- = \phi_{ij}^+ = \epsilon, \qquad \nu = 0, \qquad w^A = v = w = 0, \\ &s_{ij} \in \mathrm{ri}(\mathcal{K}_4). \end{aligned}$$

Then the relaxed attacker box and budget constraint (46b), the cone/nonnegativity constraints (45b)–(45c) (except $\mu^{\pm}$), and the Big-M constraints (43a)–(43b) are all satisfied strictly for sufficiently large M. The remaining stationarity equality (44) is affine in the nonnegative variables $\mu^{\pm}$. So given the other values, one can always choose some $\mu_i^{\pm} > 0$ so that the equation holds exactly ($\cdot$ for any $t \in \mathbb{R}$, choose $\mu_i^- = 1 + \max(t, 0)$ and $\mu_i^+ = 1 + \max(-t, 0)$, so that $\mu_i^- - \mu_i^+ = t$).

For $m = 0$, the attacker variables are absent. For $A = 0$, eliminate the fixed variables $a = w^A = v = 0$ and use the same construction for the remaining conic variables. Zero envelope constants can likewise be handled by eliminating their fixed expressions; otherwise choose positive valid constants. Boundedness follows by evaluating the upper Lagrangian at any feasible intact-network operator point: the box, flow, and SOC terms are nonpositive, the quadratic term has a finite supremum, and the attacker correction terms are bounded by the envelopes. Thus, the relaxed CP satisfies the relative Slater condition [31], hence strong duality holds and we get

$$\bar{\mathcal{D}}^*_{\mathrm{AC}}(d) = \bar{\mathcal{P}}^*_{\mathrm{AC}}(d).$$

□

E.5 Proof of Theorem 5.1

Re-statement: Let S^* be the optimal value of the tri-level problem (TL) and S_{SL} the optimal value of the (SL-MILP) (or (SL-MICP)). Then we have $S_{\mathrm{SL}} \geqslant S^*$.

Proof. By Lemmas 2.1 and 2.2, the inner operator problem satisfies

$$\mathcal{P}^*(z) = \mathcal{D}^*(z), \qquad \forall z \in \{0,1\}^m.$$

For each fixed d, the first Big-M linearization is exact, hence

$$\widetilde{\mathcal{D}}^*(d) = \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{D}^*(z(a,d)).$$

Relaxing $a \in \mathcal{A}_{\mathrm{bin}}$ to $a \in \mathcal{A}_{\mathrm{rel}}$ enlarges the attacker's feasible set, so

$$\bar{\mathcal{D}}^*(d) \geqslant \widetilde{\mathcal{D}}^*(d).$$

Next, by Lemmas 4.1 and 4.3,

$$\bar{\mathcal{D}}^*(d) = \bar{\mathcal{P}}^*(d).$$

The second Big-M linearization is again exact, and therefore

$$S_{\mathrm{SL}} = \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \bar{\mathcal{P}}^*(d) = \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \bar{\mathcal{D}}^*(d).$$

Combining the above,

$$\begin{aligned} S_{\mathrm{SL}} &= \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \bar{\mathcal{D}}^*(d) \\ &\geqslant \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \widetilde{\mathcal{D}}^*(d) \\ &= \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{D}^*(z(a,d)) \\ &= \min_{\substack{d \in \{0,1\}^m \\ \mathbf{1}^\top d \leqslant D}} \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a,d)) \\ &= S^*. \end{aligned}$$

Hence $S_{\mathrm{SL}} \geqslant S^*$. □

E.6 Proof of Theorem 5.2

Re-statement: Let d^*_{SL} be an optimal defence from the (SL-MILP) (or (SL-MICP)), and define the *realized worst-case shedding*

$$S_{\mathrm{true}}(d^*_{\mathrm{SL}}) := \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a, d^*_{\mathrm{SL}})).$$

Then $S_{\mathrm{SL}} \geqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \geqslant S^*$.

Proof. From the proof of Theorem 5.1, for every feasible defense d we have

$$\bar{\mathcal{P}}^*(d) = \bar{\mathcal{D}}^*(d) \geqslant \widetilde{\mathcal{D}}^*(d) = \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a, d)).$$

Evaluating at $d = d^*_{\mathrm{SL}}$ and using $S_{\mathrm{SL}} = \bar{\mathcal{P}}^*(d^*_{\mathrm{SL}})$ gives

$$S_{\mathrm{SL}} \geqslant \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a, d^*_{\mathrm{SL}})) = S_{\mathrm{true}}(d^*_{\mathrm{SL}}).$$

On the other hand, by definition of the tri-level optimum,

$$S^* = \min_{d} \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a, d)).$$

Since d^*_{SL} is one feasible defense,

$$S_{\mathrm{true}}(d^*_{\mathrm{SL}}) = \max_{a \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(a, d^*_{\mathrm{SL}})) \geqslant S^*.$$

Combining the two inequalities yields

$$S_{\mathrm{SL}} \geqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \geqslant S^*.$$

□

E.7 Proof of Theorem 5.4

Re-statement: Under the defence d^*_{SL}, no binary attack $a \in \mathcal{A}_{\mathrm{bin}}$ can cause shedding exceeding S_{SL}: $\mathcal{P}^*(z(a, d^*_{\mathrm{SL}})) \leqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \leqslant S_{\mathrm{SL}}$.

Proof. For any fixed $a \in \mathcal{A}_{\mathrm{bin}}$,

$$\mathcal{P}^*(z(a, d^*_{\mathrm{SL}})) \leqslant \max_{\hat{a} \in \mathcal{A}_{\mathrm{bin}}} \mathcal{P}^*(z(\hat{a}, d^*_{\mathrm{SL}})) = S_{\mathrm{true}}(d^*_{\mathrm{SL}}).$$

Further, by Theorem 5.2,

$$S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \leqslant S_{\mathrm{SL}}.$$

Combining the two inequalities gives

$$\mathcal{P}^*(z(a, d^*_{\mathrm{SL}})) \leqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \leqslant S_{\mathrm{SL}}.$$

□

F Additional Experimental Results

F.1 Realized binary-attack certificates

For a fixed returned defense d^*_{SL}, we solve the exact binary attacker-operator problem once to obtain $S_{\mathrm{true}}(d^*_{\mathrm{SL}})$. This is the quantity that certifies the deployed defense without the fractional-attacker conservatism. Exact tri-level enumeration, when tractable, additionally gives S^* and hence Δ_0.

The large Δ_C values at $D = 3, 4$ in Table 7 make the fractional-attacker conservatism explicit. They do not invalidate the certificate: the post-hoc value is smaller and is exact for the returned defense within the SOCP model.

Table 5: **Certificate-gap decomposition for the IEEE 14-bus DC case with $A = 2$.** The SL-MILP runtime is included for reference, all load shed values are in MW.

D	S_{SL}	$S_{\mathrm{true}}(d^*_{\mathrm{SL}})$	Δ_C	S^*	Δ_0	Runtime (s)
0	14.9	14.9	0.0	14.9	0.0	0.007
1	12.5	12.5	0.0	12.5	0.0	0.018
2	12.5	9.0	3.5	6.1	2.9	0.042
3	6.2	6.1	0.1	3.5	2.6	0.048

Table 6: **IEEE 118-bus DC extension.** All reported instances satisfy the chain $S_{\mathrm{SL}} \geqslant S_{\mathrm{true}}(d^*_{\mathrm{SL}}) \geqslant S^*$, all load shed values are in MW.

D	S_{SL}	$S_{\mathrm{true}}(d^*_{\mathrm{SL}})$	Δ_C	S^*	Δ_0
0	110.0	110.0	0.0	110.0	0.0
1	105.2	104.8	0.4	104.0	0.8
2	51.6	49.3	2.3	48.1	1.2
3	45.0	43.6	1.4	42.1	1.5
4	44.8	43.1	1.7	42.0	1.1
5	43.8	42.3	1.5	41.0	1.3

Table 7: **Realized load-shedding gaps for the WSCC 9-bus SL-MICP defenses in Table 4.** The post-hoc and tri-level values are obtained by binary enumeration under the SOCP operator model. All load shed values are in MW.

D	S_{SL}	$S_{\mathrm{true}}(d^*_{\mathrm{SL}})$	Δ_C	S^*	Δ_0	Worst attack on d^*_{SL}	Enumerated tri-level d^*
0	125.0	125.00	0.00	125.00	0.00	8–9, 9–4	—
1	101.3	100.00	1.30	100.00	0.00	6–7, 7–8	8–9
2	97.3	90.00	7.30	90.00	0.00	4–5, 5–6	6–7, 8–9
3	66.1	12.08	54.02	12.08	0.00	3–6, 8–2	4–5, 6–7, 8–9
4	65.7	12.08	53.62	0.00	12.08	3–6, 8–2	4–5, 3–6, 6–7, 8–9
5	0.0	0.00	0.00	0.00	0.00	—	4–5, 3–6, 6–7, 8–9

Table 8: **Load-shedding-only SL-MICP on larger IEEE systems.** Runtime is the single-level solve time, all load shed values are in MW.

System	D	Protected lines d^*_{SL}	S_{SL}	$S_{\mathrm{true}}(d^*_{\mathrm{SL}})$	Δ_C	S^*	Δ_0	Worst-case attack	Runtime (s)
14-bus	0	—	14.9	14.9	0.0	14.9	0.0	9–14, 13–14	0.50
	1	9–14	12.5	12.5	0.0	12.5	0.0	6–11, 9–10	0.53
	2	9–10, 9–14	12.5	9.0	3.5	6.1	2.9	6–12, 12–13	0.77
	3	6–12, 9–10, 9–14	6.2	6.1	0.1	3.5	2.6	6–11, 10–11	0.91
30-bus	0	—	30.0	30.0	0.0	30.0	0.0	6–8, 8–28	2.63
	1	6–8	24.0	22.8	1.2	21.8	1.0	2–5, 6–7	2.98
	2	6–7, 6–8	19.5	17.5	2.0	17.3	0.2	10–21, 21–22	3.01
	3	6–7, 6–8, 10–21	16.5	14.9	1.6	12.8	2.1	15–18, 10–20	3.83
57-bus	0	—	43.0	43.0	0.0	43.0	0.0	1–16, 12–16	10.07
	1	1–16	42.5	42.0	0.5	41.3	0.7	1–17, 12–17	11.06
	2	1–16, 1–17	40.5	39.0	1.5	37.1	1.9	49–50, 10–51	11.70
	3	1–16, 1–17, 49–50	37.5	35.8	1.7	32.8	3.0	29–52, 9–55	12.55

F.2 Larger SOCP-AC systems

We additionally run the load-shedding-only SL-MICP on IEEE 14, 30, and 57-bus systems with $A = 2$. In this experiment the generation-cost coefficients are set to zero, so all reported certificates are in MW. Exact enumeration is used for the columns $S_{\mathrm{true}}(d^*_{\mathrm{SL}})$ and S^* in these cases.

F.3 SOCP versus nonlinear AC-OPF on the 9-bus base case

We compare the SOCP operator with a standard nonlinear AC-OPF solve on the intact WSCC 9-bus base case. Both return zero shedding and generation cost \$5296.67 per hour. The nonlinear solve is local rather than a certified global solution, so this comparison assesses empirical tightness only and does not extend Theorems 5.1–5.4 to nonconvex AC-OPF. Active-power dispatch, voltage magnitudes, and phase angles are close; reactive-power dispatch shows larger differences at Generators 2 and 3.

Table 9: **Generator dispatch: nonlinear AC-OPF versus SOCP.**

Generator	NLP p_g (MW)	NLP q_g (MVar)	SOCP p_g (MW)	SOCP q_g (MVar)
Gen 1	89.80	12.97	89.80	13.00
Gen 2	134.32	0.03	134.33	5.11
Gen 3	94.19	-22.63	94.18	2.78

Table 10: **Voltage profiles: nonlinear AC-OPF versus SOCP.**

Bus	NLP $\lvert v_i \rvert$	SOCP $\lvert v_i \rvert$	NLP θ_i (rad)	SOCP θ_i (rad)
1	1.100	1.100	0.000	0.000
2	1.097	1.099	0.085	0.084
3	1.087	1.093	0.057	0.059
4	1.094	1.094	-0.043	-0.043
5	1.084	1.084	-0.069	-0.069
6	1.100	1.100	0.011	0.013
7	1.089	1.089	-0.021	-0.018
8	1.100	1.100	0.016	0.014
9	1.072	1.072	-0.081	-0.081

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction explicitly state the single-level reformulation pipeline, and the certified robustness guarantees, all of which are supported by theorems and experiments in the paper.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses the controlled attacker relaxation as the approximation step, the conservativeness of the resulting bound, AC-relaxations in Section 7.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: All theorems and results include the required assumptions, and complete proofs are provided in the appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: The IEEE benchmark systems and the MATPOWER library are openly available, and the SL-MILP and SL-MICP formulations are described in full (Section 4 and the appendix derivations), so they can be implemented with any standard solver (e.g., Gurobi, MOSEK).

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: The benchmark data are publicly available (IEEE systems via MATPOWER), all optimization formulations and solver settings are described in the paper, and specific parts of the code are available upon justified request.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: This work does not involve statistical training; all solver configurations, benchmark cases, attack/defense budgets, and evaluation protocols are explicitly specified.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The experiments are deterministic mixed-integer optimization runs on fixed IEEE benchmark systems rather than stochastic learning trials, so error bars are not applicable; we instead report exact runtimes and objective values.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The experimental section reports the optimization solver used, hardware platform, and runtime measurements.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The work involves mathematical optimization, benchmark simulation, and power-grid defense-security analysis, and does not involve deceptive, unsafe, or unethical experimental practices.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Section 7 discusses the positive implications for resilient grid security planning as well as possible dual-use concerns associated with adversarial infrastructure analysis.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release generative models, scraped datasets, or other high-risk assets requiring controlled-access safeguards.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: All IEEE benchmark test systems, optimization solvers, and prior formulations used for comparison are properly cited and credited.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.

- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The submission does not introduce a new dataset, pretrained model, or standalone software asset for public release.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing, annotation labor, or experiments with human participants.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: No human subjects are involved in this research.

    Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research?

    Answer: [Yes]

    Justification: LLM is only used for minor editing e.g., grammar, spelling, word choice.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.