Les Cahiers du GERAD ISSN: 0711–2440

Hierarchical constraint reduction for the penalized security-constrained optimal power flow

V. Darleguy, A. Lesage-Landry

G–2026–14

March 2026

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2026
– Bibliothèque et Archives Canada, 2026

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2026
– Library and Archives Canada, 2026

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Hierarchical constraint reduction for the penalized security-constrained optimal power flow

Victor Darleguy [a, b, c]

Antoine Lesage-Landry [a, b, c]

[a] *Polytechnique Montréal, Montréal (Qc), Canada, H3T 1J4*

[b] *Mila – Québec Artificial Intelligence Institute, Montréal, (Qc), Canada*

[c] *GERAD, Montréal (Qc), Canada, H3T 1J4*

victor.darleguyl@polymtl.ca
antoine.lesage-landry@gerad.ca

March 2026
Les Cahiers du GERAD
G–2026–14

Abstract : We consider the security-constrained optimal power flow (SCOPF) problem in a linearized form where thermal line limits are enforced as soft constraints to reflect operational flexibility. We propose a constraint-reduction method for the penalized SCOPF based on a nested hierarchy of non-redundant constraint subsets. Each hierarchical level defines a reduced SCOPF that is equivalent to the full problem while involving fewer constraints. We develop an extraction algorithm to construct this hierarchy from the unpenalized problem and introduce a cumulative encoding coupled with a supervised learning approach to predict the minimal hierarchy level corresponding to a load profile for the penalized SCOPF. Exactness is then guaranteed through an optional correction step. The method is illustrated in numerical simulations and hierarchical constraint subsets are presented. Using `LightGBM` predictors on the IEEE 39-bus system, we achieve a 94.28% hierarchy level prediction accuracy, resulting in an average computational time reduction of 19.72%, including the correction step.

Keywords: Electric power systems; security-constrained optimal power flow; soft constraints; constraint reduction

1 Introduction

Modern power systems face increasing complexity due to their large scale, the important integration of intermittent renewable resources, and the need for real-time decision-making compatible with market-based mechanisms [12]. In this context, the security-constrained optimal power flow (SCOPF) problem plays a central role in ensuring both economic efficiency and operational reliability [5]. The SCOPF is widely adopted in its linearized formulation based on power transfer distribution factors (PTDFs) to evaluate power flows subject to thermal line limits under different contingency scenarios [15, 17]. However, this formulation suffers from a large number of dense constraints that have a global impact on the optimization variables [12]. These yield an overly constrained representation of power system operations and important resolution times [2, 16]. Moreover, line limits are commonly enforced via hard constraints in the SCOPF, thereby limiting the ability of such models to capture the operational flexibility inherent to electricity markets. Contrastingly, industry practices often rely on soft rather than hard constraints [7, 15, 18], i.e., allowing for controlled power flow violations at explicit penalty costs.

Motivated by market applications, we consider the penalized SCOPF problem and propose a constraint-reduction method to improve its resolution time. We first define a nested hierarchy of non-redundant constraint subsets from the unpenalized problem that holds under any loading profile and propose an algorithm to extract them. We identify the minimal hierarchical level needed for equivalence with the full penalized SCOPF from its primal-dual optimum. We then devise a supervised learning methodology to infer to which hierarchical level a load profile corresponds to without the need for ex-post optimum information, so a reduced number of constraints is considered when solving the penalized problem. Finally, a correction step is included to ensure the resulting minimum matches the full penalized SCOPF's.

Constraint-reduction or screening methods for SCOPF limit the contingency-induced constraints to only ones that affect its minimum. Umbrella contingencies, i.e., constraints from contingency scenarios that ensure security with respect to all other scenarios, are defined in [4]. The authors of [2] proposes approaches to find all umbrella contingencies using a set of linear programs. In [13], an extension is proposed for further computational improvements in addition to considering the security-constrained unit commitment (SCUC) problem. The set of umbrella contingencies is estimated as a function of the loading profile using a multilayer perceptron in [1]. In [16], interval bounds on the optimization variables of the SCUC are first constructed using a parallelizable set of linear programs. Constraints that do not intersect with the bounds are then considered redundant. The Clarkson algorithm [8] which combines the linear programming test (`LP-test`) [9] with a `RayShoot` step, is employed in [20] to identify the minimal set of non-redundant constraints for the multi-period SCOPF. Nodal injection bounds are then used to remove additional constraints.

The aforementioned methods all rely on the assumption that thermal limit constraints are hard and cannot be violated, which is contrary to market practices [7, 15]. To the best of our knowledge, no existing method addresses constraint reduction for the penalized SCOPF. In this context, the set of non-redundant constraints becomes dependent on the loading profile as illustrated later in Figure 1. For this reason, we first propose a hierarchical non-redundant hard constraint set extraction algorithm and then an inference process to match a load profile to its corresponding minimal hierarchical level. We ensure exactness through an iterative correction step.

Next, we introduce the penalized SCOPF problem in Section 2. We present our constraint-reduction method in Section 3 and provide numerical examples in Section 4. Section 5 wraps up with closing remarks.

2 Penalized-SCOPF

We consider the linearized optimal power flow (OPF) problem expressed in terms of PTDFs. A detailed derivation of this form, including the computation of PTDFs, is provided in [6]. Consider a power grid modelled by the set of buses $\mathcal{N} \subset \mathbb{N}$ and the set of transmission lines $\mathcal{L} \subset \mathbb{N}$ indexed from 1 to $\operatorname{card}\mathcal{L}$. For all bus $i \in \mathcal{N}$, we define p_i as the nodal power injection or absorption, and $\underline{p}_i$ and $\overline{p}_i$ as the minimum and maximum generation, respectively, or as the load if $\underline{p}_i = \overline{p}_i < 0$. For all lines $j \in \mathcal{L}$, we let $\overline{s}_j$ be the thermal line limit. We express all nodal and line parameters and variables as the vectors $\mathbf{p}$, $\overline{\mathbf{p}}$, $\underline{\mathbf{p}} \in \mathbb{R}^{\operatorname{card}\mathcal{N}}$, and $\overline{\mathbf{s}} \in \mathbb{R}^{\operatorname{card}\mathcal{L}}$ where subscript $i \in \mathcal{N}$ or $j \in \mathcal{L}$ correspond to a bus or a line, respectively. For a line and bus couple $j, i \in \mathcal{L} \times \mathcal{N}$, we denote the corresponding PTDF as $\Phi_{j,i}$ and use $\mathbf{\Phi} \in \mathbb{R}^{\operatorname{card}\mathcal{L} \times \operatorname{card}\mathcal{N}}$ to denote all lines and bus sensitivity factors in matrix form. The linearized OPF problem is:

$$\min_{\mathbf{p} \in \mathbb{R}^{\operatorname{card}\mathcal{N}}} \quad f(\mathbf{p}) \tag{1a}$$

$$\text{s.t.} \quad \mathbf{1}^\top \mathbf{p} = 0 \tag{1b}$$

$$\underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} \tag{1c}$$

$$-\overline{\mathbf{s}} \leq \mathbf{\Phi}\mathbf{p} \leq \overline{\mathbf{s}}, \tag{1d}$$

where (1a) minimizes the active power generation with $f : \mathbb{R}^{\operatorname{card}\mathcal{N}} \to \mathbb{R}$ assumed to be convex, (1b) is the nodal power balance, (1c) are the nodal active power limits, and (1d) are the thermal transmission line limits.

The SCOPF follows from considering contingency scenarios, e.g., the loss of lines or generating units, within (1). Consider the $(N-k)$-SCOPF where $k \in \mathbb{N}$ lines are simultaneously lost in the network. Let $\mathcal{C} \subset \mathbb{N}$ be the set of contingencies. The power flow for each contingency scenario $c \in \mathcal{C}$ can be computed through the post-contingency PTDF matrix $\mathbf{\Phi}^{(c)}$. We refer the reader to [10] for a detailed, closed-form methodology on how to compute $\mathbf{\Phi}^{(c)}$. When a contingency occurs, the system operator can respond with several post-contingency mechanisms like reconfiguring its network topology and allowing for temporary line thermal rating violations, e.g., through emergency ratings immediately after the event [17]. We model the network operating condition changes for contingency $c \in \mathcal{C}$ through the modified thermal line limit vector $\overline{\mathbf{s}}^{(c)}$. The $(N-k)$-SCOPF is then:

$$\min_{\mathbf{p} \in \mathbb{R}^{\operatorname{card}\mathcal{N}}} \quad f(\mathbf{p}) \tag{2a}$$

$$\text{s.t.} \quad \mathbf{1}^\top \mathbf{p} = 0, \quad \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} \tag{2b}$$

$$-\overline{\mathbf{s}}^{(c)} \leq \mathbf{\Phi}^{(c)}\mathbf{p} \leq \overline{\mathbf{s}}^{(c)}, \quad c \in \mathcal{C} \cup \{0\}, \tag{2c}$$

where scenario $c = 0$ corresponds to the nominal network.

In practice, independent system operators (ISOs) implements soft rather than hard constraints for thermal line limits [7, 15, 18]. This is done to ensure that feasible dispatches are possible even in highly stressed network conditions. Constraint violations are penalized in the objective and thus translate into financial impacts on the market. We denote the penalty cost by $M > 0$. For example, MISO lists a \$1,500/MW penalty for thermal constraint violations, either under nominal operations or contingency [18]. For line $j \in \mathcal{L}$ and contingency scenario $c \in \mathcal{C}$, let $\xi_{j,+}^{(c)}$ and $\xi_{j,-}^{(c)}$ be the upper and lower thermal constraint violation, respectively. We express both terms in vector form, respectively, as $\boldsymbol{\xi}_+^{(c)}, \boldsymbol{\xi}_-^{(c)} \in \mathbb{R}^{\operatorname{card}\mathcal{L}}$ for all c. The penalized SCOPF is:

$$\min_{\mathbf{p}, \boldsymbol{\xi}_+^{(c)}, \boldsymbol{\xi}_-^{(c)}} \quad f(\mathbf{p}) + \sum_{j \in \mathcal{L}} M \left(\xi_{j,+}^{(c)} + \xi_{j,-}^{(c)} \right) \tag{3a}$$

$$\text{s.t.} \quad \mathbf{1}^\top \mathbf{p} = 0, \quad \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} \tag{3b}$$

$$\mathbf{\Phi}^{(c)}\mathbf{p} \leq \overline{\mathbf{s}}^{(c)} + \boldsymbol{\xi}_+^{(c)}, \quad c \in \mathcal{C} \cup \{0\} \tag{3c}$$

$$-\boldsymbol{\Phi}^{(c)}\mathbf{p} \leq \overline{\mathbf{s}}^{(c)} + \boldsymbol{\xi}_{-}^{(c)}, \quad c \in \mathcal{C} \cup \{0\} \tag{3d}$$

$$\boldsymbol{\xi}_{-}^{(c)} \geq \mathbf{0}, \quad \boldsymbol{\xi}_{+}^{(c)} \geq \mathbf{0}. \tag{3e}$$

We re-express (3) in a concise form as:

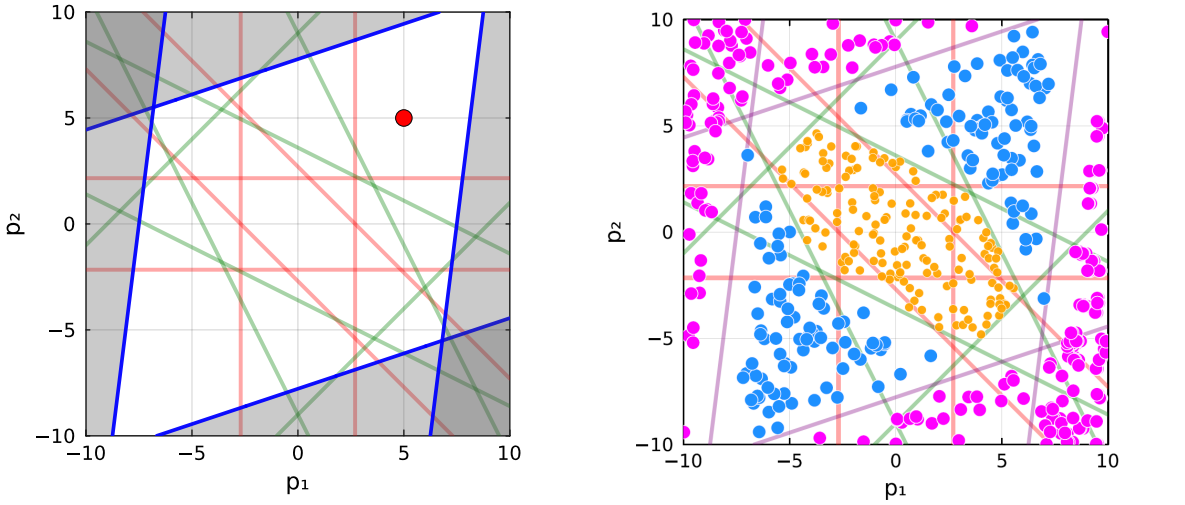
$$\min_{\substack{\mathbf{p}\in\mathbb{R}^{\operatorname{card}\mathcal{N}} \\ \boldsymbol{\xi}\in\mathbb{R}^{2(1+\operatorname{card}\mathcal{C})\operatorname{card}\mathcal{L}}\geq\mathbf{0}}} \quad f(\mathbf{p}) + M\mathbf{1}^{\top}\boldsymbol{\xi} \tag{4a}$$

$$\text{s.t.} \quad \mathbf{1}^{\top}\mathbf{p} = 0, \quad \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} \tag{4b}$$

$$\boldsymbol{\Phi}^{\mathrm{C}}\mathbf{p} \leq \overline{\mathbf{s}}^{\mathrm{C}} + \boldsymbol{\xi}, \tag{4c}$$

where $\boldsymbol{\Phi}^{\mathrm{C}} \in \mathbb{R}^{2(1+\operatorname{card}\mathcal{C})\operatorname{card}\mathcal{L}\times\operatorname{card}\mathcal{N}}$ and $\overline{\mathbf{s}}^{\mathrm{C}} \in \mathbb{R}^{2(1+\operatorname{card}\mathcal{C})\operatorname{card}\mathcal{L}}$ are the concatenations of the positive and negative PTDF matrices $\begin{pmatrix}\boldsymbol{\Phi}^{(c)\top} & -\boldsymbol{\Phi}^{(c)\top}\end{pmatrix}^{\top}$ and twice the line thermal limit vectors $\begin{pmatrix}\overline{\mathbf{s}}^{(c)\top} & \overline{\mathbf{s}}^{(c)\top}\end{pmatrix}^{\top}$ for all $c \in \mathcal{C} \cup \{0\}$. Problem (4) is akin to (1) but now comprises a largely increased number of constraints and variables.

3 Constraint reduction for penalized SCOPF



(a) Non-redundant constraints for hierarchical level $k^{\star} = 2$ (red point: optimum, orange lines: level 1, green lines: level 2, blue lines outward: redundant levels)

(b) Optimum for different loading profile and its corresponding $k^{\star}$ (points are optima from orange: level 1, blue: level 2, magenta: level 3, lines as in (a))

Figure 1: Hierarchical constraint levels for the 9-bus test system.

We now present our exact constraint-reduction method for the penalized SCOPF. The method is illustrated in Figure 1. Let $\mathcal{I} = \{1, 2, \ldots, 2\,(1 + \operatorname{card}\mathcal{C})\operatorname{card}\mathcal{L}\}$ be the set of indices of thermal limit constraints for contingency scenarios $c \in \mathcal{C} \cup \{0\}$. Let $\boldsymbol{\phi}_i \in \mathbb{R}^{\operatorname{card}\mathcal{N}}$ and $\overline{s}_i^{\mathrm{C}} \in \mathbb{R}_+$ correspond to row i of the affine inequalities (4c), i.e., to the thermal limit constraint parameters for a line under a given contingency scenario. We observe that while for the hard-constrained case (2), constraint i is active or non-redundant if $\left\{\mathbf{p} \in \mathbb{R}^{|\mathcal{N}|} \;\middle|\; \boldsymbol{\phi}_i^{\top}\mathbf{p} \leq \overline{s}_i^{\mathrm{C}}, i \in \mathcal{I} \setminus \{i\}\right\} \neq \left\{\mathbf{p} \in \mathbb{R}^{|\mathcal{N}|} \;\middle|\; \boldsymbol{\phi}_i^{\top}\mathbf{p} \leq \overline{s}_i^{\mathrm{C}}, i \in \mathcal{I}\right\}$, i.e., when it contributes to delimiting the feasible polytope, this does not translate to (4) because its feasibility is guaranteed by construction.

Let $\boldsymbol{\lambda} \in \mathbb{R}^{\operatorname{card}\mathcal{I}}$ be the dual variables associated to constraints (4c). Let the superscript * denote an optimum. We identify a redundant constraint $i \in \mathcal{I}$ for the penalized problem by its zero primal-dual optimal pair $\xi_i^{\star}$ and $\lambda_i^{\star}$.

Theorem 1. Let $f : \mathbb{R}^{\operatorname{card}\mathcal{N}} \to]-\infty, +\infty]$ be a proper, lower semicontinuous, convex function. Consider

$$\mathcal{X} := \left\{\mathbf{x} \in \mathbb{R}^{|\mathcal{N}|} \;\middle|\; \mathbf{1}^{\top}\mathbf{x} = 0,\ \underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}}\right\},$$

and assume that $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(\mathcal{X}) \neq \emptyset$. For each $i \in \mathcal{I}$, define the affine mapping $g_i : \mathbb{R}^{\operatorname{card}\mathcal{N}} \to \mathbb{R}$ and the penalty $\psi_i : \mathbb{R}^{\operatorname{card}\mathcal{N}} \to \mathbb{R}$ as $g_i(\mathbf{x}) := \boldsymbol{\phi}_i^\top \mathbf{x} - \overline{s}_i^{\mathrm{C}}$, and $\psi_i(\mathbf{x}) := [g_i(\mathbf{x})]_+ \equiv \max\{g_i(\mathbf{x}), 0\}$, respectively. For any subset $\mathcal{J} \subseteq \mathcal{I}$, consider the penalized objective function

$$F_{\mathcal{J}}(\mathbf{x}) := f(\mathbf{x}) + M \sum_{i \in \mathcal{J}} \psi_i(\mathbf{x}) + \mathbb{I}_{\mathcal{X}}(\mathbf{x}),$$

where the indicator function $\mathbb{I}_{\mathcal{X}} : \mathbb{R}^{\operatorname{card}\mathcal{N}} \to \{0, +\infty\}$ is 0 if $\mathbf{x} \in \mathcal{X}$ and $+\infty$ otherwise, and assume $F_{\mathcal{I}}(\mathbf{x}) < +\infty$ for some $\mathbf{x} \in \mathcal{X}$. Let $\mathcal{S} := \mathcal{I} \setminus \mathcal{J}$ be the set of removed constraints. Then, the following assertions are equivalent:

1) $\min_{\mathbf{x}} F_{\mathcal{J}}(\mathbf{x}) = \min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x})$;

2) There exist $\mathbf{x}^\star \in \arg\min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x})$ and $\lambda_i^\star \in \mathbb{R}_+$, $i \in \mathcal{I}$, such that

1. $\mathbf{0} \in \partial f(\mathbf{x}^\star) + \sum_{i \in \mathcal{I}} \lambda_i^\star \boldsymbol{\phi}_i + N_{\mathcal{X}}(\mathbf{x}^\star)$, where $\partial\cdot$ denotes the subdifferential of $\cdot$ and $N_{\mathcal{X}}(\mathbf{x}) = \partial \mathbb{I}_{\mathcal{X}}(\mathbf{x})$ is the normal cone to $\mathcal{X}$, with $\lambda_i^\star \in M\, \partial [g_i(\mathbf{x}^\star)]_+$ for all $i \in \mathcal{I}$, and;
2. $g_i(\mathbf{x}^\star) \leq 0$ and $\lambda_i^\star = 0$, for all $i \in \mathcal{S}$.

Letting $\xi_i^\star = [g_i(\mathbf{x}^\star)]_+$, the conditions on $\mathcal{S}$ are equivalently provided by $\xi_i^\star = 0$ and $\lambda_i^\star = 0$ for all $i \in \mathcal{S}$.

Proof. We first prove **2)** $\Longrightarrow$ **1)**. Suppose that **2)** holds. We have $[g_i(\mathbf{x}^\star)]_+ = 0$ for all $i \in \mathcal{S}$ because $g_i(\mathbf{x}^\star) \leq 0$, and thus, $F_{\mathcal{J}}(\mathbf{x}^\star) = F_{\mathcal{I}}(\mathbf{x}^\star)$. Then, because $\lambda_i^\star = 0$ for all $i \in \mathcal{S}$, the stationarity inclusion for $\min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x})$ reduces to:

$$\mathbf{0} \in \partial f(\mathbf{x}^\star) + \sum_{i \in \mathcal{J}} \lambda_i^\star \boldsymbol{\phi}_i + N_{\mathcal{X}}(\mathbf{x}^\star). \tag{5}$$

The subdifferential of the reduced objective $F_{\mathcal{J}}(\mathbf{x})$ satisfies:

$$\partial f(\mathbf{x}^\star) + \sum_{i \in \mathcal{J}} \partial(M\psi_i)(\mathbf{x}^\star) + N_{\mathcal{X}}(\mathbf{x}^\star) \subseteq \partial F_{\mathcal{J}}(\mathbf{x}^\star). \tag{6}$$

Because $g_i(\mathbf{x})$ is an affine mapping, applying the convex chain rule [19, Th. 23.9] to the scalar function $t \mapsto \max(0, t)$ yields the subdifferential of the penalty:

$$\partial \psi_i(\mathbf{x}^\star) = \{c_i \boldsymbol{\phi}_i \mid c_i \in \partial [\cdot]_+ (g_i(\mathbf{x}^\star))\}. \tag{7}$$

Hence, we have $\lambda_i^\star \boldsymbol{\phi}_i \in \partial(M\psi_i)(\mathbf{x}^\star)$ for all $i \in \mathcal{J}$, where $\lambda_i^\star = Mc_i \in [0, M]$. Substituting this into (5), we have that (5) belongs to (6)'s left-hand side. Then, it follows that $\mathbf{0} \in \partial F_{\mathcal{J}}(\mathbf{x}^\star)$. By [3, Prop. 4.7.2], $\mathbf{x}^\star \in \arg\min_{\mathbf{x}} F_{\mathcal{J}}(\mathbf{x})$, and we have $\min_{\mathbf{x}} F_{\mathcal{J}}(\mathbf{x}) = F_{\mathcal{J}}(\mathbf{x}^\star) = F_{\mathcal{I}}(\mathbf{x}^\star) = \min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x})$, which establishes **1)**.

We now show **1)** $\Longrightarrow$ **2)**. Assume $\min_{\mathbf{x}} F_{\mathcal{J}}(\mathbf{x}) = \min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x})$, and let $\mathbf{x}^\star \in \arg\min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x})$. Because $F_{\mathcal{J}}(\mathbf{x}) \leq F_{\mathcal{I}}(\mathbf{x})$ for all $\mathbf{x}$, we have:

$$F_{\mathcal{I}}(\mathbf{x}^\star) \geq F_{\mathcal{J}}(\mathbf{x}^\star) \geq \min_{\mathbf{x}} F_{\mathcal{J}}(\mathbf{x}) = \min_{\mathbf{x}} F_{\mathcal{I}}(\mathbf{x}) = F_{\mathcal{I}}(\mathbf{x}^\star).$$

Hence, $F_{\mathcal{J}}(\mathbf{x}^\star) = F_{\mathcal{I}}(\mathbf{x}^\star)$, which implies that the removed constraints are satisfied at $\mathbf{x}^\star$, i.e.,

$$M \sum_{i \in \mathcal{S}} [g_i(\mathbf{x}^\star)]_+ = 0 \iff g_i(\mathbf{x}^\star) \leq 0,$$

for all $i \in \mathcal{S}$. Because $\mathbf{x}^\star \in \arg\min_{\mathbf{x}} F_{\mathcal{J}}(\mathbf{x})$, [3, Prop. 4.7.2] yields $\mathbf{0} \in \partial F_{\mathcal{J}}(\mathbf{x}^\star)$. Under the constraint qualification $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(\mathcal{X}) \neq \emptyset$, we have that the inclusion (6) holds with equality [19, Th. 23.8]. Recalling (7), there exist scalars $c_i^\star \in \partial [\cdot]_+ (g_i(\mathbf{x}^\star))$ and corresponding $\lambda_i^\star = Mc_i^\star \in [0, M]$ for $i \in \mathcal{J}$,

such that $\mathbf{0} \in \partial f(\mathbf{x}^\star)+\sum_{i\in\mathcal{J}} \lambda_i^\star \boldsymbol{\phi}_i+N_{\mathcal{X}}(\mathbf{x}^\star)$. Finally, for $i \in \mathcal{S}$, we have already shown that $g_i(\mathbf{x}^\star) \leq 0$, which implies $0 \in \partial[\cdot]_+(g_i(\mathbf{x}^\star))$. By setting $c_i^\star = 0 = \lambda_i^\star$ for all $i \in \mathcal{S}$, we obtain:

$$\mathbf{0} \in \partial f(\mathbf{x}^\star) + \sum_{i\in\mathcal{I}} \lambda_i^\star \boldsymbol{\phi}_i + N_{\mathcal{X}}(\mathbf{x}^\star),$$

with $g_i(\mathbf{x}^\star) \leq 0$ and $\lambda_i^\star = 0$ for all $i \in \mathcal{S}$, proving **2)**. □

For penalized problems like (4), the geometrical position of a constraint with respect to the feasible polytope do not provide the necessary information to determine if it is redundant or not. To this end, we can use the constraint's impact on the minimum, characterized by its associated penalty and dual optimum being zero. Next, we present our method that first uses geometrical properties to build a nested hierarchy of constraint subsets based on the unpenalized problem and second, employs Theorem 1 to identify the smallest hierarchical level corresponding to the penalized problem's set of non-redundant constraints for a load profile.

3.1 Non-redundant constraint hierarchy

We first define a nested hierarchy of non-redundant constraint subsets from the unpenalized problem. The subset extraction is specific to the collection of $\boldsymbol{\Phi}^{(c)}$-matrices corresponding to all contingencies and, as such, is agnostic to the demand profile and can be computed once prior to deployment. Section 3.2 then identifies the penalized problem's set of non-redundant constraints for a demand profile based on these subsets. Let $k \in \{1, 2, \ldots, k_{\max}\} \subset \mathbb{N}$ be the constraint subset hierarchy level. We define recursively $\mathcal{I}_k \subset \mathcal{I}$ to be the subset of non-redundant hard-constraint indices constructed from all constraints except the ones already included in $\mathcal{I}_{k-1}, \mathcal{I}_{k-2}, \ldots, \mathcal{I}_1, \mathcal{I}_0 = \emptyset$. When all constraints have been assigned to a subset $\mathcal{I}_k$, we let $k = k_{\max}$. We obtain an ensemble of disjoint subsets such that $\mathcal{I} = \bigcup_{k=1}^{k_{\max}} \mathcal{I}_k$, where $\mathcal{I}_k$ collects all constraint indices that are structurally relevant for hierarchy level k with respect to the lower levels. The feasible set for the unpenalized problem induced by level k's thermal limits becomes:

$$\left\{ \mathbf{p} \in \mathbb{R}^{|\mathcal{N}|} \;\middle|\; \boldsymbol{\phi}_i^\top \mathbf{p} \leq \overline{s}_i^{\mathrm{C}}, i \in \mathcal{I}_{1:k} \right\}, \tag{8}$$

where $\mathcal{I}_{1:k} \equiv \bigcup_{j=1}^{k} \mathcal{I}_j$. This construction yields the strict inclusion relation between feasible sets (8) in terms of k.

Algorithm 1 presents our method to build the nested hierarchical constraint subset $\mathcal{I}_k$, $k = 1, 2, \ldots, k_{\max}$ from the unpenalized problem. First, we employ a filtering step (Lines $1-5$) to discard exact duplicates of the constraints. This step ensures that: (i) the first hierarchical set $\mathcal{I}_1$ is only made of distinct constraints and (ii) the intersection of any subsets is empty. Next, we iteratively build the constraint subsets $\mathcal{I}_k$, $k = 1, 2, \ldots, k_{\max}$, by applying a constraint redundancy removal algorithm on the remaining set of constraints $\mathcal{I} \setminus \bigcup_{j=1}^{k-1} \mathcal{I}_j$ (Lines $8-11$), yielding the constraint hierarchy subsets. We adopt [20]'s constraint redundancy removal for unpenalized problem algorithm which is itself based on the linear programming test [9, Proposition 8.5] [20, (11)] and the `RayShoot` algorithms [8] [20, Algorithm 2].

Algorithm 1 yields a nested hierarchy of constraint subsets where their union up to level k, $\mathcal{I}_{1:k}$, leads to a sequence of tighter feasible sets for the unpenalized problem. These unions are then used to reduce the resolution time of the penalized problem (4) by considering only non-redundant constraints from primal-dual optimum information. For level $k \in \{1, 2, \ldots, k_{\max}\}$, consider:

$$\min_{\mathbf{p},\, \boldsymbol{\xi} \geq \mathbf{0}} \quad f(\mathbf{p}) + \sum_{i\in\mathcal{I}_{1:k}} M\xi_i \tag{9a}$$

$$\text{s.t.} \quad \mathbf{1}^\top \mathbf{p} = 0, \quad \underline{\mathbf{p}} \leq \mathbf{p} \leq \overline{\mathbf{p}} \tag{9b}$$

$$\boldsymbol{\phi}_i^\top \mathbf{p} \leq \overline{s}_i^{\mathrm{C}} + \xi_i, \quad i \in \mathcal{I}_{1:k}, \tag{9c}$$

where both the number of thermal limit constraints and penalty variables have been reduced to $\operatorname{card}\mathcal{I}_{1:k}$. Then, for a load profile $\mathbf{p}^{\mathrm{d}} \in \mathbb{R}^{\operatorname{card}\mathcal{N}}$, determining the smallest hierarchical level $k^\star$ for which (9) is equivalent to (4) allows us to more efficiently solve the penalized SCOPF. Figure 1a illustrates $k^\star$ and its corresponding non-redundant constraint subset. Contrarily to the SCOPF (2) and as shown in Figure 1b, the non-redundant constraint subset is not fixed and depends on the load because thermal limit constraint violations may be desirable under certain conditions, e.g., high local demand. From Theorem 1, we can calculate $k^\star$ after solving (4) using:

$$k^\star = \max_{k \in \{1,2,\ldots,k_{\max}\}} \{k \mid \exists i \in \mathcal{I}_k \text{ s.t. } \xi_i^\star + \lambda_i^\star > 0\}, \tag{10}$$

Next, we design a supervised learning pipeline to infer $k^\star$ as a function of $\mathbf{p}^{\mathrm{d}}$ without the need to solve (4), followed by a correction step to preserve exactness. Recall that our notation defines $p_i^{\mathrm{d}} = \overline{p}_i$ if $\underline{p}_i = \overline{p}_i < 0$ and $p_i^{\mathrm{d}} = 0$ otherwise.

Algorithm 1: Hierarchical constraint extraction

Input: $\mathbf{\Phi}^{\mathrm{C}}$, $\overline{\mathbf{s}}^{\mathrm{C}}$, $\mathcal{I}$, and $\mathbf{z} = \mathbf{0}$.
Output: $\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_{k_{\max}}$ such that $\bigcup_{k=1}^{k_{\max}} \mathcal{I}_k = \mathcal{I}$.

Duplicate constraint filtering:
foreach $i \in \mathcal{I}$ **do**
  if $\nexists\, j < i$ *such that* $\mathbf{\Phi}_i^{\mathrm{C}} = \mathbf{\Phi}_j^{\mathrm{C}}$ *and* $\overline{\mathbf{s}}_i = \overline{\mathbf{s}}_j$ **then**
    $\mathcal{I}^{\text{filtered}} \leftarrow \mathcal{I}^{\text{filtered}} \cup \{i\}$;
$\mathcal{I} \leftarrow \mathcal{I}^{\text{filtered}}$;
Hierarchical constraint subset identification:
$k \leftarrow 1$;
while $\mathcal{I} \neq \emptyset$ **do**
  $\mathcal{I}_k \leftarrow$ `RedundancyRemoval`$(\mathbf{\Phi}^{\mathrm{C}}, \overline{\mathbf{s}}^{\mathrm{C}}, \mathbf{z}, \mathcal{I})$;
  $\mathcal{I} \leftarrow \mathcal{I} \setminus \mathcal{I}_k$, $k \leftarrow k + 1$;
$k_{\max} \leftarrow k - 1$;
return $\{\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_{k_{\max}}\}$

3.2 Hierarchical constraint level inference

We introduce a supervised learning process that captures the hierarchical properties of the subsets $\mathcal{I}_k$ to identify the minimal hierarchical level $k^\star \in \{1, 2, \ldots, k_{\max}\}$ associated with a load profile $\mathbf{p}^{\mathrm{d}}$ based on Theorem 1. Consider the dataset $\left\{\mathbf{p}_t^{\mathrm{d}}, k_t^\star\right\}_{t=1}^{T}$ where $\mathbf{p}_t^{\mathrm{d}} \in \mathbb{R}^{\operatorname{card}\mathcal{N}}$ is a load profile and $k_t^\star$ is computed ex-post via (10), with the subscript t referring to the t^{th} data point of the dataset of size $T \in \mathbb{N}$. Let $\overline{k} = \max_{t=1,2,\ldots,T} k_t^\star$. We re-express the label $k_t^\star$ as the cumulative binary label vector $\mathbf{z}_t \in \{0,1\}^{\overline{k}}$ such that $\mathbf{z}_t(k) = 1_{k \leq k_t^\star}$ for $k = 1, 2, \ldots, \overline{k}$, where 1_x is the indicator function that returns 1 if x is true and 0 otherwise.

We define the collection of predictors $\left\{\mathfrak{C}_k : \mathbb{R}^{\operatorname{card}\mathcal{N}} \to [0,1]\right\}_{k=1}^{\overline{k}}$ each trained to approximate the probability of success $\Pr\left[\mathbf{z}(k) = 1 \middle| \mathbf{p}^{\mathrm{d}}\right]$ of a Bernoulli random variable modelling the inclusion of $\mathcal{I}_k$ within the feasible set. Let the predicted cumulative binary label vector be $\hat{\mathbf{z}}(k) = 1_{\mathfrak{C}_k(\mathbf{p}^{\mathrm{d}}) \geq \tau}$ for all k, where $\tau \in]0,1[$ is a threshold to be tuned. We infer the smallest hierarchical level from a load profile using $\hat{k}^\star := \max\left\{k \in \{2, \ldots, \overline{k}\} \,\middle|\, \hat{\mathbf{z}}(k) = 1\right\}$, and ensure that lower levels are included by construction of $\mathcal{I}_{1:\hat{k}^\star}$. Because level $k = 1$ is always active given the hierarchical nature of the constraint subsets, the predictor $\mathfrak{C}_1$ can be omitted and we always set $\hat{\mathbf{z}}(1) = 1$. This encoding and inference process permits us to train $\overline{k} - 1$ predictors independently while modelling the hierarchical nature of the constraint subsets. In Section 4, we use the tree-based learning approach `LightGBM` for $\mathfrak{C}_k$.

Finally, we include an additional step to ensure the exactness of our method with respect to (4). We remark that if $\hat{k}^\star > k^\star$, the resolution of (9) using $\hat{k}^\star$ remains exact with respect to the full problem, albeit at a higher computational cost given the inclusion of redundant constraints and variables.

Conversely, $\hat{k}^\star < k^\star$ yields a relaxation of (4) with no feasibility guarantee with respect to (4). To address this, we implement an iterative correction for level underestimation. Let $\mathbf{p}^\star_{\hat{k}^\star} \in \mathbb{R}^{\operatorname{card}\mathcal{N}}$ be an optimum of the reduced problem (9) where $\hat{k}^\star$ is used to define the thermal limit constraint set. Using Theorem 1's notation, let $F_{\mathcal{I}_{1:\hat{k}^\star}}(\mathbf{p})$ and $F_{\mathcal{I}}(\mathbf{p})$ denote the penalized objective functions of the reduced problem (9) and the full problem (4), respectively. Exactness is guaranteed if $\mathbf{p}^\star_{\hat{k}^\star}$ satisfies all the omitted physical limits:

$$\boldsymbol{\phi}_i^\top \mathbf{p}^\star_{\hat{k}^\star} \leq \overline{s}_i^{\mathrm{C}}, \quad \forall\, i \in \mathcal{I}_{\hat{k}^\star+1:k_{\max}} = \bigcup\nolimits_{j=\hat{k}^\star+1}^{k_{\max}} \mathcal{I}_j. \tag{11}$$

If (11) holds, the omitted penalties vanish, yielding $F_{\mathcal{I}_{1:\hat{k}^\star}}(\mathbf{p}^\star_{\hat{k}^\star}) = F_{\mathcal{I}}(\mathbf{p}^\star_{\hat{k}^\star})$. Because the reduced problem omits non-negative penalty terms, we have $F_{\mathcal{I}_{1:\hat{k}^\star}}(\mathbf{p}) \leq F_{\mathcal{I}}(\mathbf{p})$ for all $\mathbf{p}$. It directly follows that $\min_{\mathbf{p}\in\mathcal{X}} F_{\mathcal{I}}(\mathbf{p}) \leq F_{\mathcal{I}}(\mathbf{p}^\star_{\hat{k}^\star}) = F_{\mathcal{I}_{1:\hat{k}^\star}}(\mathbf{p}^\star_{\hat{k}^\star}) = \min_{\mathbf{p}\in\mathcal{X}} F_{\mathcal{I}_{1:\hat{k}^\star}}(\mathbf{p}) \leq \min_{\mathbf{p}\in\mathcal{X}} F_{\mathcal{I}}(\mathbf{p})$. It then follows that $\mathbf{p}^\star_{\hat{k}^\star}$ is a global optimum of the full problem (4).

If (11) is not met, the constraint hierarchy level is incremented and (9) is solved again with the next level. The process is repeated for $k = \hat{k}^\star + n$, $n = 1, 2, \ldots, \overline{k} - \hat{k}^\star$, until a solution $\mathbf{p}^\star_k$ meets (11). We note that (11) only implements the equivalent of a matrix product and an element-wise comparison. The number of correction steps is bounded, thus the exact minimum is always achieved at a reasonable cost.

4 Numerical examples

We illustrate our method in numerical simulations and evaluate its performance for reducing the computational time of solving the penalized SCOPF. All implementations are done in `Julia` on a i7-12700H, 2.30 GHz-laptop computer.

4.1 Hierarchical constraint subset extraction

We consider the IEEE 9, 39, and 118 bus systems [21] and formulate the penalized $(N-1)$-SCOPF problem where contingency scenarios model each line being removed one at the time from $\mathcal{L}$. We run Algorithm 1 and extract the hierarchical constraint subsets $\mathcal{I}_1$, $\mathcal{I}_2$, …, $\mathcal{I}_{k_{\max}}$. Table 1 provides the cardinality of each set, i.e., the number of non-redundant hard constraints for each level k, together with the number of constraints of the original problem, $\operatorname{card}\mathcal{I}$. Our approach effectively regroups non-redundant hard constraints with respect to the $k_{\max}$ regimes observed during resolution. Next, we exemplify the computational speedup and exactness of the hierarchical constraint subset extraction. We sample a load profile $\mathbf{p}^{\mathrm{d}}$ for each test system. We compute the profile's k-level by solving the full penalized problem (4) and then by evaluating (10). Table 2 compares the resolution of the full problems to the reduced instances with $k^\star = 1$, 2, and 5, for, respectively, the IEEE 9-bus, 39-bus, and 118-bus systems. It illustrates both the potential for resolution time reduction when $k^\star$ is identified for a load profile and the exactness of the hierarchical constraint levels.

Table 1: Description of the hierarchical constraint subsets $\mathcal{I}_k$.

Test system	$\operatorname{card}\mathcal{I}_1$	$\operatorname{card}\mathcal{I}_2$	$\operatorname{card}\mathcal{I}_3$	$\operatorname{card}\mathcal{I}_4$	$\operatorname{card}\mathcal{I}_5$	$\operatorname{card}\mathcal{I}_6$	$\operatorname{card}\mathcal{I}_7$	$\operatorname{card}\mathcal{I}_8$	$\operatorname{card}\mathcal{I}_9$	$\operatorname{card}\mathcal{I}_{10}$	$\operatorname{card}\mathcal{I}_{11}$	$\operatorname{card}\mathcal{I}$
9 bus	35	39	12	8	2	–	–	–	–	–	–	96
39 bus	719	793	273	113	88	–	–	–	–	–	–	1,986
118 bus	2,613	17,671	22,350	5,100	1,491	2,120	863	192	23	6	1	52,430

Henceforth, we focus on the 39-bus system as it allows to consider a large number of problem instances in a limited time while offering an interesting perspective into the benefits of our method. We consider 2,125 load profiles with $k^\star$ ranging from 1 to $k_{\max} = 5$ and provide the relative computational

Table 2: Comparison between the full SCOPF and the hierarchically reduced formulations.

Test system	$\text{time}_{\text{full}}$ [s]	$\text{time}_{\text{reduced}}$ [s]	Δtime [s]	Δtime [%]	opt_{full}	$\text{opt}_{\text{reduced}}$	$\text{card}\,\mathcal{I}$	$\text{card}\,\mathcal{I}_{1:k^\star}$	$\Delta\mathcal{I}$	$\Delta\mathcal{I}$ [%]	$k_{\max}$	$k^\star$
9 bus	0.008	0.004	0.003	37.5	5,781	5,781	96	35	61	63.5	5	1
39 bus	0.305	0.179	0.126	41.31	215,151	215,151	1,986	1,512	474	23.87	5	2
118 bus	2.296	1.930	0.366	15.94	3,602	3,602	52,430	49,225	3,205	6.11	11	5

time speedup in Figure 2. As shown in Figure 2, significant speedups are achieved at lower hierarchical constraint levels, viz., 1, 2, and 3. However, at levels 4 and 5, the speedup is limited, if not negative. This illustrates the solver's numerical instability for grids operating near their limits even if the number of constraints has been slightly reduced. This fact also highlights the benefit of inferring $k^\star$ to guide the resolution process.

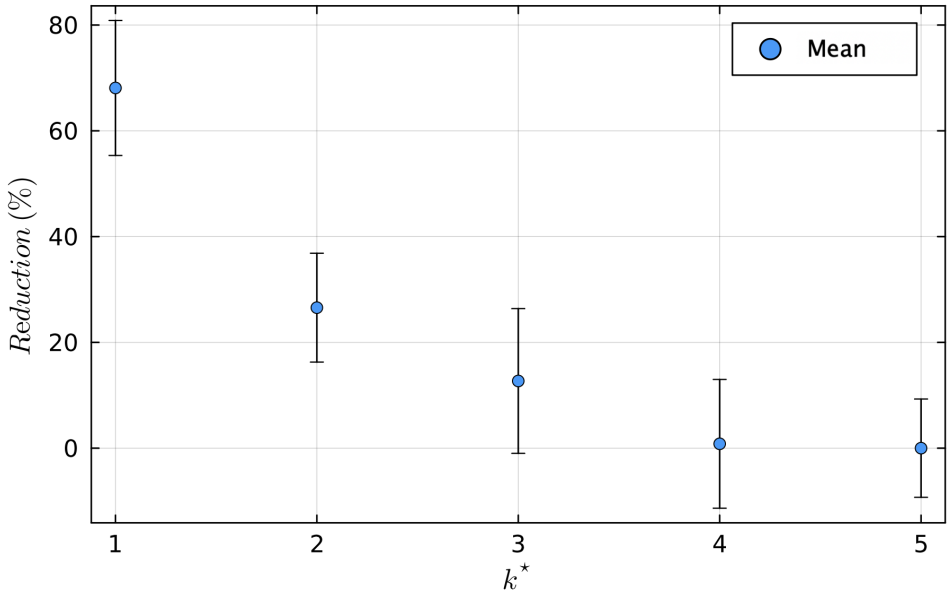


Figure 2: Solve time reduction average and standard deviation.

4.2 Hierarchical constraint level inference

We now showcase the constraint level inference methodology on the 39-bus system. We generate load profiles, compute their $k^\star$ using (10), and encode it as $\mathbf{z}$. 20,890 pairs of load profiles and $k^\star$ up to 4 are used to ensure that at least 5% of the dataset is of each level. If $k^\star = k_{\max} = 5$ is to be needed during resolution, the correction step would guarantee it is eventually used.

The dataset is split into training and test samples comprising 80% and 20% of the data, respectively. In-line with our computational efficiency objective, we employ the decision-tree predictor `LighGBM` [14] for $\mathfrak{C}_k$, $k = 2, 3, 4$, with $\mathbf{p}^{\text{d}}$ and $\mathbf{z}(k)$ as feature and label, respectively. The binary loss function weighted by the relative frequency of the labels is used to help during training given that a label might be rare [11]. The predictors are trained independently on the dataset and a 5-fold cross-validation is used for hyperparameter tuning, viz., the learning rate, maximum number of leaves, maximum depth, the feature and bagging fraction, and the ℓ_1- and ℓ_2-regularizer weights. We set $\tau = 0.5$ to balance over- and underestimation. We note that this hyperparameter could be tuned in future work, e.g., to reduce underestimating $k^\star$ over overestimating it, the former of which leads to thermal limits violation and the use of the correction step.

Table 3 aggregates the performance of the `LightGBM`-based inference approach. Table 5 further decomposes the results by hierarchical levels. It shows that our method successfully estimates the smallest hierarchy level $k^\star$ in most instances with an average error of about a single level, if not. The

impact of both over- and under-estimating $k^\star$ is thus lessen given that only the next $\mathcal{I}_k$ is added to the constraint subset or a single correction step is utilized. Table 4 emphasizes that the predictor is best when it yields the highest computational improvement at $k^\star = 1$. Level 2 is the precision bottleneck with an over-estimation error of 10.14% meaning that further speedup could be obtained if these problem instances were correctly labelled as level 1.

Table 3: Hierarchical level inference with `LightGBM`.

Performance indicator	Value
Precision	94.28%
Under-estimation ratio	3.48%
Over-estimation ratio	2.25%
Mean absolute error	0.061 level
Over-estimation mean error	1.063 levels
Under-estimation mean error	1.061 levels

Table 4: Inference performance for each level.

$k^\star$	**Data ratio** [%]	**Precision** [%]	**Under** [%]	**Over** [%]
1	22.34	99.58	–	0.42
2	11.42	89.86	0.00	10.14
3	52.91	93.30	4.83	1.88
4	13.33	93.09	6.91	–

Finally, we estimate the computational gain of our approach, i.e., of inferring the hierarchical constraint level with the proposed encoding and `LightGBM`. We compare it to solving the full problem (4) in Table 5. We remark that the bottom half of Table 5 refers to the total resolution time from solving the reduced problem (9) at $\hat{k}^\star$ and any next k-value(s) from the correction step. Overall, our method yields a 19.72% time reduction on this limited-size test case without compromising the exactness of the solution. Corollary to the time reduction is a lower memory requirement due to, namely, the reduced number of constraints and variables considered when solving the problem.

Table 5: Resolution time reduction for the IEEE 39-bus.

Performance indicator	Value
Problem instances	4265
Total recourse steps	852
Average full problem resolution time	0.121 s
Average reduced problem & recourse resolution time	0.096 s
Average absolute time gain	0.024 s
Average relative time gain	19.72%

5 Conclusion

We propose a constraint-reduction method dedicated to the penalized security-constrained optimal power flow (SCOPF) problem. We first define a nested hierarchy of non-redundant constraint subsets from the unpenalized problem. The minimal hierarchical level is then retrieved from the optimal penalty and associated dual variables and selects the constraint subsets to be used such that its minimum and the full penalized SCOPF's match. We infer the minimal level using a cumulative encoding and a supervised learning process. The exactness of the estimated level-based reduced problem is assessed through a lightweight verification, and the level is incremented until the condition is satisfied. We illustrate our method on the IEEE 9, 39, and 118 bus test systems and exemplify the hierarchical constraint subset extraction outcome. We implement a `LightGBM`-based predictor on the 39-bus system and identify the hierarchical level with a 94.28% accuracy and an underestimation rate below

3.5%. This results in an average computational time reduction for the SCOPF of 19.7%, including necessary correction step(s). Future work includes the deployment on larger test cases, e.g., via the parallelization of Algorithm 1, and the reduction of the underestimation rates to minimize the use of the correction step using an asymmetrical training loss function.

References

[1] Ali Jahanbani Ardakani and François Bouffard. Prediction of umbrella constraints. In 2018 Power systems computation conference (PSCC), pages 1–7. IEEE.

[2] Ali Jahanbani Ardakani and François Bouffard. Identification of umbrella constraints in DC-based security-constrained optimal power flow. IEEE Transactions on Power Systems, 28(4):3924–3934, 2013.

[3] Dimitri P. Bertsekas, Angelia Nedić, and Asuman E. Ozdaglar. Convex Analysis and Optimization. Athena Scientific, Belmont, MA, 2003.

[4] François Bouffard, Federico D. Galiana, and José M. Arroyo. Umbrella contingencies in security-constrained optimal power flow. In 15th Power Systems Computation Conference (PSCC), pages 1–7, Liège, Belgium, 2005.

[5] Florin Capitanescu, JL Martinez Ramos, Patrick Panciatici, Daniel Kirschen, A Marano Marcolini, Ludovic Platbrood, and Louis Wehenkel. State-of-the-art, challenges, and future trends in security constrained optimal power flow. Electric Power Systems Research, 81(8):1731–1741, 2011.

[6] Spyros Chatzivasileiadis. Lecture notes on optimal power flow (OPF). arXiv preprint arXiv:1811.00943, 2018.

[7] Wenbo Chen, Mathieu Tanneau, and Pascal Van Hentenryck. End-to-end feasible optimization proxies for large-scale economic dispatch. IEEE Transactions on Power Systems, 39(2):4723–4734, 2023.

[8] Kenneth L Clarkson. More output-sensitive geometric algorithms. In Proceedings 35th Annual Symposium on Foundations of Computer Science, pages 695–702. IEEE, 1994.

[9] Komei Fukuda. Lecture: Polyhedral computation, spring 2016. Institute for Operations Research and Institute of Theoretical Computer Science, ETH Zurich. https://inf.ethz.ch/personal/fukudak/lect/pclect/notes2015/PolyComp2015.pdf, 2016.

[10] Jiachun Guo, Yong Fu, Zuyi Li, and Mohammad Shahidehpour. Direct calculation of line outage distribution factors. IEEE Transactions on Power Systems, 24(3):1633–1634, 2009.

[11] Haibo He and Edwardo A Garcia. Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 21(9):1263–1284, 2009.

[12] Jonas Hörsch, Henrik Ronellenfitsch, Dirk Witthaut, and Tom Brown. Linear optimal power flow using cycle flows. Electric Power Systems Research, 158:126–135, 2018.

[13] Ali Jahanbani Ardakani and François Bouffard. Acceleration of umbrella constraint discovery in generation scheduling problems. IEEE Transactions on Power Systems, 30(4):2100–2109, 2015.

[14] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, et al. LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30, 2017.

[15] Xiaonan Ma, Hongming Song, Ming Hong, Jun Wan, Yifan Chen, and Edward Zak. The security-constrained commitment and dispatch for Midwest ISO day-ahead co-optimized energy and ancillary service market. In 2009 IEEE Power & Energy Society General Meeting, pages 1–8.

[16] Ramtin Madani, Javad Lavaei, and Ross Baldick. Constraint screening for security analysis of power networks. IEEE Transactions on Power Systems, 32(3):1828–1838, 2016.

[17] Midcontinent Independent System Operator. Energy and operating reserve markets business practices manual (BPM-002-r24). MISO, 2023. Effective date: September 30, 2023.

[18] Midcontinent Independent System Operator. Schedule 28 – demand curves for operating reserves, 2023. Accessed: 2025-07-10.

[19] Ralph Tyrrell Rockafellar. Convex Analysis. Princeton University Press, Princeton, NJ, 1970.

[20] Richard Weinhold and Robert Mieth. Fast security-constrained optimal power flow through low-impact and redundancy screening. IEEE Transactions on Power Systems, 35(6):4574–4584, 2020.

[21] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE Transactions on Power Systems, 26(1):12–19, 2010.