Alternating Methods for Large-Scale AC Optimal Power Flow with Unit Commitment

Matthew Brun[*]   Thomas Lee[†]   Dirk Lauinger[‡]   Xin Chen[§]   Xu Andy Sun[¶]

Abstract

Security-constrained unit commitment with alternating current optimal power flow (SCUC-ACOPF) is a central problem in power grid operations that optimizes commitment and dispatch of generators under a physically accurate power transmission model while encouraging robustness against component failures. SCUC-ACOPF requires solving large-scale problems that involve multiple time periods and networks with thousands of buses within strict time limits. In this work, we study a detailed SCUC-ACOPF model with a rich set of features of modern power grids, including price-sensitive load, reserve products, transformer controls, and energy-limited devices. We propose a decomposition scheme and a penalty alternating direction method to find high-quality solutions to this model. Our methodology leverages spatial and temporal decomposition, separating the problem into a set of mixed-integer linear programs for each bus and a set of continuous nonlinear programs for each time period. To improve the performance of the algorithm, we introduce a variety of heuristics, including restrictions of temporal linking constraints, a second-order cone relaxation, and a contingency screening algorithm. We quantify the quality of feasible solutions through a dual bound from a convex second-order cone program. To evaluate our algorithm, we use large-scale test cases from Challenge 3 of the U.S. Department of Energy's Grid Optimization Competition that resemble real power grid data under a variety of operating conditions and decision horizons. The experiments yield feasible solutions with an average optimality gap of 1.33%, demonstrating that this approach generates near-optimal solutions within stringent time limits.

Key words: alternating current optimal power flow, unit commitment, decomposition, penalty alternating direction method

1 Introduction

Unit commitment (UC) is a key problem in power systems used to schedule the operations of generators. This problem seeks minimal cost dispatch solutions that satisfy operational constraints and meet estimated load. Underlying the problem are a set of physical constraints that dictate how power can be transmitted over the electric grid. These alternating current optimal power flow (ACOPF) constraints introduce complex nonconvexities and form a mixed-integer nonlinear program (MINLP; Castillo et al.

---

[*]Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA (brunm@mit.edu).

[†]Institute for Data, Systems, & Society, Massachusetts Institute of Technology, Cambridge, MA (t_lee@mit.edu).

[‡]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA (lauinger@mit.edu).

[§]Department of Electrical & Computer Engineering, Texas A&M University, College Station, TX (xin_chen@tamu.edu).

[¶]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA (sunx@mit.edu).

2016). Current practice necessitates that commitment solutions be updated as often as every few minutes, requiring that the models be solved within strict time limits (ISO New England 2023).

The class of ACOPF problems is strongly NP-hard (Bienstock and Verma 2019) and is NP-hard even on tree networks (Lehmann et al. 2016). A standard approach for addressing this complexity is to use the direct current optimal power flow (DCOPF), which replaces the nonconvex AC power flow constraints with a linearized approximation (Stott et al. 2009). However, even under mild assumptions, the solutions for DCOPF are infeasible for the AC problem (Baker 2021). In practice, AC feasibility can be restored from the DC solution via iterative procedures (Molzahn and Hiskens 2019 Ch. 6, Fang et al. 2022). Despite common use, studies have shown significant divergence in dispatch solutions between the true AC model and the approximated DC variant (Castillo et al. 2016). Additionally, DC approximations may significantly increase carbon emissions associated with the generation solution (Winner et al. 2023) and can introduce artificial grid congestion not found in AC-feasible decisions (Bichler and Knörr 2023).

Historically, the adoption of mixed-integer linear optimization for UC problems has led to savings for system operators on the order of billions of dollars per year (O'Neill et al. 2011, Carlson et al. 2012). Improvements to solution quality from the adoption of AC-based models are estimated to yield annual savings of \$6 – \$19 billion in the United States (Cain et al. 2012). However, a methodology for solving industry-scale UC problems with ACOPF constraints (UC-ACOPF) under realistic time limits has not yet been demonstrated. The value of improved dispatch efficiency and stringent limits on solution time make algorithmic improvements a critical research topic.

To facilitate developments on large-scale UC-ACOPF, the U.S. Department of Energy's Advanced Research Projects Agency-Energy (ARPA-E) conducted the Grid Optimization Competition. The competition consisted of three optimization challenges spanning nine years. Challenge 1 (Aravena et al. 2023) focuses on single-period ACOPF with security constraints, which encourage operational robustness against component failures. Challenge 2 (Elbert et al. 2024a) increases the realism of the power flow model and introduces single-period commitment decisions for fast-starting units. The competition culminated in Grid Optimization Competition Challenge 3 (GOC3; Holzer et al. 2024a), which adopts a multi-period model with full unit commitment decisions, reserve products, and security constraints, yielding a large-scale security-constrained UC-ACOPF model (SCUC-ACOPF). This model considers a rich set of realistic features of modern power grids, including price-sensitive load, various real and reactive reserve products, transformer and transmission line controls, and energy-limited devices (e.g., energy storage and demand response), presenting a level of detail beyond that previously considered in the literature.

To evaluate competing algorithms while complying with critical infrastructure information regulations, synthetic networks and multi-period datasets were generated to resemble realistic grid data under a variety of operating conditions (Birchfield et al. 2017, Li et al. 2018). In GOC3, the networks range in size up to 8,000 nodes, and instances of the SCUC-ACOPF models contain millions of binary variables and nonconvex constraints. Three operational time scales are considered: a day-ahead problem for wholesale markets spanning two days with a one-hour resolution, a real-time problem to account for revised forecasts spanning 8 hours with a 15-minute to one-hour resolution, and a planning process for severe weather events spanning 7 days with a 4-hour resolution (Holzer et al. 2025). In this work, we detail an iterative

decomposition-based algorithm for the GOC3 model and evaluate our algorithm on the numerous datasets from the competition.

1.1 Literature Review

An early Benders' decomposition approach for UC-ACOPF was proposed by Fu et al. (2005), who enforce feasibility for the AC power flow constraints by adding cuts to a master UC problem. Subsequently, alternating methods have been applied to facilitate decomposition. Gholami et al. (2023) apply a bilevel alternating direction method of multipliers (ADMM) to single-period SC-ACOPF, allowing the security constraints to be separated from the ACOPF problem. Zhang et al. (2023) solve UC-ACOPF on small networks (≤ 300 buses) using bilevel ADMM, which facilitates decomposition over the discrete UC and continuous ACOPF portions of the model.

Other work has applied spatial decompositions to ACOPF. Tu et al. (2021) separate subnetworks from a master network (e.g., distribution grids from the transmission grid) and solve a master problem via an interior point method which queries subproblems for gradient and Hessian information. Sun and Sun (2021) decompose a network into regions and use bilevel ADMM to identify a feasible solution. Many other spatial decompositions appear in the literature on distributed computing; for a survey of these techniques, see Molzahn et al. (2017).

Parker and Coffrin (2024) propose a benchmark algorithm for the GOC3 model that separates the nonlinear AC power flow constraints from the discrete unit commitment decisions. The benchmark first solves a copper-plate multi-period unit commitment problem, which does not contain any network power flows. Then, the commitment decisions from this model are fixed in a set of temporally-decomposed AC power flow problems, which ignore reserve and ramping requirements. Feasibility for ramping bounds is restored by projecting the power injection solution onto these constraints. Finally, reserve products are re-dispatched by solving a set of linear programs, with the power injection solution fixed. Chevalier (2024) introduces an alternate approach for the GOC3 problem that combines gradient-based solvers for unconstrained optimization with problem-specific heuristics, demonstrating competitive performance on some instances.

1.2 Contributions

This paper demonstrates that realistic SCUC-ACOPF can be solved to near-optimality at industry scale within stringent time limits. We introduce an iterative algorithm to identify high-quality feasible solutions to the SCUC-ACOPF model proposed in GOC3 and evaluate its performance on a diverse set of test cases. Specifically, our contributions are summarized as follows:

1. We identify a decomposition scheme that allows for the separation of the problem into spatially-independent mixed-integer linear programs (MILPs) for unit commitment, and temporally-independent continuous nonlinear programs for AC power flow. We then propose a penalty alternating direction method that utilizes this decomposition;

2. We prove the convergence of this algorithm to a partial optimal solution, strengthening existing convergence results for penalty alternating direction methods;

3

3. We introduce heuristics to tailor the algorithm to the SCUC-ACOPF problem, including:

   (a) a final solve of AC problems with fixed unit commitment decisions to ensure feasibility,

   (b) restrictions on ramping and energy output that allow for temporal decomposition of the final AC solve while maintaining flexibility in power injection decisions,

   (c) a second-order cone relaxation of the AC constraints to speed up iterations, and

   (d) a screening algorithm to identify high-impact security constraints;

4. We construct a convex relaxation that yields a dual bound on the optimal objective value, allowing for characterization of the quality of feasible solutions; and

5. We evaluate our algorithm on many large-scale test cases with up to 8,000 buses and 48 time periods with time limits as low as 10 minutes. Our algorithm finds high-quality feasible solutions with an average optimality gap of 1.33% relative to the dual bound.

The remainder of this paper is organized as follows. Section 2 introduces a compact representation of the problem formulation. Section 3 describes a penalized model that facilitates spatial and temporal decomposition. Section 4 presents a basic penalty alternating direction method from this decomposition and provides convergence results. Section 5 proposes a set of heuristics which tailor this algorithm to the SCUC-ACOPF setting. Section 6 defines a dual bound on the optimal objective value. Section 7 provides computational results on the GOC3 test cases, demonstrating the performance of our algorithm and the benefits of the heuristics. Section 8 concludes the paper. All proofs are provided in the appendix.

2 The SCUC-ACOPF Model

We introduce a compact representation of the SCUC-ACOPF model proposed in Holzer et al. (2024a). The model includes device-level unit commitment, net power injection, and AC branch control decisions, subject to bus power balance, zonal reserve constraints, device ramping and reserve limits, and security constraints.

We denote $[\cdot]^+ := \max\{\cdot, 0\}$ and $[\![\cdot]\!] := \{1, \ldots, \cdot\}$. The model spans $\mathcal{T} := [\![T]\!]$ time periods, where $t = 0$ represents the initial state of the system. The length of period t is given by $d_t > 0$.

2.1 Device Unit Commitment

Denote by $\mathcal{J}^{\mathrm{pr}}$ (resp. $\mathcal{J}^{\mathrm{cs}}$) the set of producing (resp. consuming) devices, and dispatchable devices by the union $\mathcal{J}^{\mathrm{sd}} := \mathcal{J}^{\mathrm{pr}} \cup \mathcal{J}^{\mathrm{cs}}$. Producing devices generate power and consuming devices consume power. Let u_{jt}^{on}, u_{jt}^{su}, and u_{jt}^{sd} indicate whether device j is on, starting up, or shutting down in period t, respectively. These variables are subject to the following constraints:

$$u_{jt}^{\mathrm{on}} - u_{j,t-1}^{\mathrm{on}} = u_{jt}^{\mathrm{su}} - u_{jt}^{\mathrm{sd}} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \tag{1a}$$

$$u_{jt}^{\mathrm{su}} + u_{jt}^{\mathrm{sd}} \leq 1 \qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \tag{1b}$$

$$\{u_{jt}^{\mathrm{on}}, u_{jt}^{\mathrm{su}}, u_{jt}^{\mathrm{sd}}\}_{t \in \mathcal{T}} \in \mathcal{X}_j^{\mathrm{u}} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \tag{1c}$$

$$(u_{jt}^{\mathrm{on}},\ u_{jt}^{\mathrm{su}},\ u_{jt}^{\mathrm{sd}}) \in \{0,1\}^3 \qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}. \tag{1d}$$

Constraints (1a)–(1b) connect the device startups and shutdowns to their on status. The initial state u_{j0}^{on} is provided as data. Constraint (1c) enforces device-level commitment logic, including must-run and must-outage, minimum downtime and uptime, and maximum startup requirements. These constraints are given by the set $\mathcal{X}_j^{\mathrm{u}}$, which is defined by linear constraints and described in detail in the appendix.

The cost associated with a commitment decision is denoted by $C_j^{\mathrm{uc}}(u)$, which contains fixed shutdown and online costs, as well as downtime-dependent startup costs:

$$C_j^{\mathrm{uc}}(u) := \sum_{t \in \mathcal{T}} \left(c_{jt}^{\mathrm{on}} u_{jt}^{\mathrm{on}} + c_j^{\mathrm{su}} u_{jt}^{\mathrm{su}} + c_j^{\mathrm{sd}} u_{jt}^{\mathrm{sd}} - \sum_{\substack{t' \in \mathcal{T} \\ t' < t}} c_{jtt'}^{\mathrm{dd}} [u_{jt}^{\mathrm{su}} + u_{jt'}^{\mathrm{on}} - 1]^+ \right) \quad \forall j \in \mathcal{J}^{\mathrm{sd}}.$$

The parameter c_{jt}^{on} gives fixed costs associated with committed devices. Parameters $c_j^{\mathrm{su}} \geq 0$ and $c_j^{\mathrm{sd}} \geq 0$ give baseline startup and shutdown costs. Parameters $c_{jtt'}^{\mathrm{dd}} \geq 0$ give downtime-dependent startup cost adjustments that reduce the cost of startup for devices that were recently on. Constraints (1) and the cost functions C_j^{uc} are both representable in a mixed-integer linear setting and contain interactions across time periods, but are separable across devices.

2.2 Device Net Power Injection and Reserves

Let p_{jt}^{tot} and q_{jt}^{tot} be the total real and reactive power injection of device j at time t, and let p_{jt}^{on} give the dispatchable real power. We use the convention that these variables represent injections for producing devices and withdrawals for consuming devices. Total power injection and dispatchable power may differ due to injection during device startup or shutdown. Variables p_{jt}^{su} (resp. p_{jt}^{sd}) give the quantity of injection during startup (resp. shutdown). The parameter $p_{jtt'}^{\mathrm{supc}} \geq 0$ gives the quantity of power injection at time t if a startup is scheduled at future time t'. The set $\mathcal{T}_{jt}^{\mathrm{supc}} := \{t' > t \ : \ p_{jtt'}^{\mathrm{supc}} > 0\}$ contains the periods in which a startup will cause power to be produced at period t. Parameters $p_{jtt'}^{\mathrm{sdpc}}$ and sets $\mathcal{T}_{jt}^{\mathrm{sdpc}} := \{t' \leq t \ : \ p_{jtt'}^{\mathrm{sdpc}} > 0\}$ are defined similarly for shutdowns.

Let $\mathcal{R}$ be the set of real power reserve products. Variable p_{jtr}^{res} gives the quantity of reserve type $r \in \mathcal{R}$ provided by a device. Net real power injection and reserves are constrained as follows:

$$p_{jt}^{\mathrm{tot}} = p_{jt}^{\mathrm{on}} + p_{jt}^{\mathrm{su}} + p_{jt}^{\mathrm{sd}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}, \tag{2a}$$

$$p_{jt}^{\mathrm{su}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} u_{jt'}^{\mathrm{su}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}, \tag{2b}$$

$$p_{jt}^{\mathrm{sd}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} p_{jtt'}^{\mathrm{sdpc}} u_{jt'}^{\mathrm{sd}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}, \tag{2c}$$

$$(u_{jt}^{\mathrm{on}}, p_{jt}^{\mathrm{tot}}, p_{jt}^{\mathrm{on}}, \{p_{jtr}^{\mathrm{res}}\}_{r \in \mathcal{R}}) \in \mathcal{X}_{jt}^{\mathrm{p}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}, \tag{2d}$$

$$(p_{jt}^{\mathrm{tot}}, p_{jt}^{\mathrm{on}}, \{p_{jtr}^{\mathrm{res}}\}_{r \in \mathcal{R}}) \geq 0 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}. \tag{2e}$$

Constraints (2a)–(2c) define total real power in terms of dispatchable and startup/shutdown power. Constraint (2d) enforces device-level constraints that limit real power and reserve provision. The set $\mathcal{X}_{jt}^{\mathrm{p}}$ bounds power injection and reserve quantities with linear constraints that link the commitment status of the device to its real power and reserve variables; this set is described in the appendix.

Reactive power up and down reserves are given by the variables $q_{jt\uparrow}^{\mathrm{res}}$ and $q_{jt\downarrow}^{\mathrm{res}}$, respectively. Reactive power injection and reserve products are constrained by the following logic:

$$q_{jt}^{\mathrm{tot}} \leq q_{jt}^{\max}\left(u_{jt}^{\mathrm{on}} + \sum_{t'\in\mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} + \sum_{t'\in\mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}}\right) + \beta_j^{\max}p_{jt}^{\mathrm{tot}} - \begin{cases} q_{jt\uparrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{pr}} \\ q_{jt\downarrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{cs}} \end{cases} \quad \forall j \in \mathcal{J}^{\mathrm{sd}},\, t \in \mathcal{T}, \quad (3\mathrm{a})$$

$$q_{jt}^{\mathrm{tot}} \geq q_{jt}^{\min}\left(u_{jt}^{\mathrm{on}} + \sum_{t'\in\mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} + \sum_{t'\in\mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}}\right) + \beta_j^{\min}p_{jt}^{\mathrm{tot}} + \begin{cases} q_{jt\downarrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{pr}} \\ q_{jt\uparrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{cs}} \end{cases} \quad \forall j \in \mathcal{J}^{\mathrm{sd}},\, t \in \mathcal{T}, \quad (3\mathrm{b})$$

$$(q_{jt\uparrow}^{\mathrm{res}}, q_{jt\downarrow}^{\mathrm{res}}) \geq 0 \qquad\qquad\qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\, t \in \mathcal{T}. \quad (3\mathrm{c})$$

Constraints (3a)–(3b) give bounds on reactive power injection and reserve products, where reserve products count against the appropriate bound. Parameters $q_{jt}^{\max}$ and $q_{jt}^{\min}$ give the maximum and minimum reactive power bounds. If a device cannot produce real power (i.e., it is not online or producing startup/shutdown power), it may not provide any reactive power or reserves. The reactive power bounds may depend linearly on real power output with coefficients $\beta_j^{\min}$ and $\beta_j^{\max}$.

The cost associated with a power injection decision for a producing device is given by the convex piecewise linear function $C_{jt}^{\mathrm{pow}}(p_{jt}^{\mathrm{tot}})$. Similarly, the utility associated with energy consumption by a consuming device is given by the concave piecewise linear function $R_{jt}^{\mathrm{pow}}(p_{jt}^{\mathrm{tot}})$. These functions only depend on total real power injection for a device and time period. When maximizing utility R_{jt}^{pow} and minimizing cost C_{jt}^{pow}, these functions and the constraints (2)–(3) are representable in a linear setting and are separable by device. However, the constraints (2)–(3) are not separable across time due to interactions with startup and shutdown unit commitment variables across periods.

2.3 Device Energy Injection Limits

Devices may incur a penalty for exceeding maximum or minimum net energy injection limits over some time interval. A set $W \in \mathcal{W}_j^{\mathrm{en,min}}$ is a set of time periods subject to a minimum energy limit ($W \subseteq \mathcal{T}$), and $e_j^{\min}(W)$ gives the value of the corresponding limit. Similarly, $\mathcal{W}_j^{\mathrm{en,max}}$ and $e_j^{\max}(W)$ give data for maximum energy limits. The energy violation cost function is

$$C_j^{\mathrm{e}}(p) := c^{\mathrm{e}}\left(\sum_{W\in\mathcal{W}_j^{\mathrm{en,min}}}\left[e_j^{\min}(W) - \sum_{t\in W} d_t p_{jt}^{\mathrm{tot}}\right]^+ + \sum_{W\in\mathcal{W}_j^{\mathrm{en,max}}}\left[-e_j^{\max}(W) + \sum_{t\in W} d_t p_{jt}^{\mathrm{tot}}\right]^+\right) \quad \forall j \in \mathcal{J}^{\mathrm{sd}},$$

where $c^{\mathrm{e}} \geq 0$ is the energy limit penalty coefficient. When minimizing cost, C_j^{e} is linearly representable and separable across devices but couples time periods.

2.4 Device Real Power Ramping

Changes in device real power injection are subject to ramp rate bounds:

$$p_{jt}^{\mathrm{tot}} - p_{j,t-1}^{\mathrm{tot}} \geq -d_t\left(p_j^{\mathrm{rd}}u_{jt}^{\mathrm{on}} + p_j^{\mathrm{rd,sd}}(1 - u_{jt}^{\mathrm{on}})\right) \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\, t \in \mathcal{T}, \qquad (4\mathrm{a})$$

$$p_{jt}^{\mathrm{tot}} - p_{j,t-1}^{\mathrm{tot}} \le d_t \left(p_j^{\mathrm{ru}}(u_{jt}^{\mathrm{on}} - u_{jt}^{\mathrm{su}}) + p_j^{\mathrm{ru,su}}(u_{jt}^{\mathrm{su}} - u_{jt}^{\mathrm{on}} + 1) \right) \qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}. \qquad (4\mathrm{b})$$

The parameters p_j^{ru} (resp. p_j^{rd}) give the maximum rates for increasing (resp. decreasing) power injection. The parameters $p_j^{\mathrm{ru,su}}$ and $p_j^{\mathrm{rd,sd}}$ give the adjusted rates during startup and shutdown. Initial power injection p_{j0}^{tot} is provided as data. These constraints are separable across devices but couple consecutive time periods.

2.5 Neighborhood Reserve Requirements

Reserve products are subject to minimum requirements across neighborhoods of devices. We denote the set of neighborhoods for real power reserves by $\mathcal{N}^{\mathrm{p}}$, where $N \in \mathcal{N}^{\mathrm{p}}$ is a set of devices j in the same neighborhood ($N \subseteq \mathcal{J}^{\mathrm{sd}}$). We similarly define $\mathcal{N}^{\mathrm{q}}$ as the set of reactive power neighborhoods. A reserve requirement applies to some subset of reserve products $R \subset \mathcal{R}$. Violation of the minimum reserve requirement within a neighborhood is subject to a penalty. The sum of reserve requirement penalties and the cost of providing reserves is given by the functions $C_t^{\mathrm{res}}(p, q)$:

$$
\begin{aligned}
C_t^{\mathrm{res}}(p,q) := \quad & d_t \left(\sum_{\substack{j \in \mathcal{J}^{\mathrm{sd}} \\ r \in \mathcal{R}}} c_{jtr}^{\mathrm{res,p}} p_{jtr}^{\mathrm{res}} + \sum_{R \subset \mathcal{R}} c_R^{\mathrm{res}} \sum_{N \in \mathcal{N}^{\mathrm{p}}} \left[p_t^{\mathrm{res,min}}(p; R, N) - \sum_{r \in R} \sum_{j \in N} p_{jtr}^{\mathrm{res}} \right]^+ \right. \\
& \left. + \sum_{r \in \{\uparrow, \downarrow\}} \left(\sum_{j \in \mathcal{J}^{\mathrm{sd}}} c_{jtr}^{\mathrm{res,q}} q_{jtr}^{\mathrm{res}} + c_r^{\mathrm{res}} \sum_{N \in \mathcal{N}^{\mathrm{q}}} \left[q_t^{\mathrm{res,min}}(q; r, N) - \sum_{j \in N} q_{jtr}^{\mathrm{res}} \right]^+ \right) \right) \quad \forall t \in \mathcal{T}.
\end{aligned}
$$

Variable costs for reserve products are given by parameters $c_{jtr}^{\mathrm{res,p}}$ and $c_{jtr}^{\mathrm{res,q}}$. The functions $p_t^{\mathrm{res,min}}$ and $q_t^{\mathrm{res,min}}$ compute the reserve requirement for the set of products R and neighborhood N. These functions take a constant value, or a proportion of the total ($\sum_{j \in N} p_{jt}^{\mathrm{tot}}$) or maximum ($\max_{j \in N} p_{jt}^{\mathrm{tot}}$) power injection in the neighborhood. Reserve requirement violations for a product set R are penalized at rate c_R^{res}. When minimizing cost, these functions are linearly representable for the given choices of functions $p_t^{\mathrm{res,min}}$ and $q_t^{\mathrm{res,min}}$, and are separable over time but couple devices.

2.6 Alternating Current Optimal Power Flow

We now model the optimal power flow constraints over the transmission network. For a thorough introduction to ACOPF formulations, see Bienstock et al. (2022). Let $\mathcal{I}$ be the set of buses and $\mathcal{J}^{\mathrm{br}}$ the set of branches. These branches are either DC lines ($\mathcal{J}^{\mathrm{dc}}$) or AC branches ($\mathcal{J}^{\mathrm{ac}}$), where transformers are considered AC branches. Each branch j connects an origin bus i_j to a destination bus i_j'. The network may also include shunt devices ($\mathcal{J}^{\mathrm{sh}}$), with shunt j located at bus i_j. Each dispatchable device j is located at bus i_j.

Each bus i has a voltage magnitude control v_{it} and a phase angle control θ_{it}. Transformer j has a tap ratio control τ_{jt} and a phase shift control ϕ_{jt}. For AC lines, these controls are fixed to 1 and 0, respectively. Power flows on DC lines are lossless and bounded. We denote real power flow at the origin and destination nodes of branch j by p_{jt}^{fr} and p_{jt}^{to}, oriented from the bus into the branch. Similarly,

reactive power flows are denoted by q_{jt}^{fr} and q_{jt}^{to}. Shunt device j is controlled by activating some integral number of steps, selected by variable u_{jt}^{sh}. Real and reactive power over the shunt are given by p_{jt}^{sh} and q_{jt}^{sh}. Branch and shunt power flows satisfy the following relations:

$$\Delta_{jt} = \theta_{i_j t} - \theta_{i'_j t} - \phi_{jt} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}, \quad (5\mathrm{a})$$

$$p_{jt}^{\mathrm{fr}} = v_{i_j t}\left(\frac{g_{i_j j} v_{i_j t}}{\tau_{jt}^2} + \frac{v_{i'_j t}}{\tau_{jt}}\left(-g_j \cos(\Delta_{jt}) - b_j \sin(\Delta_{jt})\right)\right) \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}, \quad (5\mathrm{b})$$

$$p_{jt}^{\mathrm{to}} = v_{i'_j t}\left(g_{i'_j j} v_{i'_j t} + \frac{v_{i_j t}}{\tau_{jt}}\left(-g_j \cos(\Delta_{jt}) + b_j \sin(\Delta_{jt})\right)\right) \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}, \quad (5\mathrm{c})$$

$$q_{jt}^{\mathrm{fr}} = v_{i_j t}\left(\frac{-b_{i_j j} v_{i_j t}}{\tau_{jt}^2} + \frac{v_{i'_j t}}{\tau_{jt}}\left(b_j \cos(\Delta_{jt}) - g_j \sin(\Delta_{jt})\right)\right) \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}, \quad (5\mathrm{d})$$

$$q_{jt}^{\mathrm{to}} = v_{i'_j t}\left(-b_{i'_j j} v_{i'_j t} + \frac{v_{i_j t}}{\tau_{jt}}\left(b_j \cos(\Delta_{jt}) + g_j \sin(\Delta_{jt})\right)\right) \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}, \quad (5\mathrm{e})$$

$$p_{jt}^{\mathrm{fr}} + p_{jt}^{\mathrm{to}} = 0 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{dc}},\ t \in \mathcal{T}, \quad (5\mathrm{f})$$

$$p_{jt}^{\mathrm{sh}} = g_j^{\mathrm{sh}} u_{jt}^{\mathrm{sh}} v_{i_j t}^2 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}},\ t \in \mathcal{T}, \quad (5\mathrm{g})$$

$$q_{jt}^{\mathrm{sh}} = -b_j^{\mathrm{sh}} u_{jt}^{\mathrm{sh}} v_{i_j t}^2 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}},\ t \in \mathcal{T}, \quad (5\mathrm{h})$$

$$v_{it} \in [v_i^{\mathrm{min}}, v_i^{\mathrm{max}}] \qquad\qquad \forall i \in \mathcal{I},\ t \in \mathcal{T}, \quad (5\mathrm{i})$$

$$\phi_{jt} \in [\phi_j^{\mathrm{min}}, \phi_j^{\mathrm{max}}],\ \tau_{jt} \in [\tau_j^{\mathrm{min}}, \tau_j^{\mathrm{max}}] \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}, \quad (5\mathrm{j})$$

$$(p_{jt}^{\mathrm{fr}}, p_{jt}^{\mathrm{to}}) \in [-p_j^{\mathrm{max}}, p_j^{\mathrm{max}}]^2,\ (q_j^{\mathrm{fr}}, q_j^{\mathrm{to}}) \in [q_j^{\mathrm{fr,min}}, q_j^{\mathrm{fr,max}}] \times [q_j^{\mathrm{to,min}}, q_j^{\mathrm{to,max}}] \qquad \forall j \in \mathcal{J}^{\mathrm{dc}},\ t \in \mathcal{T}, \quad (5\mathrm{k})$$

$$u_{jt}^{\mathrm{sh}} \in [u_j^{\mathrm{sh,min}}, u_j^{\mathrm{sh,max}}] \cap \mathbb{Z}_+ \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}},\ t \in \mathcal{T}. \quad (5\mathrm{l})$$

Constraints (5a)–(5e) define AC branch real and reactive power flows via branch and bus controls. Parameters g_j and b_j give the series conductance and susceptance of branch j. Parameter g_{ij} is the sum of g_j and the conductance of the branch shunt at bus i, and b_{ij} is the sum of b_j, the charging susceptance of the line at bus i, and the susceptance of the branch shunt at bus i. Constraint (5f) enforces lossless real power balance over DC lines. Constraints (5g)–(5h) define power flows over shunts, where g_j^{sh} and b_j^{sh} give the per-step conductance and susceptance of the shunt. Constraints (5i)–(5l) establish appropriate variable domains, where tap ratio bounds $(\tau_j^{\mathrm{min}}, \tau_j^{\mathrm{max}})$ and voltage bounds $(v_i^{\mathrm{min}}, v_i^{\mathrm{max}})$ are positive and shunt step bounds $(u_j^{\mathrm{sh,min}}, u_j^{\mathrm{sh,max}})$ are integral.

AC branches are subject to a penalty for violating apparent power limits, given by $C_{jt}^{\mathrm{ac}}(p, q)$:

$$C_{jt}^{\mathrm{ac}}(p, q) := c^{\mathrm{s}} d_t \left[\max\left\{ \left\| (p_{jt}^{\mathrm{fr}}, q_{jt}^{\mathrm{fr}}) \right\|_2,\ \left\| (p_{jt}^{\mathrm{to}}, q_{jt}^{\mathrm{to}}) \right\|_2 \right\} - s_j^{\mathrm{max}} \right]^+ \quad \forall j \in \mathcal{J}^{\mathrm{ac}},\ t \in \mathcal{T}.$$

The parameter s_j^{max} gives the apparent power limit for branch j and $c^{\mathrm{s}} \geq 0$ is the penalty coefficient for branch limit violations. The function C_{jt}^{ac} is convex and separable both by branch and time. The AC branch power flow definitions (5b)–(5e) are nonconvex and spatially coupled by bus control decisions. The shunt power flow constraints (5g)–(5h) and (5l) are also nonconvex. Overall, the constraints (5) are spatially coupled but separable over time.

2.7 Nodal Power Balance

The transmission network is subject to real and reactive power balance requirements at each node, which are enforced by penalty functions $C_{it}^{\mathrm{bal}}(p, q)$:

$$
\begin{aligned}
C_{it}^{\mathrm{bal}}(p, q) := d_t \Bigg(c^{\mathrm{p}} \Bigg| &\sum_{\substack{j \in \mathcal{J}^{\mathrm{pr}} \\ i_j = i}} p_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{cs}} \\ i_j = i}} p_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{sh}} \\ i_j = i}} p_{jt}^{\mathrm{sh}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i_j = i}} p_{jt}^{\mathrm{fr}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i'_j = i}} p_{jt}^{\mathrm{to}} \Bigg| \\
+ c^{\mathrm{q}} \Bigg| &\sum_{\substack{j \in \mathcal{J}^{\mathrm{pr}} \\ i_j = i}} q_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{cs}} \\ i_j = i}} q_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{sh}} \\ i_j = i}} q_{jt}^{\mathrm{sh}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i_j = i}} q_{jt}^{\mathrm{fr}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i'_j = i}} q_{jt}^{\mathrm{to}} \Bigg| \Bigg) \quad \forall i \in \mathcal{I}, \ t \in \mathcal{T}.
\end{aligned}
$$

The penalty coefficients c^{p} and c^{q} are nonnegative. When minimizing cost, C_{it}^{bal} is linearly representable and separable over buses and time periods.

2.8 Security Constraints

The active and reactive power set points are optimized in consideration of apparent power flow overloads that may result from branch outages, i.e., contingencies. Such overloads cost

$$
C_{kt}^{\mathrm{ctg}}(p, q) := c^{\mathrm{s}} d_t \sum_{\substack{j \in \mathcal{J}^{\mathrm{ac}} \\ j \neq j_k}} \left[\left\| (f_{kjt}(p), \max\{|q_{jt}^{\mathrm{fr}}|, |q_{jt}^{\mathrm{to}}|\}) \right\|_2 - s_j^{\mathrm{max,ctg}} \right]^+ \quad \forall k \in \mathcal{K}, \ t \in \mathcal{T},
$$

where $\mathcal{K} = [\![K]\!]$ is a set of contingencies, j_k is the branch outaged by contingency k, and the post-contingency apparent power limit on branch j is given by $s_j^{\mathrm{max,ctg}}$. Post-contingency branch flows are computed under a DC power flow model, which reallocates real power, while reactive power flows remain fixed at pre-contingency levels. The affine function $f_{kjt}(p) := \sum_{i \in \mathcal{I}} F_{kji} \left(p_{it}^{\mathrm{inj}}(p) - \frac{p_t^{\mathrm{sl}}(p)}{|\mathcal{I}|} \right)$ computes the new real power flow on branch j after contingency k under the DC model. The function p_{it}^{inj} computes the net real power injection at bus i from all elements except AC branches,

$$
p_{it}^{\mathrm{inj}}(p) := \sum_{\substack{j \in \mathcal{J}^{\mathrm{pr}} \\ i_j = i}} p_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{cs}} \\ i_j = i}} p_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{sh}} \\ i_j = i}} p_{jt}^{\mathrm{sh}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{dc}} \\ i_j = i}} p_{jt}^{\mathrm{fr}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{dc}} \\ i'_j = i}} p_{jt}^{\mathrm{to}},
$$

and the function $p_t^{\mathrm{sl}}(p) := \sum_{i \in \mathcal{I}} p_{it}^{\mathrm{inj}}(p)$ computes the systemwide imbalance in net injections, which is allocated evenly across buses when computing post-contingency flows. Parameters F_{kij} are generation shift factors (GSFs) from bus i to line j under contingency k. These factors are defined explicitly in the appendix; for details, see Wood et al. (2013, App. 8C) and Holzer et al. (2024b). The total contingency cost is the sum of average and worst-case contingency costs:

$$
C_t^{\mathrm{ctg}}(p, q) := \frac{1}{K} \sum_{k \in \mathcal{K}} C_{kt}^{\mathrm{ctg}}(p, q) + \max_{k \in \mathcal{K}} \ C_{kt}^{\mathrm{ctg}}(p, q) \quad \forall t \in \mathcal{T}.
$$

	(1)	(2) − (4)	(5)	C_j^{uc}	$C_{jt}^{\mathrm{pow}}, R_{jt}^{\mathrm{pow}}$	C_j^{e}	C_t^{res}	C_{jt}^{ac}	C_{it}^{bal}	C_t^{ctg}	R_j^{J}	R_t^{T}
Temporal Coupling	✓	✓		✓		✓					✓	
Spatial Coupling			✓				✓		∼	✓		✓
Nonlinearity			✓					✓		✓		✓
Nonconvexity			✓									
Integrality	✓		✓	✓							✓	

Table 1: Sources of complexity by model component. The objective terms C_{it}^{bal} couple devices at the same bus but do not couple different buses; this limited spatial coupling is indicated by the symbol $\sim$.

The functions C_t^{ctg} are convex and separable over time periods.

2.9 The Complete SCUC-ACOPF Model

We combine the cost and utility terms into grouped objective functions:

$$R_t^{\mathrm{T}}(p,q) := -C_t^{\mathrm{res}}(p,q) - C_t^{\mathrm{ctg}}(p,q) - \sum_{j \in \mathcal{J}^{\mathrm{ac}}} C_{jt}^{\mathrm{ac}}(p,q) - \sum_{i \in \mathcal{I}} C_{it}^{\mathrm{bal}}(p,q) \qquad \forall t \in \mathcal{T},$$

$$R_j^{\mathrm{J}}(p,u) := -C_j^{\mathrm{uc}}(u) - C_j^{\mathrm{e}}(p) + \begin{cases} \sum_{t \in \mathcal{T}} R_{jt}^{\mathrm{pow}}(p) & \text{if } j \in \mathcal{J}^{\mathrm{cs}} \\ \sum_{t \in \mathcal{T}} -C_{jt}^{\mathrm{pow}}(p) & \text{if } j \in \mathcal{J}^{\mathrm{pr}} \end{cases} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}}.$$

This partition of the objective functions emphasizes how various terms permit temporal or spatial decomposition. We now construct the full SCUC-ACOPF model:

$$Z^* := \max_{\substack{p,q,u,v \\ \Delta,\theta,\tau,\phi}} \quad \sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p,q) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(p,u) \qquad \text{(SCUC-ACOPF)}$$
$$\text{s.t.} \quad (1) - (5).$$

The model identifies feasible unit commitments, real and reactive power injections, and transmission network controls that maximize utility from delivered power minus operational costs and penalties. The problem is a nonconvex MINLP due to unit commitment decisions and AC power flow constraints. Table 1 categorizes the sources of model complexity, including temporal and spatial coupling, nonlinearity, nonconvexity, and integrality, by constraints and objective terms.

2.10 Assumptions

From now on, we make the following assumption on the problem data. The assumption accompanies the nonnegativity and integrality assumptions introduced previously.

Assumption 1. The problem data is such that:

1. For all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$, it holds that $c_j^{\mathrm{su}} \geq \sum_{\substack{t' \in \mathcal{T} \\ t' < t}} c_{jtt'}^{\mathrm{dd}}$;

2. For all u satisfying (1),

$$u_{jt}^{\mathrm{on}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} \leq 1, \quad u_{jt}^{\mathrm{on}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}} \leq 1 \quad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}; \qquad (6)$$

3. The sets $\mathcal{X}_{jt}^{\mathrm{p}}$ are closed, convex, bounded, and polyhedral for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$; and

4. (SCUC-ACOPF) is feasible with optimal objective value $Z^* \geq 0$.

First, the startup cost adjustments cannot exceed the total startup cost; second, devices cannot simultaneously be committed and produce power related to a future startup or past shutdown; third, the feasible region for real power injection is bounded and defined by linear constraints; and fourth, the model is feasible with a nonnegative optimal objective value.

3   Spatial-Temporal Reformulation

Problem (SCUC-ACOPF) scales with the number of planning periods and the size of the network. Time periods are linked by unit commitment, ramping, and energy limits, and buses are linked by power flows and neighborhood reserves. In this section, we reformulate the model to facilitate a decomposition into continuous nonconvex temporally separate subproblems and mixed-integer linear spatially separate subproblems. We use the notation u_j to denote the subset of variables u that contain index j; for instance, $u_j = \{(u_{jt}^{\mathrm{on}}, u_{jt}^{\mathrm{su}}, u_{jt}^{\mathrm{sd}})\}_{t \in \mathcal{T}}$ if $j \in \mathcal{J}^{\mathrm{sd}}$. Subscripts t are used similarly. We construct the following sets:

- $\mathcal{X}_j^{\mathrm{uc}} := \{(p_j, q_j, u_j) : (1) - (4)\} \quad \forall j \in \mathcal{J}^{\mathrm{sd}}$;

- $\mathcal{X}_t^{\mathrm{ac}} := \{(p_t, q_t, u_t^{\mathrm{sh}}) : \exists(v_t, \Delta_t, \theta_t, \tau_t, \phi_t) \text{ s.t. } (5\mathrm{a}) - (5\mathrm{k})\} \quad \forall t \in \mathcal{T}$;

- $\mathcal{X}_{jt}^{\mathrm{sh}} := \{u_{jt}^{\mathrm{sh}} : (5\mathrm{l})\} \quad \forall j \in \mathcal{J}^{\mathrm{sh}}, \ t \in \mathcal{T}$.

The sets $\mathcal{X}_j^{\mathrm{uc}}$ impose temporal coupling on the decision variables, the sets $\mathcal{X}_t^{\mathrm{ac}}$ impose spatial coupling, and the sets $\mathcal{X}_{jt}^{\mathrm{sh}}$ impose neither type of coupling.

We also construct relaxed versions of the sets that are compatible with the desired decompositions. Specifically, these relaxations of $\mathcal{X}_j^{\mathrm{uc}}$ and $\mathcal{X}_{jt}^{\mathrm{sh}}$ are continuous and separable over time. To facilitate this construction, we introduce the following constraints which relax (2b)–(2c):

$$0 \leq p_{jt}^{\mathrm{su}} \leq \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} u_{jt'}^{\mathrm{su}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}, \tag{7a}$$

$$0 \leq p_{jt}^{\mathrm{sd}} \leq \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} p_{jtt'}^{\mathrm{sdpc}} u_{jt'}^{\mathrm{sd}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}. \tag{7b}$$

This relaxation will be used to aggregate the variables u_{jt}^{su} and u_{jt}^{sd} in Section 5.4. The relaxed sets are defined for all $t \in \mathcal{T}$ as follows:

- $\mathcal{Y}_{jt}^{\mathrm{uc}} := \{(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) : \exists(u_j^{\mathrm{su}}, u_j^{\mathrm{sd}}) \in [0,1]^{2T} \text{ s.t. } (2\mathrm{a}), (2\mathrm{d}), (2\mathrm{e}), (3), (6), (7), u_{jt}^{\mathrm{on}} \geq 0\} \ \forall j \in \mathcal{J}^{\mathrm{sd}}$;

- $\mathcal{Y}_{jt}^{\mathrm{sh}} := \{u_{jt}^{\mathrm{sh}} \in [u_j^{\mathrm{sh,min}}, u_j^{\mathrm{sh,max}}]\} \quad \forall j \in \mathcal{J}^{\mathrm{sh}}$.

Sets $\mathcal{Y}_{jt}^{\mathrm{uc}}$ relax sets $\mathcal{X}_j^{\mathrm{uc}}$ by replacing constraints (1) with (6) and relaxing the binary constraints (1d) to $(u_{jt}^{\mathrm{on}}, u_{jt}^{\mathrm{sd}}, u_{jt}^{\mathrm{su}}) \in [0,1]^3$. Additionally, constraints (2b)–(2c) are relaxed to (7) and constraints (4) are removed. Sets $\mathcal{Y}_{jt}^{\mathrm{sh}}$ relax the integrality restrictions of $\mathcal{X}_{jt}^{\mathrm{sh}}$. Lemma 1 shows that the sets $\mathcal{Y}$ are valid relaxations of their counterparts $\mathcal{X}$.

11

Lemma 1. *It holds that*

1. $\mathcal{X}_j^{\mathrm{uc}} \subseteq \{(p_j, q_j, u_j) : (p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}} \ \forall t \in \mathcal{T}\} \quad \forall j \in \mathcal{J}^{\mathrm{sd}};$ *and*

2. $\mathcal{X}_{jt}^{\mathrm{sh}} \subseteq \mathcal{Y}_{jt}^{\mathrm{sh}} \quad \forall j \in \mathcal{J}^{\mathrm{sh}}, t \in \mathcal{T}.$

From these set representations of the constraints, we construct a copy-constrained model:

$$\max_{\substack{p,q,u \\ \overline{p},\overline{q},\overline{u}}} \ \sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p_t, q_t) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j, \overline{u}_j) \tag{COPY.a}$$

$$\text{s.t.} \quad (p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}} \qquad\qquad\qquad \forall t \in \mathcal{T}, \tag{COPY.b}$$

$$(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, t \in \mathcal{T}, \tag{COPY.c}$$

$$u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}} \qquad\qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}}, t \in \mathcal{T}, \tag{COPY.d}$$

$$(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}} \qquad\qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \tag{COPY.e}$$

$$\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}} \qquad\qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}}, t \in \mathcal{T}, \tag{COPY.f}$$

$$p^{\mathrm{tot}} = \overline{p}^{\mathrm{tot}}, \ p^{\mathrm{su}} = \overline{p}^{\mathrm{su}}, \ p^{\mathrm{sd}} = \overline{p}^{\mathrm{sd}}, \ q^{\mathrm{tot}} = \overline{q}^{\mathrm{tot}}, \ q^{\mathrm{res}} = \overline{q}^{\mathrm{res}}, \ u^{\mathrm{on}} = \overline{u}^{\mathrm{on}}, \ u^{\mathrm{sh}} = \overline{u}^{\mathrm{sh}}. \tag{COPY.g}$$

This model duplicates the variables into blocks (p, q, u) and $(\overline{p}, \overline{q}, \overline{u})$. The variables (p, q, u) are subject to the spatially linked power flow constraints $\mathcal{X}_t^{\mathrm{ac}}$, while the variables $(\overline{p}, \overline{q}, \overline{u})$ are subject to the temporally linked unit commitment, power injection, and ramping constraints $\mathcal{X}_j^{\mathrm{uc}}$ and shunt step constraints $\mathcal{X}_{jt}^{\mathrm{sh}}$. To tighten the formulation in an eventual decomposition, variables (p, q, u) are also constrained by sets $\mathcal{Y}$, which impose a relaxed form of the constraints assigned to $(\overline{p}, \overline{q}, \overline{u})$ such that the temporal separability remains valid. Using the decomposable structure of the objective, the functions R_t^{T} and R_j^{J} take only the local variables (p_t, q_t) and $(\overline{p}_j, \overline{u}_j)$ as input. The final set of constraints are copy constraints, applied to a set of variables chosen so that the model (COPY) is equivalent to (SCUC-ACOPF). This equivalence is given in Proposition 1.

Proposition 1. *The models* (SCUC-ACOPF) *and* (COPY) *are equivalent; that is, given a feasible solution for one model, there exists a corresponding feasible solution for the other model that has the same objective value and the same values for the copied variables.*

As a final step towards decomposition, we penalize the copy constraints in the objective with penalty parameter $\rho \geq 0$ and degree $\ell \in \{1, 2\}$. We define the penalty function

$$\Gamma_\ell(p, q, u, \overline{p}, \overline{q}, \overline{u}) := \left\| (p^{\mathrm{tot}}, p^{\mathrm{su}}, p^{\mathrm{sd}}, q^{\mathrm{tot}}, q^{\mathrm{res}}, u^{\mathrm{on}}, u^{\mathrm{sh}}) - (\overline{p}^{\mathrm{tot}}, \overline{p}^{\mathrm{su}}, \overline{p}^{\mathrm{sd}}, \overline{q}^{\mathrm{tot}}, \overline{q}^{\mathrm{res}}, \overline{u}^{\mathrm{on}}, \overline{u}^{\mathrm{sh}}) \right\|_\ell^\ell$$

and introduce the penalized model:

$$\max_{\substack{p,q,u \\ \overline{p},\overline{q},\overline{u}}} \ \sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p_t, q_t) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j, \overline{u}_j) - \rho \Gamma_\ell(p, q, u, \overline{p}, \overline{q}, \overline{u})$$

$$\text{s.t.} \quad (\text{COPY.b}) - (\text{COPY.f}). \tag{PEN}$$

4 Basic Penalty Alternating Direction Method and Decomposition

Penalizing violations of relaxed constraints in the objective is leveraged in popular algorithms, including penalty methods and augmented Lagrangian methods (Nocedal and Wright 2006, Ch. 17). To allow for decomposition, we apply an alternating direction method to the penalized formulation (PEN). This method alternates between fixing the the second block of variables, i.e., $(\bar{p}, \bar{q}, \bar{u})$, and optimizing over the first block, i.e., (p, q, u), and fixing the first block of variables and optimizing over the second block. This style of algorithm, applied to a problem with constraints penalized in the objective, is referred to as a penalty alternating direction method (Geißler et al. 2017) and has been used to solve MINLPs for optimal gas transport (Geißler et al. 2015). These methods have structural similarities to the feasibility pump, a heuristic for identifying feasible solutions to problems with integrality constraints (Fischetti et al. 2005, Bonami et al. 2009). In this section, we propose a basic version of the algorithm and describe its convergence.

As the first block of variables are not coupled across time periods, we define a set of temporally-decomposed block subproblems. For each $t \in \mathcal{T}$, the block subproblem is given by

$$
\begin{aligned}
Z_t^{\mathrm{T}}(\bar{p}_t, \bar{q}_t, \bar{u}_t; \rho) \;\; := \;\; \max_{p_t, q_t, u_t} \;\; & R_t^{\mathrm{T}}(p_t, q_t) - \rho \Gamma_2(p_t, q_t, u_t, \bar{p}_t, \bar{q}_t, \bar{u}_t) \\
\text{s.t.} \;\; & (p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}, \\
& (p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \\
& u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}}.
\end{aligned}
$$

These models select degree $\ell = 2$ so that auxiliary variables are not needed to model the L_1 norm. As the constraints and function R_t^{T} contain nonlinearities, this choice does not increase the problem complexity. The temporally-decomposed block subproblems are continuous nonconvex programs, which can be solved to stationarity by interior point methods (Wächter and Biegler 2005).

Next, we define the block subproblems for the second block of variables, which are coupled temporally but not spatially. This block also does not contain interactions between the shunt step variables $\bar{u}_{jt}^{\mathrm{sh}}$ and other decisions, so the variables $\bar{u}_{jt}^{\mathrm{sh}}$ can be separated into their own blocks. For this reason, we revise the copy penalty function to exclude these variables:

$$
\Gamma_\ell'(p, q, u, \bar{p}, \bar{q}, \bar{u}) := \left\| (p^{\mathrm{tot}}, p^{\mathrm{su}}, p^{\mathrm{sd}}, q^{\mathrm{tot}}, q^{\mathrm{res}}, u^{\mathrm{on}}) - (\bar{p}^{\mathrm{tot}}, \bar{p}^{\mathrm{su}}, \bar{p}^{\mathrm{sd}}, \bar{q}^{\mathrm{tot}}, \bar{q}^{\mathrm{res}}, \bar{u}^{\mathrm{on}}) \right\|_\ell^\ell .
$$

Then, the spatially-decomposed device-level block subproblems are given, for each $j \in \mathcal{J}^{\mathrm{sd}}$, by

$$
\begin{aligned}
Z_j^{\mathrm{J}}(p_j, q_j, u_j; \rho) \;\; := \;\; \max_{\bar{p}_j, \bar{q}_j, \bar{u}_j} \;\; & R_j^{\mathrm{J}}(\bar{p}_j, \bar{u}_j) - \rho \Gamma_1'(p_j, q_j, u_j, \bar{p}_j, \bar{q}_j, \bar{u}_j) \\
\text{s.t.} \;\; & (\bar{p}_j, \bar{q}_j, \bar{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}.
\end{aligned}
$$

These problems select degree $\ell = 1$ so that each problem can be represented as a mixed-integer linear program. The shunt variables $\bar{u}_{jt}^{\mathrm{sh}}$ are only subject to interval and integrality constraints and thus can be

---

Algorithm 1: Basic Penalty Alternating Direction Method (pADM)

Input: Penalty schedule $\{\rho_\tau\}_{\tau=1}^{\overline{\tau}} \subset \mathbb{R}_>$

1 **for** $\tau \in \{1, \ldots, \overline{\tau}\}$ **do**

2   **if** $\tau = 1$ **then**

3   For $t \in \mathcal{T}$, solve $Z_t^{\mathrm{T}}(0,0,0;0)$ to solution $(p_t^{(\tau)}, q_t^{(\tau)}, u_t^{(\tau)})$;

4   **else**

5   For $t \in \mathcal{T}$, solve $Z_t^{\mathrm{T}}(\overline{p}_t^{(\tau-1)}, \overline{q}_t^{(\tau-1)}, \overline{u}_t^{(\tau-1)}; \rho_\tau)$ to solution $(p_t^{(\tau)}, q_t^{(\tau)}, u_t^{(\tau)})$;

6   **end**

7   For $j \in \mathcal{J}^{\mathrm{sd}}$, solve $Z_j^{\mathrm{J}}(p_j^{(\tau)}, q_j^{(\tau)}, u_j^{(\tau)}; \rho_\tau)$ to solution $(\overline{p}_j^{(\tau)}, \overline{q}_j^{(\tau)}, \overline{u}_j^{(\tau)})$;

8   For $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$, compute $\overline{u}_{jt}^{\mathrm{sh}(\tau)} = Z_{jt}^{\mathrm{SH}}(u_{jt}^{\mathrm{sh}(\tau)})$;

9 **end**

10 **return** $(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})})$

---

optimized in closed form via rounding:

$$
Z_{jt}^{\mathrm{SH}}(u_{jt}^{\mathrm{sh}}) \ := \ \left\lfloor \frac{1}{2} + \operatorname*{proj}_{\left[u_j^{\mathrm{sh,min}}, u_j^{\mathrm{sh,max}}\right]}(u_{jt}^{\mathrm{sh}}) \right\rceil \ \in \ \operatorname*{arg\,max}_{\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}} \quad -\rho \left\| u_{jt}^{\mathrm{sh}} - \overline{u}_{jt}^{\mathrm{sh}} \right\|_1 .
$$

We now define the basic penalty alternating direction method (pADM), given as Algorithm 1. The method takes as input a set of penalty coefficients $\{\rho_\tau\}_{\tau=1}^{\overline{\tau}}$ and outputs a feasible solution to the device-level subproblems Z_j^{J} in $\overline{\tau}$ iterations. A feasible solution to the device-level subproblems corresponds to a feasible solution to (COPY); this property is formalized in Proposition 2. Note that, in the first iteration, the subproblems Z_t^{T} are solved with $\rho = 0$, eliminating the penalty term.

Proposition 2. *Let $(\overline{p}, \overline{q}, \overline{u})$ satisfy $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. Then, there is some (p, q, u) such that $(p, q, u, \overline{p}, \overline{q}, \overline{u})$ is feasible for (COPY).*

Proposition 2, combined with Proposition 1, implies that a feasible solution to the device-level subproblems corresponds to a feasible solution to (SCUC-ACOPF) where the copied variables are fixed to their values in the device-level solution. As the output of Algorithm 1 is a feasible solution to the device-level problems, this output corresponds to a feasible solution of (SCUC-ACOPF).

Theorem 1 characterizes the convergence of the algorithm, showing that the infeasibility of the iterates for the copy constraint is bounded by the inverse square root of the penalty parameter. The result also shows that a similar bound holds on the distance between the copied variables in consecutive iterations. To state the result, we define D_t as the diameter of the (bounded) feasible region of Z_t^{T} and L_t as the Lipschitz constant of the functions R_t^{T} over this region, so that every subgradient of R_t^{T} on the region is bounded in magnitude by L_t. Similarly, R_j is the maximum difference in the value of the function R_j^{J} over the (bounded) feasible region of Z_j^{J}, and $R := \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j$. Finally, let d be the number of copied variables, i.e., $(p^{\mathrm{tot}}, p^{\mathrm{su}}, p^{\mathrm{sd}}, q^{\mathrm{tot}}, q^{\mathrm{res}}, u^{\mathrm{on}}, u^{\mathrm{sh}}) \in \mathbb{R}^d$. We require that the nonlinear subproblems Z_t^{T} are solved to stationarity (i.e., the solution satisfies the Karush-Kuhn-Tucker conditions, see Nocedal and Wright 2006 Thm. 12.1) and the mixed-integer linear subproblems Z_j^{J} are solved to global optimality; these are natural assumptions for each problem class.

Theorem 1. *Let* $(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$ *be a sequence of iterates generated by Algorithm 1, where the subproblems* Z_t^{T} *are solved to stationarity. Then,*

$$\Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)}) \leq \sqrt{\frac{C}{2\rho_\tau}} \quad \forall \tau \in [\![\tau]\!] \setminus \{1\},$$

where $C := \sum_{t \in \mathcal{T}} dL_t D_t$. *If the subproblems* Z_j^{J} *are solved to global optimality, then*

$$\begin{aligned}
\Gamma_1(\overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)}) &\leq \frac{R}{\rho_\tau} + \sqrt{\frac{2C}{\rho_\tau}} \\
\Gamma_1(p^{(\tau+1)}, q^{(\tau+1)}, u^{(\tau+1)}, p^{(\tau)}, q^{(\tau)}, u^{(\tau)}) &\leq \frac{R}{\rho_\tau} + \sqrt{\frac{2C}{\min\{\rho_\tau, \rho_{\tau+1}\}}}
\end{aligned} \quad \forall \tau \in [\![\tau]\!] \setminus \{1\}.$$

Theorem 2 gives that, as the penalty parameter grows arbitrarily large, the iterates converge to a partial optimum for (COPY). A *partial optimum* is a solution $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$ that is (globally) optimal for (COPY) when each block of variables is fixed at the solution value, i.e., $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$ is optimal for (COPY) with the variables (p, q, u) fixed to (p^*, q^*, u^*) and is optimal for (COPY) with the variables $(\overline{p}, \overline{q}, \overline{u})$ fixed to $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$. Convergence to partial optima is common in the analysis of alternating methods (Gorski et al. 2007). Existing results for penalty alternating direction methods require fixing a penalty parameter and alternating between blocks until convergence is observed; after convergence, the penalty coefficient is increased (Geißler et al. 2017). Our result leverages the structure of (COPY) to allow for changes to the penalty parameter in every iteration, reducing the number of required subproblem evaluations. Theorem 2 requires that both sets of subproblems are solved to global optimality, which may be impractical for the nonconvex subproblems Z_t^{T}. Nevertheless, the result gives an improved theoretical understanding of the algorithm in this setting.

Theorem 2. *Suppose that* $\overline{\tau} = \infty$ *and* $\lim_{\tau \to \infty} \rho_\tau = \infty$. *Let* $(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$ *be iterates generated by Algorithm 1, where the subproblems* Z_t^{T} *and* Z_j^{J} *are solved to global optimality. Then, any limit point* $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$ *of the sequence of iterates is a partial optimum for* (COPY).

Although the basic pADM algorithm satisfies desirable convergence properties, it often fails to efficiently generate high-quality feasible solutions for (SCUC-ACOPF). In Section 5, we highlight several insufficiencies of this algorithm and propose approaches to improve performance.

5 Tailoring the pADM: Improvements and Heuristics

Algorithm 1 gives a basic penalty alternating direction framework that allows for spatial and temporal decomposition of (SCUC-ACOPF). In this section, we adapt the algorithm to improve the solution quality and reduce the computation time. Section 5.1 adds a set of final subproblems that fix the discrete variables and reoptimize over the continuous variables. Section 5.2 accelerates the iterations by replacing the nonconvex temporally-decomposed block subproblems with their second-order cone (SOC) relaxations. Section 5.3 reduces the number of contingencies included in the subproblems via a screening algorithm. Section 5.4 adds terms to the block subproblem objectives to more closely approximate the true objective function.

5.1 Exact Feasibility with Final Subproblems

The convergence property in Theorem 1 establishes that, as the penalty coefficient ρ grows, the iterates of Algorithm 1 converge towards feasibility for the copy constraints in the formulation (COPY). However, this convergence may be slow in practice, requiring many iterations and excessively large values for ρ, which may introduce computational difficulties and numerical instability. We ensure exact feasibility of the final solution by fixing the discrete unit commitment variables u to their value from the last-iteration device-level block subproblems Z_j^{J} and solving the original formulation (SCUC-ACOPF). The resulting restricted problem is temporally coupled only by the ramping constraints (4) and energy violation penalty functions C_j^{e}. Using the device-level power injection solution from the last iteration of Z_j^{J}, we construct restrictions of the linking elements that allow for temporal decomposition of the final problem.

5.1.1 Ramping Constraint Restriction

When solving the final models with fixed unit commitment decisions $\bar{u}$, fixing the real power injection variables p_{jt}^{tot} to the solution $\bar{p}_{jt}^{\mathrm{tot}}$ from the device-level block subproblems Z_j^{J} ensures feasibility for the ramping constraints (4) and allows for temporal separability but empirically leads to poor solution quality. Allowing real power injection to deviate from the device-level subproblem solution introduces flexibility, which leads to better solutions and improves the computational performance of the interior point solver. We introduce such flexibility by restricting p_{jt}^{tot} to lie in an interval around the solution $\bar{p}_{jt}^{\mathrm{tot}}$:

$$\bar{p}_{jt}^{\mathrm{tot}} - \delta_{jt}^- \leq p_{jt}^{\mathrm{tot}} \leq \bar{p}_{jt}^{\mathrm{tot}} + \delta_{jt}^+ \qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}. \tag{9}$$

For $j \in \mathcal{J}^{\mathrm{sd}}$, the parameters δ_j are optimal solutions to a set of auxiliary linear programs:

$$H_j(\bar{p}_j, \bar{u}_j) := \arg\max_{\delta_j^-, \delta_j^+} \ \sum_{t \in \mathcal{T}} (\delta_{jt}^- + \delta_{jt}^+) \tag{10a}$$

$$\text{s.t.} \quad \delta_{jt}^- + \delta_{j,t-1}^+ \leq r_{jt}^{\mathrm{d}} + (\bar{p}_{jt}^{\mathrm{tot}} - \bar{p}_{j,t-1}^{\mathrm{tot}}) \qquad \forall t \in \mathcal{T}, \tag{10b}$$

$$\delta_{j,t-1}^- + \delta_{jt}^+ \leq r_{jt}^{\mathrm{u}} - (\bar{p}_{jt}^{\mathrm{tot}} - \bar{p}_{j,t-1}^{\mathrm{tot}}) \qquad \forall t \in \mathcal{T}, \tag{10c}$$

$$\delta_{jt}^- \geq \frac{1}{2} \min\left\{ r_{jt}^{\mathrm{d}} + (\bar{p}_{jt}^{\mathrm{tot}} - \bar{p}_{j,t-1}^{\mathrm{tot}}), \ r_{j,t+1}^{\mathrm{u}} - (\bar{p}_{j,t+1}^{\mathrm{tot}} - \bar{p}_{jt}^{\mathrm{tot}}) \right\} \qquad \forall t \in \mathcal{T}, \tag{10d}$$

$$\delta_{jt}^+ \geq \frac{1}{2} \min\left\{ r_{jt}^{\mathrm{u}} - (\bar{p}_{jt}^{\mathrm{tot}} - \bar{p}_{j,t-1}^{\mathrm{tot}}), \ r_{j,t+1}^{\mathrm{d}} + (\bar{p}_{j,t+1}^{\mathrm{tot}} - \bar{p}_{jt}^{\mathrm{tot}}) \right\} \qquad \forall t \in \mathcal{T}, \tag{10e}$$

$$r_{jt}^{\mathrm{d}} := d_t \left(p_j^{\mathrm{rd}} \bar{u}_{jt}^{\mathrm{on}} + p_j^{\mathrm{rd,sd}} (1 - \bar{u}_{jt}^{\mathrm{on}}) \right) \qquad \forall t \in \mathcal{T}, \tag{10f}$$

$$r_{jt}^{\mathrm{u}} := d_t \left(p_j^{\mathrm{ru}} (\bar{u}_{jt}^{\mathrm{on}} - \bar{u}_{jt}^{\mathrm{su}}) + p_j^{\mathrm{ru,su}} (\bar{u}_{jt}^{\mathrm{su}} - \bar{u}_{jt}^{\mathrm{on}} + 1) \right) \qquad \forall t \in \mathcal{T}. \tag{10g}$$

These problems maximize the cumulative length of all intervals while enforcing that any solution within the intervals is feasible for the ramping constraints. Constraints (10b)–(10c) enforce feasibility of the interval for the original ramping constraints (4), where parameters r_{jt}^{d} and r_{jt}^{u} give the maximum ramp down and up power under unit commitment solution $\bar{u}$. To ease notation, we fix $\delta_{j0}^- = \delta_{j0}^+ = 0$. Constraints (10d)–(10e) implement a heuristic that requires the interval lengths above and below the solution $\bar{p}_{jt}^{\mathrm{tot}}$ to be at least as large as in the solution that distributes the ramping slack evenly across time periods.

Proposition 3 states that the problems (10) are feasible and the optimal parameters (δ_j^-, δ_j^+) are nonnegative. Proposition 4 shows that, when the variables u are fixed to $\overline{u}$, the constraint (9) is a restriction of (4), and thus any solution $(p, \overline{p}, \overline{u}, \delta)$ satisfying (9)–(10) is feasible for (4).

Proposition 3. *Let $j \in \mathcal{J}^{\mathrm{sd}}$ and $(\overline{p}, \overline{q}, \overline{u})$ be a solution such that $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$. Then, problem (10) is feasible and bounded, and every (δ_j^-, δ_j^+) satisfying (10b)–(10e) is nonnegative.*

Proposition 4. *Let $(\overline{p}, \overline{u}, \delta)$ be a solution that satisfies (10b)–(10c) for all $j \in \mathcal{J}^{\mathrm{sd}}$. Then, for any p satisfying (9), the solution $(p, \overline{u})$ is feasible for (4).*

By Proposition 4, the heuristic cuts (10d)–(10e) are not needed to enforce a restriction of (4). In practice, without these cuts, problem (10) yields many near-zero interval lengths, which does not improve device controllability. One alternative is to maximize $\log(\prod_t (\delta_{jt}^- + \delta_{jt}^+ + 1)) = \sum_t \log(\delta_{jt}^- + \delta_{jt}^+ + 1)$, representing the *volume* generated by the intervals. This objective decreases the fraction of near-zero interval lengths but renders the problem nonlinear. Adding the heuristic cuts yields solutions that achieve most of the volume and non-zero interval lengths obtained by the volume objective and preserves linearity; see Section 7.4.1 for quantitative results.

5.1.2 Energy Injection Limit Restriction

We similarly design a restriction of the energy violation penalty functions C_j^{e} defined in Section 2.3. As (SCUC-ACOPF) minimizes the penalty, we define a restriction as an upper bound on the function. The bound will be the sum of time-separable functions $\overline{C}_{jt}^{\mathrm{e}}$ centered around some candidate injection $\overline{p}$:

$$
\overline{C}_{jt}^{\mathrm{e}}(p; \overline{p}) := c^{\mathrm{e}} d_t \left(\sum_{W \in \mathcal{W}_{jt}^{\mathrm{en,min}}} \left[\frac{e_j^{\min}(W) - \sum_{t' \in W} d_{t'} \overline{p}_{jt'}^{\mathrm{tot}}}{\sum_{t' \in W} d_{t'}} + \overline{p}_{jt}^{\mathrm{tot}} - p_{jt}^{\mathrm{tot}} \right]^+ \right.
$$

$$
\left. + \sum_{W \in \mathcal{W}_{jt}^{\mathrm{en,max}}} \left[\frac{-e_j^{\max}(W) + \sum_{t' \in W} d_{t'} \overline{p}_{jt'}^{\mathrm{tot}}}{\sum_{t' \in W} d_{t'}} + p_{jt}^{\mathrm{tot}} - \overline{p}_{jt}^{\mathrm{tot}} \right]^+ \right) \quad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T},
$$

where $\mathcal{W}_{jt}^{\mathrm{en,min}} := \{W \in \mathcal{W}_j^{\mathrm{en,min}} : t \in W\}$ is the set of minimum energy periods that intersect with time t; $\mathcal{W}_{jt}^{\mathrm{en,max}}$ is defined analogously. Proposition 5 states that the functions $\overline{C}_{jt}^{\mathrm{e}}$ provide the desired bound. In practice, the injection $\overline{p}$ will be a feasible solution to the device-level subproblems.

Proposition 5. *For any $\overline{p}$, p, and $j \in \mathcal{J}^{\mathrm{sd}}$, it holds that $C_j^{\mathrm{e}}(p) \leq \sum_{t \in \mathcal{T}} \overline{C}_{jt}^{\mathrm{e}}(p; \overline{p})$.*

5.1.3 Final Subproblem Formulation

Let $(\overline{p}, \overline{q}, \overline{u})$ be a feasible solution to the device-level block subproblems, i.e., $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. Let δ satisfy (10). We define a final objective function R_t^{F} that excludes unit commitment and contingency costs and applies the energy restriction:

$$
R_t^{\mathrm{F}}(p, q; \overline{p}) := \sum_{j \in \mathcal{J}^{\mathrm{cs}}} R_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{pr}}} C_{jt}^{\mathrm{pow}}(p) - C_t^{\mathrm{res}}(p, q) - \sum_{j \in \mathcal{J}^{\mathrm{ac}}} C_{jt}^{\mathrm{ac}}(p, q) - \sum_{i \in \mathcal{I}} C_{it}^{\mathrm{bal}}(p, q) - \sum_{j \in \mathcal{J}^{\mathrm{sd}}} \overline{C}_{jt}^{\mathrm{e}}(p; \overline{p}).
$$

For $t \in \mathcal{T}$, the temporally-separate final subproblems that incorporate the ramping and energy restrictions are defined by

$$
Z_t^{\mathrm{F}}(\bar{p}, \bar{q}, \bar{u}, \emptyset; \delta) := \max_{p_t, q_t, u_t} \quad R_t^{\mathrm{F}}(p_t, q_t; \bar{p})
$$

$$
\text{s.t.} \quad (2)-(3), (9), (p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}, \ u = \bar{u}.
$$

The final constraint fixes the unit commitment variables to the device-level solution $\bar{u}$.

Propositions 1 and 2 imply that there is some feasible solution for (SCUC-ACOPF) with the copied variables fixed to the device-level solution; such a solution satisfies the ramping restriction (9) and thus the final subproblems Z_t^{F} are feasible. The final argument δ_t of Z_t^{F} denotes the dependence of the ramping restriction on the interval parameters from the auxiliary problems (10). The argument $\emptyset$ indicates that the final subproblems exclude contingency costs. Section 5.3 expands this notation to incorporate contingencies.

5.2 Second-Order Cone Relaxation

As the final subproblems Z_t^{F} only require a feasible solution to the device-level subproblems Z_j^{J}, we can use a proxy in place of the temporally-decomposed subproblems Z_t^{T} in the iterative algorithm. We choose an SOC relaxation similar to those by Jabr (2006) and Kocuk et al. (2016). The relaxation accelerates each iteration and yields subproblems that can be solved to global optimality. Feasibility for the original ACOPF constraints is restored by solving the final subproblems. The relaxation is applied to the constraints (5a)–(5l). We define the change of variables

$$
c_{jt} = \frac{1}{\tau_{jt}} v_{i_j t} v_{i'_j t} \cos(\Delta_{jt}), \quad s_{jt} = \frac{1}{\tau_{jt}} v_{i_j t} v_{i'_j t} \sin(\Delta_{jt}), \quad \omega_{it} = v_{it}^2, \quad \mu_{jt} = \frac{v_{i_j t}^2}{\tau_{jt}^2}, \quad \text{and} \quad \mu_{jt}^{\mathrm{sh}} = u_{jt}^{\mathrm{sh}} v_{i_j t}^2.
$$

This transformation is invertible if and only if $c_{jt}^2 + s_{jt}^2 = \mu_{jt}\omega_{i'_j t}$ and $\Delta_{jt} = \arctan(s_{jt}/c_{jt})$. To convexify these conditions, we relax the first constraint to $c_{jt}^2 + s_{jt}^2 \leq \mu_{jt}\omega_{i'_j t}$ and remove the second constraint, eliminating the angle differences Δ_{jt} from the model. Then, the following defines the SOC relaxation:

$$
p_{jt}^{\mathrm{fr}} = g_{i_j j}\mu_{jt} - g_j c_{jt} - b_j s_{jt} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}}, \ t \in \mathcal{T}, \tag{11a}
$$

$$
p_{jt}^{\mathrm{to}} = g_{i'_j j}\omega_{i'_j t} - g_j c_{jt} + b_j s_{jt} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}}, \ t \in \mathcal{T}, \tag{11b}
$$

$$
q_{jt}^{\mathrm{fr}} = -b_{i_j j}\mu_{jt} + b_j c_{jt} - g_j s_{jt} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}}, \ t \in \mathcal{T}, \tag{11c}
$$

$$
q_{jt}^{\mathrm{to}} = -b_{i'_j j}\omega_{i'_j t} + b_j c_{jt} + g_j s_{jt} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}}, \ t \in \mathcal{T}, \tag{11d}
$$

$$
p_{jt}^{\mathrm{sh}} = g_j^{\mathrm{sh}} \mu_{jt}^{\mathrm{sh}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}}, \ t \in \mathcal{T}, \tag{11e}
$$

$$
q_{jt}^{\mathrm{sh}} = -b_j^{\mathrm{sh}} \mu_{jt}^{\mathrm{sh}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}}, \ t \in \mathcal{T}, \tag{11f}
$$

$$
c_{jt}^2 + s_{jt}^2 \leq \mu_{jt}\omega_{i'_j t} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}}, \ t \in \mathcal{T}, \tag{11g}
$$

$$
\frac{\omega_{i_j t}}{(\tau_j^{\max})^2} \leq \mu_{jt} \leq \frac{\omega_{i_j t}}{(\tau_j^{\min})^2} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{ac}}, \ t \in \mathcal{T}, \tag{11h}
$$

$$
u_j^{\mathrm{sh,min}}\omega_{i_j t} \leq \mu_{jt}^{\mathrm{sh}} \leq u_j^{\mathrm{sh,max}}\omega_{i_j t} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sh}}, \ t \in \mathcal{T}, \tag{11i}
$$

$$
\omega_{it} \in [(v_i^{\min})^2, (v_i^{\max})^2] \qquad\qquad \forall i \in \mathcal{I}, \ t \in \mathcal{T}, \tag{11j}
$$

$$(5f), (5k). \tag{11k}$$

We denote the region defined by these constraints as $\mathcal{X}_t^{\mathrm{soc}} := \{(p_t, q_t, \mu_t, \omega_t) : \exists(c_t, s_t) \text{ s.t. } (11)\}$. As the constraints (11g) are SOC-representable and the other constraints are linear, the sets $\mathcal{X}_t^{\mathrm{soc}}$ are convex and SOC-representable. Proposition 6 states that these sets relax constraints (5).

Proposition 6. *For all $t \in \mathcal{T}$, the constraints (11) are a convex relaxation of the power flow constraints (5). In particular, for all $t \in \mathcal{T}$, the set $\mathcal{X}_t^{\mathrm{soc}}$ satisfies*

$$\left\{ (p_t, q_t, u_t^{\mathrm{sh}}) : \exists(\mu_t, \omega_t) \text{ s.t. } (p_t, q_t, \mu_t, \omega_t) \in \mathcal{X}_t^{\mathrm{soc}}, \ u_{jt}^{\mathrm{sh}} = \mu_{jt}^{\mathrm{sh}}/\omega_{i_jt} \ \forall j \in \mathcal{J}^{\mathrm{sh}} \right\}$$
$$\supseteq \{(p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}} : u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}} \ \forall j \in \mathcal{J}^{\mathrm{sh}}\}.$$

When constructing $\mathcal{X}_t^{\mathrm{soc}}$, the variables u_{jt}^{sh} are projected out. However, the corresponding value for these variables can be recovered by $u_{jt}^{\mathrm{sh}} = \frac{\mu_{jt}^{\mathrm{sh}}}{\omega_{i_jt}}$. The quadratic penalty terms for these variables, which appear in the penalty functions Γ_2, can instead be written as

$$(u_{jt}^{\mathrm{sh}} - \overline{u}_{jt}^{\mathrm{sh}})^2 = (\mu_{jt}^{\mathrm{sh}}/\omega_{i_jt} - \overline{u}_{jt}^{\mathrm{sh}})^2 = (\mu_{jt}^{\mathrm{sh}} - \overline{u}_{jt}^{\mathrm{sh}}\omega_{i_jt})^2/\omega_{i_jt}^2 \approx (\mu_{jt}^{\mathrm{sh}} - \overline{u}_{jt}^{\mathrm{sh}}\omega_{i_jt})^2,$$

where the approximate equality is exact when voltages are 1. The approximating term is convex quadratic in (μ, ω). Any error from $\omega_{i_jt} \neq 1$ simply rescales the penalty coefficients of this term.

5.3 Contingency Screening

Large numbers of contingencies may significantly increase subproblem solution times, as each contingency k requires the incorporation of its nonlinear (albeit convex) cost function C_{kt}^{ctg} into the objective (see Section 2.8). Even computing the GSFs F_{kji} requires a matrix inversion, which may be costly. Empirically, many contingencies incur minimal cost and thus have little impact on the optimal solution, which motivates contingency screening. We identify a set of contingencies $\overline{\mathcal{K}} \subset \mathcal{K}$ and branches $\overline{\mathcal{J}}_{kt} \subset \mathcal{J}^{\mathrm{ac}}$ for $k \in \overline{\mathcal{K}}$ that incur a large contingency cost at a candidate solution $(\overline{p}, \overline{q})$.

To avoid evaluating every contingency, we first identify a set of contingencies that are likely to incur penalties. Specifically, this set of prescreened contingencies is denoted by $\overline{\mathcal{K}}$ and contains the N^{ctg} contingencies with largest real power flows on the outaged line j_k in the candidate solution. We then evaluate the post-contingency line flows for prescreened contingencies and select $\overline{\mathcal{J}}_{kt}$ as the set of the N^{br} branches with the largest post-contingency apparent power limit violations exceeding some threshold $s^{\mathrm{thr}} \geq 0$. Post-contingency line flows are evaluated by computing the GSFs. To compute these parameters efficiently, we update the required matrix inverse by a Sherman-Morrison rank-1 update on a base matrix (Hager 1989), avoiding the inversion of a new matrix for each contingency (Holzer et al. 2024b). See the appendix for additional details.

We approximate the post-contingency real power flow functions by

$$\overline{f}_{kjt}(p; \overline{\mathcal{I}}_{kjt}) := \sum_{i \in \overline{\mathcal{I}}_{kjt}} F_{kji} \left(p_{it}^{\mathrm{inj}}(p) - \frac{p_t^{\mathrm{sl}}(p)}{|\mathcal{I}|} \right).$$

Algorithm 2: Contingency Screening

Input: Candidate power injection solution $(\overline{p}, \overline{q})$, thresholds $N^{\text{ctg}}, N^{\text{br}}, N^{\text{bus}}, s^{\text{thr}}$

1 Compute the base-case GSF matrix;

2 Define $\overline{\mathcal{K}} := \underset{k \in \mathcal{K}}{\arg\max}{}^{N^{\text{ctg}}} \ \underset{t \in \mathcal{T}}{\max} \ \max\{|\overline{p}_{j_k t}^{\text{fr}}|, |\overline{p}_{j_k t}^{\text{to}}|\};$

3 for $k \in \overline{\mathcal{K}}, \ t \in \mathcal{T}$ **do**

4 $\quad$ Compute post-contingency GSF matrix F_k via a rank-1 update and construct functions f_{kjt} for all $j \in \mathcal{J}^{\text{ac}}$;

5 $\quad$ Define $\mathcal{J}_{kt}^{\text{viol}} := \{j \in \mathcal{J}^{\text{ac}} \setminus \{j_k\} : \left\| (f_{kjt}(\overline{p}), \max\{|\overline{q}_{jt}^{\text{fr}}|, |\overline{q}_{jt}^{\text{to}}|\}) \right\|_2 - s_j^{\max,\text{ctg}} \geq s^{\text{thr}}\};$

6 $\quad$ Define $\overline{\mathcal{J}}_{kt} := \underset{j \in \mathcal{J}_{kt}^{\text{viol}}}{\arg\max}{}^{N^{\text{br}}} \left\| (f_{kjt}(\overline{p}), \max\{|\overline{q}_{jt}^{\text{fr}}|, |\overline{q}_{jt}^{\text{to}}|\}) \right\|_2 - s_j^{\max,\text{ctg}};$

7 $\quad$ Define $\overline{\mathcal{I}}_{kjt} := \underset{i \in \mathcal{I}}{\arg\max}{}^{N^{\text{bus}}} \left| F_{kji} \left(p_{it}^{\text{inj}}(\overline{p}) - \frac{p_t^{\text{sl}}(\overline{p})}{|\mathcal{I}|} \right) \right|$ for all $j \in \overline{\mathcal{J}}_{kt};$

8 end

9 return $\overline{\mathcal{N}} := (\overline{\mathcal{I}}, \overline{\mathcal{J}}, \overline{\mathcal{K}})$

The subset $\overline{\mathcal{I}}_{kjt} \subseteq \mathcal{I}$ consists of the N^{bus} buses with the largest absolute contributions to the post-contingency line flow at the incumbent solution, $\left| F_{kji} \left(p_{it}^{\text{inj}}(\overline{p}) - \frac{p_t^{\text{sl}}(\overline{p})}{|\mathcal{I}|} \right) \right|$. This approximation increases the sparsity in the power injection variables, improving computational performance. Algorithm 2 details the contingency screening approach. To ease notation, we denote by $\arg\max^k$ the operator that selects the k largest elements, with ties broken arbitrarily. The algorithm returns the subsets $\overline{\mathcal{N}} := (\overline{\mathcal{I}}, \overline{\mathcal{J}}, \overline{\mathcal{K}})$. Based on these subsets, we construct new cost functions:

$$\overline{C}_{kt}^{\text{ctg}}(p, q; \overline{\mathcal{N}}) := c^s d_t \sum_{j \in \overline{\mathcal{J}}_{kt}} \left[\left\| (\overline{f}_{kjt}(p; \overline{\mathcal{I}}_{kjt}), \max\{|q_{jt}^{\text{fr}}|, |q_{jt}^{\text{to}}|\}) \right\|_2 - s_j^{\max,\text{ctg}} \right]^+ \quad \forall k \in \overline{\mathcal{K}}, \ t \in \mathcal{T},$$

$$\overline{C}_t^{\text{ctg}}(p, q; \overline{\mathcal{N}}) := \frac{\sum_{k \in \overline{\mathcal{K}}_t} C_{kt}^{\text{ctg}}(\overline{p}, \overline{q})}{N^{\text{ctg}} \sum_{k \in \overline{\mathcal{K}}_t} \overline{C}_{kt}^{\text{ctg}}(\overline{p}, \overline{q}; \overline{\mathcal{N}})} \left(\sum_{k \in \overline{\mathcal{K}}_t} \overline{C}_{kt}^{\text{ctg}}(p, q; \overline{\mathcal{N}}) \right) + \max_{k \in \overline{\mathcal{K}}} \overline{C}_{kt}^{\text{ctg}}(p, q; \overline{\mathcal{N}}) \quad \forall t \in \mathcal{T}.$$

These functions are based on fewer contingencies and branches and thus are easier to construct and optimize. The coefficient in the second function averages over the N^{ctg} selected contingencies and scales the cost by the proportion of the apparent power limit violations identified by the modified functions $\overline{f}_{kjt}$ to reduce bias in the average cost after screening.

In practice, the candidate solution $(\overline{p}, \overline{q})$ comes from the solution to the final subproblems Z_t^{F}. After contingency screening around this solution, we solve a version of the final subproblems that includes the approximated contingency objectives:

$$Z_t^{\text{F}}(\overline{p}, \overline{q}, \overline{u}, \overline{\mathcal{N}}; \delta) := \max_{p_t, q_t, u_t} \quad R_t^{\text{F}}(p_t, q_t; \overline{p}) - \overline{C}_t^{\text{ctg}}(p_t, q_t; \overline{\mathcal{N}})$$

$$\text{s.t.} \quad (2) - (3), \ (9), \ (p_t, q_t, u_t^{\text{sh}}) \in \mathcal{X}_t^{\text{ac}}, \ u = \overline{u}.$$

For a fixed unit commitment solution, this allows the power injection decisions to adapt to impactful contingencies under the same restrictions as in the original final subproblems.

5.4 Objective Function Approximation and Device Subproblem Aggregation

The objective function of (SCUC-ACOPF) is partitioned into the subproblems used in Algorithm 1, such that every objective term is contained in exactly one subproblem. This precludes each subproblem from considering terms assigned to other problems, even if these terms are relevant to its decision variables. Here, we modify the subproblem objectives by reintroducing relevant terms.

In the device subproblems Z_j^{J}, we introduce an approximation of the power balance violation penalty functions C_{it}^{bal} (see Section 2.7). To facilitate approximation, we aggregate devices colocated at the same bus i. In practice, this strategy is effective because the device-level subproblems are MILPs of relatively small size, so the benefit of decomposition at the device level, compared to the bus level, is minimal. This aggregation allows for the total net power injection at a bus to be modeled in the approximation. As the spatially-decomposed subproblems Z_j^{J} do not consider any power flows, we fix the line and shunt power flow variables $(p^{\mathrm{fr}}, p^{\mathrm{to}}, p^{\mathrm{sh}}, q^{\mathrm{fr}}, q^{\mathrm{to}}, q^{\mathrm{sh}})$ to a solution from the temporally-decomposed subproblems Z_t^{T}. We then define the approximated power balance penalty functions at a fixed power flow solution $(\tilde{p}, \tilde{q})$ by

$$
\begin{aligned}
\overline{C}_{it}^{\mathrm{bal}}(p, q; \tilde{p}, \tilde{q}) := d_t \Bigg(c^{\mathrm{p}} &\left| \sum_{\substack{j \in \mathcal{J}^{\mathrm{pr}} \\ i_j = i}} p_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{cs}} \\ i_j = i}} p_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{sh}} \\ i_j = i}} \tilde{p}_{jt}^{\mathrm{sh}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i_j = i}} \tilde{p}_{jt}^{\mathrm{fr}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i'_j = i}} \tilde{p}_{jt}^{\mathrm{to}} \right| \\
&+ c^{\mathrm{q}} \left| \sum_{\substack{j \in \mathcal{J}^{\mathrm{pr}} \\ i_j = i}} q_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{cs}} \\ i_j = i}} q_{jt}^{\mathrm{tot}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{sh}} \\ i_j = i}} \tilde{q}_{jt}^{\mathrm{sh}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i_j = i}} \tilde{q}_{jt}^{\mathrm{fr}} - \sum_{\substack{j \in \mathcal{J}^{\mathrm{br}} \\ i'_j = i}} \tilde{q}_{jt}^{\mathrm{to}} \right| \Bigg) \quad \forall i \in \mathcal{I}, \ t \in \mathcal{T}.
\end{aligned}
$$

Next, the temporal problems Z_t^{T} do not consider the cost or utility of a power injection p_{jt}^{tot}. We add the corresponding functions R_{jt}^{pow} and C_{jt}^{pow} to the objective. Similarly, the device-level problems Z_j^{J} do not have costs for reserve products; we add the corresponding linear cost terms $c_{jtr}^{\mathrm{res,p}} p_{jtr}^{\mathrm{res}}$ and $c_{jtr}^{\mathrm{res,q}} q_{jtr}^{\mathrm{res}}$ to the objective.

The startup and shutdown variables u_{jt}^{su} and u_{jt}^{sd} are implicitly included in the sets $\mathcal{Y}_{jt}^{\mathrm{uc}}$ to model feasible real and reactive power injections during startup and shutdown (see Section 3). However, since no objective cost is associated with these variables in the models Z_t^{T}, they offer a path to generate power without incurring unit commitment cost. For instance, by constraint (7a), real power at time t may be produced by starting up in a future time period, which will incur an online and startup cost. These costs are not considered in the objectives R_t^{T}. In the penalty framework, producing power in such a way encourages device startups at a later time in the device-level models Z_j^{J}; this solution may be suboptimal, especially for devices with high commitment costs.

To remedy this issue, we reimpose a cost on startups and shutdowns by introducing aggregate variables $\chi_{jt}^{\mathrm{su}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}}$ and $\chi_{jt}^{\mathrm{sd}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}}$ and rewriting relevant constraints in $\mathcal{Y}_{jt}^{\mathrm{uc}}$:

$$ 0 \leq p_{jt}^{\mathrm{su}} \leq \chi_{jt}^{\mathrm{su}} \max_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}}, \ t \in \mathcal{T}, \tag{12a} $$

$$0 \le p_{jt}^{\mathrm{sd}} \le \chi_{jt}^{\mathrm{sd}} \max_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} p_{jtt'}^{\mathrm{sdpc}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (12\mathrm{b})$$

$$q_{jt}^{\mathrm{tot}} \le q_{jt}^{\max}\left(u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} + \chi_{jt}^{\mathrm{sd}}\right) + \beta_j^{\max} p_{jt}^{\mathrm{tot}} - \begin{cases} q_{jt\uparrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{pr}} \\ q_{jt\downarrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{cs}} \end{cases} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (12\mathrm{c})$$

$$q_{jt}^{\mathrm{tot}} \ge q_{jt}^{\min}\left(u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} + \chi_{jt}^{\mathrm{sd}}\right) + \beta_j^{\min} p_{jt}^{\mathrm{tot}} + \begin{cases} q_{jt\downarrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{pr}} \\ q_{jt\uparrow}^{\mathrm{res}} & \text{if } j \in \mathcal{J}^{\mathrm{cs}} \end{cases} \qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (12\mathrm{d})$$

$$u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} \le 1 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (12\mathrm{e})$$

$$u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{sd}} \le 1 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (12\mathrm{f})$$

$$(u_{jt}^{\mathrm{on}}, \chi_{jt}^{\mathrm{su}}, \chi_{jt}^{\mathrm{sd}}) \in [0,1]^3 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}. \qquad (12\mathrm{g})$$

We denote the feasible region for this formulation by the set $\overline{\mathcal{Y}}_{jt}^{\mathrm{uc}} := \{(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt}) : (2\mathrm{a}), (2\mathrm{d}) - (2\mathrm{e}), (3\mathrm{c}), (12)\}$. Proposition 7 gives that these regions are equivalent to the sets $\mathcal{Y}_{jt}^{\mathrm{uc}}$.

Proposition 7. *For all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$, the sets $\mathcal{Y}_{jt}^{\mathrm{uc}}$ and $\overline{\mathcal{Y}}_{jt}^{\mathrm{uc}}$ are equivalent; that is,*

$$\mathcal{Y}_{jt}^{\mathrm{uc}} = \{(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) : \exists \chi_{jt} \text{ s.t. } (p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt}) \in \overline{\mathcal{Y}}_{jt}^{\mathrm{uc}}\}.$$

As nonzero χ_{jt}^{su} (resp. χ_{jt}^{sd}) implies a future startup (resp. past shutdown) and therefore implies that the device must be on in some other time period, we associate online and startup (resp. shutdown) costs with these variables. The revised unit commitment cost functions are

$$\overline{C}_t^{\mathrm{uc}}(u, \chi) := \sum_{j \in \mathcal{J}^{\mathrm{sd}}} c_{jt}^{\mathrm{on}} u_{jt}^{\mathrm{on}} + \left(\max_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} c_{jt'}^{\mathrm{on}} + c_j^{\mathrm{su}}\right)\chi_{jt}^{\mathrm{su}} + \left(\max_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} c_{j,t'-1}^{\mathrm{on}} + c_j^{\mathrm{sd}}\right)\chi_{jt}^{\mathrm{sd}} \quad \forall t \in \mathcal{T}.$$

These functions overestimate the cost coefficients of χ_{jt}^{su} and χ_{jt}^{sd} in two ways: first, downtime-dependent cost adjustments $c_{jtt'}^{\mathrm{dd}}$ are ignored; and second, the largest plausible commitment cost c_{jt}^{on} for future (resp. past) online status is used.

In summary, we define updated subproblem objective functions that implement these changes:

$$\overline{R}_t^{\mathrm{T}}(p, q, u, \chi) := \sum_{j \in \mathcal{J}^{\mathrm{cs}}} R_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{pr}}} C_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{ac}}} C_{jt}^{\mathrm{ac}}(p, q) - \sum_{i \in \mathcal{I}} C_{it}^{\mathrm{bal}}(p, q) - \overline{C}_t^{\mathrm{uc}}(u, \chi) - C_t^{\mathrm{res}}(p, q),$$

$$\overline{R}_i^{\mathrm{I}}(p, q, u; \tilde{p}, \tilde{q}) := \sum_{\substack{j \in \mathcal{J}^{\mathrm{sd}} \\ i_j = i}} \left(R_j^{\mathrm{J}}(p, u) - \sum_{\substack{t \in \mathcal{T} \\ r \in \mathcal{R}}} d_t c_{jtr}^{\mathrm{res,p}} p_{jtr}^{\mathrm{res}} - \sum_{\substack{t \in \mathcal{T} \\ r \in \{\uparrow, \downarrow\}}} d_t c_{jtr}^{\mathrm{res,q}} q_{jtr}^{\mathrm{res}}\right) - \sum_{t \in \mathcal{T}} \overline{C}_{it}^{\mathrm{bal}}(p, q; \tilde{p}, \tilde{q}).$$

Contingency screening adds contingency costs to the final subproblem objectives, so these costs are removed from the intermediate subproblem objectives $\overline{R}_t^{\mathrm{T}}$. Additionally, reserve requirement and branch limit penalties are withheld from $\overline{R}_i^{\mathrm{I}}$ and energy penalties are withheld from $\overline{R}_t^{\mathrm{T}}$. Computational experiments demonstrate that these omissions do not degrade solution quality.

Algorithm 3: Tailored Penalty Alternating Direction Method

Input: Penalty schedule $\{\rho_\tau\}_{\tau=1}^{\overline{\tau}} \subset \mathbb{R}_>$, contingency thresholds $N^{\text{ctg}}, N^{\text{br}}, N^{\text{bus}}, s^{\text{thr}}$

1 for $\tau \in \{1, \dots, \overline{\tau}\}$ **do**

2   **if** $\tau = 1$ **then**

3   For $t \in \mathcal{T}$, solve $\overline{Z}_t^{\text{T}}(0, 0, 0; 0)$ to solution $(p_t^{(\tau)}, q_t^{(\tau)}, u_t^{(\tau)}, \chi_t^{(\tau)}, \mu_t^{(\tau)}, \omega_t^{(\tau)})$;

4   **else**

5   For $t \in \mathcal{T}$, solve $\overline{Z}_t^{\text{T}}(\overline{p}_t^{(\tau-1)}, \overline{q}_t^{(\tau-1)}, \overline{u}_t^{(\tau-1)}; \rho_\tau)$ to solution $(p_t^{(\tau)}, q_t^{(\tau)}, u_t^{(\tau)}, \chi_t^{(\tau)}, \mu_t^{(\tau)}, \omega_t^{(\tau)})$;

6   **end**

7   For $i \in \mathcal{I}$, solve $\overline{Z}_i^{\text{I}}(p_i^{(\tau)}, q_i^{(\tau)}, u_i^{(\tau)}; \rho_\tau)$ to solution $(\overline{p}_i^{(\tau)}, \overline{q}_i^{(\tau)}, \overline{u}_i^{(\tau)})$;

8   For $j \in \mathcal{J}^{\text{sh}}$ and $t \in \mathcal{T}$, compute $\overline{u}_{jt}^{\text{sh}(\tau)} = Z_{jt}^{\text{SH}}(\mu_{jt}^{\text{sh}(\tau)}/\omega_{i_jt}^{(\tau)})$;

9 end

10 For $j \in \mathcal{J}^{\text{sd}}$, solve $H_j(\overline{p}_j^{(\overline{\tau})}, \overline{u}_j^{(\overline{\tau})})$ to solution (δ_j^-, δ_j^+);

11 For $t \in \mathcal{T}$, solve $Z_t^{\text{F}}(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})}, \emptyset; \delta)$ to solution $(p_t^{\text{F}}, q_t^{\text{F}}, u_t^{\text{F}})$;

12 Compute $\overline{\mathcal{N}} = \text{Contingency Screening}(p^{\text{F}}, q^{\text{F}}, N^{\text{ctg}}, N^{\text{br}}, N^{\text{bus}}, s^{\text{thr}})$, using Algorithm 2;

13 For $t \in \mathcal{T}$, solve $Z_t^{\text{F}}(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})}, \overline{\mathcal{N}}; \delta)$ to solution (p_t, q_t, u_t);

14 return (p, q, u)

5.5 The Tailored pADM

Now, we combine the adaptations discussed in the previous sections and present Algorithm 3. The algorithm uses new versions of the subproblems Z_t^{T} and Z_j^{J} defined as follows, where u_i denotes the subset of variables u that contain index j with $i_j = i$:

$$
\begin{aligned}
\overline{Z}_t^{\text{T}}(\overline{p}_t, \overline{q}_t, \overline{u}_t; \rho) &:= \max_{\substack{p_t, q_t, u_t, \\ \chi_t, \mu_t, \omega_t}} \quad \overline{R}_t^{\text{T}}(p_t, q_t, u_t, \chi_t) - \rho\Gamma_2'(p_t, q_t, u_t, \overline{p}_t, \overline{q}_t, \overline{u}_t) - \rho \sum_{j \in \mathcal{J}^{\text{sh}}} (\mu_{jt}^{\text{sh}} - \overline{u}_{jt}^{\text{sh}} \omega_{i_jt})^2 \\
&\quad \text{s.t.} \quad (p_t, q_t, \mu_t, \omega_t) \in \mathcal{X}_t^{\text{soc}}, \\
&\qquad\qquad (p_{jt}, q_{jt}, u_{jt}^{\text{on}}, \chi_{jt}) \in \overline{\mathcal{Y}}_{jt}^{\text{uc}} \qquad \forall j \in \mathcal{J}^{\text{sd}}; \\
\overline{Z}_i^{\text{I}}(p_i, q_i, u_i; \rho) &:= \max_{\overline{p}_i, \overline{q}_i, \overline{u}_i} \quad \overline{R}_i^{\text{I}}(\overline{p}_i, \overline{q}_i, \overline{u}_i; p_i, q_i) - \rho\Gamma_1'(p_i, q_i, u_i, \overline{p}_i, \overline{q}_i, \overline{u}_i) \\
&\quad \text{s.t.} \quad (\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\text{uc}} \quad \forall j \in \mathcal{J}^{\text{sd}} \text{ with } i_j = i.
\end{aligned}
$$

First, the algorithm alternates between solving the SOC relaxations of the temporally decomposed subproblems and the spatially decomposed bus-level subproblems. Then, the power injection and unit commitment decisions from the final-iteration bus-level models $\overline{Z}_i^{\text{I}}$ are used to compute ramping intervals for constraints (9). These bounds and the last iterate generate the final ACOPF subproblems, which are solved without contingencies. In the last step, the contingencies are screened based on the final subproblem solution. The final subproblems are solved again, incorporating costs from the selected contingencies, to generate a candidate solution to (SCUC-ACOPF). We show the validity of this algorithm in Theorem 3.

Theorem 3. *Let (p, q, u) be a solution returned by Algorithm 3. Then, there is some $(v, \Delta, \theta, \tau, \phi)$ such that $(p, q, u, v, \Delta, \theta, \tau, \phi)$ is feasible for* (SCUC-ACOPF).

6   A Temporally-Decomposed Dual Bound

The temporally-decomposed subproblems Z_t^{T}, combined with the SOC relaxation proposed in Section 5.2, yield a set of reasonably sized convex subproblems that can be efficiently solved to global optimality. We slightly adapt these subproblems so that they generate an upper (dual) bound on the objective value of the problem (SCUC-ACOPF). Access to such a bound on the primal objective is a useful tool for computational analysis of our algorithms. Specifically, the upper bound allows for certification of the quality of a primal solution, in that feasible solutions with objective values near the upper bound cannot be significantly improved. We define the upper bound as

$$
Z^{\mathrm{UB}} \; := \; \sum_{t \in \mathcal{T}} \; \max_{\substack{p_t, q_t, u_t, \\ \mu_t, \omega_t}} \quad C_t^{\mathrm{UB}}(p_t, q_t, u_t) \tag{UB}
$$
$$
\text{s.t.} \quad (p_t, q_t, \mu_t, \omega_t) \in \mathcal{X}_t^{\mathrm{soc}}, \;\; (p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}} \quad \forall j \in \mathcal{J}^{\mathrm{sd}},
$$

where

$$
C_t^{\mathrm{UB}}(p, q, u) := \sum_{j \in \mathcal{J}^{\mathrm{cs}}} R_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{pr}}} C_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{ac}}} C_{jt}^{\mathrm{ac}}(p, q) - \sum_{i \in \mathcal{I}} C_{it}^{\mathrm{bal}}(p, q) - \sum_{j \in \mathcal{J}^{\mathrm{sd}}} c_{jt}^{\mathrm{on}} u_{jt}^{\mathrm{on}} - C_t^{\mathrm{res}}(p, q).
$$

As the feasible set relaxes that of (SCUC-ACOPF) and the objective overestimates the true objective, Z^{UB} is an upper bound on Z^*. Theorem 4 formalizes this result. This bound relaxes several aspects of the primal problem. Specifically, penalties from contingencies, energy limit violations, and device startup/shutdown costs are ignored, ramping constraints are removed, integrality constraints are relaxed, and AC power flows are replaced with their SOC relaxation.

Theorem 4. *The problem* (UB) *generates an upper bound on* (SCUC-ACOPF); *that is,* $Z^* \leq Z^{\mathrm{UB}}$.

7   Computational Results

We implement Algorithms 2 and 3 and evaluate them on the test cases provided in the fourth and final event of GOC3 (Elbert et al. 2024b). Exact details of our implementation are provided in Section 7.1. The test cases mimic realistic grid data under a diverse set of operating conditions. The dataset contains 584 cases across 8 networks which vary in scale from 73 buses to 8,316 buses. We refer to each network by the number of buses it contains. The cases are separated into three divisions, each with a different lookahead window. Cases in Division 1 (D1) are real-time market optimization problems with an 8 hour lookahead, Division 2 (D2) contains day-ahead problems with a two day lookahead, and Division 3 (D3) contains advisory problems with a one week lookahead. D1 cases span 18 time periods ($T = 18$), D2 cases 48 periods, and D3 cases 42 periods. Table 2 shows the number of cases by network size and division. Cases with the same network and division may differ in other parameters, such as generation and load profiles.

Computing Resources

Our experiments are conducted in Julia 1.9 with modeling language JuMP (Lubin et al. 2023) on the MIT SuperCloud (Reuther et al. 2018). Each case runs on a single 48-core Intel Xeon Platinum 8260@2.40 GHz

# Buses	73	617	1,576	2,000	4,224	6,049	6,717	8,316	Total
D1	24	38	24	18	24	36	36	48	248
D2	40	24	0	18	24	18	18	24	166
D3	24	24	24	3	24	24	3	44	170
Total	88	86	48	39	72	78	57	116	584

Table 2: Number of test cases by network size and division.

Network	73	617	1,576	2,000	4,224	6,049	6,717	8,316	Average
D1	10 / 0.83	10 / 1	10 / 1	9.6 / 0.72	10 / 0	4.8 / 0	2.8 / 0	3.7 / 0	7 / 0.38
D2	10 / 0.90	10 / 1	— / —	10 / 1	10 / 1	10 / 0.94	10 / 0.72	10 / 0	10 / 0.80
D3	10 / 1	10 / 1	10 / 1	10 / 1	10 / 1	10 / 1	10 / 1	10 / 0.02	10 / 0.75
Average	10 / 0.91	10 / 1	10 / 1	9.8 / 0.87	10 / 0.67	7.6 / 0.53	5.5 / 0.28	7.4 / 0.01	8.7 / 0.61

Table 3: Average number of iterations / proportion of solutions that consider screened contingencies by network size and division. No 1,576-bus D2 cases are in the dataset.

processor with 192 GB of RAM. The (mixed-integer) linear bus-level models $\overline{Z}_i^{\mathrm{I}}$ and ramping restriction problems (10) are solved with Gurobi 10.0, the temporally-decomposed SOC relaxations $\overline{Z}_t^{\mathrm{T}}$ are solved with Mosek 10.0, and the final subproblems Z_t^{F} are solved with Ipopt 3.14.10 (Wächter and Biegler 2006) using linear solver `ma57`. Feasibility is evaluated with a tolerance of 10^{-8}; all subproblems are solved with this tolerance.

7.1 Implementation Details

The decomposed models are solved in parallel on 48 threads. This allows for the temporally decomposed models ($\overline{Z}_t^{\mathrm{T}}$ or Z_t^{F}) to be solved simultaneously across all time periods, while bus-level subproblems $\overline{Z}_i^{\mathrm{I}}$ are batched into threads. The default scheme runs 10 iterations of Algorithm 3 ($\overline{\tau} = 10$) with $\rho_\tau = 10^{\tau-1}$. We adapt the number of iterations and penalty coefficients to conform with time limits, as discussed in the following paragraph.

In accordance with GOC3, a strict runtime limit is imposed: 10 minutes for D1, two hours for D2, and four hours for D3. As our algorithm runs, we estimate the maximum number of iterations that can be completed within the remaining time and distribute the remaining penalty coefficients evenly on a log scale between the current coefficient value and 10^9. The iteration count may not exceed the default value of 10. We report the average number of completed iterations by network and division in Table 3. For every case in D2 and D3, there is sufficient time to run 10 iterations. In D1, for networks with over 6,000 buses, the iteration count is often decreased, whereas smaller cases typically run all 10 iterations. If the time limit is reached while solving the final subproblems Z_t^{F}, we return the current iterate and evaluate its feasibility. Solutions with negative objective values are considered infeasible.

The contingency screening thresholds are selected by validation on a set of cases from previous competition events. The values differ between large and small networks and are set as follows:

Small Networks (73 and 617 buses): $\quad N^{\mathrm{ctg}} = 30, \quad N^{\mathrm{br}} = 20, \quad N^{\mathrm{bus}} = 50, \quad s^{\mathrm{thr}} = 10^{-3},$

Large Networks ($\geq 1{,}576$ buses): $\quad N^{\mathrm{ctg}} = 30, \quad N^{\mathrm{br}} = 20, \quad N^{\mathrm{bus}} = 20, \quad s^{\mathrm{thr}} = 10^{-2}.$

If sufficient time is available, we solve the final models Z_t^{F} with screened contingencies in step 13 of Algorithm 3; otherwise, we evaluate the solution to the contingency-free models from step 11. Table 3 gives the proportion of cases which include screened contingencies in the evaluated solution. Cases that do not include contingencies do not have sufficient remaining time for contingency screening and post-screening optimization, do not yield any valid contingencies via Algorithm 2, or fail to demonstrate a sufficient objective increase in the post-contingency solution relative to the contingency-free solution. In D1, time limits are binding for large cases, so these cases typically do not consider contingencies. In D2 and D3, most cases consider contingencies. Across all divisions, we identify impactful contingencies on only a few instances from the 8,316-bus network.

7.2 Solution Quality

In this section, we investigate the quality of the feasible solution returned by Algorithm 3. Our decomposition scheme yields dual bounds via problem (UB). We evaluate feasible solutions to (SCUC-ACOPF) against this upper bound, quantifying the optimality gap of the solution. From the primal perspective, we compare our solutions against those generated by the benchmark algorithm of Parker and Coffrin (2024, "Benchmark"). The benchmark is subject to the time limits described in Section 7.1. It solves (MI)LPs with Gurobi and nonlinear problems with Ipopt, using linear solver `ma27`. We highlight several key differences between the benchmark and our algorithm: first, we iterate between unit commitment and power flow problems to fix a unit commitment solution instead of relying on a single pass; second, we use *ex-ante* bounds to incorporate ramping constraints into the final subproblems instead of projecting onto the feasible set; third, we include reserve and security constraints in the final models.

Figure 1 compares the quality of feasible solutions to the upper bound. We compute the percent optimality gap of a feasible solution with objective z by $100 \times \frac{Z^{\mathrm{UB}}-z}{Z^{\mathrm{UB}}}$, where infeasible solutions are assigned a gap of 100%. The figure shows the empirical probability that a test case has a gap no larger than some value, with results reported on the solutions returned by the tailored pADM (Algorithm 3) and the benchmark. The figure also shows distributions for modified versions of our algorithm; these results are discussed in Section 7.4

The average optimality gap for feasible solutions returned by Algorithm 3 is 1.33%, compared to an average gap of 6.93% for benchmark solutions. Our algorithm fails to identify a feasible solution within the time limit for 19 of the 584 cases, while the benchmark gives infeasible solutions for 13 cases. This infeasibility is due to early termination by the time limit; for our algorithm, this only occurs in D1. The 80th percentile gap for our algorithm is 1.06%, compared to 11.83% for the benchmark. At the 90th percentile, the corresponding values are 1.54% and 21.88%. On average, our solution objective improves over the benchmark by 6.94%. We note that there are four cases with gaps $\geq 80\%$ for both algorithms. These cases are among those used to evaluate the impact of switching off AC lines, which is not permitted in our model. Removing these cases reduces the average gap for our algorithm to 0.70% and for the benchmark to 6.34%. Overall, these results demonstrate that our algorithm identifies a near-optimal solution in almost every case and significantly improves upon the benchmark solutions.
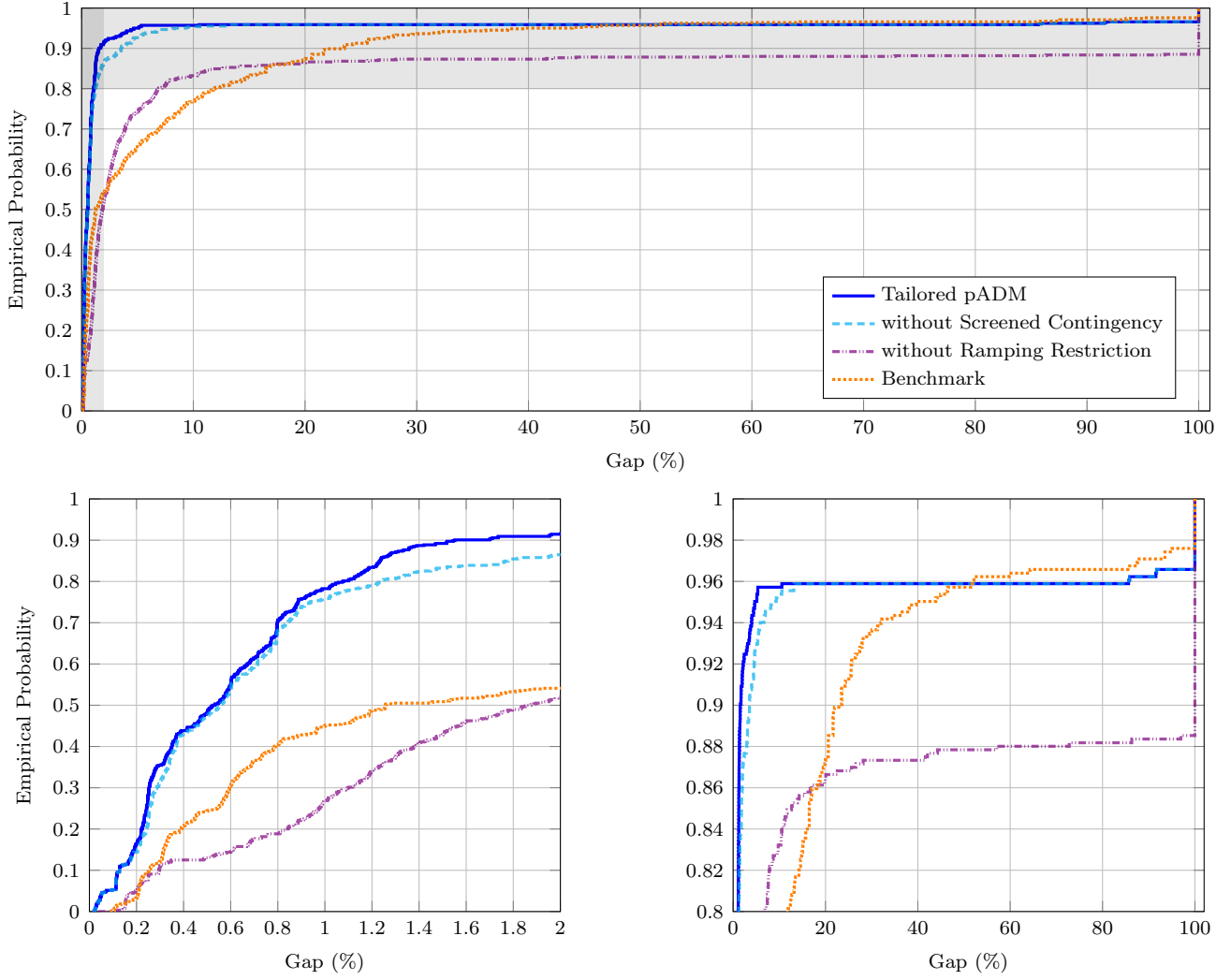
Figure 1: (top) Empirical cumulative distribution function of the optimality gap. Distributions are shown for feasible solutions obtained from the tailored pADM, tailored pADM without adding contingencies, tailored pADM without the ramping restriction, and the benchmark algorithm. (left) Distribution on the gap interval of $0\% - 2\%$, corresponds to the leftmost shaded region of (top). (right) Distribution on the upper quintile, corresponds to the uppermost shaded region of (top).
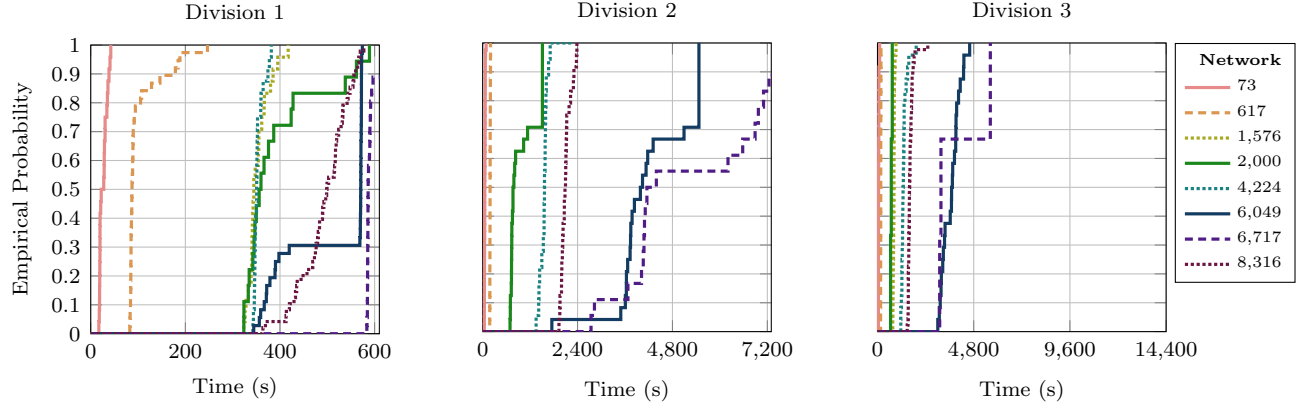
Figure 2: Empirical cumulative distribution function of solution time by division and network.
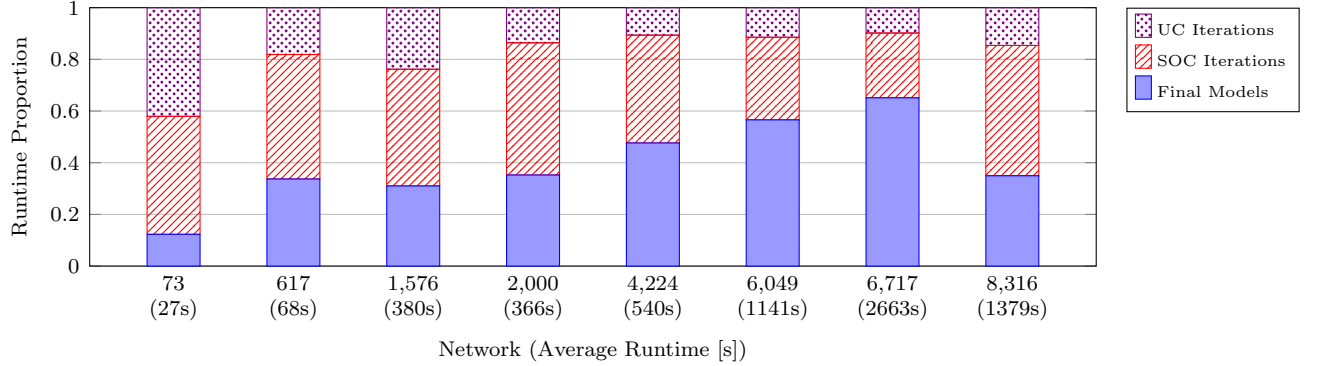
Figure 3: Proportion of runtime allocated to UC iterations, SOC iterations, and final models on cases for which the time limit is not binding and 10 iterations are run ($N = 457$). Runtime is totaled across cases with the same network and across iterations for UC and SOC models. Average runtime for each network is given in parentheses.

7.3   Algorithm Runtime

Although we enforce a maximum runtime, our algorithm often terminates before this limit. Figure 2 shows the empirical probability that a case terminates within some amount of time, by division and network. For small and medium size networks ($\leq 4{,}224$ buses), the time limits are rarely binding; that is, the cases can be solved well within the imposed time limit. For large cases in D1, especially the 6,049 and 6,717-bus networks, the 10 minute time limit is often binding and causes early termination. In D2, the time limit is only binding for about 15% of cases on the 6,717-bus network; for other networks, the time limit is not binding. In D3, the time limits are never binding; all cases are solved in under half of the allotted time.

Figure 3 shows the distribution of time spent solving each subproblem class as a proportion of total time. This runtime includes the total time for the bus-level models $\overline{Z}_i^{\mathrm{I}}$ in the iterative portion (UC Iterations), for the temporal models $\overline{Z}_t^{\mathrm{T}}$ in the iterative portion (SOC Iterations), and for the final models Z_t^{F} (Final Models). The figure shows data only for cases that run all 10 iterations and are not terminated early due to the time limit. Runtime includes neither the time for contingency screening nor for solving final models with contingencies.

Model	Time (s)	$\|\{j, t : \delta_{jt} \leq 10^{-8}\}\|$	$\sum_{j,t} \delta_{jt} \times 10^{-3}$	$\sum_{j,t} v(\delta_{jt}) \times 10^{-3}$
$\max \sum_{j,t} v(\delta_{jt})$ s.t. (10b)–(10c)	9.1	354 (0.5%)	99.87 (100%)	42.35 (100%)
LP (10a)–(10c)	0.2	7,874 (11.7%)	99.87 (100%)	36.67 (86.6%)
Heuristic (10d)–(10e)	0.1	325 (0.5%)	89.70 (89.8%)	42.28 (99.9%)
LP (10a)–(10e)	0.2	319 (0.5%)	99.80 (99.9%)	42.31 (99.9%)

Table 4: Comparison of ramping interval quality on a representative 6,049-bus case. Interval lengths are defined as $\delta_{jt} = \delta_{jt}^- + \delta_{jt}^+$, and the volume objective is given by $v(\delta_{jt}) = \log(\delta_{jt} + 1)$. Percentages are relative to the total number of elements or the global maxima of $\sum_{j,t} \delta_{jt}$ and $\sum_{j,t} v(\delta_{jt})$.

UC iterations occupy a small portion of the total runtime, especially on large networks. Large networks experience a growth in the proportion of time spent solving the final models; that is, the time spent on UC and SOC iterations does not grow as rapidly as the time spent on the final models. This suggests that efforts to reduce the runtime of our algorithm on large networks are best spent on accelerating the final subproblems. On the other hand, for small networks, a majority of the time is spent in the UC and SOC iterations; runtime reductions for these cases can be achieved by decreasing the total number of iterations.

7.4 Benefits of the Tailored Algorithm

We now evaluate the impact of the ramping restriction (Section 5.1.1), the energy limit restriction (Section 5.1.2), the SOC relaxation (Section 5.2), and contingency screening (Section 5.3).

7.4.1 Analysis of the Ramping Constraint Restriction

First, we demonstrate the quality of the ramping intervals generated by the auxiliary linear programs (10). Table 4 compares the linear programs to other models that yield valid restrictions of the ramping constraint. These models are evaluated on the number of devices and periods for which the ramping bounds yield a near-empty interval, the total length of the intervals, and a proxy for the volume generated by the intervals, computed as the sum of log interval lengths. Directly maximizing the volume (i.e., the first row of Table 4) gives the best realization of the volume metric; however, this (convex) nonlinear problem has a nontrivial solve time. Maximizing the sum of interval lengths with the linear program (10a)–(10c) speeds up computation but generates an overly sparse solution with many near-empty intervals, even when using an interior point solver without crossover. Such a solution reduces the amount of power injection flexibility in the final models. The solution at equality for the heuristic expressions (10d)–(10e) distributes slack to the ramping bounds evenly between contiguous periods. This yields intervals which are relatively non-sparse and a volume that is within 99.9% of the upper bound, with a reduced total interval length. Adding the heuristic constraints to the linear program (10a)–(10e) combines the benefits: a relatively fast method that generates near-optimal intervals by the sum and volume metrics and yields few near-empty intervals.

Next, we compare Algorithm 3 to an approach that fixes real power injection p^{tot} to the final bus-level solution $\bar{p}^{\text{tot}(\bar{\tau})}$ in the final subproblems Z_t^{F}. This is equivalent to taking $\delta = 0$ and demonstrates the impact of the flexibility provided by the ramping restriction. Fixing power injection significantly degrades the solution quality, as shown in Figure 1. With fixed power injection, the number of infeasible

| | | | without Restriction | | with Restriction | |
Division	Scenario	Z^{UB}	Objective	Penalty	Objective	Penalty
D2	027	572,227,184	498,453,819	58,480,001	564,196,225	0
D2	031	484,996,329	411,492,289	37,520,166	474,577,847	599,096
D2	032	650,963,270	588,681,039	37,680,001	640,863,341	375,629
D2	033	949,087,700	899,846,594	37,680,001	941,670,329	418,576
D2	036	652,367,566	607,441,569	37,680,001	648,867,311	360,291
D2	041	484,735,515	418,757,179	37,553,010	475,422,381	887,888
D2	042	650,664,039	587,119,183	37,680,001	636,952,776	382,855
D2	043	948,728,190	899,529,862	37,680,001	937,687,552	395,104
D3	027	1,920,855,345	1,849,812,212	20,800,000	1,906,039,779	0

Table 5: Performance on 1,576-bus cases from event 3 with and without the energy restriction included in the final subproblem objective. The SOC upper bound, feasible solution objective, and energy penalty are shown.

cases increases from 19 to 66. Relative to fixed injection, the ramping restriction increases the objective for feasible cases by an average of 3.09% with a maximum improvement of 2.6x. The total objective improvement across all cases (including infeasible cases) is 12.68%.

7.4.2 Analysis of the Energy Limit Restriction

Even with the energy limit penalty ignored in the final subproblems, none of the test cases incur any penalty at our reported solution. To demonstrate the benefit of the restriction, we consider 9 cases on the 1,576-bus network, taken from the third event of GOC3. These are the only cases from this event that have nonzero energy limit penalty in our reported solution when the penalty is ignored in the final subproblems. In these 9 cases, including the energy limit penalty restriction in the final objective reduces the total penalty $\sum_{j \in \mathcal{J}^{\mathrm{sd}}} C_j^{\mathrm{e}}(p)$ by an average of 98.99% and improves the solution objective by an average of 8.68%. Table 5 shows the solution objectives and penalty contributions for these cases.

7.4.3 Analysis of the Second-Order Cone Relaxation

The main benefit of the SOC relaxation is to reduce the computational cost of each iteration. We quantify this benefit by comparing the solve time of the final models Z_t^{F}, which contain exact nonconvex ACOPF constraints, to that of the SOC relaxations $\overline{Z}_t^{\mathrm{T}}$. Although the final models differ in other areas, including the ramping restriction and fixing of the unit commitment variables, the comparison indicates the relative scale of the runtimes. Figure 3 shows that the time spent on 10 iterations of the SOC models is on the same order as the time spent on the final supbroblems. Comparing the time to solve the first-iteration SOC models (which have penalty coefficient $\rho = 0$) against the final subproblems, we find that the SOC models solve on average 76.62% faster than the AC models. Across all cases, the total time spent on the first SOC iteration is 88.92% less than time spent on the final subproblems. This indicates that the SOC relaxation yields large reductions in solve times and that the reduction is more significant in cases with long solve times. The speed-up allows more iterations to be run and permits large cases to be solved within the time limits.

The quality of the SOC relaxation can be observed from the results in Section 7.2, which demonstrate the existence of AC-feasible solutions with small optimality gaps relative to the dual bounds. As these bounds are derived from the SOC relaxation, small gaps suggest that this relaxation does not significantly reduce accuracy relative to the AC model.

7.4.4 Analysis of the Contingency Screening Heuristic

We compare the final solution from Algorithm 3 to the solution $(p^{\mathrm{F}}, q^{\mathrm{F}}, u^{\mathrm{F}})$ generated by the contingency-free final subproblems in step 11 of the algorithm. Figure 1 plots an empirical cumulative distribution function for the optimality gap of these contingency-free solutions. Per Table 3, 39% of reported solutions for the full algorithm already are from the contingency-free solve. For the remaining cases, Figure 1 shows improvement in the gap when contingencies are considered. In cases where contingency screening is completed, the post-contingency solutions improve the total objective by an average of 0.45% and reduce the value of the contingency penalty $\sum_{t \in \mathcal{T}} C_t^{\mathrm{ctg}}(p, q)$ by an average of 24.45%. Across all cases, the post-contingency solutions reduce the total contingency penalty by 71.47%. These results demonstrate that contingency screening eliminates most high-impact contingency costs.

7.5 Differences from GO Competition Evaluation

The GOC3 test cases allow for some AC branches to be dynamically switched off. In this work, we describe our algorithm and evaluate its performance in a switching-free environment, where all lines remain on. We convert all cases to a switching-free version, and duplicate cases (included in the dataset with and without switching) are consolidated.

GOC3 evaluated software from 14 teams. For reproducibility, we compare only against the open-source benchmark solver (Parker and Coffrin 2024). Elbert et al. (2024b) provide the results of an independent evaluation of submissions from every team, including the solver described in this work. These results include additional experiments on a 6,708-bus network generated from industry data that was not released publicly and a 23,643-bus network withheld from this work due to binding memory constraints.

8 Conclusion

In this paper, we demonstrate that high-quality solutions to realistic industry-scale SCUC-ACOPF problems can be found within stringent time limits. We present a decomposition scheme that separates the SCUC-ACOPF problem into a set of device-level mixed-integer linear programs and a set of temporally-separate nonlinear programs. We apply a penalty alternating direction method to this decomposition and prove its convergence. We introduce a number of heuristics to improve the solution quality and computational cost of the algorithm, including restrictions of time-coupling constraints, a second-order cone relaxation, and a contingency screening algorithm. The algorithm tailored with these heuristics demonstrates superior performance on large-scale test cases and satisfies strict time limits, identifying near-optimal solutions relative to a dual bound.

Acknowledgement

References

Aravena I, Molzahn DK, Zhang S, Petra CG, Curtis FE, Tu S, Wächter A, Wei E, Wong E, Gholami A, Sun XA, Elbert ST, Holzer JT, Veeramany A (2023) Recent developments in security-constrained AC optimal power flow: Overview of Challenge 1 in the ARPA-E Grid Optimization Competition. *Operations Research* 71(6):1997–2014.

Baker K (2021) Solutions of DC OPF are never AC feasible. *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, 264–268.

Bichler M, Knörr J (2023) Getting prices right on electricity spot markets: On the economic impact of advanced power flow models. *Energy Economics* 126:106968.

Bienstock D, Escobar M, Gentile C, Liberti L (2022) Mathematical programming formulations for the alternating current optimal power flow problem. *Annals of Operations Research* 314(1):277–315.

Bienstock D, Verma A (2019) Strong NP-hardness of AC power flows feasibility. *Operations Research Letters* 47(6):494–501.

Birchfield AB, Xu T, Gegner KM, Shetye KS, Overbye TJ (2017) Grid structural characteristics as validation criteria for synthetic networks. *IEEE Transactions on Power Systems* 32(4):3258–3265.

Bonami P, Cornuéjols G, Lodi A, Margot F (2009) A feasibility pump for mixed integer nonlinear programs. *Mathematical Programming* 119(2):331–352.

Cain MB, O'Neill RP, Castillo A (2012) History of optimal power flow and formulations. Technical report, Federal Energy Regulatory Commission.

Carlson B, Chen Y, Hong M, Jones R, Larson K, Ma X, Nieuwesteeg P, Song H, Sperry K, Tackett M, Taylor D, Wan J, Zak E (2012) MISO unlocks billions in savings through the application of operations research for energy and ancillary services markets. *Interfaces* 42(1):58–73.

Castillo A, Laird C, Silva-Monroy CA, Watson JP, O'Neill RP (2016) The unit commitment problem with AC optimal power flow constraints. *IEEE Transactions on Power Systems* 31(6):4853–4866.

Chevalier S (2024) A parallelized, Adam-based solver for reserve and security constrained AC unit commitment. *Electric Power Systems Research* 235:110685.

Elbert S, Holzer J, Veeramany A, Coffrin C, DeMarco C, Duthu R, Greene S, Kuchar O, Lesieutre B, Li H, Mak T, Mittelmann H, O'Neill R, Overbye T, Tbaileh A, Van Hentenryck P, Wert J (2024a) ARPA-E Grid Optimization (GO) Competition Challenge 2. URL `http://dx.doi.org/10.25984/2448433`.

Elbert S, Holzer J, Veeramany A, O'Neill R, Mittelmann H, Coffrin C, Garcia M, Parker R, Elgindy T, Hale E, Palmintier B, Mak T, Overbye T, Safdarian F, DeMarco C, Greene S, Lesieutre B, Eldridge B, Oh H (2024b) ARPA-E Grid Optimization (GO) Competition Challenge 3. URL `http://dx.doi.org/10.25984/2426334`.

Fang X, Yang Z, Yu J, Wang Y (2022) AC feasibility restoration in market clearing: Problem formulation and improvement. *IEEE Transactions on Industrial Informatics* 18(11):7597–7607.

Fischetti M, Glover F, Lodi A (2005) The feasibility pump. *Mathematical Programming* 104:91–104.

Fu Y, Shahidehpour M, Li Z (2005) Security-constrained unit commitment with AC constraints. *IEEE Transactions on Power Systems* 20(2):1001–1013.

Geißler B, Morsi A, Schewe L, Schmidt M (2015) Solving power-constrained gas transportation problems using an MIP-based alternating direction method. *Computers & Chemical Engineering* 82:303–317.

Geißler B, Morsi A, Schewe L, Schmidt M (2017) Penalty alternating direction methods for mixed-integer optimization: A new view on feasibility pumps. *SIAM Journal on Optimization* 27(3):1611–1636.

Gholami A, Sun K, Zhang S, Sun XA (2023) An ADMM-based distributed optimization method for solving security-constrained alternating current optimal power flow. *Operations Research* 71(6):2045–2060.

Gorski J, Pfeuffer F, Klamroth K (2007) Biconvex sets and optimization with biconvex functions: a survey and extensions. *Mathematical Methods of Operations Research* 66(3):373–407.

Hager WW (1989) Updating the inverse of a matrix. *SIAM Review* 31(2):221–239.

Holzer J, Coffrin C, DeMarco C, Duthu R, Elbert S, Eldridge B, Elgindy T, Garcia M, Greene S, Guo N, Hale E, Lesieutre B, Mak T, McMillan C, Mittelmann H, Oh H, O'Neill R, Overbye T, Plamintier B, Paker R, Safdarian F, Tbaileh A, Van Hentenryck P, Veeramany A, Wert J (2024a) Grid Optimization Competition Challenge 3 problem formulation. URL `http://dx.doi.org/10.25984/2426334`.

Holzer JT, Chen Y, Wu Z, Pan F, Veeramany A (2024b) Fast simultaneous feasibility test for security constrained unit commitment. *IEEE Transactions on Power Systems* 39(1):1068–1078.

Holzer JT, Elbert S, Mittelmann H, O'Neill R, Oh H (2025) GO Competition Challenge 3: Problem, solvers, and solution analysis. *Energy Systems* 1–49.

ISO New England (2023) An overview of New England's wholesale electricity markets.

Jabr RA (2006) Radial distribution load flow using conic programming. *IEEE Transactions on Power Systems* 21(3):1458–1459.

Kocuk B, Dey SS, Sun XA (2016) Strong SOCP relaxations for the optimal power flow problem. *Operations Research* 64(6):1177–1196.

Lehmann K, Grastien A, Van Hentenryck P (2016) AC-feasibility on tree networks is NP-hard. *IEEE Transactions on Power Systems* 31(1):798–801.

Li H, Bornsheuer AL, Xu T, Birchfield AB, Overbye TJ (2018) Load modeling in synthetic electric grids. *2018 IEEE Texas Power and Energy Conference (TPEC)*, 1–6.

Lubin M, Dowson O, Dias Garcia J, Huchette J, Legat B, Vielma JP (2023) JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation* 15:581–589.

Molzahn DK, Dörfler F, Sandberg H, Low SH, Chakrabarti S, Baldick R, Lavaei J (2017) A survey of distributed optimization and control algorithms for electric power systems. *IEEE Transactions on Smart Grid* 8(6):2941–2962.

Molzahn DK, Hiskens IA (2019) A survey of relaxations and approximations of the power flow equations. *Foundations and Trends in Electric Energy Systems* 4(1-2):1–221.

Nocedal J, Wright SJ (2006) *Numerical Optimization* (Springer).

O'Neill RP, Dautel T, Krall E (2011) Recent ISO software enhancements and future software and modeling plans. Technical report, Federal Energy Regulatory Commission.

Parker R, Coffrin C (2024) Managing power balance and reserve feasibility in the AC unit commitment problem. *Electric Power Systems Research* 234:110670.

Reuther A, Kepner J, Byun C, Samsi S, Arcand W, Bestor D, Bergeron B, Gadepally V, Houle M, Hubbell M, Jones M, Klein A, Milechin L, Mullen J, Prout A, Rosa A, Yee C, Michaleas P (2018) Interactive supercomputing

on 40,000 cores for machine learning and data analysis. *2018 IEEE High Performance Extreme Computing Conference (HPEC)*, 1–6 (IEEE).

Stott B, Jardim J, Alsaç O (2009) DC power flow revisited. *IEEE Transactions on Power Systems* 24(3):1290–1300.

Sun K, Sun XA (2021) A two-level ADMM algorithm for AC OPF with global convergence guarantees. *IEEE Transactions on Power Systems* 36(6):5271–5281.

Tu S, Wächter A, Wei E (2021) A two-stage decomposition approach for AC optimal power flow. *IEEE Transactions on Power Systems* 36(1):303–312.

Wächter A, Biegler LT (2005) Line search filter methods for nonlinear programming: Motivation and global convergence. *SIAM Journal on Optimization* 16(1):1–31.

Wächter A, Biegler LT (2006) On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106:25–57.

Winner C, Garland J, Crozier C, Baker K (2023) Carbon emissions resulting from different power flow models for dispatch. *2023 IEEE Power & Energy Society General Meeting (PESGM)*, 1–5.

Wood AJ, Wollenberg BF, Sheblé GB (2013) *Power Generation, Operation, and Control* (John Wiley & Sons).

Zhang W, Kim Y, Kim K (2023) On solving unit commitment with alternating current optimal power flow on GPU. arXiv:2310.13145v1 [math.OC].

A   Appendix

This appendix is organized as follows. Section A.1 provides a formulation for the feasible region of device-level unit commitment decisions. Section A.2 describes the feasible region for device-level real power and reserve decisions. Section A.3 defines generalized shift factors and describes the scheme for efficiently computing them across contingencies. Section A.4 introduces compact notation for the temporally-decomposed subproblems, which is used in analysis. Section A.5 gives a pair of relevant technical lemmas. Section A.6 provides proofs of our results.

A.1   Device-Level Unit Commitment Constraints

In this section, we describe the constraints which define the sets $\mathcal{X}_j^{\mathrm{u}}$. These constraints are as follows:

$$u_{jt}^{\mathrm{on}} = 1 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}_j^{\mathrm{mr}}, \qquad (\text{A.1a})$$

$$u_{jt}^{\mathrm{on}} = 0 \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}_j^{\mathrm{mo}}, \qquad (\text{A.1b})$$

$$u_{jt}^{\mathrm{su}} \leq 1 - \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{dn,min}}} u_{jt'}^{\mathrm{sd}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (\text{A.1c})$$

$$u_{jt}^{\mathrm{sd}} \leq 1 - \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{up,min}}} u_{jt'}^{\mathrm{su}} \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}, \qquad (\text{A.1d})$$

$$\sum_{t \in W} u_{jt}^{\mathrm{su}} \leq u_j^{\mathrm{su,max}}(W) \qquad\qquad \forall j \in \mathcal{J}^{\mathrm{sd}},\ W \in \mathcal{W}_j^{\mathrm{su,max}}. \qquad (\text{A.1e})$$

Constraints (A.1a)–(A.1b) enforce must-run and must-outage requirements, where $\mathcal{T}_j^{\mathrm{mr}}$ (resp. $\mathcal{T}_j^{\mathrm{mo}}$) gives the set of time periods during which device j must be on (resp. off). Constraints (A.1c)–(A.1d) enforce minimum downtime and uptime requirements, and (A.1e) limits the number of device startups over some time interval. The sets $\mathcal{T}_{jt}^{\mathrm{dn,min}}$ (resp. $\mathcal{T}_{jt}^{\mathrm{up,min}}$) contain prior time periods $t' < t$ within the minimum downtime (resp. uptime) ranges relative to period t for device j. A set $W \in \mathcal{W}_j^{\mathrm{su,max}}$ is a set of time periods subject to a startup limit ($W \subseteq \mathcal{T}$), and $u_j^{\mathrm{su,max}}(W)$ gives the corresponding limit. From these constraints, we define $\mathcal{X}_j^{\mathrm{u}} := \{\{u_{jt}^{\mathrm{on}}, u_{jt}^{\mathrm{su}}, u_{jt}^{\mathrm{sd}}\}_{t \in \mathcal{T}} : (\text{A.1})\}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$, relying on the separability of the constraints over devices.

A.2   Device-Level Real Power and Reserve Constraints

Here, we detail the device-level real power constraints of the sets $\mathcal{X}_{jt}^{\mathrm{p}}$. We consider three reserve categories: up-reserve products for online devices, down-reserve products for online devices, and up-reserve products for offline devices. We let $\mathcal{R}_{\mathrm{on\uparrow}}$, $\mathcal{R}_{\mathrm{on\downarrow}}$, and $\mathcal{R}_{\mathrm{off\uparrow}}$ be the subsets of $\mathcal{R}$ containing products assigned to the three categories. These reserve subsets are totally ordered on the quality of the product, e.g., Regulation > Synchronized > Ramping for elements of $\mathcal{R}_{\mathrm{on\uparrow}}$. For a complete description of modeled reserve products, see Holzer et al. (2024a). Real power injection and reserve products are constrained by the following logic:

$$\sum_{\substack{r' \in \mathcal{R}_a \\ r' \geq r}} p_{jtr'}^{\mathrm{res}} \leq p_{jr}^{\mathrm{res,max}} u_{jt}^{\mathrm{on}} \qquad\qquad \forall a \in \{\mathrm{on}\uparrow, \mathrm{on}\downarrow\},\ j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T},\ r \in \mathcal{R}_a, \qquad (\text{A.2a})$$

$$\sum_{\substack{r' \in \mathcal{R}_{\text{off}\uparrow} \\ r' \geq r}} p_{jtr'}^{\text{res}} \leq p_{jr}^{\text{res,max}}(1 - u_{jt}^{\text{on}}) \qquad\qquad \forall j \in \mathcal{J}^{\text{sd}}, \ t \in \mathcal{T}, \ r \in \mathcal{R}_{\text{off}\uparrow}, \qquad \text{(A.2b)}$$

$$p_{jt}^{\text{on}} \leq p_{jt}^{\max} u_{jt}^{\text{on}} - \begin{cases} \displaystyle\sum_{r \in \mathcal{R}_{\text{on}\uparrow}} p_{jtr}^{\text{res}} & \text{if } j \in \mathcal{J}^{\text{pr}} \\ \displaystyle\sum_{r \in \mathcal{R}_{\text{on}\downarrow}} p_{jtr}^{\text{res}} & \text{if } j \in \mathcal{J}^{\text{cs}} \end{cases} \qquad \forall j \in \mathcal{J}^{\text{sd}}, \ t \in \mathcal{T}, \qquad \text{(A.2c)}$$

$$p_{jt}^{\text{on}} \geq p_{jt}^{\min} u_{jt}^{\text{on}} + \begin{cases} \displaystyle\sum_{r \in \mathcal{R}_{\text{on}\downarrow}} p_{jtr}^{\text{res}} & \text{if } j \in \mathcal{J}^{\text{pr}} \\ \displaystyle\sum_{r \in \mathcal{R}_{\text{on}\uparrow}} p_{jtr}^{\text{res}} & \text{if } j \in \mathcal{J}^{\text{cs}} \end{cases} \qquad \forall j \in \mathcal{J}^{\text{sd}}, \ t \in \mathcal{T}, \qquad \text{(A.2d)}$$

$$p_{jt}^{\text{tot}} \leq p_{jt}^{\max}(1 - u_{jt}^{\text{on}}) - \sum_{r \in \mathcal{R}_{\text{off}\uparrow}} p_{jtr}^{\text{res}} \qquad\qquad \forall j \in \mathcal{J}^{\text{sd}}, \ t \in \mathcal{T}. \qquad \text{(A.2e)}$$

Constraints (A.2a)–(A.2b) enforce absolute limits on reserve products, where $p_{jr}^{\text{res,max}}$ gives the maximum quantity of reserves at least as good as product r. Whether a device can provide a reserve product depends on if the device has the appropriate commitment status for the reserve category. Constraints (A.2c)–(A.2e) enforce reserve limits relative to dispatched real power. The parameters $p_{jt}^{\max}$ and $p_{jt}^{\min}$ give maximum and minimum total real power. These limits count reserves that cut into headroom to the minimum or maximum; that is, reserve products that could increase real power injection count against maximum power, and those that decrease injection against the minimum. Whether a product increases or decreases power injection depends on whether the device produces or consumes power. We define $\mathcal{X}_{jt}^{\text{p}} := \{(u_{jt}^{\text{on}}, p_{jt}^{\text{tot}}, p_{jt}^{\text{on}}, \{p_{jtr}^{\text{res}}\}_{r \in \mathcal{R}}) \ : \ (\text{A.2})\}$ for all $j \in \mathcal{J}^{\text{sd}}$ and $t \in \mathcal{T}$. This construction relies on the separability of the constraints over devices and time.

A.3 Generalized Shift Factors and Fast Contingency Evaluation

This section defines the generalized shift factors F_{kji} for any connected subnetwork $\mathcal{J} \subseteq \mathcal{J}^{\text{ac}}$.

Let $M \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{J}^{\text{ac}}|}$ be the bus-branch incidence matrix, where

$$M_{ij} = \begin{cases} 1 & \text{if } i_j = i \\ -1 & \text{if } i'_j = i \\ 0 & \text{otherwise.} \end{cases}$$

Let $B(\mathcal{J}) \in \mathbb{R}^{|\mathcal{J}^{\text{ac}}| \times |\mathcal{J}^{\text{ac}}|}$ be the diagonal branch susceptance matrix, where

$$B(\mathcal{J})_{jj} = \begin{cases} -b_j & \text{if } j \in \mathcal{J} \\ 0 & \text{otherwise.} \end{cases}$$

Under the DC approximation, the real power flows $p^{\text{fl}} \in \mathbb{R}^{|\mathcal{J}^{\text{ac}}|}$ over AC branches are computed as a linear function of the bus phase angles $\theta \in \mathbb{R}^{|\mathcal{I}|}$:

$$p^{\text{fl}} = B(\mathcal{J})M^{\mathsf{T}}\theta.$$

For simplicity, we ignore the dependence on phase shifts ϕ. Then, given a vector of net nodal real power injections $p^{\text{inj}} \in \mathbb{R}^{|\mathcal{I}|}$, nodal power balance requires

$$p^{\text{inj}} = Mp^{\text{fl}} = MB(\mathcal{J})M^{\mathsf{T}}\theta.$$

The matrix $MB(\mathcal{J})M^{\mathsf{T}}$ is the bus admittance matrix. For some reference bus $i_0 \in \mathcal{I}$, we construct

$$H(\mathcal{J})_{ii'} = \begin{cases} (MB(\mathcal{J})M^{\mathsf{T}})_{ii'} & \text{if } i_0 \notin \{i, i'\} \\ 1 & \text{if } i = i' = i_0 \\ 0 & \text{otherwise.} \end{cases}$$

If the branches $\mathcal{J}$ form a connected network, $H(\mathcal{J})$ is invertible. Feasible bus angles can then be computed by inverting $H(\mathcal{J})$ and setting $\theta_{i_0} = 0$; that is,

$$\theta = SH(\mathcal{J})^{-1}p^{\text{inj}},$$

where $S = I - \mathbf{e}_{i_0}\mathbf{e}_{i_0}^{\mathsf{T}}$ sets the i_0-th entry to 0. Here, I denotes the identity matrix and $\mathbf{e}_{i_0}$ is the i_0-th standard basis vector. Now, the real power flows are given by

$$p^{\text{fl}} = B(\mathcal{J})M^{\mathsf{T}}SH(\mathcal{J})^{-1}p^{\text{inj}}.$$

The coefficient $(B(\mathcal{J})M^{\mathsf{T}}SH(\mathcal{J})^{-1})_{ji}$ gives the GSF from bus i to line j under network $\mathcal{J}$. Therefore, the GSF from bus i to line j under contingency k is defined by

$$F_{kji} := (B(\mathcal{J}^{\text{ac}} \setminus \{j_k\})M^{\mathsf{T}}SH(\mathcal{J}^{\text{ac}} \setminus \{j_k\})^{-1})_{ji}.$$

Computing factors F_{kji} requires the inverse of matrix $H(\mathcal{J}^{\text{ac}} \setminus \{j_k\})$, which may be difficult to compute for all contingencies k. Instead, we express this matrix as a rank-1 update to the matrix $H(\mathcal{J}^{\text{ac}})$. If $i_0 \notin \{i_{j_k}, i'_{j_k}\}$, then

$$H(\mathcal{J}^{\text{ac}} \setminus \{j_k\}) = H(\mathcal{J}^{\text{ac}}) + b_{j_k}(\mathbf{e}_{i_{j_k}} - \mathbf{e}_{i'_{j_k}})(\mathbf{e}_{i_{j_k}} - \mathbf{e}_{i'_{j_k}})^{\mathsf{T}}.$$

If $i_0 \in \{i_{j_k}, i'_{j_k}\}$, a similar rank-1 update can be derived. We leverage this structure by first evaluating $H(\mathcal{J}^{\text{ac}})^{-1}$, then computing $H(\mathcal{J}^{\text{ac}} \setminus \{j_k\})^{-1}$ for relevant k by the Sherman-Morrison formula (Hager 1989). This method for the efficient computation of GSFs across multiple contingencies is similar to that of Holzer et al. (2024b).

A.4 Stylized Notation for Temporally-Decomposed Subproblems

To aid in the analysis of Algorithm 1, we introduce a generic form of the models $Z_t^{\mathrm{T}}(\bar{p}_t, \bar{q}_t, \bar{u}_t; \rho)$:

$$
\begin{aligned}
\max_{x,y,z} \quad & F(x, y, z) - \rho\Gamma(x, \bar{x}) \\
\text{s.t.} \quad & x \in \mathcal{X}, \\
& y = A_z(x), \\
& z \in \mathcal{Z}.
\end{aligned}
\tag{A.3}
$$

The objects that define this model have the following properties:

1. The feasible set of (A.3) is bounded with diameter D;

2. $\mathcal{X}$ is polyhedral;

3. A_z is a linear function for all $z \in \mathcal{Z}$;

4. $F(x, y, z)$ is Lipschitz continuous with parameter L over the feasible set of (A.3), so that all $\nabla F \in \partial F(x, A_z(x), z)$ with $x \in \mathcal{X}$ and $z \in \mathcal{Z}$ satisfy $\|\nabla F\|_2 \leq L$;

5. Penalty function $\Gamma(x, \bar{x}) = \|E(x - \bar{x})\|_2^2$ where $E^{\mathsf{T}}E = E$; and

6. If $(\bar{p}_j, \bar{q}_j, \bar{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $\bar{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$, then $\bar{x} \in \mathcal{X}$.

We demonstrate that, for any $t \in \mathcal{T}$ and solution $(\bar{p}, \bar{q}, \bar{u})$, the model $Z_t^{\mathrm{T}}(\bar{p}_t, \bar{q}_t, \bar{u}_t; \rho)$ is of the form (A.3). Let

$$
\mathcal{Z} = \{(p_t^{\mathrm{fr}}, p_t^{\mathrm{to}}, q_t^{\mathrm{fr}}, q_t^{\mathrm{to}}, v_t) \; : \; \exists(\Delta_t, \theta_t, \tau_t, \phi_t) \text{ s.t. } (5a) - (5f), (5i) - (5k)\}
$$

and

$$
\mathcal{X} = \left\{ (p_t, q_t, u_t) \backslash (p_t^{\mathrm{fr}}, p_t^{\mathrm{to}}, p_t^{\mathrm{sh}}, q_t^{\mathrm{fr}}, q_t^{\mathrm{to}}, q_t^{\mathrm{sh}}) \; : \;
\begin{array}{ll}
(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}} & \forall j \in \mathcal{J}^{\mathrm{sd}}, \\
u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}} & \forall j \in \mathcal{J}^{\mathrm{sh}}
\end{array}
\right\},
$$

where we abuse the notation $(p_t, q_t, u_t) \backslash (p_t^{\mathrm{fr}}, p_t^{\mathrm{to}}, p_t^{\mathrm{sh}}, q_t^{\mathrm{fr}}, q_t^{\mathrm{to}}, q_t^{\mathrm{sh}})$ to represent the projection of the solution (p_t, q_t, u_t) onto all coordinates except the branch and shunt power flow variables. The variables x lie in $\mathcal{X}$ and the variables z lie in $\mathcal{Z}$. Let

$$
A_z(x) = \begin{bmatrix} g_j^{\mathrm{sh}} v_{i_j t}^2 u_{jt}^{\mathrm{sh}} \\ -b_j^{\mathrm{sh}} v_{i_j t}^2 u_{jt}^{\mathrm{sh}} \end{bmatrix} \qquad \forall j \in \mathcal{J}^{\mathrm{sh}} = \begin{bmatrix} p_t^{\mathrm{sh}} \\ q_t^{\mathrm{sh}} \end{bmatrix}.
$$

For fixed z, the function A_z is linear in x as the variables $v_{i_j t}$ are contained in z and the variables u_{jt}^{sh} are contained in x. Any $x \in \mathcal{X}$ contains all the elements of (p_t, q_t, u_t) except for branch and shunt power flows, while $y = A_z(x)$ contains the shunt power flow variables $(p_t^{\mathrm{sh}}, q_t^{\mathrm{sh}})$. Any (x, y, z) with $y = A_z(x)$ satisfies constraints (5g)–(5h).

Under these constructions, the feasible set of (A.3) corresponds to the feasible set of Z_t^{T}, which is bounded under Assumption 1. We call the diameter of this set D. Under Assumption 1, $\mathcal{X}$ is polyhedral because the sets $\mathcal{Y}_{jt}^{\mathrm{uc}}$ and $\mathcal{Y}_{jt}^{\mathrm{sh}}$ are polyhedral.

With the sets $(\mathcal{X}, \mathcal{Z})$ and functions A_z constructed in this way, any feasible (x, y, z) contains exactly one copy of each variable in (p_t, q_t, u_t). We define $F(x, y, z) = R_t^{\mathrm{T}}(p_t, q_t)$, which is Lipschitz continuous over any bounded domain. To ensure that all subgradients on the domain of (A.3) are properly bounded, we define the constant L to be the Lipschitz constant of F over some bounded set that strictly contains this domain.

We construct

$$\overline{x} = (\overline{p}_t, \overline{q}_t, \overline{u}_t) \setminus (\overline{p}_t^{\mathrm{fr}}, \overline{p}_t^{\mathrm{to}}, \overline{p}_t^{\mathrm{sh}}, \overline{q}_t^{\mathrm{fr}}, \overline{q}_t^{\mathrm{to}}, \overline{q}_t^{\mathrm{sh}}),$$

and accordingly define

$$\Gamma(x, \overline{x}) = \|E(x - \overline{x})\|_2^2 = \Gamma_2(p_t, q_t, u_t, \overline{p}_t, \overline{q}_t, \overline{u}_t),$$

where E is a diagonal matrix with an entry of 1 if the corresponding element of x is a copied variable (i.e., penalized in Γ_2) and an entry of 0 otherwise. This matrix clearly satisfies $E^{\mathrm{T}} E = E$. In the generic form, $x \in \mathcal{X}$ contains every copied variable, so we can represent Γ_2 as a function of Ex and $E\overline{x}$.

Finally, suppose that $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. By Lemma 1, it directly follows that $\overline{x} \in \mathcal{X}$. This demonstrates that $Z_t^{\mathrm{T}}(\overline{p}_t, \overline{q}_t, \overline{u}_t; \rho)$ satisfies the structure of problem (A.3).

A.5 Technical Lemmas

This section presents two technical lemmas that are used to prove results from the body of this paper. Lemma A.1 bounds the magnitude of the penalty term for the device-level block subproblems by the magnitude of the penalty term from the previous iteration solution to the temporally-decomposed block subproblems. Lemma A.2 establishes a fundamental property on translating the limit point of a convergent sequence contained in a polyhedron.

Lemma A.1. *Let* $(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$ *be a sequence of iterates generated by Algorithm 1, where the subproblems* Z_j^{J} *are solved to global optimality. Let R be the same constant as in the statement of Theorem 1. Then,*

$$\Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)}) \leq \frac{R}{\rho_\tau} + \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)}) \quad \forall \tau \in [\![\tau]\!] \setminus \{1\}.$$

Proof. Take any $j \in \mathcal{J}^{\mathrm{sd}}$ and iteration $\tau > 1$. Then, $(\overline{p}_j^{(\tau)}, \overline{q}_j^{(\tau)}, \overline{u}_j^{(\tau)})$ is globally optimal for $Z_j^{\mathrm{J}}(p_j^{(\tau)}, q_j^{(\tau)}, u_j^{(\tau)}; \rho_\tau)$. Under Assumption 1, the variables which appear in the objective function R_j^{J} have bounded domain, so $\underline{R}_j \leq R_j^{\mathrm{J}}(p, u) \leq \overline{R}_j$ for some bounds $(\underline{R}_j, \overline{R}_j)$ and the constants $R_j = \overline{R}_j - \underline{R}_j$ used to define R are well defined.

As the feasible region of Z_j^{J} does not change between iterations, $(\overline{p}_j^{(\tau-1)}, \overline{q}_j^{(\tau-1)}, \overline{u}_j^{(\tau-1)})$ remains feasible. Applying global optimality for the maximization, we have for all $j \in \mathcal{J}^{\mathrm{sd}}$ that

$$R_j^{\mathrm{J}}(\overline{p}_j^{(\tau-1)}, \overline{u}_j^{(\tau-1)}) - \rho_\tau \Gamma_1'(p_j^{(\tau)}, q_j^{(\tau)}, u_j^{(\tau)}, \overline{p}_j^{(\tau-1)}, \overline{q}_j^{(\tau-1)}, \overline{u}_j^{(\tau-1)})$$
$$\leq R_j^{\mathrm{J}}(\overline{p}_j^{(\tau)}, \overline{u}_j^{(\tau)}) - \rho_\tau \Gamma_1'(p_j^{(\tau)}, q_j^{(\tau)}, u_j^{(\tau)}, \overline{p}_j^{(\tau)}, \overline{q}_j^{(\tau)}, \overline{u}_j^{(\tau)}).$$

Additionally, for $j \in \mathcal{J}^{\text{sh}}$, we have that

$$\rho_\tau \left\| u_j^{\text{sh}(\tau)} - \overline{u}_j^{\text{sh}(\tau-1)} \right\|_1 \geq \rho_\tau \left\| u_j^{\text{sh}(\tau)} - \overline{u}_j^{\text{sh}(\tau)} \right\|_1$$

by definition of subproblems Z_j^{SH}. These properties yield the relation

$$\rho_\tau \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$$

$$= \rho_\tau \left(\sum_{j \in \mathcal{J}^{\text{sd}}} \Gamma_1'(p_j^{(\tau)}, q_j^{(\tau)}, u_j^{(\tau)}, \overline{p}_j^{(\tau)}, \overline{q}_j^{(\tau)}, \overline{u}_j^{(\tau)}) + \sum_{j \in \mathcal{J}^{\text{sh}}} \left\| u_j^{\text{sh}(\tau)} - \overline{u}_j^{\text{sh}(\tau)} \right\|_1 \right)$$

$$\leq \sum_{j \in \mathcal{J}^{\text{sd}}} \left(R_j^{\text{J}}(\overline{p}_j^{(\tau)}, \overline{u}_j^{(\tau)}) - R_j^{\text{J}}(\overline{p}_j^{(\tau-1)}, \overline{u}_j^{(\tau-1)}) + \rho_\tau \Gamma_1'(p_j^{(\tau)}, q_j^{(\tau)}, u_j^{(\tau)}, \overline{p}_j^{(\tau-1)}, \overline{q}_j^{(\tau-1)}, \overline{u}_j^{(\tau-1)}) \right)$$

$$+ \rho_\tau \sum_{j \in \mathcal{J}^{\text{sh}}} \left\| u_j^{\text{sh}(\tau)} - \overline{u}_j^{\text{sh}(\tau-1)} \right\|_1$$

$$\leq \sum_{j \in \mathcal{J}^{\text{sd}}} R_j + \rho_\tau \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)})$$

where the first line follows from the definitions of Γ_1 and Γ_1', the second from global optimality for Z_j^{J} and Z_j^{SH}, and the third from the objective function bounds and again the definitions of Γ_1 and Γ_1'. As $R = \sum_{j \in \mathcal{J}^{\text{sd}}} R_j$, this proves the result. $\qquad\square$

Lemma A.2. *Let $\mathcal{X}$ be a polyhedron with $\{x^*, \tilde{x}\} \in \mathcal{X}$. Consider a sequence $\{x_i\}_{i=1}^\infty \subset \mathcal{X}$ such that $\lim_{i\to\infty} x_i = x^*$. Then, there exists some $\{\lambda_i\}_{i=1}^\infty \subset \mathbb{R}_\geq$ such that the sequence $\tilde{x}_i := x_i + \lambda_i(\tilde{x} - x^*)$ satisfies $\{\tilde{x}_i\}_{i=1}^\infty \subset \mathcal{X}$ and $\lim_{i\to\infty} \tilde{x}_i = \tilde{x}$.*

Proof. Define (A, b) with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ such that $\mathcal{X} = \{x : Ax \leq b\}$. Let $M := \{j \in \llbracket m \rrbracket : a_j^\mathsf{T}(\tilde{x} - x^*) > 0\}$, where a_j is the j-th row of matrix A. We define a sequence of multipliers λ_i by

$$\lambda_i := \begin{cases} \min \left\{ \min_{j \in M} \frac{b_j - a_j^\mathsf{T} x_i}{a_j^\mathsf{T}(\tilde{x} - x^*)}, 1 \right\} & \text{if } M \neq \emptyset \\ 1 & \text{if } M = \emptyset. \end{cases}$$

These multipliers yield the sequence of iterates $\tilde{x}_i := x_i + \lambda_i(\tilde{x} - x^*)$. Note that $a_j^\mathsf{T} x_i \leq b_j$ as $x_i \in \mathcal{X}$ and $a_j^\mathsf{T}(\tilde{x} - x^*) > 0$ for $j \in M$, so $\lambda_i \geq 0$.

First, we establish that the sequence $\tilde{x}_i$ lies in $\mathcal{X}$. Consider some index i. For $j \in \llbracket m \rrbracket \setminus M$,

$$a_j^\mathsf{T} \tilde{x}_i = a_j^\mathsf{T} x_i + \lambda_i a_j^\mathsf{T}(\tilde{x} - x^*) \leq a_j^\mathsf{T} x_i \leq b,$$

as $\lambda_i \geq 0$ and $a_j^\mathsf{T}(\tilde{x} - x^*) \leq 0$. For $j \in M$, we have that $\lambda_i \leq \frac{b_j - a_j^\mathsf{T} x_i}{a_j^\mathsf{T}(\tilde{x} - x^*)}$ and $a_j^\mathsf{T}(\tilde{x} - x^*) > 0$, so

$$a_j^\mathsf{T} \tilde{x}_i = a_j^\mathsf{T} x_i + \lambda_i a_j^\mathsf{T}(\tilde{x} - x^*)$$

$$\leq a_j^\mathsf{T} x_i + \left(\frac{b_j - a_j^\mathsf{T} x_i}{a_j^\mathsf{T}(\tilde{x} - x^*)} \right) a_j^\mathsf{T}(\tilde{x} - x^*)$$

$$= b_j.$$

40

These results demonstrate that $A\tilde{x}_i \leq b$ for all i, and thus $\{\tilde{x}_i\}_{i=1}^{\infty} \subset \mathcal{X}$.

Next, we consider the limit of the iterates λ_i. If $M = \emptyset$, clearly $\lim_{i\to\infty} \lambda_i = 1$. Otherwise,

$$
\begin{aligned}
\lim_{i\to\infty} \lambda_i &= \lim_{i\to\infty} \min \left\{ \min_{j\in M} \frac{b_j - a_j^\mathsf{T} x_i}{a_j^\mathsf{T}(\tilde{x} - x^*)}, 1 \right\} \\
&= \min \left\{ \min_{j\in M} \lim_{i\to\infty} \frac{b_j - a_j^\mathsf{T} x_i}{a_j^\mathsf{T}(\tilde{x} - x^*)}, 1 \right\} \\
&= \min \left\{ \min_{j\in M} \frac{b_j - a_j^\mathsf{T} x^*}{a_j^\mathsf{T}(\tilde{x} - x^*)}, 1 \right\} \\
&\geq \min \left\{ \min_{j\in M} \frac{b_j - a_j^\mathsf{T} x^*}{b_j - a_j^\mathsf{T} x^*}, 1 \right\} \\
&= 1,
\end{aligned}
$$

where the second line follows from continuity of the min function, the third line from the convergence of $\{x_i\}_{i=1}^{\infty}$ to x^*, and the fourth from the feasibility of $\tilde{x}$ and x^* for $\mathcal{X}$. As $\lambda_i \leq 1$, we have shown that $\lim_{i\to\infty} \lambda_i = 1$. Therefore, $\lim_{i\to\infty} \tilde{x}_i = x^* + (\tilde{x} - x^*) = \tilde{x}$. $\qquad\square$

A.6 Proofs of Results

Proof of Lemma 1

Proof. We first consider property 1. Consider some $j \in \mathcal{J}^{\mathrm{sd}}$ and $(p_j, q_j, u_j) \in \mathcal{X}_j^{\mathrm{uc}}$. By Assumption 1, such a u_j satisfies (6). Further, (u_j, p_j, q_j) satisfy (2)–(3) by definition of $\mathcal{X}_j^{\mathrm{uc}}$. As $p_{jtt'}^{\mathrm{supc}} \geq 0$, we have by (2b) that

$$
p_{jt}^{\mathrm{su}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} u_{jt'}^{\mathrm{su}} \geq 0,
$$

and thus (7a) is satisfied. Similarly, we see that (7b) holds. Therefore, with certificate $(u_j^{\mathrm{su}}, u_j^{\mathrm{sd}})$, it holds that $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$ for all $t \in \mathcal{T}$, proving property 1. Next, consider property 2. Clearly, as any $u_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ satisfies (5l), $u_{jt}^{\mathrm{sh}} \in [u_j^{\mathrm{sh,min}}, u_j^{\mathrm{sh,max}}]$ and thus $u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}}$. $\qquad\square$

Proof of Proposition 1

Proof. We will prove that feasible solutions for each model can be mapped to feasible solutions for the other model with the same objective value and the same values for the copied variables.

First, consider a feasible solution $(p, q, u, v, \Delta, \theta, \tau, \phi)$ for (SCUC-ACOPF). Define $(\bar{p}, \bar{q}, \bar{u}) = (p, q, u)$. By this definition, the last constraint of (COPY) is satisfied for solution $(p, q, u, \bar{p}, \bar{q}, \bar{u})$ and the copied variables have the same values in both solutions. As $(p, q, u, v, \Delta, \theta, \tau, \phi)$ satisfies (5), it holds that $(p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$. Additionally, for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$, it holds that $u_{jt}^{\mathrm{sh}} = \bar{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}} \subseteq \mathcal{Y}_{jt}^{\mathrm{sh}}$ by Lemma 1. Similarly, as (p, q, u) satisfy (1)–(4), for all $j \in \mathcal{J}^{\mathrm{sd}}$ it holds that $(p_j, q_j, u_j) = (\bar{p}_j, \bar{q}_j, \bar{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ and further, by Lemma 1, $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$ for all $t \in \mathcal{T}$. This demonstrates that $(p, q, u, \bar{p}, \bar{q}, \bar{u})$ is feasible for (COPY). Considering the objective, $\sum_{t\in\mathcal{T}} R_t^{\mathrm{T}}(p_t, q_t) + \sum_{j\in\mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(p_j, u_j) =$

$\sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p_t, q_t) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j, \overline{u}_j)$, and the solutions have the same objective value in their respective models.

Next, consider a feasible solution $(p, q, u, \overline{p}, \overline{q}, \overline{u})$ for (COPY). First, $\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$ and thus satisfies (5l). By the copy constraints, $u^{\mathrm{sh}} = \overline{u}^{\mathrm{sh}}$. Then, $(p_t, q_t, \overline{u}_t^{\mathrm{sh}}) = (p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$ and we can construct a corresponding solution $(v_t, \Delta_t, \theta_t, \tau_t, \phi_t)$ such that constraints (5a)–(5k) are satisfied by the solution $(p, q, \overline{u}, v, \Delta, \theta, \tau, \phi)$. In the remainder of the proof, we refer to this solution as the *candidate solution*.

The candidate solution is constructed so that the copied variables have the same values as in $(p, q, u, \overline{p}, \overline{q}, \overline{u})$. As $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$, these variables satisfy (1)–(4). As constraints (1) only contain variables $\overline{u}$, the candidate solution satisfies these constraints. By the copy constraints, $p^{\mathrm{su}} = \overline{p}^{\mathrm{su}}$ and $p^{\mathrm{sd}} = \overline{p}^{\mathrm{sd}}$, so $(p, \overline{u})$ satisfies constraints (2b)–(2c). For the remaining constraints in (2), specifically (2a), (2d), and (2e), note that $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$, so there are some $(\tilde{u}_j^{\mathrm{su}}, \tilde{u}_j^{\mathrm{sd}})$ that allow (p, u^{on}) to be feasible for (2). These remaining constraints depend on u only through u^{on} and do not contain $\tilde{u}^{\mathrm{su}}$ or $\tilde{u}^{\mathrm{sd}}$. As $u^{\mathrm{on}} = \overline{u}^{\mathrm{on}}$ by the copy constraints, $(p, \overline{u})$ also satisfies (2). As the copy constraints enforce $(p^{\mathrm{tot}}, q^{\mathrm{tot}}, q^{\mathrm{res}}) = (\overline{p}^{\mathrm{tot}}, \overline{q}^{\mathrm{tot}}, \overline{q}^{\mathrm{res}})$, the solution $(p, q, \overline{u})$ satisfies constraints (3) because constraints (3) depend on (p, q) only through $(p^{\mathrm{tot}}, q^{\mathrm{tot}}, q^{\mathrm{res}})$. Similarly, again as $p^{\mathrm{tot}} = \overline{p}^{\mathrm{tot}}$, the solution $(p, q, \overline{u})$ satisfies constraints (4).

Therefore, the candidate solution $(p, q, \overline{u}, v, \Delta, \theta, \tau, \phi)$ is feasible for (SCUC-ACOPF). Comparing the objective, the terms in R_j^{J} that contain variable p are R_{jt}^{pow}, C_{jt}^{pow}, and C_j^{e}, all of which only contain variables p^{tot}. As the copy constraint enforces $p^{\mathrm{tot}} = \overline{p}^{\mathrm{tot}}$, it holds that $R_j^{\mathrm{J}}(\overline{p}_j, \overline{u}_j) = R_j^{\mathrm{J}}(p_j, \overline{u}_j)$. Then, $\sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p_t, q_t) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j, \overline{u}_j) = \sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p_t, q_t) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(p_j, \overline{u}_j)$ and the solutions have the same objective value in their respective models. $\qquad\square$

Proof of Proposition 2

Proof. Under Assumption 1, there exists some solution $(\tilde{p}, \tilde{q}, \tilde{u}, \tilde{v}, \tilde{\Delta}, \tilde{\theta}, \tilde{\tau}, \tilde{\phi})$ that is feasible for (SCUC-ACOPF). Take any $(\overline{p}, \overline{q}, \overline{u})$ such that $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $\overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. We construct the solution (p, q, u) by the following rules. Take $(p^{\mathrm{fr}}, p^{\mathrm{to}}, q^{\mathrm{fr}}, q^{\mathrm{to}}) = (\tilde{p}^{\mathrm{fr}}, \tilde{p}^{\mathrm{to}}, \tilde{q}^{\mathrm{fr}}, \tilde{q}^{\mathrm{to}})$ and $u = \overline{u}$. Define $p_{jt}^{\mathrm{sh}} = g_j^{\mathrm{sh}} \overline{u}_{jt}^{\mathrm{sh}} \tilde{v}_{i_j t}^2$ and $q_{jt}^{\mathrm{sh}} = -b_j^{\mathrm{sh}} \overline{u}_{jt}^{\mathrm{sh}} \tilde{v}_{i_j t}^2$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. All other values in (p, q) take the corresponding values from $(\overline{p}, \overline{q})$.

This construction ensures that the solution $(p, q, u, \overline{p}, \overline{q}, \overline{u})$ is feasible for the copy constraints of (COPY). Further, the set constraints applied to the variables $(\overline{p}, \overline{q}, \overline{u})$ in (COPY) are satisfied by the selection of $(\overline{p}, \overline{q}, \overline{u})$. By Lemma 1, $u_{jt}^{\mathrm{sh}} = \overline{u}_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}} \subseteq \mathcal{Y}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$ and $(\overline{p}_{jt}, \overline{q}_{jt}, \overline{u}_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$. As $u = \overline{u}$ satisfies (1), by Assumption 1, u satisfies (6). Further, as (p, q, u) does not differ from $(\overline{p}, \overline{q}, \overline{u})$ in the variables that appear in constraints (2)–(3) and (7), we conclude that $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$.

Now, consider the solution $(p_t, q_t, \overline{u}_t^{\mathrm{sh}}, \tilde{v}_t, \tilde{\Delta}_t, \tilde{\theta}_t, \tilde{\tau}_t, \tilde{\phi}_t)$. The components of (p, q) which appear in constraints (5a)–(5k), namely $(p^{\mathrm{fr}}, p^{\mathrm{to}}, p^{\mathrm{sh}}, q^{\mathrm{fr}}, q^{\mathrm{to}}, q^{\mathrm{sh}})$ are constructed to ensure feasibility for these constraints with respect to this solution. As a result, we see that $(p_t, q_t, \overline{u}_t^{\mathrm{sh}}) = (p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$. This shows that $(p, q, u, \overline{p}, \overline{q}, \overline{u})$ is feasible for (COPY). $\qquad\square$

Proof of Theorem 1

Proof. To prove the first result, consider any iteration $\tau > 1$ and time period $t \in \mathcal{T}$. We analyze the generic form (A.3) of the model $Z_t^{\mathsf{T}}(\bar{p}_t^{(\tau-1)}, \bar{q}_t^{(\tau-1)}, \bar{u}_t^{(\tau-1)}; \rho_\tau)$.

Let the polyhedron $\mathcal{X}$ be described by (B, b), so that $\mathcal{X} = \{x : Bx \leq b\}$. Let (x, y, z) be a stationary point of (A.3), in that it satisfies the Karush-Kuhn-Tucker (KKT) conditions. There must exist some (λ_1, λ_2) such that (x, y, z) satisfies the following conditions, which are a subset of the KKT conditions:

$$0 \in - \begin{bmatrix} \partial_x(F(x, y, z) - \rho\Gamma(x, \bar{x})) \\ \partial_y F(x, y, z) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix}^{\mathsf{T}} \lambda_1 + \begin{bmatrix} A(z) \\ -I \end{bmatrix}^{\mathsf{T}} \lambda_2, \tag{A.4a}$$

$$\lambda_1^{\mathsf{T}}(b - Bx) = 0, \tag{A.4b}$$

$$\lambda_1 \geq 0, \tag{A.4c}$$

where I represents the identity matrix and $A(z)$ gives the matrix form of the linear map A_z. Note that $\partial_x(F(x, y, z) - \rho\Gamma(x, \bar{x})) = \partial_x F(x, y, z) - \rho\nabla_x\Gamma(x, \bar{x})$ as Γ is a smooth function (Rockafellar and Wets 2009, Exercise 8.8).

Fix (λ_1, λ_2) that satisfies (A.4) and denote by $\nabla F \in \begin{bmatrix} \partial_x F(x, y, z) \\ \partial_y F(x, y, z) \end{bmatrix}$ the (partial) subgradient of F at (x, y, z) that satisfies (A.4a); that is,

$$\nabla F = \begin{bmatrix} 2\rho E(x - \bar{x}) + B^{\mathsf{T}}\lambda_1 + A(z)^{\mathsf{T}}\lambda_2 \\ -\lambda_2 \end{bmatrix}.$$

Observe that the iterates $(\bar{p}_j^{(\tau-1)}, \bar{q}_j^{(\tau-1)}, \bar{u}_j^{(\tau-1)})$ are feasible for the subproblems Z_j^{J} for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $\bar{u}_{jt}^{\mathrm{sh}(\tau-1)}$ are feasible for Z_{jt}^{SH} for all $j \in \mathcal{J}^{\mathrm{sh}}$. Therefore, $\bar{x} \in \mathcal{X}$ by the structure of model (A.3). With $\bar{y} = A(z)\bar{x}$, we have that $(\bar{x}, \bar{y}, z)$ is feasible for (A.3).

By the above properties,

$$\nabla F^{\mathsf{T}} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} = \begin{bmatrix} 2\rho E(x - \bar{x}) + B^{\mathsf{T}}\lambda_1 + A(z)^{\mathsf{T}}\lambda_2 \\ -\lambda_2 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix}$$

$$= 2\rho(x - \bar{x})^{\mathsf{T}} E^{\mathsf{T}}(x - \bar{x}) + \lambda_1^{\mathsf{T}} B(x - \bar{x}) + \lambda_2^{\mathsf{T}}(A(z)(x - \bar{x}) - (y - \bar{y}))$$

$$= 2\rho(x - \bar{x})^{\mathsf{T}} E^{\mathsf{T}}(x - \bar{x}) + \lambda_1^{\mathsf{T}}(b - B\bar{x})$$

$$\geq 2\rho(x - \bar{x})^{\mathsf{T}} E^{\mathsf{T}}(x - \bar{x})$$

$$= 2\rho(x - \bar{x})^{\mathsf{T}} E^{\mathsf{T}} E(x - \bar{x}) = 2\rho \left\| E(x - \bar{x}) \right\|_2^2.$$

where the third relation follows from (A.4b) and the definitions $[y, \bar{y}] = A(z)[x, \bar{x}]$, the fourth relation follows from the feasibility of $\bar{x}$ for $\mathcal{X}$ and (A.4c), and the fifth relation follows from the property $E^{\mathsf{T}}E = E$.

We note that $\|\nabla F\|_2 \leq L$, and that (x, y, z) and $(\bar{x}, \bar{y}, z)$ are feasible for (A.3), so

$$2\rho \left\| E(x - \bar{x}) \right\|_2^2 \leq \nabla F^{\mathsf{T}} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \leq \|\nabla F\|_2 \left\| \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \right\|_2 \leq LD,$$

43

where the second relation follows from the Cauchy-Schwarz inequality. Therefore, $\Gamma(x, \overline{x}) = \|E(x - \overline{x})\|_2^2 \leq \frac{LD}{2\rho}$.

We now apply this result to the problems $Z_t^{\mathrm{T}}(\overline{p}_t^{(\tau-1)}, \overline{q}_t^{(\tau-1)}, \overline{u}_t^{(\tau-1)}; \rho_\tau)$ in each period $t \in \mathcal{T}$, where the feasible region diameters are given by D_t and the Lipschitz constants by L_t. Recall that d is the number of copied variables. By the separability of Γ_2 and the standard inequality relating the L_1 and L_2 norm in $\mathbb{R}^d$,

$$\Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)})$$
$$\leq \sqrt{d\Gamma_2(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)})}$$
$$= \sqrt{d \sum_{t \in \mathcal{T}} \Gamma_2(p_t^{(\tau)}, q_t^{(\tau)}, u_t^{(\tau)}, \overline{p}_t^{(\tau-1)}, \overline{q}_t^{(\tau-1)}, \overline{u}_t^{(\tau-1)})}$$
$$\leq \sqrt{\frac{\sum_{t \in \mathcal{T}} dL_t D_t}{2\rho_\tau}}.$$

This proves the first result.

Now, we consider the second result. It holds that

$$\Gamma_1(\overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)})$$
$$\leq \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)}) + \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$$
$$\leq 2\Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)}) + \frac{R}{\rho_\tau}$$
$$\leq \frac{R}{\rho_\tau} + \sqrt{\frac{2 \sum_{t \in \mathcal{T}} dL_t D_t}{\rho_\tau}},$$

where the first inequality follows from the triangle inequality, the second from Lemma A.1, and the third from the first result of this theorem. This application of Lemma A.1 requires the global optimality of subproblems Z_j^{J}. This proves the second result.

Last,

$$\Gamma_1(p^{(\tau+1)}, q^{(\tau+1)}, u^{(\tau+1)}, p^{(\tau)}, q^{(\tau)}, u^{(\tau)})$$
$$\leq \Gamma_1(p^{(\tau+1)}, q^{(\tau+1)}, u^{(\tau+1)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)}) + \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$$
$$\leq \sqrt{\frac{\sum_{t \in \mathcal{T}} dL_t D_t}{2\rho_{\tau+1}}} + \frac{R}{\rho_\tau} + \Gamma_1(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau-1)}, \overline{q}^{(\tau-1)}, \overline{u}^{(\tau-1)})$$
$$\leq \frac{R}{\rho_\tau} + \sqrt{\frac{2 \sum_{t \in \mathcal{T}} dL_t D_t}{\min\{\rho_\tau, \rho_{\tau+1}\}}}$$

where the first inequality follows from the triangle inequality and the remaining inequalities follow from the first result of this theorem and Lemma A.1. This proves the third result. $\square$

Proof of Theorem 2

Proof. Consider some limit point $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$ of the sequence $(p^{(\tau)}, q^{(\tau)}, u^{(\tau)}, \overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$ generated by Algorithm 1, and let $\{\tau_i\}_{i=1}^\infty$ give the subsequence of indices which converge to this limit point.

We first demonstrate that this point is feasible for (COPY). By Lemma A.1 and Theorem 1,

$$0 \leq \lim_{i \to \infty} \Gamma_1(p^{(\tau_i)}, q^{(\tau_i)}, u^{(\tau_i)}, \overline{p}^{(\tau_i)}, \overline{q}^{(\tau_i)}, \overline{u}^{(\tau_i)})$$

$$\leq \lim_{i \to \infty} \frac{R}{\rho_{\tau_i}} + \Gamma_1(p^{(\tau_i)}, q^{(\tau_i)}, u^{(\tau_i)}, \overline{p}^{(\tau_i-1)}, \overline{q}^{(\tau_i-1)}, \overline{u}^{(\tau_i-1)})$$

$$\leq \lim_{i \to \infty} \frac{R}{\rho_{\tau_i}} + \sqrt{\frac{C}{2\rho_{\tau_i}}} = 0.$$

As Γ_1 is continuous, this gives that

$$\Gamma_1(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*) = \lim_{i \to \infty} \Gamma_1(p^{(\tau_i)}, q^{(\tau_i)}, u^{(\tau_i)}, \overline{p}^{(\tau_i)}, \overline{q}^{(\tau_i)}, \overline{u}^{(\tau_i)}) = 0.$$

This is equivalent to

$$(p^{\text{tot}*}, p^{\text{su}*}, p^{\text{sd}*}, q^{\text{tot}*}, q^{\text{res}*}, u^{\text{on}*}, u^{\text{sh}*}) = (\overline{p}^{\text{tot}*}, \overline{p}^{\text{su}*}, \overline{p}^{\text{sd}*}, \overline{q}^{\text{tot}*}, \overline{q}^{\text{res}*}, \overline{u}^{\text{on}*}, \overline{u}^{\text{sh}*}),$$

and thus the limit point satisfies the copy constraint of (COPY).

Algorithm 1 guarantees that the solution $(p_t^{(\tau_i)}, q_t^{(\tau_i)}, u_t^{(\tau_i)})$ is feasible for Z_t^{T} for all $t \in \mathcal{T}$, the solution $(\overline{p}_j^{(\tau_i)}, \overline{q}_j^{(\tau_i)}, \overline{u}_j^{(\tau_i)})$ is feasible for Z_j^{J} for all $j \in \mathcal{J}^{\text{sd}}$, and $\overline{u}_{jt}^{\text{sh}(\tau_i)}$ is feasible for Z_{jt}^{SH} for all $j \in \mathcal{J}^{\text{sh}}$ and $t \in \mathcal{T}$. Under Assumption 1, the subproblem feasible regions are closed, and thus the limit $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$ is contained in the appropriate sets. This demonstrates that the limit point is feasible for (COPY).

Next, we show that the limit point is optimal for (COPY) with the first variable block (p, q, u) fixed to (p^*, q^*, u^*). Consider the model (COPY) with the additional constraint that $(p, q, u) = (p^*, q^*, u^*)$, which can be written as

$$\max_{\overline{p}, \overline{q}, \overline{u}} \quad \sum_{t \in \mathcal{T}} R_t^{\text{T}}(p_t^*, q_t^*) + \sum_{j \in \mathcal{J}^{\text{sd}}} R_j^{\text{J}}(\overline{p}_j, \overline{u}_j)$$

$$\text{s.t.} \quad (\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\text{uc}} \qquad \forall j \in \mathcal{J}^{\text{sd}}, \tag{A.5}$$

$$\overline{p}^{\text{tot}} = p^{\text{tot}*}, \ \overline{p}^{\text{su}} = p^{\text{su}*}, \ \overline{p}^{\text{sd}} = p^{\text{sd}*}, \ \overline{q}^{\text{tot}} = q^{\text{tot}*}, \ \overline{q}^{\text{res}} = q^{\text{res}*}, \ \overline{u}^{\text{on}} = u^{\text{on}*}, \ \overline{u}^{\text{sh}} = u^{\text{sh}*}.$$

As above, $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$ is feasible for this problem. Consider any feasible solution $(\tilde{p}, \tilde{q}, \tilde{u})$ to (A.5).

Theorem 1 implies that the maximum difference between the copied variables of $(\overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$ becomes arbitrarily small with increasing ρ_τ, i.e., the copied variables of the iterates form a Cauchy sequence. Thus, the iterates converge to $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$ in the copied variables. Binary variables admit only unit changes in value, which implies that there is some iteration τ' after which the copied binary variables $\overline{u}^{\text{on}}$ and $\overline{u}^{\text{sh}}$ never deviate from their values $\overline{u}^{\text{on}*}$ and $\overline{u}^{\text{sh}*}$ in the limit point. By constraints (1a)–(1b), and (1d), the variables $\overline{u}^{\text{su}}$ and $\overline{u}^{\text{sd}}$ are uniquely defined by $\overline{u}^{\text{on}}$, and therefore, after iteration τ', all binary variables in solution $\overline{u}^{(\tau)}$ remain at their value in solution $\overline{u}^*$. Similarly, as feasibility for (A.5) requires $\tilde{u}^{\text{on}} = u^{\text{on}*}$ and $\tilde{u}^{\text{sh}} = u^{\text{sh}*}$, all binary variables in $\tilde{u}$ have the same value as in $\overline{u}^*$.

Define by

$$\mathcal{Z}^{\text{uc}}(\overline{u}^*) := \{(\overline{p}, \overline{q}, \overline{u}) \ : \ \overline{u} = \overline{u}^*, \ (\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\text{uc}} \ \forall j \in \mathcal{J}^{\text{sd}}\}.$$

Clearly, $(\overline{p}^*, \overline{q}^*, \overline{u}^*) \in \mathcal{Z}^{\mathrm{uc}}(\overline{u}^*)$. Also, by the logic above, $(\tilde{p}, \tilde{q}, \tilde{u}) \in \mathcal{Z}^{\mathrm{uc}}(\overline{u}^*)$ and $(\overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)}) \in \mathcal{Z}^{\mathrm{uc}}(\overline{u}^*)$ for all $\tau \geq \tau'$. As all binary variables are fixed, $\mathcal{Z}^{\mathrm{uc}}(\overline{u}^*)$ is polyhedral under Assumption 1.

Consider some subsequence of $\{\tau_i\}_{i=1}^{\infty}$, denoted by $\{\tau_i'\}_{i=1}^{\infty}$, such that $\tau_1' \geq \tau'$. We apply Lemma A.2 to polyhedron $\mathcal{Z}^{\mathrm{uc}}(\overline{u}^*)$, solutions $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$ and $(\tilde{p}, \tilde{q}, \tilde{u})$, and sequence $(\overline{p}^{(\tau_i')}, \overline{q}^{(\tau_i')}, \overline{u}^{(\tau_i')})$. Denote by $\{\lambda_i\}_{i=1}^{\infty}$ a sequence of multipliers that satisfies the statement of Lemma A.2. We then define

$$(\tilde{p}^{(\tau_i')}, \tilde{q}^{(\tau_i')}, \tilde{u}^{(\tau_i')}) = (\overline{p}^{(\tau_i')}, \overline{q}^{(\tau_i')}, \overline{u}^{(\tau_i')}) + \lambda_i((\tilde{p}, \tilde{q}, \tilde{u}) - (\overline{p}^*, \overline{q}^*, \overline{u}^*)).$$

By application of Lemma A.2, the sequence $(\tilde{p}^{(\tau_i')}, \tilde{q}^{(\tau_i')}, \tilde{u}^{(\tau_i')})$ is feasible for $\mathcal{Z}^{\mathrm{uc}}(\overline{u}^*)$ and thus is feasible for subproblems Z_j^{J}, and has $\lim_{i \to \infty}(\tilde{p}^{(\tau_i')}, \tilde{q}^{(\tau_i')}, \tilde{u}^{(\tau_i')}) = (\tilde{p}, \tilde{q}, \tilde{u})$.

By feasibility for (A.5), it holds that solutions $(\tilde{p}, \tilde{q}, \tilde{u})$ and $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$ have the same values for the copied variables (which appear in Γ_1), and thus $(\tilde{p}^{(\tau_i')}, \tilde{q}^{(\tau_i')}, \tilde{u}^{(\tau_i')})$ and $(\overline{p}^{(\tau_i')}, \overline{q}^{(\tau_i')}, \overline{u}^{(\tau_i')})$ have the same values in those variables for all i. As a result, for all $i \in \mathbb{Z}_{\geq}$ and $j \in \mathcal{J}^{\mathrm{sd}}$ and any solution (p_j, q_j, u_j),

$$\Gamma_1'(p_j, q_j, u_j, \overline{p}_j^{(\tau_i')}, \overline{q}_j^{(\tau_i')}, \overline{u}_j^{(\tau_i')}) = \Gamma_1'(p_j, q_j, u_j, \tilde{p}_j^{(\tau_i')}, \tilde{q}_j^{(\tau_i')}, \tilde{u}_j^{(\tau_i')}).$$

To conclude, we observe that

$$\sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j^*, \overline{u}_j^*) = \lim_{i \to \infty} \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j^{(\tau_i')}, \overline{u}_j^{(\tau_i')})$$

$$= \lim_{i \to \infty} \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\overline{p}_j^{(\tau_i')}, \overline{u}_j^{(\tau_i')}) -$$

$$\rho_{\tau_i'} \left(\Gamma_1'(p_j^{(\tau_i')}, q_j^{(\tau_i')}, u_j^{(\tau_i')}, \overline{p}_j^{(\tau_i')}, \overline{q}_j^{(\tau_i')}, \overline{u}_j^{(\tau_i')}) - \Gamma_1'(p_j^{(\tau_i')}, q_j^{(\tau_i')}, u_j^{(\tau_i')}, \tilde{p}_j^{(\tau_i')}, \tilde{q}_j^{(\tau_i')}, \tilde{u}_j^{(\tau_i')}) \right)$$

$$\geq \lim_{i \to \infty} \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\tilde{p}_j^{(\tau_i')}, \tilde{u}_j^{(\tau_i')}) -$$

$$\rho_{\tau_i'} \left(\Gamma_1'(p_j^{(\tau_i')}, q_j^{(\tau_i')}, u_j^{(\tau_i')}, \tilde{p}_j^{(\tau_i')}, \tilde{q}_j^{(\tau_i')}, \tilde{u}_j^{(\tau_i')}) - \Gamma_1'(p_j^{(\tau_i')}, q_j^{(\tau_i')}, u_j^{(\tau_i')}, \tilde{p}_j^{(\tau_i')}, \tilde{q}_j^{(\tau_i')}, \tilde{u}_j^{(\tau_i')}) \right)$$

$$= \lim_{i \to \infty} \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\tilde{p}_j^{(\tau_i')}, \tilde{u}_j^{(\tau_i')})$$

$$= \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\tilde{p}_j, \tilde{u}_j),$$

where the inequality follows from the global optimality of iterates $(\overline{p}^{(\tau)}, \overline{q}^{(\tau)}, \overline{u}^{(\tau)})$ for subproblems Z_j^{J}. This proves that the solution $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$ is optimal for (A.5), and thus satisfies the global partial optimality condition for its block.

Last, we consider the other block of variables and show that the limit point is optimal for (COPY) with the second variable block $(\overline{p}, \overline{q}, \overline{u})$ fixed to $(\overline{p}^*, \overline{q}^*, \overline{u}^*)$. We again analyze the generic form (A.3), which has limit point $(x^*, y^*, z^*, \overline{x}^*, \overline{y}^*, \overline{z}^*)$ that corresponds to $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$. Problem (COPY)

with the additional constraint $(\bar{p}, \bar{q}, \bar{u}) = (\bar{p}^*, \bar{q}^*, \bar{u}^*)$ can be written in this generic form as

$$
\begin{aligned}
\max_{x,y,z} \quad & \sum_{t \in \mathcal{T}} F_t(x_t, y_t, z_t) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(\bar{p}_j^*, \bar{u}_j^*) \\
\text{s.t.} \quad & x_t \in \mathcal{X}_t && \forall t \in \mathcal{T}, \\
& y_t = A_{z_t}(x_t) && \forall t \in \mathcal{T}, \\
& z_t \in \mathcal{Z}_t && \forall t \in \mathcal{T}, \\
& Ex_t = E\bar{x}_t^* && \forall t \in \mathcal{T},
\end{aligned}
\tag{A.6}
$$

where the objects $(F, \mathcal{X}, \mathcal{Z})$ for each time period are now indexed by t, and Ex_t selects the variables which appear in the copy constraint. We prove the result by demonstrating that (x^*, y^*, z^*) is optimal for (A.6).

Consider any $(\tilde{x}, \tilde{y}, \tilde{z})$ feasible for (A.6). For each $t \in \mathcal{T}$, we apply Lemma A.2 to polyhedron $\mathcal{X}_t$, solutions x_t^* and $\tilde{x}_t$, and sequence $x_t^{(\tau_i)}$. Note that $(x^{(\tau_i)}, y^{(\tau_i)}, z^{(\tau_i)})$ corresponds to the sequence $(p^{(\tau_i)}, q^{(\tau_i)}, u^{(\tau_i)})$. Denote by $\{\lambda_{it}\}_{i=1}^{\infty}$ a sequence of multipliers that satisfies the statement of Lemma A.2. We then define

$$
\tilde{x}_t^{(\tau_i)} = x_t^{(\tau_i)} + \lambda_{it}(\tilde{x}_t - x_t^*) \qquad \text{and} \qquad \tilde{y}_t^{(\tau_i)} = A_{\tilde{z}_t}(\tilde{x}_t^{(\tau_i)}).
$$

By Lemma A.2, the sequence $\tilde{x}_t^{(\tau_i)} \in \mathcal{X}_t$, so $(\tilde{x}_t^{(\tau_i)}, \tilde{y}_t^{(\tau_i)}, \tilde{z}_t)$ is feasible for (A.3) at time t and thus for subproblem Z_t^{T}. Note that this sequence holds the variables $\tilde{z}_t$ constant. Lemma A.2 gives that $\lim_{i \to \infty} \tilde{x}_t^{(\tau_i)} = \tilde{x}_t$. As $\tilde{y}_t = A_{\tilde{z}_t}(\tilde{x}_t)$ and $A_{\tilde{z}_t}$ is continuous, $\lim_{i \to \infty} \tilde{y}_t^{(\tau_i)} = \tilde{y}_t$.

By feasibility for (A.6), the solutions $\tilde{x}_t$ and x_t^* have the same values for the copied variables Ex_t that appear in penalty function Γ. As a result, $\tilde{x}_t^{(\tau_i)}$ and $x_t^{(\tau_i)}$ have the same values in these variables for all i. Thus, for all $i \in \mathbb{Z}_{\geq}$ and $t \in \mathcal{T}$ and any solution $\bar{x}_t$,

$$
\Gamma(x_t^{(\tau_i)}, \bar{x}_t) = \Gamma(\tilde{x}_t^{(\tau_i)}, \bar{x}_t).
$$

To conclude, we have that

$$
\begin{aligned}
\sum_{t \in \mathcal{T}} F_t(x_t^*, y_t^*, z_t^*) = \; & \lim_{i \to \infty} \sum_{t \in \mathcal{T}} F_t(x_t^{(\tau_i)}, y_t^{(\tau_i)}, z_t^{(\tau_i)}) \\
= \; & \lim_{i \to \infty} \sum_{t \in \mathcal{T}} F_t(x_t^{(\tau_i)}, y_t^{(\tau_i)}, z_t^{(\tau_i)}) - \rho_{\tau_i} \left(\Gamma(x_t^{(\tau_i)}, \bar{x}_t^{(\tau_i - 1)}) - \Gamma(\tilde{x}_t^{(\tau_i)}, \bar{x}_t^{(\tau_i - 1)}) \right) \\
\geq \; & \lim_{i \to \infty} \sum_{t \in \mathcal{T}} F_t(\tilde{x}_t^{(\tau_i)}, \tilde{y}_t^{(\tau_i)}, \tilde{z}_t) - \rho_{\tau_i} \left(\Gamma(\tilde{x}_t^{(\tau_i)}, \bar{x}_t^{(\tau_i - 1)}) - \Gamma(\tilde{x}_t^{(\tau_i)}, \bar{x}_t^{(\tau_i - 1)}) \right) \\
= \; & \lim_{i \to \infty} \sum_{t \in \mathcal{T}} F_t(\tilde{x}_t^{(\tau_i)}, \tilde{y}_t^{(\tau_i)}, \tilde{z}_t) \\
= \; & \sum_{t \in \mathcal{T}} F_t(\tilde{x}_t, \tilde{y}_t, \tilde{z}_t),
\end{aligned}
$$

where the inequality follows from the global optimality of iterates $(x_t^{(\tau)}, y_t^{(\tau)}, z_t^{(\tau)})$ for subproblems Z_t^{T}. This demonstrates the optimality of (x^*, y^*, z^*) for (A.6), and therefore (p^*, q^*, u^*) satisfies the partial

optimality condition for its block. This demonstrates that $(p^*, q^*, u^*, \overline{p}^*, \overline{q}^*, \overline{u}^*)$ is a partial optimum for (COPY). □

Proof of Proposition 3

Proof. Take any $j \in \mathcal{J}^{\mathrm{sd}}$ and $(\overline{p}_j, \overline{q}_j, \overline{u}_j)$ such that $(\overline{p}_j, \overline{q}_j, \overline{u}_j) \in \mathcal{X}_j^{\mathrm{uc}}$. For all $t \in \mathcal{T}$, construct

$$\delta_{jt}^- = \frac{1}{2} \min \left\{ r_{jt}^{\mathrm{d}} + (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}),\ r_{j,t+1}^{\mathrm{u}} - (\overline{p}_{j,t+1}^{\mathrm{tot}} - \overline{p}_{jt}^{\mathrm{tot}}) \right\},$$

$$\delta_{jt}^+ = \frac{1}{2} \min \left\{ r_{jt}^{\mathrm{u}} - (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}),\ r_{j,t+1}^{\mathrm{d}} + (\overline{p}_{j,t+1}^{\mathrm{tot}} - \overline{p}_{jt}^{\mathrm{tot}}) \right\},$$

which satisfies (10d)–(10e) and

$$\delta_{jt}^- + \delta_{j,t-1}^+ \leq \frac{1}{2} \left(r_{jt}^{\mathrm{d}} + (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}) + r_{jt}^{\mathrm{d}} + (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}) \right) = r_{jt}^{\mathrm{d}} + (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}),$$

$$\delta_{j,t-1}^- + \delta_{jt}^+ \leq \frac{1}{2} \left(r_{jt}^{\mathrm{u}} - (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}) + r_{jt}^{\mathrm{u}} - (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}) \right) = r_{jt}^{\mathrm{u}} - (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}}).$$

The construction thus satisfies the constraints (10b)–(10c) and is feasible for (10).

Next, we demonstrate nonnegativity and boundedness. Consider any (δ_j^-, δ_j^+) that satisfies (10b)–(10e). Since $(\overline{p}_j, \overline{q}_j, \overline{u}_j)$ satisfies the ramping constraints (4), we have $-r_{jt}^{\mathrm{d}} \leq \overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}} \leq r_{jt}^{\mathrm{u}}$ for all $t \in \mathcal{T}$. Then, constraints (10d)–(10e) imply nonnegativity. Constraint (10b) provides an upper bound on δ_{jt}^-. Combining $\delta_{j,t-1}^+ \geq 0$ and (10b) gives that $\delta_{jt}^- \leq r_{jt}^{\mathrm{d}} + (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}})$. Similarly, from constraint (10c), $\delta_{jt}^+ \leq r_{jt}^{\mathrm{u}} - (\overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}})$. Thus, the set of feasible $(\delta_{jt}^-, \delta_{jt}^+)$ is bounded. □

Proof of Proposition 4

Proof. Consider any $(\overline{p}, \overline{u}, \delta)$ that satisfies (10b)–(10c) for all $j \in \mathcal{J}^{\mathrm{sd}}$. Take any p such that p_{jt}^{tot} satisfies (9), i.e.,

$$\overline{p}_{jt}^{\mathrm{tot}} - \delta_{jt}^- \leq p_{jt}^{\mathrm{tot}} \leq \overline{p}_{jt}^{\mathrm{tot}} + \delta_{jt}^+ \quad \forall j \in \mathcal{J}^{\mathrm{sd}},\ t \in \mathcal{T}.$$

Constraints (9), (10b), and definition (10f) yield

$$p_{jt}^{\mathrm{tot}} - p_{j,t-1}^{\mathrm{tot}} \geq \overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}} - (\delta_{jt}^- + \delta_{j,t-1}^+) \geq -r_{jt}^{\mathrm{d}} = -d_t \left(p_j^{\mathrm{rd}} \overline{u}_{jt}^{\mathrm{on}} + p_j^{\mathrm{rd,sd}} (1 - \overline{u}_{jt}^{\mathrm{on}}) \right),$$

and thus $(p, \overline{u})$ satisfies (4a). Similarly, constraints (9), (10c), and definition (10g) yield

$$p_{jt}^{\mathrm{tot}} - p_{j,t-1}^{\mathrm{tot}} \leq \overline{p}_{jt}^{\mathrm{tot}} - \overline{p}_{j,t-1}^{\mathrm{tot}} + \delta_{j,t-1}^- + \delta_{jt}^+ \leq r_{jt}^{\mathrm{u}} = d_t \left(p_j^{\mathrm{ru}} (\overline{u}_{jt}^{\mathrm{on}} - \overline{u}_{jt}^{\mathrm{su}}) + p_j^{\mathrm{ru,su}} (\overline{u}_{jt}^{\mathrm{su}} - \overline{u}_{jt}^{\mathrm{on}} + 1) \right),$$

and thus $(p, \overline{u})$ satisfies (4b). □

Proof of Proposition 5

Proof. The result relies on the inequality $[\sum_i x_i]^+ \le \sum_i [x_i]^+$ for any x. Then,

$$\sum_{t\in\mathcal{T}} \sum_{W\in\mathcal{W}_{jt}^{\text{en,min}}} d_t \left[\frac{e_j^{\min}(W) - \sum_{t'\in W} d_{t'}\bar{p}_{jt'}^{\text{tot}}}{\sum_{t'\in W} d_{t'}} + \bar{p}_{jt}^{\text{tot}} - p_{jt}^{\text{tot}} \right]^+$$

$$= \sum_{W\in\mathcal{W}_j^{\text{en,min}}} \sum_{t\in W} d_t \left[\frac{e_j^{\min}(W) - \sum_{t'\in W} d_{t'}\bar{p}_{jt'}^{\text{tot}}}{\sum_{t'\in W} d_{t'}} + \bar{p}_{jt}^{\text{tot}} - p_{jt}^{\text{tot}} \right]^+$$

$$\ge \sum_{W\in\mathcal{W}_j^{\text{en,min}}} \left[\sum_{t\in W} d_t \left(\frac{e_j^{\min}(W) - \sum_{t'\in W} d_{t'}\bar{p}_{jt'}^{\text{tot}}}{\sum_{t'\in W} d_{t'}} + \bar{p}_{jt}^{\text{tot}} - p_{jt}^{\text{tot}} \right) \right]^+$$

$$= \sum_{W\in\mathcal{W}_j^{\text{en,min}}} \left[e_j^{\min}(W) - \sum_{t\in W} d_t p_{jt}^{\text{tot}} \right]^+.$$

The first relation follows from the definition of $\mathcal{W}_{jt}^{\text{en,min}}$ and the second from the aforementioned property of $[\cdot]^+$ and nonnegativity of d_t. Similarly, for the energy maximum terms,

$$\sum_{t\in\mathcal{T}} \sum_{W\in\mathcal{W}_{jt}^{\text{en,max}}} d_t \left[\frac{-e_j^{\max}(W) + \sum_{t'\in W} d_{t'}\bar{p}_{jt'}^{\text{tot}}}{\sum_{t'\in W} d_{t'}} + p_{jt}^{\text{tot}} - \bar{p}_{jt}^{\text{tot}} \right]^+$$

$$= \sum_{W\in\mathcal{W}_j^{\text{en,max}}} \sum_{t\in W} d_t \left[\frac{-e_j^{\max}(W) + \sum_{t'\in W} d_{t'}\bar{p}_{jt'}^{\text{tot}}}{\sum_{t'\in W} d_{t'}} + p_{jt}^{\text{tot}} - \bar{p}_{jt}^{\text{tot}} \right]^+$$

$$\ge \sum_{W\in\mathcal{W}_j^{\text{en,max}}} \left[\sum_{t\in W} d_t \left(\frac{-e_j^{\max}(W) + \sum_{t'\in W} d_{t'}\bar{p}_{jt'}^{\text{tot}}}{\sum_{t'\in W} d_{t'}} + p_{jt}^{\text{tot}} - \bar{p}_{jt}^{\text{tot}} \right) \right]^+$$

$$= \sum_{W\in\mathcal{W}_j^{\text{en,max}}} \left[-e_j^{\max}(W) + \sum_{t\in W} d_t p_{jt}^{\text{tot}} \right]^+.$$

Together, as $c^{\text{e}} \ge 0$, these properties yield the bound $C_j^{\text{e}}(p) \le \sum_{t\in\mathcal{T}} \overline{C}_{jt}^{\text{e}}(p;\bar{p})$. $\qquad\square$

Proof of Proposition 6

Proof. Consider $t \in \mathcal{T}$ and $(p_t, q_t, u_t^{\text{sh}}) \in \mathcal{X}_t^{\text{ac}}$ with $u_{jt}^{\text{sh}} \in \mathcal{Y}_{jt}^{\text{sh}}$ for all $j \in \mathcal{J}^{\text{sh}}$. Then, there is some solution $(v_t, \Delta_t, \theta_t, \tau_t, \phi_t)$ such that this solution and $(p_t, q_t, u_t^{\text{sh}})$ are feasible for (5a)–(5k). Construct the following objects:

$$c_{jt} = \frac{v_{i_jt}v_{i'_jt}}{\tau_{jt}}\cos(\Delta_{jt}), \ \ s_{jt} = \frac{v_{i_jt}v_{i'_jt}}{\tau_{jt}}\sin(\Delta_{jt}), \ \ \mu_{jt} = \frac{v_{i_jt}^2}{\tau_{jt}^2} \quad \forall j \in \mathcal{J}^{\text{ac}},$$

$$\omega_{it} = v_{it}^2 \quad \forall i \in \mathcal{I}, \quad \mu_{jt}^{\text{sh}} = u_{jt}^{\text{sh}}v_{i_jt}^2 \quad \forall j \in \mathcal{J}^{\text{sh}}.$$

We consider the feasibility of the solution $(c_t, s_t, p_t, q_t, \mu_t, \omega_t)$ for constraints (11). First, by (5b)–(5e) we have for all $j \in \mathcal{J}^{\mathrm{ac}}$ that

$$g_{i_j j}\mu_{jt} - g_j c_{jt} - b_j s_{jt} = \frac{g_{i_j j} v_{i_j t}^2}{\tau_{jt}^2} + \frac{v_{i_j t} v_{i'_j t}}{\tau_{jt}} \left(-g_j \cos(\Delta_{jt}) - b_j \sin(\Delta_{jt})\right) = p_{jt}^{\mathrm{fr}},$$

$$g_{i'_j j}\omega_{i'_j t} - g_j c_{jt} + b_j s_{jt} = g_{i'_j j} v_{i'_j t}^2 + \frac{v_{i_j t} v_{i'_j t}}{\tau_{jt}} \left(-g_j \cos(\Delta_{jt}) + b_j \sin(\Delta_{jt})\right) = p_{jt}^{\mathrm{to}},$$

$$-b_{i_j j}\mu_{jt} + b_j c_{jt} - g_j s_{jt} = \frac{-b_{i_j j} v_{i_j t}^2}{\tau_{jt}^2} + \frac{v_{i_j t} v_{i'_j t}}{\tau_{jt}} \left(b_j \cos(\Delta_{jt}) - g_j \sin(\Delta_{jt})\right) = q_{jt}^{\mathrm{fr}},$$

$$-b_{i'_j j}\omega_{i'_j t} + b_j c_{jt} + g_j s_{jt} = -b_{i'_j j} v_{i'_j t}^2 + \frac{v_{i_j t} v_{i'_j t}}{\tau_{jt}} \left(b_j \cos(\Delta_{jt}) + g_j \sin(\Delta_{jt})\right) = q_{jt}^{\mathrm{to}}.$$

Next, using (5g)–(5h), we have for all $j \in \mathcal{J}^{\mathrm{sh}}$ that

$$g_j^{\mathrm{sh}}\mu_{jt}^{\mathrm{sh}} = g_j^{\mathrm{sh}} u_{jt}^{\mathrm{sh}} v_{i_j t}^2 = p_{jt}^{\mathrm{sh}} \quad \text{and} \quad -b_j^{\mathrm{sh}}\mu_{jt}^{\mathrm{sh}} = -b_j^{\mathrm{sh}} u_{jt}^{\mathrm{sh}} v_{i_j t}^2 = q_{jt}^{\mathrm{sh}}.$$

This establishes that the solution satisfies constraints (11a)–(11f).

We now consider the remaining constraints. It holds for all $j \in \mathcal{J}^{\mathrm{ac}}$ that

$$c_{jt}^2 + s_{jt}^2 = \frac{v_{i_j t}^2 v_{i'_j t}^2}{\tau_{jt}^2} \left(\cos^2(\Delta_{jt}) + \sin^2(\Delta_{jt})\right) = \frac{v_{i_j t}^2 v_{i'_j t}^2}{\tau_{jt}^2} = \mu_{jt}\omega_{i'_j t},$$

and (11g) holds. By constraint (5j), nonnegativity of τ^{min} and τ^{max}, and $\omega_{it} = v_{it}^2 \geq 0$,

$$\frac{\omega_{i_j t}}{(\tau_j^{\mathrm{max}})^2} \leq \frac{\omega_{i_j t}}{\tau_{jt}^2} = \mu_{jt} = \frac{\omega_{i_j t}}{\tau_{jt}^2} \leq \frac{\omega_{i_j t}}{(\tau_j^{\mathrm{min}})^2},$$

and (11h) holds. By definition of $\mathcal{Y}_{jt}^{\mathrm{sh}}$, for all $j \in \mathcal{J}^{\mathrm{sh}}$ we have $u_j^{\mathrm{sh,min}} \leq u_{jt}^{\mathrm{sh}} \leq u_j^{\mathrm{sh,max}}$. Thus,

$$u_j^{\mathrm{sh,min}}\omega_{i_j t} \leq u_{jt}^{\mathrm{sh}}\omega_{i_j t} = u_{jt}^{\mathrm{sh}} v_{i_j t}^2 = \mu_{jt}^{\mathrm{sh}} = u_{jt}^{\mathrm{sh}} v_{i_j t}^2 = u_{jt}^{\mathrm{sh}}\omega_{i_j t} \leq u_j^{\mathrm{sh,max}}\omega_{i_j t},$$

again using $\omega \geq 0$, and (11i) is satisfied. For all $i \in \mathcal{I}$, due to (5i) and $(v^{\mathrm{min}}, v^{\mathrm{max}}) \geq 0$,

$$(v_i^{\mathrm{min}})^2 \leq v_{it}^2 = \omega_{it} = v_{it}^2 \leq (v_i^{\mathrm{max}})^2,$$

and (11j) holds. The solution trivially satisfies (11k). Therefore, $(p_t, q_t, \mu_t, \omega_t) \in \mathcal{X}_t^{\mathrm{soc}}$. By construction, $\frac{\mu_{jt}^{\mathrm{sh}}}{\omega_{i_j t}} = u_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and the result holds. $\qquad \square$

Proof of Proposition 7

Proof. Fix some $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$. First, consider a solution $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$. Then, let $(u_j^{\mathrm{su}}, u_j^{\mathrm{sd}}) \geq 0$ certify inclusion in $\mathcal{Y}_{jt}^{\mathrm{uc}}$. Define

$$\chi_{jt}^{\mathrm{su}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} \quad \text{and} \quad \chi_{jt}^{\mathrm{sd}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}}.$$

50

By (6), we have $u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} \leq 1$ and $u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{sd}} \leq 1$. Using these properties and $u_{jt}^{\mathrm{on}} \geq 0$, we have $\chi_{jt}^{\mathrm{su}} \leq 1$ and $\chi_{jt}^{\mathrm{sd}} \leq 1$. Further, we have $u_{jt'}^{\mathrm{su}} \geq 0$, so $\chi_{jt}^{\mathrm{su}} \geq 0$. Similarly, $\chi_{jt}^{\mathrm{sd}} \geq 0$. Thus, the solution $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt})$ satisfies (12e)–(12g). By constraint (7a) and nonnegativity of $u_{jt'}^{\mathrm{su}}$ and $p_{jtt'}^{\mathrm{supc}}$,

$$0 \leq p_{jt}^{\mathrm{su}} \leq \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} u_{jt'}^{\mathrm{su}} \leq \max_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} = \chi_{jt}^{\mathrm{su}} \max_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}},$$

and thus the solution satisfies (12a). Similarly, constraint (7b) implies (12b). Next, let $a = \uparrow$ if $j \in \mathcal{J}^{\mathrm{pr}}$ and otherwise $a = \downarrow$. By constraint (3a), we have

$$q_{jt}^{\mathrm{tot}} \leq q_{jt}^{\max} \left(u_{jt}^{\mathrm{on}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}} \right) + \beta_j^{\max} p_{jt}^{\mathrm{tot}} - q_{jta}^{\mathrm{res}} = q_{jt}^{\max} \left(u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} + \chi_{jt}^{\mathrm{sd}} \right) + \beta_j^{\max} p_{jt}^{\mathrm{tot}} - q_{jta}^{\mathrm{res}},$$

and (12c) is satisfied. Similarly, by (3b), we see that (12d) holds. Therefore, $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt})$ satisfies constraints (12). As constraints (2a), (2d), (2e), and (3c) do not contain u_{jt}^{su} and u_{jt}^{sd}, the solution also satisfies these constraints, and $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt}) \in \overline{\mathcal{Y}}_{jt}^{\mathrm{uc}}$.

In the other direction, take $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}})$ such that $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt}) \in \overline{\mathcal{Y}}_{jt}^{\mathrm{uc}}$ for some χ_{jt}. Let

$$\overline{t}^{\mathrm{su}} \in \arg\max_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} \quad \text{and} \quad \overline{t}^{\mathrm{sd}} \in \arg\max_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} p_{jtt'}^{\mathrm{sdpc}}.$$

We construct the solution $(u_j^{\mathrm{su}}, u_j^{\mathrm{sd}})$ for all $t' \in \mathcal{T}$ by the following rule:

$$u_{jt'}^{\mathrm{su}} = \begin{cases} \chi_{jt}^{\mathrm{su}} & \text{if } t' = \overline{t}^{\mathrm{su}} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad u_{jt'}^{\mathrm{sd}} = \begin{cases} \chi_{jt}^{\mathrm{sd}} & \text{if } t' = \overline{t}^{\mathrm{sd}} \\ 0 & \text{otherwise.} \end{cases}$$

Consider the solution $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, u_j^{\mathrm{su}}, u_j^{\mathrm{sd}})$. By construction, we have $\sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} = \chi_{jt}^{\mathrm{su}}$ and $\sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}} = \chi_{jt}^{\mathrm{sd}}$. Further, $(u_{jt}^{\mathrm{on}}, u_j^{\mathrm{su}}, u_j^{\mathrm{sd}}) \in [0,1]^{2T+1}$ by constraint (12g). Let $a = \uparrow$ if $j \in \mathcal{J}^{\mathrm{pr}}$ and otherwise $a = \downarrow$. Then, by constraint (12c) we have

$$q_{jt}^{\mathrm{tot}} \leq q_{jt}^{\max} \left(u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} + \chi_{jt}^{\mathrm{sd}} \right) + \beta_j^{\max} p_{jt}^{\mathrm{tot}} - q_{jta}^{\mathrm{res}} = q_{jt}^{\max} \left(u_{jt}^{\mathrm{on}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}} \right) + \beta_j^{\max} p_{jt}^{\mathrm{tot}} - q_{jta}^{\mathrm{res}},$$

and (3a) is satisfied. Similarly, (12d) implies (3b). By constraint (12e),

$$u_{jt}^{\mathrm{on}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} u_{jt'}^{\mathrm{su}} = u_{jt}^{\mathrm{on}} + \chi_{jt}^{\mathrm{su}} \leq 1,$$

and similarly by (12f), $u_{jt}^{\mathrm{on}} + \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{sdpc}}} u_{jt'}^{\mathrm{sd}} \leq 1$, so constraints (6) are satisfied. By definition of $\overline{\mathcal{Y}}_{jt}^{\mathrm{uc}}$, the solution satisfies (2a), (2d), (2e), and (3c). Next, by (12a) and construction of $u_{jt'}^{\mathrm{su}}$, we have

$$0 \leq p_{jt}^{\mathrm{su}} \leq \chi_{jt}^{\mathrm{su}} \max_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} = u_{j\overline{t}^{\mathrm{su}}}^{\mathrm{su}} p_{jt\overline{t}^{\mathrm{su}}}^{\mathrm{supc}} = \sum_{t' \in \mathcal{T}_{jt}^{\mathrm{supc}}} p_{jtt'}^{\mathrm{supc}} u_{jt'}^{\mathrm{su}},$$

51

and constraint (7a) is satisfied. Similarly, (12b) and the construction of $u_{jt'}^{\mathrm{sd}}$ imply (7b). This establishes, with certificate $(u_j^{\mathrm{su}}, u_j^{\mathrm{sd}})$, that $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$. $\qquad \square$

Proof of Theorem 3

Proof. We first demonstrate that the subproblems $\overline{Z}_t^{\mathrm{T}}$, $\overline{Z}_i^{\mathrm{I}}$, and Z_{jt}^{SH} are feasible at every iteration. Under Assumption 1, there is some feasible solution $(p, q, u, v, \Delta, \theta, \tau, \phi)$ for (SCUC-ACOPF). By definition, $(p_j, q_j, u_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$, $(p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$, and $u_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. By Lemma 1, $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$, and $u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. By Proposition 6, there are some (μ_t, ω_t) such that $(p_t, q_t, \mu_t, \omega_t) \in \mathcal{X}_t^{\mathrm{soc}}$ for all $t \in \mathcal{T}$. By Proposition 7, there is some χ such that $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}, \chi_{jt}) \in \overline{\mathcal{Y}}_{jt}^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$. Therefore, the subproblems $\overline{Z}_t^{\mathrm{T}}$, $\overline{Z}_i^{\mathrm{I}}$, and Z_{jt}^{SH} are always feasible.

Next, we demonstrate that the subproblems H_j and Z_t^{F} are feasible at steps 10, 11, and 13 of the algorithm. The iterate $(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})})$ is feasible for the bus-level subproblems $\overline{Z}_i^{\mathrm{I}}$, so $(\overline{p}_j^{(\overline{\tau})}, \overline{q}_j^{(\overline{\tau})}, \overline{u}_j^{(\overline{\tau})}) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$. Then, Proposition 3 gives that problems $H_j(\overline{p}_j^{(\overline{\tau})}, \overline{u}_j^{(\overline{\tau})})$ are feasible for all $j \in \mathcal{J}^{\mathrm{sd}}$ and the solutions δ are nonnegative. By Propositions 1 and 2, there is some feasible solution $(p, q, u, v, \Delta, \theta, \tau, \phi)$ for (SCUC-ACOPF) with $u = \overline{u}^{(\overline{\tau})}$ and $p^{\mathrm{tot}} = \overline{p}^{\mathrm{tot}(\overline{\tau})}$. Note that u^{on} and p^{tot} are copied variables, and u^{su} and u^{sd} are defined implicitly by u^{on} from constraints (1a)–(1b) and (1d). This solution satisfies constraints (2)–(3) and (5), and thus $(p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$. Further, as δ is nonnegative, the solution is trivially feasible for (9). Therefore, this solution is feasible for problems $Z_t^{\mathrm{F}}(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})}, \emptyset; \delta)$ and $Z_t^{\mathrm{F}}(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})}, \overline{\mathcal{N}}; \delta)$. This establishes the validity of the algorithm.

Finally, consider the output (p, q, u). This solution satisfies constraints (2)–(3). As $u = \overline{u}^{(\overline{\tau})}$ and $(\overline{p}^{(\overline{\tau})}, \overline{q}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})}) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$, it holds that u satisfies constraints (1). Additionally, $(p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$, so there is some $(v, \Delta, \theta, \tau, \phi)$ such that the solution satisfies (5a)–(5k). As $u = \overline{u}^{(\overline{\tau})}$ and $\overline{u}_{jt}^{\mathrm{sh}(\overline{\tau})}$ is an integer on $[u_j^{\mathrm{sh,min}}, u_j^{\mathrm{sh,max}}]$ by definition of the subproblems Z_{jt}^{SH}, the solution satisfies (5l). By Proposition 4, as (p, u) satisfies constraint (9) and δ is feasible for $H_j(\overline{p}^{(\overline{\tau})}, \overline{u}^{(\overline{\tau})}) = H_j(\overline{p}^{(\overline{\tau})}, u)$, the solution is feasible for constraint (4). Therefore, the solution $(p, q, u, v, \Delta, \theta, \tau, \phi)$ is feasible for (SCUC-ACOPF). $\qquad \square$

Proof of Theorem 4

Proof. Consider any feasible solution $(p, q, u, v, \Delta, \theta, \tau, \phi)$ for (SCUC-ACOPF). By definition, $(p_j, q_j, u_j) \in \mathcal{X}_j^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$, $(p_t, q_t, u_t^{\mathrm{sh}}) \in \mathcal{X}_t^{\mathrm{ac}}$ for all $t \in \mathcal{T}$, and $u_{jt}^{\mathrm{sh}} \in \mathcal{X}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. By Lemma 1, $(p_{jt}, q_{jt}, u_{jt}^{\mathrm{on}}) \in \mathcal{Y}_{jt}^{\mathrm{uc}}$ for all $j \in \mathcal{J}^{\mathrm{sd}}$ and $t \in \mathcal{T}$, and $u_{jt}^{\mathrm{sh}} \in \mathcal{Y}_{jt}^{\mathrm{sh}}$ for all $j \in \mathcal{J}^{\mathrm{sh}}$ and $t \in \mathcal{T}$. Further, by Proposition 6, there are some (μ_t, ω_t) such that $(p_t, q_t, \mu_t, \omega_t) \in \mathcal{X}_t^{\mathrm{soc}}$ for all $t \in \mathcal{T}$. Thus, we have a corresponding feasible solution $(p_t, q_t, u_t, \mu_t, \omega_t)$ for each problem of (UB).

We now establish fundamental properties of the objective functions. For any p, as $c^{\mathrm{e}} \geq 0$, it holds that $C_j^{\mathrm{e}}(p) \geq 0$ for all $j \in \mathcal{J}^{\mathrm{sd}}$. Similarly, for any (p, q), as $c^{\mathrm{s}} \geq 0$ and $d_t \geq 0$, it holds that $C_{kt}^{\mathrm{ctg}}(p, q) \geq 0$ for all $k \in \mathcal{K}$ and $t \in \mathcal{T}$, and therefore $C_t^{\mathrm{ctg}}(p, q) \geq 0$ for all $t \in \mathcal{T}$. Finally, for u satisfying (1), we have

for all $j \in \mathcal{J}^{\mathrm{sd}}$ that

$$
\begin{aligned}
C_j^{\mathrm{uc}}(u) - \sum_{t \in \mathcal{T}} c_{jt}^{\mathrm{on}} u_{jt}^{\mathrm{on}} &= \sum_{t \in \mathcal{T}} \left(c_j^{\mathrm{su}} u_{jt}^{\mathrm{su}} + c_j^{\mathrm{sd}} u_{jt}^{\mathrm{sd}} - \sum_{\substack{t' \in \mathcal{T} \\ t' < t}} c_{jtt'}^{\mathrm{dd}} [u_{jt}^{\mathrm{su}} + u_{jt'}^{\mathrm{on}} - 1]^+ \right) \\
&\geq \sum_{t \in \mathcal{T}} \left(c_j^{\mathrm{su}} u_{jt}^{\mathrm{su}} + c_j^{\mathrm{sd}} u_{jt}^{\mathrm{sd}} - \sum_{\substack{t' \in \mathcal{T} \\ t' < t}} c_{jtt'}^{\mathrm{dd}} u_{jt}^{\mathrm{su}} \right) \geq \sum_{t \in \mathcal{T}} c_j^{\mathrm{sd}} u_{jt}^{\mathrm{sd}} \geq 0,
\end{aligned}
$$

where the first inequality follows from $[u_{jt}^{\mathrm{su}} + u_{jt'}^{\mathrm{on}} - 1]^+ \leq [u_{jt}^{\mathrm{su}}]^+ = u_{jt}^{\mathrm{su}}$ and nonnegativity of $c_{jtt'}^{\mathrm{dd}}$, the second inequality from Assumption 1 ($c_j^{\mathrm{su}} \geq \sum_{\substack{t' \in \mathcal{T} \\ t' < t}} c_{jtt'}^{\mathrm{dd}}$) and $u_{jt}^{\mathrm{su}} \geq 0$, and the third inequality from nonnegativity of c_j^{sd} and u_{jt}^{sd}. Considering the objective function for (SCUC-ACOPF),

$$
\begin{aligned}
& \sum_{t \in \mathcal{T}} R_t^{\mathrm{T}}(p, q) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} R_j^{\mathrm{J}}(p, u) \\
&= \sum_{t \in \mathcal{T}} \left(\sum_{j \in \mathcal{J}^{\mathrm{cs}}} R_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{pr}}} C_{jt}^{\mathrm{pow}}(p) - \sum_{j \in \mathcal{J}^{\mathrm{ac}}} C_{jt}^{\mathrm{ac}}(p, q) - \sum_{i \in \mathcal{I}} C_{it}^{\mathrm{bal}}(p, q) - C_t^{\mathrm{res}}(p, q) - C_t^{\mathrm{ctg}}(p, q) \right) \\
&\quad - \sum_{j \in \mathcal{J}^{\mathrm{sd}}} \left(C_j^{\mathrm{uc}}(u) + C_j^{\mathrm{e}}(p) \right) \\
&= \sum_{t \in \mathcal{T}} \left(C_t^{\mathrm{UB}}(p, q, u) + \sum_{j \in \mathcal{J}^{\mathrm{sd}}} c_{jt}^{\mathrm{on}} u_{jt}^{\mathrm{on}} - C_t^{\mathrm{ctg}}(p, q) \right) - \sum_{j \in \mathcal{J}^{\mathrm{sd}}} \left(C_j^{\mathrm{uc}}(u) + C_j^{\mathrm{e}}(p) \right) \leq \sum_{t \in \mathcal{T}} C_t^{\mathrm{UB}}(p, q, u),
\end{aligned}
$$

where the inequality follows from the previously observed properties of the objective functions. This implies that the feasible solution (p, q, u, μ, ω) has an objective value in (UB) at least as large as the objective value for $(p, q, u, v, \Delta, \theta, \tau, \phi)$ in (SCUC-ACOPF). Because we are maximizing and this property applies for any feasible solution to (SCUC-ACOPF), the optimal value of (UB) must be at least as high as the optimal value of (SCUC-ACOPF). $\qquad \square$

References

Hager WW (1989) Updating the inverse of a matrix. *SIAM Review* 31(2):221–239.

Holzer J, Coffrin C, DeMarco C, Duthu R, Elbert S, Eldridge B, Elgindy T, Garcia M, Greene S, Guo N, Hale E, Lesieutre B, Mak T, McMillan C, Mittelmann H, Oh H, O'Neill R, Overbye T, Plamintier B, Paker R, Safdarian F, Tbaileh A, Van Hentenryck P, Veeramany A, Wert J (2024a) Grid Optimization Competition Challenge 3 problem formulation. URL http://dx.doi.org/10.25984/2426334.

Holzer JT, Chen Y, Wu Z, Pan F, Veeramany A (2024b) Fast simultaneous feasibility test for security constrained unit commitment. *IEEE Transactions on Power Systems* 39(1):1068–1078.

Rockafellar RT, Wets RJB (2009) *Variational Analysis* (Springer).