

Distributed AC Optimal Power Flow for Generic Integrated Transmission-Distribution Systems

Xinliang Dai[†], Junyi Zhai[†], Yuning Jiang*, Colin N. Jones, Veit Hagenmeyer

Abstract

Coordination of transmission and distribution power systems is increasingly critical in the context of the ongoing energy transition. However, traditional centralized energy management faces challenges related to privacy and/or sovereignty concerns, leading to growing research interests in distributed approaches. Nevertheless, solving distributed AC optimal power flow (OPF) problems encounters difficulties due to their nonlinearity and nonconvexity, making it challenging for state-of-the-art distributed approaches. To solve this issue, the present paper focuses on investigating the distributed AC OPF problem of generic integrated transmission-distribution (ITD) systems, considering complex grid topology, by employing a new variant of Augmented Lagrangian based Alternating Direction Inexact Newton method (ALADIN). In contrast to the standard ALADIN, we introduce a second-order correction into ALADIN to enhance its numerical robustness and properly convexify distribution subproblems within the ALADIN framework for computing efficiency. Moreover, a rigorous proof shows that the locally quadratic convergence rate can be preserved for solving the resulting distributed nonconvex problems. Extensive numerical simulations with varying problem sizes and grid topologies demonstrate the effectiveness of the proposed algorithm, outperforming state-of-the-art approaches in terms of numerical robustness, convergence speed, and scalability.

Index Terms

Generic integrated transmission-distribution (ITD) systems, AC optimal power flow (OPF), distributed nonconvex optimization, second-order correction, Augmented Lagrangian based Alternating Direction Inexact Newton method (ALADIN)

[†] equal contribution, *corresponding (yuning.jiang@ieee.org)

X. Dai and V. Hagenmeyer are with Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Germany. J. Zhai is with the College of New Energy, China University of Petroleum (East China), China. Y. Jiang and C. N. Jones are with Automatic Control Laboratory, EPFL, Switzerland.

I. INTRODUCTION

The increasing integration of distributed energy resources has intensified the interaction between transmission and distribution grids, challenging the current paradigm of managing separately the transmission and distribution networks [1]–[3]. Despite their physical connections, the transmission and distribution systems are typically operated separately by transmission system operators (TSOs) and distribution system operators (DSOs). Coordinating integrated transmission-distribution (ITD) systems is becoming essential for effective power system operation. However, system operators do not prefer centralized frameworks or are even forbidden by the respective regulation, as they require sharing detailed grid data with a centralized entity. In the United Kingdom, such centralized operation between TSOs and DSOs becomes nearly impossible under the deregulated electricity market environment, while in Germany, new legislation and the ongoing rapid energy transition toward more renewable energies force the German TSO to focus on new vertical cooperation with the numerous DSOs [4]. Hence, distributed operation frameworks serve as an efficient alternative for the coordination of ITD systems [2], enabling TSOs and DSOs to operate independently and collaborate effectively by sharing limited information with a subset of other operators [1]. This kind of distributed framework can maintain data privacy and decision-making independence, leading to significant research in distributed operation problems for ITD systems, including distributed power flow [5], [6], distributed economic dispatch [2], [7], distributed coordinated restoration [8], [9], distributed optimal reactive power flow [10], distributed OPF [11]–[13], etc.

For the purpose of economic efficiency, the present paper focuses on the AC OPF problem of generic ITD systems. However, the AC OPF problem is generally NP-Hard [14], even for radial power grids [15]. The complexity lies in the nonconvexity of power flow equations, giving rise to a nonconvex feasible set of AC OPF [16]. To balance accuracy and computational tractability, convex relaxations of power flow equations are attracting considerable attention for both the bus injection model (BIM) [17], [18] and the branch flow model (BFM) [19]–[22]. These relaxations are exact for radial networks since there is a bijection between their feasible sets [23]. Moreover, by avoiding the subtraction of variables with similar values, the second-order cone (SOC) relaxation in the BFM is superior to the others in the aspects of numerical reliability and computing time [24], [25].

Previous studies on distributed AC OPF focused on the BIM for meshed power grids, utiliz-

ing distributed approaches such as Optimality Condition Decomposition (OCD) [26], Auxiliary Problem Principle (APP) [27], diagonal quadratic approximation (DQA) [12], and Alternating Direction Method of Multipliers (ADMM) [28]–[31]. However, these approaches lack guaranteed convergence in general. Exceptions include a two-level variant of ADMM [32], a l_1 proximal surrogate Lagrangian method [3] and a heterogeneous decomposition algorithm [11]. Nonetheless, these are first-order algorithms and exhibit slow numerical convergence with modest accuracy.

The main challenge in distributed approaches, similar to centralized approaches, lies in the nonconvexity due to the power flow equations. Addressing this issue, [13] proposed a new distributed AC OPF of ITD system. The distribution subproblems are formulated as an AC OPF in the BFM with the SOC relaxation, while only the transmission subproblem is formulated in the BIM. Then, a two-layer Distribution-Cost-Correction (DCC) framework is proposed to solve transmission and distribution subproblems in a sequential manner: the conic optimization subproblems of the distribution grids are solved in the lower layer, and then the nonconvex subproblem of the transmission grid is solved in the upper layer regarding distribution cost by using quadratic approximations of distribution subproblems. This approach ensures convergence but is limited to a particular network topology where multiple radial distribution grids are connected to only one transmission grid in a star-shaped configuration. This two-layer DCC algorithm fails to handle the more generic ITD systems with multiple transmission grids, meshed distribution grids, or meshed topology of subsystems.

In contrast to the problem-specific DCC framework, Augmented Lagrangian based Alternating Direction Inexact Newton method (ALADIN) is proposed for generic nonconvex distributed problems in [33]. ALADIN can be viewed as a second-order version of ADMM and provide several advantages over the aforementioned distributed approaches. It provides a local convergence guarantee with a locally quadratic convergence rate for generic distributed nonconvex optimization problems when suitable Hessian approximations are employed. Additionally, ALADIN can achieve a global convergence by implementing the globalization strategy [33, Alg. 3]. Inspired by Sequential Quadratic Programming (SQP), an equality-constrained quadratic program (QP) is solved in a coupled step of ALADIN. The coupled QP is constructed based on curvature information derived from local subproblems provided by corresponding local computing agents in the parallelizable decoupled step. This framework eliminates the need for exchanging the original grid data, preserving information privacy. Recently, ALADIN has been successfully applied to solve the AC OPF of medium-scale transmission systems [34] and of AC/DC hybrid

systems [35], [36], demonstrating the potential of ALADIN for handling distributed problems with heterogeneous models. However, classic ALADIN for solving distributed nonconvex AC OPF problems may face numerical issues when the problem size goes large—these studies have been limited by problem sizes (less than 300 buses) and the number of subgrids (less than half a dozen).

In summary, there are several challenges in effectively solving the distributed AC OPF of generic ITD systems:

- 1) Most existing distributed algorithms [11], [12], [26]–[32] are first-order algorithms, lacking convergence guarantee or exhibiting slow numerical convergence with modest accuracy for solving the nonconvex AC OPF problem.
- 2) Although the recent work [13] successfully introduced SOC relaxation to distribution subproblems such that the proposed two-layer DCC has convergence guarantee and the computational burden of distribution subproblems is significantly reduced, this approach is strictly limited to star-shaped ITD systems, where multiple radial distribution grids are connected to one single transmission grid.
- 3) ALADIN [33], designed for generic distributed nonconvex optimization problems, has shown potential for handling AC OPF. However, by employing the active-set method, the standard ALADIN (ALADIN-STD) suffers from the following numerical issue when the problem size goes large: The number of possible active sets grows exponentially with the number of inequalities, leading to the combinatorial difficulty of active-set [37, Ch. 15.2]. This becomes especially critical when introducing SOC relaxation, as the resulting inequalities are weakly active, making it challenging to estimate the active sets accurately. Moreover, similar to SQP, the linearization error resulting from linearizing the active constraints in ALADIN-STD can lead to poor initial guesses for subsequent iterations. This issue is particularly significant for augmented-Lagrangian-type algorithms in preliminary iterations and can potentially undermine the numerical robustness of ALADIN.

The present paper investigates the distributed AC OPF of generic ITD systems and proposes a new variant of ALADIN with a second-order correction method to break through these challenges, including the convergence guarantees, the computing time, the solution accuracy and the grid topology flexibility for distributed nonconvex AC OPF problems. The main contributions are as follows:

- 1) Building upon the work of [13], we present a modified formulation of distributed AC OPF, tailored to handle the more generic ITD systems and employ a new variant of ALADIN with convergence guarantee to solve it. Compared to the state-of-art two-layer DCC [13], the proposed distributed framework can handle more generic and complex ITD systems, including those with multiple transmission grids, meshed distribution grids, or meshed topology of subsystems, offering greater flexibility regarding grid topology.
- 2) In the new variant of ALADIN for solving the AC OPF of generic ITD systems, the SOC relaxation is only implemented in a specific step of ALADIN. The proper implementation of the SOC relaxation reduces the computational burden while avoiding the combinatorial difficulty of active-set [37, Ch. 15.2] arising from the additional inequalities. Furthermore, we introduce a second-order correction method [38] [37, Ch. 18.3] to improve its numerical robustness. The proposed new full-step ALADIN with second-order correction (ALADIN-COR) compensates for linearization error, enabling the ALADIN-type framework to be effectively applied to real-world large-scale power systems. Then, we provide a rigorous proof demonstrating that the local quadratic convergence rate of ALADIN-COR can still be preserved with the additional compensation step.
- 3) Numerical investigations are conducted to compare the performance of the proposed ALADIN-COR with the state-of-art distributed nonconvex optimization algorithms, i.e., the two-layer DCC [13] and the ALADIN-STD [34], [35], demonstrating that the proposed approach has high scalability potential and surpasses the others in terms of convergence rate, computing time, and solution accuracy.

The rest of this paper is organized as follows: Section II presents the system model and problem formulation. Section III introduces the proposed distributed algorithm with implementation details. Section IV elaborates on numerical results. Section V concludes this paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

This section presents the AC OPF of a generic ITD systems. Then, the entire problem is reformulated as a generic distributed form with affine consensus constraint.

A. System Model of a Generic ITD System

We describe a generic ITD systems by a tuple $\mathcal{S} = (\mathcal{R}, \mathcal{N}, \mathcal{L})$. Thereby, $\mathcal{R} = \mathcal{R}_T \cup \mathcal{R}_D$ represents the set of all regions, $\mathcal{R}_T = \{T_1, \dots, T_m\}$ the set of m transmission grids and $\mathcal{R}_D =$

$\{D_1, \dots, D_n\}$ the set of n distribution grids, \mathcal{N} denotes the set of all buses, \mathcal{L} the set of all branches, $\mathcal{L}^{\text{tie}} \in \mathcal{L}$ represents the set of all tie-lines between neighboring regions. Regarding the distributed approach, instead of cutting the tie-lines [34], we follow the idea of sharing components [39] to ensure physical consistency, i.e., share the components of transmission grid with the neighboring distribution grids. In a specific grid $\ell \in \mathcal{R}$, $\mathcal{N}_\ell = \mathcal{N}_\ell^{\text{core}} \cup \mathcal{N}_\ell^{\text{copy}}$ denotes the set of all buses in the region ℓ , $\mathcal{N}_\ell^{\text{core}}$ the set of its own buses, and $\mathcal{N}_\ell^{\text{copy}}$ the set of buses shared by its neighboring regions. Consequently, the shared component, i.e., the tie-line between neighboring regions, is also included in the set of all branches in the region ℓ , i.e., $\mathcal{L}_\ell^{\text{tie}} \subseteq \mathcal{L}_\ell$.

Consider a simple example of ITD system consisting of 6 buses within 2 regions, as illustrated in Fig. 1. Regarding TSO-DSO connections, as shown in Fig. 1(b)(c), the distribution grid D_1 encompasses core buses $\{4, 5, 6\}$ and an additional copy bus $\{3\}$ shared by the neighboring region. On the other hand, the transmission grid T has no copy bus. The sharing components involve active and reactive power flow from bus 3 to bus 4 along the tie-line $(3, 4)$ and the voltage magnitude at bus 3. Consequently, each region $\ell \in \mathcal{R}$ can establish a self-contained AC OPF subproblem. Note that the configuration for TSO-DSO connections can also be applied to DSO-DSO connections.

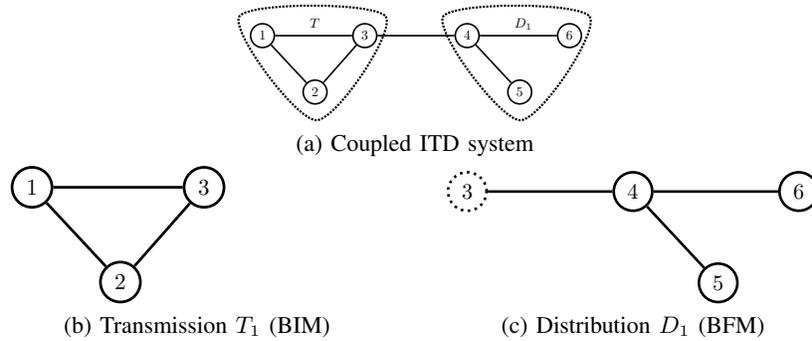


Fig. 1: Decomposition by sharing components between transmission and distribution grids

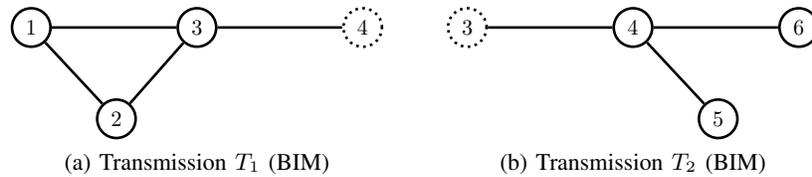


Fig. 2: Decomposition by sharing components between two transmission grids

In terms of TSO-TSO connections, as shown in Fig. 2(a)(b), the sharing components involve voltage angle and magnitude at the buses $\{3, 4\}$. Different from Fig. 1(b)(c), both transmission grids include an additional copy bus sharing by neighboring regions. By introducing additional consensus constraints to ensure the physical consistency of the shared components, the AC OPF problem of a generic ITD systems can be established in a distributed format.

B. Objective

Typically, the objective of the AC OPF for a generic ITD systems is to minimize the total operation cost of the generating units, as follows.

$$\underset{p^g}{\text{minimize}} \sum_{i \in \mathcal{N}} \left\{ a_i (p_i^g)^2 + b_i p_i^g + c_i \right\}, \quad (1)$$

where a_i , b_i , and c_i denote the cost coefficients of generator at bus i , p_i^g denotes the active power injection at bus i .

C. Constraints of Transmission Grid

The transmission part of the ITD systems is formulated based on the BIM, where the complex bus voltage V_i is represented in polar coordinates.

1) *Nodal power balance*: For the core bus in the transmission grid, the nodal power balance is expressed as

$$p_i^g - p_i^l = v_i \sum_{j \in \mathcal{N}_T} v_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}), \quad (2a)$$

$$q_i^g - q_i^l = v_i \sum_{j \in \mathcal{N}_T} v_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \quad (2b)$$

for all $i \in \mathcal{N}_T$. Thereby v_i denotes the voltage magnitude of bus i , θ_{ij} the phase angle difference between buses i and j , q_i^g the reactive power of generator at bus i , p_i^l and q_i^l denote the active and reactive load at the bus i , G_{ij} and B_{ij} are the real and reactive components of the bus admittance matrix. If bus i is not a generator bus, then $p_i^g = q_i^g = 0$, same as the load bus. If there is a tie-line $(i, k) \in \mathcal{L}^{\text{tie}}$ connected to the bus i , then p_{ik}^{core} and q_{ik}^{core} are added to the right-hand side of both the equations (2a)(2b) respectively, as power injected from the bus k in neighboring distribution grid.

2) *Branch flow limit*: Apparent power limits are added for all branches, including the tie-lines with neighboring regions.

$$p_{ij} = v_i^2 g_{ij} - v_i v_j (g_{ij} \cos \theta_{ij} + b_{ij} \sin \theta_{ij}), \quad (3a)$$

$$q_{ij} = -v_i^2 b_{ij} - v_i v_j (g_{ij} \sin \theta_{ij} - b_{ij} \cos \theta_{ij}), \quad (3b)$$

$$\bar{s}_{ij}^2 \geq p_{ij}^2 + q_{ij}^2 \quad (3c)$$

for all $(i, j) \in \mathcal{L}_T$. Thereby p_{ij} and q_{ij} represent the active and reactive power flow of branch (i, j) , g_{ij} and b_{ij} denote the conductance and the susceptance of branch (i, j) , \bar{s}_{ij} the apparent power limit of branch (i, j) .

3) *Bounds on state variables*: Except for voltage phase angle, box constraints are added on nodal voltage magnitude and active and reactive power of generators for all buses.

$$\underline{v}_i \leq v_i \leq \bar{v}_i, \quad \underline{p}_i^g \leq p_i^g \leq \bar{q}_i^g, \quad \underline{q}_i^g \leq q_i^g \leq \bar{q}_i^g \quad (4)$$

for all bus $i \in \mathcal{N}_T$, where \underline{v}_i , \bar{v}_i , \underline{p}_i^g , \bar{q}_i^g , \underline{q}_i^g , and \bar{q}_i^g denote the upper and lower bounds for the corresponding state variables.

D. Constraints of Distribution Grid

Regarding the radial distribution network in the ITD systems, the AC OPF problem of distribution part can be formulated as a QCQP based on BFM. BFM can be viewed as an angle relaxation of BIM and is exact for radial distribution grid [22]. We use the squares of the nodal voltage magnitude and the branch current magnitude as state variables in a specific distribution grid $\ell = D_m$, i.e., $u_i = |v_i|^2$ for all bus $i \in \mathcal{N}_{D_m}$ and $l_{ij} = |I_{ij}|^2$ for all branches $(i, j) \in \mathcal{L}_{D_m}$.

1) *Branch current and its limit*: The branch current magnitude can be expressed as

$$l_{ij} = \frac{p_{ij}^2 + q_{ij}^2}{u_i}, \quad (5)$$

and the current limit as

$$l_{ij} \leq \bar{l}_{ij}. \quad (6)$$

where \bar{l}_{ij} denotes the current upper bound of the branch (i, j) .

2) *Nodal power balance*: For the core bus in the distribution grid, the nodal power balance is expressed as

$$p_j^g - p_j^l = \sum_{k \in \mathcal{N}_{Dm}} p_{jk} - \sum_{i \in \mathcal{N}_{Dm}} (p_{ij} - r_{ij}l_{ij}), \quad (7a)$$

$$q_j^g - q_j^l = \sum_{k \in \mathcal{N}_{Dm}} q_{jk} - \sum_{i \in \mathcal{N}_{Dm}} (q_{ij} - x_{ij}l_{ij}) \quad (7b)$$

for all $j \in \mathcal{N}_{Dm}^{\text{core}}$. Thereby r_{ij} and x_{ij} denote the resistance and reactance of the branch (i, j) .

3) *Nodal voltage equation*: The relationship of the squares of nodal voltage magnitude is expressed as

$$u_j = u_i - 2(r_{ij}p_{ij} + x_{ij}q_{ij}) + (r_{ij}^2 + x_{ij}^2)l_{ij} \quad (8)$$

for all $(i, j) \in \mathcal{L}_{Dm}$.

4) *Bounds on state variables*: The upper and lower bounds on state variables are similar to (4). The difference is that the voltage bound is replaced by the box constraint on the square of nodal voltage magnitude.

$$\underline{u}_i \leq u_i \leq \bar{u}_i, \quad \underline{p}_i^g \leq p_i^g \leq \bar{q}_i^g, \quad \underline{q}_i^g \leq q_i^g \leq \bar{q}_i^g \quad (9)$$

for all bus $i \in \mathcal{N}_{Dm}^{\text{core}}$, where \bar{u}_i and \underline{u}_i denote the upper and lower bounds on the square of nodal voltage magnitude.

Note that the feasible set of the BFM for the distribution grid is still nonconvex due to the quadratic equality constraint (5). It can be further relaxed to a conic constraint, which will be discussed in the following section. If the distribution part of ITD systems is in a meshed form, the BIM can be adopted similarly with the transmission grid.

E. Consensus Constraint

By introducing the concept of sharing components, additional constraints on coupling variables are required to ensure the consistency between the core and copy bus. Considering the example of TSO-DSO connection shown in Fig. 1, the coupling constraint between two regions can be written as

$$u_3^{\text{copy}} = u_3^{\text{core}}, \quad p_{34}^{\text{copy}} = p_{34}^{\text{core}}, \quad q_{34}^{\text{copy}} = q_{34}^{\text{core}}, \quad (10)$$

while the coupling constraint for TSO-TSO connection (Fig. 2) can be written as

$$v_3^{\text{copy}} = v_3^{\text{core}}, v_4^{\text{copy}} = v_4^{\text{core}}, \theta_3^{\text{copy}} = \theta_3^{\text{core}}, \theta_4^{\text{copy}} = \theta_4^{\text{core}}. \quad (11)$$

Conclusively, in a specific ITD systems, the coupling constraints are linear and can be integrated into an affinely coupled form

$$\sum_{\ell \in \mathcal{R}} A_\ell x_\ell = Ax = b \quad (12)$$

with consensus matrix

$$A = (A_{T_1}, \dots, A_{T_m}, A_{D_1}, \dots, A_{D_n})$$

and the state variables

$$x = (x_{T_1}, \dots, x_{T_m}, x_{D_1}, \dots, x_{D_n}).$$

Note that b is a zero vector in the proposed model, and A has full row rank.

F. Distributed Formulation

The AC OPF of a generic ITD systems can be formulated as an affinely coupled separable form tailored to distributed optimization. The objective for each region $\ell \in \mathcal{R}$ can be written as

$$f_\ell(x_\ell) = \sum_{i \in \mathcal{N}_\ell^{\text{core}}} \left\{ a_i (p_i^g)^2 + b_i p_i^g + c_i \right\}. \quad (13)$$

Based on the discussion given above in the present section II, the AC OPF of a generic ITD systems can be formulated in the standard affinely coupled distributed form

$$\min_x f(x) := \sum_{\ell \in \mathcal{R}} f_\ell(x_\ell) \quad (14a)$$

$$\text{s.t.} \quad \sum_{\ell \in \mathcal{R}} A_\ell x_\ell = b \quad | \quad \lambda \quad (14b)$$

$$h_\ell(x_\ell) \leq 0 \quad | \quad \kappa_\ell, \ell \in \mathcal{R} \quad (14c)$$

where (14c) summarizes all local constraints (2)—(9) for local systems, and λ, κ_ℓ denote the dual variables (Lagrangian multipliers) of constraints (14b) and (14c), respectively.

III. DISTRIBUTED OPTIMIZATION FRAMEWORK

This section presents a novel variant of ALADIN for solving the distributed nonlinear and nonconvex AC OPF problem of generic ITD systems with implementation details.

A. ALADIN with Second-Order Correction

ALADIN originally proposed in [33] was developed for dealing with generic distributed non-convex optimization problems. Similar to the existing state-of-art ADMM, ALADIN is also an alternating direction-based approach. Nevertheless, the main difference between ALADIN and ADMM lies in the coordinator. ADMM updates dual variables based on first-order information in the coordinator. Additionally, ALADIN acquires both the first-order and the second-order information of local nonlinear programming (NLP) and solves an approximated QP problem to update the primal and dual together, and achieves locally quadratic convergence. However, the local convergence of the ALADIN-STD is sensitive to the linearization of constraints if the initial guess is not close enough to local minimizers or the problem size becomes large. In order to mitigate this issue, we introduce the second-order correction into ALADIN framework for compensation of the error that occurred by linearization.

Algorithm 1 outlines the ALADIN-COR for solving (14). In step 1, the original decoupled subproblems are convexified to the decoupled NLP (16) by introducing the augmented Lagrangian method, where ρ is the penalty parameter and $\Sigma_\ell \in \mathbb{R}^{x_\ell \times x_\ell}$ is the scaling matrix for the proximal term. A practical strategy to update ρ for distributed AC OPF can be found in [34].

In Step 2, all sensitivities information, i.e., Jacobian J_ℓ of active constraints h_ℓ^{act} , gradient of local objective g_ℓ , and Hessian approximation H_ℓ are computed. Thereby, the active constraint $h_\ell^{\text{act}}(x_\ell)$ at the current iteration includes the inequality constraint $[h_\ell(x_\ell)]_i$ for all

$$i \in \mathcal{S}_\ell = \{i \mid [h_\ell(x_\ell)]_i = 0\}. \quad (22)$$

Note that both steps can be executed in parallel.

After the parallelizable steps, the termination condition is checked by the coordinator. The ALADIN algorithm will be terminated if the primal and the dual residuals are smaller than the predefined tolerance ϵ

$$\|Ax - b\|_2 \leq \epsilon \text{ and } \|\Sigma(x - z)\|_2 \leq \epsilon, \quad (23)$$

Algorithm 1 Full-Step ALADIN-COR

Initialization: define L_1 penalty function

$$\Phi(x) = \sum_{\ell \in \mathcal{R}} f_\ell(x_\ell) + \zeta \left\| \sum_{\ell \in \mathcal{R}} A_\ell x_\ell - b \right\|_1 + \xi \cdot \psi(h(x)), \quad (15)$$

 with $\psi(h) = \sum_i \max\{0, [h]_i\}$ and $\zeta, \xi > 0$.

Input: $z, \lambda, \rho > 0, \mu > 0$ and scaling symmetric matrices $\Sigma_\ell \succ 0$
Repeat:

 1) solve the following decoupled NLPs for all $\ell \in \mathcal{R}$

$$\min_{x_\ell} f_\ell(x_\ell) + \lambda^\top A_\ell x_\ell + \frac{\rho}{2} \|x_\ell - z_\ell\|_{\Sigma_\ell}^2 \quad (16a)$$

$$\text{s.t. } h_\ell(x_\ell) \leq 0 \quad | \quad \kappa_\ell \quad (16b)$$

 2) compute the Jacobian matrix J_ℓ of active constraints h_ℓ^{act} at the local solution x_ℓ by

$$J_\ell = \nabla h_\ell^{\text{act}}(x_\ell), \quad (17)$$

 and gradient $g_\ell = \nabla f_\ell(x_\ell)$, choose Hessian approximation

$$H_\ell \approx \nabla^2 \{f_\ell(x_\ell) + \kappa_\ell^\top h_\ell(x_\ell)\} \succ 0, \quad (18)$$

 3) terminate if $\|Ax - b\|_2 \leq \epsilon$ and $\|\Sigma(x - z)\|_2 \leq \epsilon$ are satisfied.

 4) obtain $(z^{\text{QP}} = x + p^{\text{QP}}, \lambda^{\text{QP}})$ by solving coupled QP

$$\min_{p^{\text{QP}}, s} \sum_{\ell \in \mathcal{R}} \left\{ \frac{1}{2} (p_\ell^{\text{QP}})^\top H_\ell p_\ell^{\text{QP}} + g_\ell^\top p_\ell^{\text{QP}} \right\} + \lambda^\top s + \frac{\mu}{2} \|s\|_2^2 \quad (19a)$$

$$\text{s.t. } \sum_{\ell \in \mathcal{R}} A_\ell (x_\ell + p_\ell^{\text{QP}}) = b + s \quad | \quad \lambda^{\text{QP}} \quad (19b)$$

$$J_\ell p_\ell^{\text{QP}} = 0, \quad \ell \in \mathcal{R} \quad (19c)$$

 5) if the value of Φ increases due to the intolerant violation of active constraint h^{act} at the new iterate z^{QP} , compute $(z^{\text{SOC}} = x + p^{\text{SOC}}, \lambda^{\text{SOC}})$ by

$$\begin{bmatrix} p^{\text{SOC}} \\ \lambda^{\text{SOC}} \end{bmatrix} = \begin{bmatrix} p^{\text{QP}} \\ \lambda^{\text{QP}} \end{bmatrix} - \begin{bmatrix} I \\ \mu A \end{bmatrix} \cdot M J^\top \cdot (J M J^\top)^{-1} \cdot r \quad (20)$$

 with $r = h^{\text{act}}(x + p^{\text{QP}})$ and $M = (H + \mu A^\top A)^{-1}$.

6) update the primal and the dual variables with full step

$$(z^+, \lambda^+) = \begin{cases} (z^{\text{SOC}}, \lambda^{\text{SOC}}), & \text{step 5 executed,} \\ (z^{\text{QP}}, \lambda^{\text{QP}}), & \text{otherwise.} \end{cases} \quad (21)$$

where Σ is a block diagonal matrix consists of scaling matrix Σ_ℓ for all $\ell \in \mathcal{R}$. It indicates that the local solution x_ℓ satisfied the first-order optimality condition of the original problem (14) up to the error of $\mathcal{O}(\epsilon)$, i.e.,

$$\|\nabla \{f_\ell(x_\ell) + \kappa_\ell^\top h_\ell(x_\ell)\} + A_\ell^\top \lambda\|_2 = \mathcal{O}(\epsilon). \quad (24)$$

Remark 1. *Practically, the dual condition (23) is sufficient to ensure the small violation of the condition (24), when the predefined tolerance ϵ is small enough [40].*

In the coupled Step 4, a quadratic approximation of the original problem (14) is established based on the sensitivities at local solution x_ℓ computed in Step 2. An additional slack variable s is introduced to avoid the infeasibility of the approximated problem due to the consensus constraint (19b), and in order to increase numerical robustness. Like the SQP, the active constraints are linearized and summarized in (19c).

Remark 2. *No detailed grid data is required in the coupled Step 4, but the curvature information, including gradient, Jacobian, and approximated Hessian of the local problems (16). Therefore, data privacy wouldn't be violated by applying the new ALADIN-COR algorithm.*

The L_1 penalty function $\Phi(x)$ in Step 5 is used to measure the progress in the coordinator during the iterations. For this, we assume that the positive parameters ζ, ξ are sufficiently large such that Φ is an exact penalty function for the problem (14).

Definition 1 [37]. *A penalty function is exact if a single minimization with respect to x can yield the exact solution of the original constrained optimization problem*

Similar to the standard SQP, the indecisiveness caused by the linearized active constraints (19c) can lead the conventional ALADIN-STD algorithm to zigzag during the preliminary iterations. The issue becomes more critical and complicated when the problem size goes large, or if the initial guess is not close enough to the local minimizer. In order to overcome this deficiency, an additional second-order correction step 5 is executed if the exact penalty function (15) is increasing accompanied by an intolerant violation of the active constraint h^{act} .

Following [37, Ch. 18.3], the linearized active constraint (19c) is replaced by

$$J_\ell p_\ell^{\text{SOC}} + r_\ell = 0, \quad \ell \in \mathcal{R} \quad (25)$$

with a compensation step p_ℓ^{SOC} and a compensation vector r_ℓ computed by

$$r_\ell = h_\ell^{\text{act}}(x_\ell + p_\ell^{\text{QP}}) - J_\ell p_\ell^{\text{QP}} = h_\ell^{\text{act}}(x_\ell + p_\ell^{\text{QP}}). \quad (26)$$

The term $J_\ell p_\ell^{\text{QP}}$ can be neglected because a QP step p_ℓ^{QP} satisfied (19c). The resulting second-order correction subproblem can be written as

$$\begin{aligned} \min_{p^{\text{SOC}}, s} \quad & \sum_{\ell \in \mathcal{R}} \left\{ \frac{1}{2} (p_\ell^{\text{SOC}})^\top H_\ell p_\ell^{\text{SOC}} + g_\ell^\top p_\ell^{\text{SOC}} \right\} + \lambda^\top s + \frac{\mu}{2} \|s\|_2^2 \\ \text{s.t.} \quad & \sum_{\ell \in \mathcal{R}} A_\ell (x_\ell + p^{\text{SOC}}) = b + s \quad | \lambda^{\text{SOC}} \end{aligned} \quad (27a)$$

$$J_\ell p_\ell^{\text{SOC}} + r_\ell = 0, \quad \ell \in \mathcal{R}. \quad (27b)$$

Fortunately, the compensated step $(p^{\text{SOC}}, \lambda^{\text{SOC}})$ can be computed analytically by (20), in which all matrix inverses and the factorization for solving (19) can be reused. More details will be discussed in section III-B2.

Local convergence of the proposed ALADIN-COR algorithm is guaranteed, and its analysis will be provided in the next section, while global convergence can be achieved if the additional globalization strategy [33, Algorithm 3] is implemented.

Remark 3 (Globalization of Algorithm 1). *To enforce the global convergence of Algorithm 1 such that it converges to a local minimizer of (14), the primal-dual iterate (z, λ) is updated by*

$$z^+ = z + \alpha_1(x - z) + \alpha_2 p^{\text{QP}}, \quad (28a)$$

$$\lambda^+ = \lambda + \alpha_3(\lambda^{\text{QP}} - \lambda), \quad (28b)$$

where $(p^{\text{QP}}, \lambda^{\text{QP}})$ should be replaced by $(p^{\text{SOC}}, \lambda^{\text{SOC}})$ if second-order correction is executed, and the line search scheme [33, Algorithm 3] can be used to calculate the step sizes α_1 , α_2 and α_3 .

Remark 4 (Initialization of Algorithm 1). *OPF problems are usually initialized with a flat start, where all voltage angles are set to zero and all voltage magnitudes are set to 1.0 p.u. [41]. Besides, the dual variables of consensus constraints are set to zero for distributed OPF. For this initialization strategy, it has been demonstrated numerically that it can provide a good initial guess in practice [34], [35]. Hence, we focus on the full-step version of ALADIN, i.e., $\alpha_1 = \alpha_2 = \alpha_3 = 1$, in the present paper.*

B. Numerical Implementation

Some modifications are introduced to improve computation complexity. One is the conic relaxation of distribution subproblems in the BFM. Another is the second-order correction to compensate for the error in the linearization of the active constraints (19c) in the coupled QP step 4.

1) *Second-Order Conic Relaxation:* The feasible set of the AC OPF problem for the distribution grid defined by (5)-(9) is still nonconvex due to the quadratic equalities (5). To further reduce the computational burden and computing time of the decoupled NLP for distribution grids, we can reformulate the corresponding NLP as a conic optimization problem by relaxing the constraint (5) to a second-order cone constraint

$$l_{ij} \geq \frac{p_{ij}^2 + q_{ij}^2}{u_i}, \quad (i, j) \in \mathcal{L}_{Dm}, \quad (29)$$

which can be viewed as setting a lower bound on the current.

Theorem 1 [22]. *The conic relaxation of the OPF problem of a radial grid in the branch flow model is exact if the objective function is convex and strictly increasing in line loss.*

The solution to the conic relaxed problem is equivalent to the original problem (14). This indicates that the relaxed inequality constraint i is definitely on the boundary, i.e., i -th constraint is included in the optimal active-set

$$i \in \mathcal{S}^* = \{i \mid [h(x^*)]_i = 0\}. \quad (30)$$

In this way, detecting active-set for these relaxed inequality constraints can be evaded. By introducing augmented Lagrangian, most of the conic residuals would stay within the tolerance, while the rests would stay at a relatively low level during the preliminary iterations, and will converge to zero rapidly in practice.

2) *Second-Order Correction Step:* Based on the solution $(p^{\text{QP}}, \lambda^{\text{QP}})$ of (19), the compensation step $(p^{\text{SOC}}, \lambda^{\text{SOC}})$ can be computed analytically by using standard linear algebra. By subtracting the KKT condition of the coupled QP subproblem (19)

$$\begin{bmatrix} H & A^\top & J^\top \\ A & -\frac{I}{\mu} & 0 \\ J & 0 & 0 \end{bmatrix} \begin{bmatrix} p^{\text{QP}} \\ \lambda^{\text{QP}} - \lambda \\ \kappa^{\text{QP}} \end{bmatrix} = - \begin{bmatrix} A^\top \lambda + g \\ Ax - b \\ 0 \end{bmatrix} \quad (31)$$

with identity matrix I , and the KKT condition of the second-order correction subproblem (27)

$$\begin{bmatrix} H & A^\top & J^\top \\ A & -\frac{I}{\mu} & 0 \\ J & 0 & 0 \end{bmatrix} \begin{bmatrix} p^{\text{SOC}} \\ \lambda^{\text{SOC}} - \lambda \\ \kappa^{\text{SOC}} \end{bmatrix} = - \begin{bmatrix} A^\top \lambda + g \\ Ax - b \\ r \end{bmatrix}, \quad (32)$$

we obtain a linear system

$$\begin{bmatrix} H & A^\top & J^\top \\ A & -\frac{I}{\mu} & 0 \\ J & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta \lambda \\ \Delta \kappa \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix} \quad (33)$$

with difference of these two steps

$$(\Delta p, \Delta \lambda, \Delta \kappa) = (p^{\text{SOC}}, \lambda^{\text{SOC}}, \kappa^{\text{SOC}}) - (p^{\text{QP}}, \lambda^{\text{QP}}, \kappa^{\text{QP}}). \quad (34)$$

Under a mild assumption that the KKT point is regular, we can further reduce the system dimension by eliminating $\Delta \kappa$

$$\begin{bmatrix} H & A^\top \\ A & -\frac{1}{\mu} \end{bmatrix} \begin{bmatrix} \Delta p \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} J^\top \cdot (J M J^\top)^{-1} \\ 0 \end{bmatrix} \cdot r \quad (35)$$

with invertible matrix $M = (H + \mu A^\top A)^{-1}$.

Definition 2. A KKT point for a standard constrained optimization problem is regular [37] if linear independence constraint qualification (LICQ), strict complementarity condition and second order sufficient condition are satisfied.

As a result, the solution to the second-order correction subproblem (27) can be computed by

$$\begin{aligned} \begin{bmatrix} p^{\text{SOC}} \\ \lambda^{\text{SOC}} \end{bmatrix} &= \begin{bmatrix} p^{\text{QP}} \\ \lambda^{\text{QP}} \end{bmatrix} + \begin{bmatrix} \Delta p \\ \Delta \lambda \end{bmatrix} \\ &= \begin{bmatrix} p^{\text{QP}} \\ \lambda^{\text{QP}} \end{bmatrix} - \begin{bmatrix} I \\ \mu A \end{bmatrix} \cdot M J^\top \cdot (J M J^\top)^{-1} \cdot r. \end{aligned} \quad (36)$$

Remark 5. Locally, if the regularity condition (Definition 2) holds, the dual Hessian $J M J^\top$ is invertible. However, in practice, the LICQ might be violated such that matrix $J M J^\top$ might not be invertible. In such case, the pseudo-inverse of $J M J^\top$ has to be used.

C. Local Convergence Analysis

Here we analyze the local convergence property of Algorithm 1, in which the initial guess of primal-dual iterates is close enough to a local minimizer of (14).

Theorem 2. *Let $(z^*, \lambda^*, \kappa^*)$ be a regular KKT point for the problem (14), let f and h be twice continuously differentiable, and let ρ_{Σ_ℓ} being sufficiently large for all $\ell \in \mathcal{R}$ so that*

$$\forall \ell \in \mathcal{R}, \nabla^2 \{f_\ell(x_\ell) + \kappa_\ell^\top h_\ell(x_\ell)\} + \rho_{\Sigma_\ell} \succ 0 \quad (37)$$

are satisfied. Additionally, let the Hessian approximation H_ℓ be accurate enough so that

$$H_\ell = \nabla^2 \{f_\ell(x_\ell) + \kappa_\ell^\top h_\ell(x_\ell)\} + \mathcal{O}(\|x_\ell - z_\ell\|) \quad (38)$$

holds for all $\ell \in \mathcal{R}$. The iterate (x, λ) given by Algorithm 1 converges locally to (z^, λ^*) at a quadratic rate.*

The proof of Theorem 2 can be established the proof in two steps: we first analyze the convergence rate of Algorithm 1 without the second-order correction step 6, then we prove that the convergence rate can be preserved if the step is executed during the iterations. The detailed proof can be found in Appendix.

Remark 6. *The term $\mathcal{O}(\|x_\ell - z_\ell\|)$ in (38) is introduced to represent some regularization term used for numerical robustness. Despite these heuristic tricks for regularization, the locally quadratic convergences can be always observed in the sense of verifying the condition (38) numerically.*

IV. CASE STUDY

This section illustrates the performance of the proposed ALADIN-COR on the distributed AC OPF of generic ITD systems compared with ALADIN-STD and DCC.

A. Simulation Setting

Three ITD test cases with varying problem sizes and grid topologies are generated based on the standard IEEE test systems. In Case 1, one IEEE 39-bus transmission grid is connected by three IEEE 15-bus radial distribution grids. Case 2 comprises one IEEE 118-bus transmission grid and 20 IEEE 33-bus radial distribution grids. Both cases adopt a star-shaped topology, where the transmission grid is the central hub, and the distribution grids are connected to it.

In contrast, Case 3 explores a more complex scenario, where 4 IEEE 118-bus transmission grids are interconnected, resulting in a meshed topology. Additionally, 5 IEEE 33-bus radial distribution grids are connected to each transmission grid. The variation in problem size and grid topology enables a comprehensive analysis and comparison of different approaches under different operational conditions and system complexities.

The framework is built on Matlab-R2021a, the ITD systems are merged based on the open-source toolbox rapidPF [4]¹, and the case studies are carried out on a standard desktop computer with Intel® i5-6600K CPU @ 3.50GHz and 16.0 GB installed RAM. The CasADitoolbox [42] is used in Matlab, and ipopt [43] is used as a nonlinear solver. The centralized reference solution is obtained by solving the AC OPF problem using the default solver in Matpower. The computational time is estimated under the assumption that all subproblems are solved in parallel, and the time spent on exchanging sensitive information is not taken into consideration in the present paper.

For a fair comparison, all the algorithms are initialized with a flat start. Following [34], the quantities in the following are used to illustrate the convergence behavior

- 1) The deviation of optimization variables from the optimal value $\|x - x^*\|_2$.
- 2) The primal residual, i.e., the violation of consensus constraint $\|Ax\|_2 = \|\sum_{\ell \in \mathcal{R}} A_\ell x_\ell\|_2$.
- 3) The dual residual, i.e. the weighted euclidean distance of the ALADIN local step, $\|\Sigma(x - z)\|_2$.
- 4) The solution gap calculated as $\frac{f(x^*) - f(x)}{f(x^*)}$, where $f(x^*)$ is provided by the centralized approach.

When applying the DCC method from [13], the second and the third quantities are replaced by the mismatch of coupling variables between subproblems and the difference of upper and lower bound of the total cost for the ITD system respectively. Note that the vector b in consensus constraint is neglected since it is a zero-vector in this specific problem, c.f. (12).

The solution accuracy and the solution gap are defined by $\|x - x^*\|_2$ and $|\frac{f(x) - f(x^*)}{f(x^*)}|$ respectively, while the conic residual is defined by $\left\| \frac{p_{ij}^2 + q_{ij}^2}{u_i} - l_{ij} \right\|_2$ for all branches $(i, j) \in \mathcal{L}_{\ell \in \mathcal{R}_D}$ in distribution grids, c.f. the conic constraints (29).

¹The code is available on <https://github.com/xinliang-dai/rapidPF>

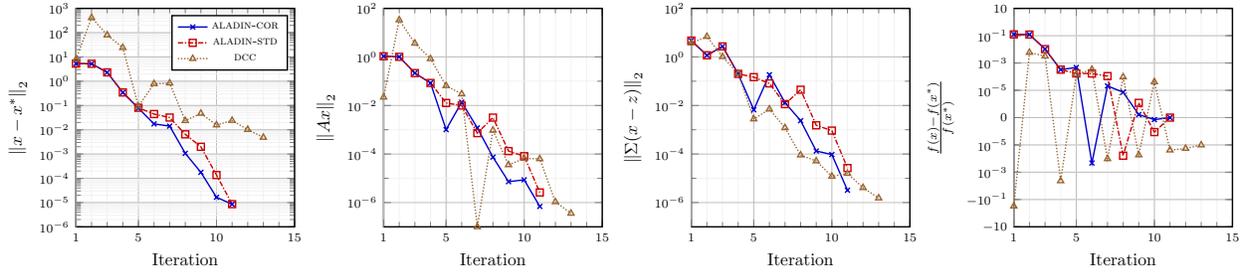


Fig. 3: Convergence behavior of different algorithms for Case 1

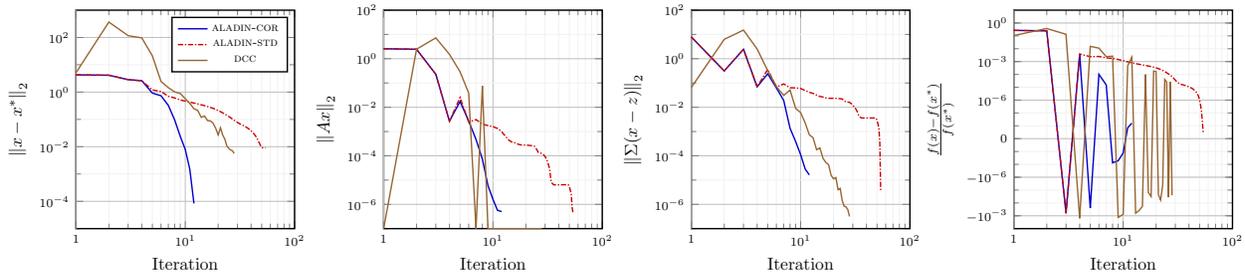


Fig. 4: Convergence behavior of different algorithms for Case 2

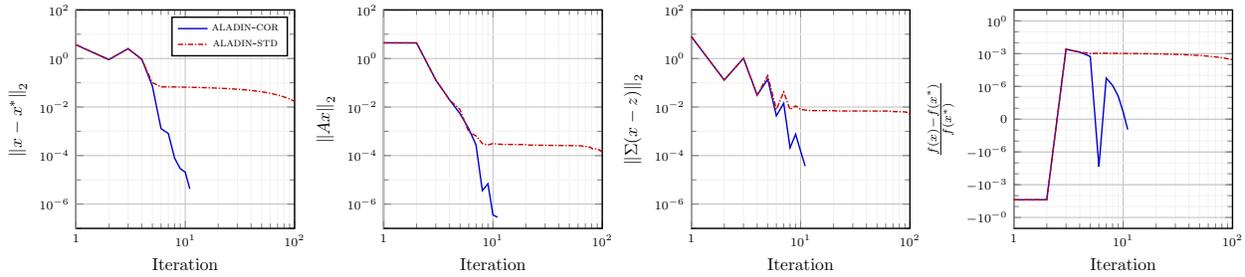
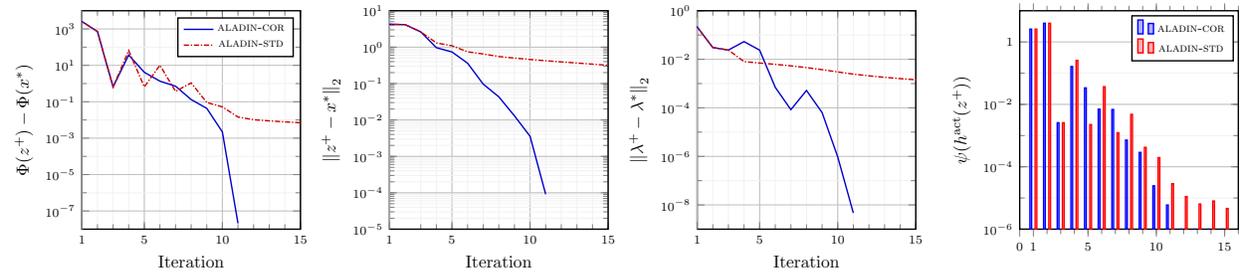


Fig. 5: Convergence behavior of different algorithms for Case 3

Fig. 6: Comparison of the updated primal-dual variables (z^+, λ^+) for Case 2

B. ALADIN-STD vs. ALADIN-COR

To demonstrate the improvement of the proposed new ALADIN-COR, we compare its convergence behavior with ALADIN-STD. In Cases 1 and 2, as illustrated in Fig. 3 and Fig. 4, the second-order correction is first carried out at the 4-th iteration of ALADIN-COR. For the smaller Case 1, the second-order correction strategy improves the convergence rate slightly. However, for the larger Cases 2 and 3, as depicted in Fig. 4 and Fig. 5, ALADIN-COR maintains the fast convergence rate and high solution quality, while convergence rate of standard ALADIN is slowed down, and it becomes difficult to achieve high accuracy. This indicates the scalability and effectiveness of the proposed ALADIN-COR for the nonconvex AC OPF of generic ITD systems with complex topology.

TABLE I: Comparison of the 4-th iterates for Case 2

	$\frac{f(z)-f(x^*)}{f(x^*)}$	$\ Az\ _1$	$\psi(h^{\text{act}}(z))$
z^{QP}	2.9×10^{-3}	1.5×10^{-17}	0.257
z^{SOC}	-1.3×10^{-5}	2.0×10^{-17}	0.163

TABLE II: Computing time by ALADIN-COR

Case	Iterations	Corrected Iterations	$t^{\text{TOTAL}}[s]$	$t^{\text{SOC}}[s]$
1	11	2	0.507	4.36×10^{-3}
2	12	4	3.222	7.61×10^{-2}
3	11	4	5.044	8.98×10^{-2}

Fig. 6 illustrates the quality of the primal-dual iterates (z^+, λ^+) for ALADIN-STD and ALADIN-COR. The comparison is presented in terms of the gap of the exact penalty function, the deviation of the primal-dual iterates to the local optimizer respectively, and the violation of the active constraint. Table I displays the solution gap, violations of the consensus constraint, and the active constraint at iterates z^{QP} and z^{SOC} at the 4-th iteration. At the 4-th iteration, the coupled QP step z^{QP} leads to the increase in the merit function (15) and significant violation of the constraints (14c). Consequently, the second-order correction implemented in ALADIN-COR is executed, and the primal-dual iterates (z^+, λ^+) are updated by re-solving the equivalent linear system (20). As a result, the objective $f(z^+)$ is improved, and the violation of the active constraints $\psi(h^{\text{act}}(z^+))$ is reduced, as shown in Table I. Compared with ALADIN-STD, the damping in the following

iterations would be avoided. A more detailed comparison between ALADIN-STD and ALADIN-COR is summarized in Table III.

The trade-off for the second-order correction strategy is the computing time of the additional step 5. As displayed in Table II, the second-order correction is only executed in a few iterations, and the corresponding computing time t^{soc} is negligible (around 5% compared with total computing time) because the corrected step can be obtained by re-solving the equivalent linear system (20).

In summary, the proposed ALADIN with second-order correction for solving the AC OPF of ITD systems can speed up the convergence of ALADIN-STD and maintain high performance at the cost of an insignificant increase in computing time. The effectiveness of this strategy would increase remarkably when the problem size goes large.

TABLE III: Comparisons of different algorithms for Cases 1 and 2

Case	Number of Buses	Number of Regions	Algorithm	Iterations	Time [s]	$\ x - x^*\ _2$	Solution Gap	Coic Residual
1	84	4	DCC	13	2.881	4.84×10^{-3}	1.02×10^{-5}	1.98×10^{-7}
			ALADIN-STD	11	0.503	8.62×10^{-6}	4.91×10^{-8}	1.08×10^{-5}
			ALADIN-COR	11	0.507	8.52×10^{-6}	1.91×10^{-8}	1.08×10^{-5}
2	778	21	DCC	28	22.545	5.65×10^{-3}	2.56×10^{-5}	2.68×10^{-7}
			ALADIN-STD	55	14.909	8.95×10^{-3}	2.84×10^{-9}	4.95×10^{-3}
			ALADIN-COR	12	3.222	8.36×10^{-5}	1.61×10^{-8}	3.02×10^{-5}
3	1132	24	DCC			Did not converge		
			ALADIN-STD	100	43.569	1.73×10^{-2}	2.78×10^{-4}	1.47×10^{-3}
			ALADIN-COR	11	5.044	4.23×10^{-6}	8.81×10^{-9}	8.32×10^{-6}

C. DCC vs. ALADIN-COR

The recent proposed DCC [13] follows a two-layer framework, where transmission and distribution subproblems are solved sequentially. In the lower layer, the conic optimization subproblems of distribution grids are solved in parallel with an additional L_1 penalty on coupling mismatch. Besides solving those subproblems, local agents also generate a lower bound of the lower layer costs, i.e., tangent plane and quadratic approximation with respect to coupling variables. Based on the lower layer solutions, the higher layer transmission problem is formulated as nonconvex AC OPF with distribution costs approximated by the estimated lower bounds. Then, the transmission provides the optimized coupling variables to local distributions. Thereby, the

transmission acts as a centralized coordinator, and privacy can be preserved since no detailed grid data is communicated between local agents.

Although the two-layer strategy provides a convergence guarantee for the AC OPF of star-shaped ITD systems, some limitations need to be acknowledged. Firstly, this methodology is limited to systems where multiple radial distributions are connected to only one transmission in a star-shaped configuration. It fails to handle a generic ITD systems with multiple transmission grids, meshed distribution grids, or meshed topology of subsystems. Secondly, the convergence of DCC heavily relies on the solvability of nonconvex subproblems in the upper layer. Thirdly, numerical issues may arise during the iterative process. With respect to small variations of the coupling variables, the quadratic approximation function is generated with the assistance of an equivalent linear coefficient matrix while neglecting second-order terms of the variation. Regarding the weakly active conic constraints (29), both the variation and the Lagrangian multipliers are small near the optimizer. By multiplication of variables with small values, the quadratic model loses numerical stability or even faces degeneracy issues in such cases. As a result, DCC would zigzag in coupling variables as it approaches high accuracy, as observed in Fig. 7 and Fig. 8.

In contrast, the proposed new ALADIN-COR is specially designed for generic distributed optimization and treats all subproblems equally. It provides convergence guarantees for AC OPF of generic ITD systems, considering more complex grid topologies. Moreover, several techniques are implemented in ALADIN-COR to ensure numerical robustness. In the decoupled step (16), the local objective is regularized by a proximal term, ensuring a descent direction, while additional slack variable s is introduced in the coupled step (19b) to prevent infeasibility caused by linearization of constraints. Furthermore, the second-order correction utilizing linear algebra for compensation improves numerical stability and maintains the quadratic convergence rate.

For a fair comparison with the original DCC [13], we provide the Root-Mean-Square Error (RMSE) of coupling variables in Fig. 7 and Fig. 8 for Cases 1 and 2 respectively, and show that DCC in the present paper reaches the same accuracy as the DCC of the original paper [13]. Compared with ALADIN-COR, DCC performs poorly in terms of the deviation of variables $\|x - x^*\|_2$, especially in the preliminary iterations for both Cases, shown in Fig. 3, Fig. 4. Besides, DCC starts zigzagging in terms of coupling variables near the local minimizer, resulting in a slower convergence rate when approaching higher accuracy, illustrated in Fig. 7 and Fig. 8.

Furthermore, DCC fails to handle the meshed connected multiple transmission grids in Case 3 within the data-preserving two-layer framework. In contrast, our ALADIN-COR converges within a dozen iterations in 5 seconds, as shown in Fig. 5. These observations highlight the superiority and scalability of the proposed ALADIN-COR for solving the AC OPF problem of generic ITD systems.

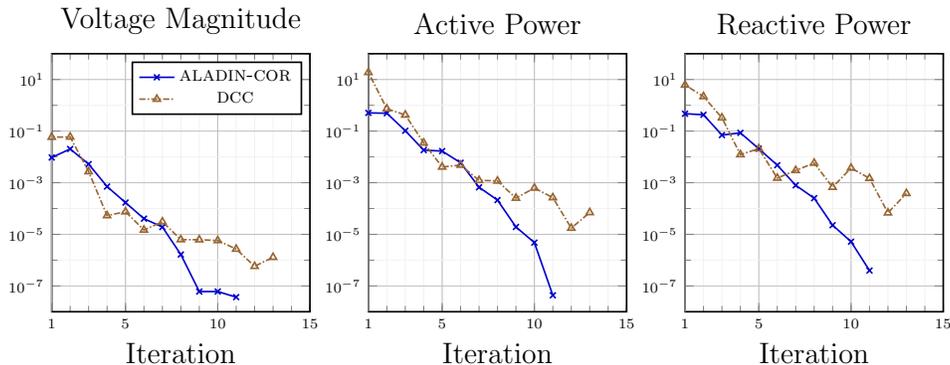


Fig. 7: RMSE of coupling variables for Case 1

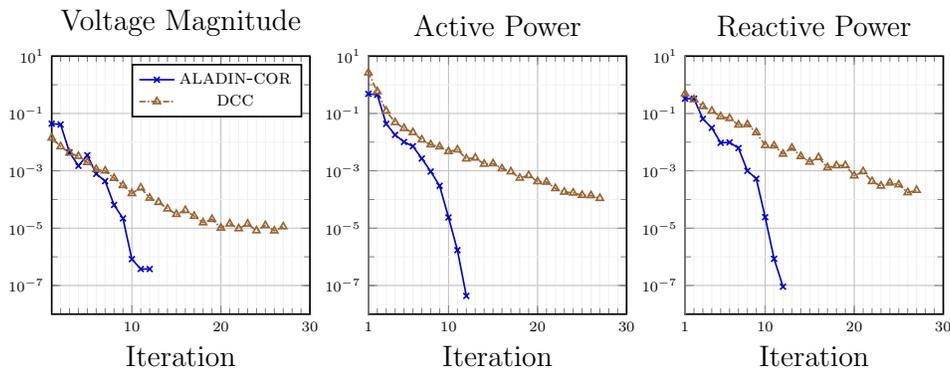


Fig. 8: RMSE of coupling variables for Case 2

V. CONCLUSIONS & OUTLOOK

The present paper proposes a novel ALADIN-COR algorithm for efficiently solving the distributed AC OPF problem of generic ITD systems regarding complex topology. The rigorous proof shows that the proposed ALADIN-COR with additional compensation step can maintain a locally quadratic convergence rate, ensuring its efficiency and numerical robustness. Numerical experiments conducted on three ITD benchmark cases, considering varying problem sizes and

grid topologies, demonstrate that the proposed distributed algorithm offers a more stable, efficient, and scalable solution compared with the state-of-art DCC and ALADIN-STD.

The proposed methodology opens several problems and extensions for future research. One direction is to investigate the impact of communication delay, packet loss, and asynchronous updates on the algorithm's performance. This investigation involves conducting experiments within a distributed computing software architecture to enhance the algorithm's robustness in real-world power systems where communication constraints are present. Additionally, further efforts can be devoted to scaling up the algorithm for large-scale power systems, introducing a distributionally robust framework to deal with renewable energy uncertainty, and solving online multi-period AC OPF with the receding horizon approach for dynamic power system operation.

APPENDIX

According to the assumptions of regularity and ρ , the local minimizer of subproblems (16), x_ℓ is parametric with (z, λ) and the solution maps are Lipschitz continuous, i.e., there exists constants $\chi_1, \chi_2 > 0$ such that

$$\|x - z^*\| \leq \chi_1 \|z - z^*\| + \chi_2 \|\lambda - \lambda^*\|. \quad (39)$$

From the local convergence analysis of Newton methods [37, Ch. 3.3], we have

$$\left\| \begin{bmatrix} z^{\text{QP}} - z^* \\ \lambda^{\text{QP}} - \lambda^* \end{bmatrix} \right\| \leq \|H - \nabla^2 \{f(x) + \kappa^\top h(x)\}\| \cdot \mathcal{O}(\|x - z^*\|) + \mathcal{O}(\|x - z^*\|^2).$$

By considering the accuracy of the Hessian approximation (38), the quadratic contractions

$$\|z^{\text{QP}} - z^*\| \leq \mathcal{O}(\|x - z^*\|^2), \quad (40a)$$

$$\|\lambda^{\text{QP}} - \lambda^*\| \leq \mathcal{O}(\|x - z^*\|^2), \quad (40b)$$

can be established. By combining (39) and (40), locally quadratic convergence of Algorithm 1 can be guaranteed if the second-order correction step 5 is never executed, i.e., $(z^{\text{QP}}, \lambda^{\text{QP}})$ is always accepted.

Then, we take the second-order correction step into consideration. Following [38], utilizing the relations $z^{\text{QP}} - x = p^{\text{QP}}$ yields an upper bound of the step p^{QP}

$$\|p^{\text{QP}}\| = \|z^{\text{QP}} - x\|$$

$$\begin{aligned}
&\leq \|z^{\text{QP}} - z^*\| + \|x - z^*\| \\
&\leq \mathcal{O}(\|x - z^*\|^2) + \mathcal{O}(\|x - z^*\|).
\end{aligned} \tag{41}$$

By Taylor series, we have

$$\begin{aligned}
r_\ell &= h_\ell^{\text{act}}(x_\ell + p_\ell^{\text{QP}}) = h_\ell^{\text{act}}(x_\ell) + J_\ell p_\ell^{\text{QP}} + \mathcal{O}(\|p_\ell^{\text{QP}}\|^2) \\
&= \mathcal{O}(\|p_\ell^{\text{QP}}\|^2), \forall \ell \in \mathcal{R},
\end{aligned} \tag{42}$$

where iterates x_ℓ satisfies $h_\ell^{\text{act}}(x_\ell) = 0$ and the step p_ℓ^{QP} satisfies the linearized equality constraint (19c). Under the regularity condition (Definition 2), the KKT matrix in the right-hand side of (33) is invertible, and the corresponding inverse matrix is bounded. Consequently, by combining (33), (41), (42), we have a quadratic contraction of primal and dual variables

$$\left\| \begin{bmatrix} p^{\text{SOC}} - p^{\text{QP}} \\ \lambda^{\text{SOC}} - \lambda^{\text{QP}} \end{bmatrix} \right\| \leq \mathcal{O}(\|r\|) \leq \mathcal{O}(\|x - z^*\|^2). \tag{43}$$

From (39), (40) and (43), it is sufficient to prove that the locally quadratic convergence can be achieved no matter whether the second-order correction step (step 5) is executed during the iterations.

REFERENCES

- [1] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, “A survey of distributed optimization and control algorithms for electric power systems,” *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.
- [2] J. Zhai, Y. Jiang, Y. Shi, C. Jones, and X. Zhang, “Distributionally robust joint chance-constrained dispatch for integrated transmission-distribution systems via distributed optimization,” *IEEE Trans. Smart Grid*, 2022.
- [3] M. A. Bragin and Y. Dvorkin, “Tso-dso operational planning coordination through “l_1– proximal” surrogate lagrangian relaxation,” *IEEE Trans. Power Syst.*, vol. 37, no. 2, pp. 1274–1285, 2021.
- [4] T. Mühlpfordt, X. Dai, A. Engelmann, and V. Hagenmeyer, “Distributed power flow and distributed optimization—formulation, solution, and open source implementation,” *Sustain. Energy, Grids Netw.*, vol. 26, p. 100471, 2021.
- [5] K. Tang, S. Dong, Y. Liu, L. Wang, and Y. Song, “Asynchronous distributed global power flow method for transmission–distribution coordinated analysis considering communication conditions,” *Electr. Power Syst. Res.*, vol. 182, p. 106256, 2020.
- [6] X. Dai, I. Yingzhao, Y. Jiang, and V. Hagenmeyer, “Hypergraph-based fast distributed ac power flow optimization,” in *2023 IEEE 62nd Conference on Decision and Control (CDC)*, 2023.
- [7] Z. Chen, Z. Li, C. Guo, J. Wang, and Y. Ding, “Fully distributed robust reserve scheduling for coupled transmission and distribution systems,” *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 169–182, 2021.

- [8] R. Roofegari nejad, W. Sun, and A. Golshani, "Distributed restoration for integrated transmission and distribution systems with DERs," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4964–4973, 2019.
- [9] J. Zhao, H. Wang, Y. Liu, Q. Wu, Z. Wang, and Y. Liu, "Coordinated restoration of transmission and distribution system using decentralized scheme," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3428–3442, 2019.
- [10] J. Zhao, Z. Zhang, J. Yao, S. Yang, and K. Wang, "A distributed optimal reactive power flow for global transmission and distribution network," *Int. J. Electr. Power Energy Syst.*, vol. 104, no. JAN., pp. 524–536, 2019.
- [11] Z. Li, Q. Guo, H. Sun, and J. Wang, "Coordinated transmission and distribution AC optimal power flow," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1228–1240, 2018.
- [12] A. Mohammadi, M. Mehrtash, and A. Kargarian, "Diagonal quadratic approximation for decentralized collaborative TSO+DSO optimal power flow," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2358–2370, 2019.
- [13] C. Lin, W. Wu, and M. Shahidehpour, "Decentralized AC optimal power flow for integrated transmission and distribution grids," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2531–2540, 2020.
- [14] D. Bienstock and A. Verma, "Strong NP-Hardness of AC power flows feasibility," *Operations Research Letters*, vol. 47, no. 6, pp. 494–501, 2019.
- [15] K. Lehmann, A. Grastien, and P. Van Hentenryck, "AC-feasibility on tree networks is NP-Hard," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 798–801, 2015.
- [16] S. H. Low, "Convex relaxation of optimal power flow—Part I: Formulations and equivalence," *IEEE Trans. Control. Netw. Syst.*, vol. 1, no. 1, pp. 15–27, 2014.
- [17] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1458–1459, 2006.
- [18] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *Int. J. Electr. Power Energy Syst.*, vol. 30, no. 6-7, pp. 383–392, 2008.
- [19] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 725–734, 1989.
- [20] M. Baran and F. F. Wu, "Optimal sizing of capacitors placed on a radial distribution system," *IEEE Trans. Power Deliv.*, vol. 4, no. 1, pp. 735–743, 1989.
- [21] M. Farivar, C. R. Clarke, S. H. Low, and K. M. Chandy, "Inverter var control for distribution systems with renewables," in *2011 IEEE SmartGridComm*, pp. 457–462, IEEE, 2011.
- [22] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification (Parts I, II)," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554–2564, 2013.
- [23] B. Subhonmesh, S. H. Low, and K. M. Chandy, "Equivalence of branch flow and bus injection models," in *2012 50th Annu. Allerton Conf. Commun. Control Comput.*, pp. 1893–1899, IEEE, 2012.
- [24] L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," in *2014 Power Systems Computation Conference*, pp. 1–9, IEEE, 2014.
- [25] D. K. Molzahn, I. A. Hiskens, *et al.*, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.
- [26] G. Hug-Glanzmann and G. Andersson, "Decentralized optimal power flow control for overlapping areas in power systems," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 327–336, 2009.
- [27] R. Baldick, B. Kim, C. Chase, and Y. Luo, "A fast distributed implementation of optimal power flow," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 858–864, 1999.
- [28] X. He, Y. Zhao, and T. Huang, "Optimizing the dynamic economic dispatch problem by the distributed consensus-based admm approach," *IEEE Trans. Ind. Inform.*, vol. 16, no. 5, pp. 3210–3221, 2020.

- [29] T. Erseghe, "Distributed optimal power flow using ADMM," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2370–2380, 2014.
- [30] B. Oh, D.-H. Lee, W.-C. Jeong, and D. Lee, "Distributed optimal power flow for distribution system using second order cone programming and consensus alternating direction method of multipliers," *J. Electr. Eng. Technol.*, vol. 17, no. 2, pp. 999–1008, 2022.
- [31] B. Huang, Y. Li, F. Zhan, Q. Sun, and H. Zhang, "A distributed robust economic dispatch strategy for integrated energy system considering cyber-attacks," *IEEE Trans. Ind. Inform.*, vol. 18, no. 2, pp. 880–890, 2022.
- [32] K. Sun and X. A. Sun, "A two-level ADMM algorithm for AC OPF with global convergence guarantees," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5271–5281, 2021.
- [33] B. Houska, J. V. Frasch, and M. Diehl, "An augmented lagrangian based algorithm for distributed nonconvex optimization," *SIAM J. Optim.*, vol. 26, no. 2, pp. 1101–1127, 2016.
- [34] A. Engelmann, Y. Jiang, T. Mühlpfordt, B. Houska, T. Faulwasser, "Toward Distributed OPF Using ALADIN," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 584–594, 2019.
- [35] J. Zhai, X. Dai, Y. Jiang, Y. Xue, V. Hagenmeyer, C. Jones, and X.-P. Zhang, "Distributed optimal power flow for VSC-MTDC meshed AC/DC grids using ALADIN," *IEEE Trans. Power Syst.*, pp. 1–1, 2022.
- [36] N. Meyer-Huebner, M. Suriyah, and T. Leibfried, "Distributed optimal power flow in hybrid AC–DC grids," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2937–2946, 2019.
- [37] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.
- [38] R. Fletcher, "Second order corrections for non-differentiable optimization," in *Numerical analysis*, pp. 85–114, Springer, 1982.
- [39] J. Zhai, M. Zhou, J. Li, X. Zhang, G. Li, C. Ni, and W. Zhang, "Hierarchical and robust scheduling approach for VSC-MTDC meshed ac/dc grid with high share of wind power," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 793–805, 2021.
- [40] B. Houska and Y. Jiang, "Distributed optimization and control with ALADIN," *Recent Advances in Model Predictive Control: Theory, Algorithms, and Applications*, p. 135–163, 2021.
- [41] S. Frank and S. Rebennack, "An introduction to optimal power flow: Theory, formulation, and examples," *IIE transactions*, vol. 48, no. 12, pp. 1172–1197, 2016.
- [42] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Math. Program. Comput.*, vol. 11, no. 1, pp. 1–36, 2019.
- [43] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, no. 1, pp. 25–57, 2006.