

Real-Time Distributed Algorithms for Nonconvex Optimal Power Flow

Yuejiang Liu, Jean-Hubert Hours, Giorgos Stathopoulos and Colin N. Jones

Abstract—The optimal power flow (OPF) problem, a fundamental problem in power systems, is generally nonconvex and computationally challenging for networks with an increasing number of smart devices and real-time control requirements. In this paper, we first investigate a fully distributed approach by means of the augmented Lagrangian and proximal alternating minimization method to solve the nonconvex OPF problem with a convergence guarantee. Given time-critical requirements, we then extend the algorithm to a distributed parametric tracking scheme with practical warm-starting and termination strategies, which aims to provide a closed-loop sub-optimal control policy while taking into account the grid information updated at the time of decision making. The effectiveness of the proposed algorithm for real-time nonconvex OPF problems is demonstrated in numerical simulations.

I. INTRODUCTION

One fundamental problem in the optimization and control of power systems is the optimal power flow (OPF) problem, which seeks to optimize an objective such as generation cost or power loss under certain network constraints such as voltage operation limits and device capacities. The OPF problem is generally nonconvex, and thus NP-hard. Various OPF algorithms have been developed based on convex relaxations, which are appealing for some particular networks where the relaxation is *exact*, i.e. the optimal solution of the convex relaxed problem is also a solution of the original OPF problem [17], [18]. However, solving the OPF problem in the case of general networks remains an open and challenging problem.

In addition to non-convexity, the size of the OPF problem poses another challenge to future grid control. Smart devices such as controllable loads and storage elements are increasingly integrated into power grids. On one hand, these devices assist to shave the peak load and compensate for the uncertainties of the electricity demands and renewable energy generations [9], but on the other hand, they inevitably increase the scale and complexity of the OPF problem.

The OPF problem is traditionally solved by a Newton-Raphson method in a centralized manner [14]. However, a distributed method that allows for more agents to simultaneously participate in the computations is highly desirable for scalability. Some distributed algorithms based on semidefinite programming (SDP) or second-order cone programming (SOCP) relaxations for OPF problems have been developed

for the cases where tight convexification holds [7], [16], [21]. Nevertheless, for the general networks where the convex relaxation is inexact, decomposition needs to take place at the nonconvex level. Recently, [19] and [12] proposed distributed nonconvex OPF algorithms via an alternating minimization method combined with sequential convex approximation or trust region method respectively, but their uses in the real-time scenario are hampered by their reliance on a large number of coordinations between subsystems.

In this work, we aim at solving the nonconvex OPF problem in real-time, as more frequent solutions are requested by power networks with increased penetration of renewable energy resources [9]. To achieve this goal, we first tackle the nonconvex OPF problem by a distributed approach with a convergence guarantee. We then further speed it up with practical warm-starting, termination and tuning strategies with the purpose of providing a closed-loop suboptimal control policy, while taking into account the updated grid information at (or close to) the time of decision making. The main contribution of the paper lies in applying the distributed optimization techniques in [2], [11], [13] to the real-time nonconvex OPF problems where the convex relaxation approach [8], [17], [18] fails.

The paper is organized as follows. Section II introduces the formulation of OPF problems. The real-time distributed OPF algorithms are then presented in Section III and tested in Section IV. We finally conclude the paper in Section V.

II. PROBLEM FORMULATION

In this section, we introduce the formulation of the nonconvex AC OPF problems considered in the paper.

A. Power Flow Model

We consider a radial power network $\mathcal{T} := (\mathcal{N}, \varepsilon)$ consisting of a set of buses $\mathcal{N} := \{0, \dots, n\}$ and a set of lines $\varepsilon := \{1, \dots, n\}$. We index the root node by 0, and denote the rest by $\mathcal{N}_+ := \mathcal{N} \setminus \{0\}$. In the network, each bus $i \in \mathcal{N}_+$ has a set of neighboring buses, including a unique ancestor A_i and a set of children $C_i := \{C_{1_i}, C_{2_i}, \dots, C_{r_i}, \dots, C_{R_i}\}$, and each line $i \in \varepsilon$ connects bus i and bus A_i . For each bus $i \in \mathcal{N}$, we denote by V_i the complex voltage, and define $v_i := |V_i|^2$. We also denote by $s_i = p_i + \mathbf{j}q_i$ the power injection at bus i where p_i and q_i represent the real and reactive power respectively. At the substation bus, the voltage V_0 is fixed while the power injection s_0 is fully flexible to feed the network. For each line $i \in \varepsilon$, let $Z_i = R_i + \mathbf{j}X_i$ denote the complex impedance and $S_i = P_i + \mathbf{j}Q_i$ denote the power flow from bus i to bus A_i . Also, let I_i denote the complex current on the line from bus i to bus A_i , and define

This work has received support from the European Research Council under the European Unions Seventh Framework Programme (FP/2007-2013)/ ERC Grant Agreement n. 307608 (BuildNet).

Yuejiang Liu, yuejiang.liu@outlook.com

Jean-Hubert Hours, jhhours@gmail.com

Giorgos Stathopoulos, georgios.stathopoulos@epfl.ch

Colin N. Jones, colin.jones@epfl.ch

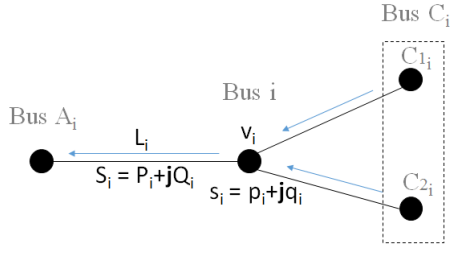


Fig. 1. Notations in a radial power network

$L_i := |I_i|^2$. The notation of variables is summarized in the table below. The lowercase corresponds to the bus while the uppercase corresponds to the line.

Bus $i \in \mathcal{N}$	v_i	p_i	q_i
Line $i \in \varepsilon$ (Bus i - A_i)	L_i	P_i	Q_i

Given the network topology \mathcal{T} , the impedance Z and the substation voltage v_0 , the power flow in a radial network can be described by the branch flow model [3], [4] as follows:

$$v_{A_i} = v_i - 2(R_i P_i + X_i Q_i) + L_i(R_i^2 + X_i^2) \quad (1a)$$

$$\sum_{C_{r_i} \in C_i} (P_{C_{r_i}} - L_{C_{r_i}} R_{C_{r_i}}) + p_i = P_i \quad (1b)$$

$$\sum_{C_{r_i} \in C_i} (Q_{C_{r_i}} - L_{C_{r_i}} X_{C_{r_i}}) + q_i = Q_i \quad (1c)$$

$$P_i^2 + Q_i^2 = v_i L_i \quad (1d)$$

B. OPF Problem

Two types of constraints commonly exist in grid operation. The first type regards the power injection constraints on the controllable devices at each bus. Suppose that a number of K_i devices such as load, generator and energy storage elements are installed at bus i , the total power injection at this bus is as follows [8]:

$$p_i = \sum_{k=1}^{K_i} p_{i,k}, \quad q_i = \sum_{k=1}^{K_i} q_{i,k} \quad (2)$$

Let $p_i^{\text{gen}}(t)$ and $q_i^{\text{gen}}(t)$ denote the active and reactive power injection, respectively, from a generator at time $t \in T$, which are bounded by:

$$\underline{p}_i^{\text{gen}} \leq p_i^{\text{gen}}(t) \leq \bar{p}_i^{\text{gen}} \quad (3a)$$

$$\underline{q}_i^{\text{gen}} \leq q_i^{\text{gen}}(t) \leq \bar{q}_i^{\text{gen}} \quad (3b)$$

For a storage element at bus i , let $e_i^{\text{bat}}(t)$ denote the amount of energy storage and $p_i^{\text{bat}}(t)$ denote the discharge rate from the battery to the network at time $t \in T$. The dynamics of the battery are modeled to follow a first-order difference equation with constraints:

$$e_i^{\text{bat}}(t+1) = e_i^{\text{bat}}(t) - p_i^{\text{bat}}(t)\Delta t \quad (4a)$$

$$\underline{e}_i^{\text{bat}} \leq e_i^{\text{bat}}(t) \leq \bar{e}_i^{\text{bat}} \quad (4b)$$

$$\underline{p}_i^{\text{bat}} \leq p_i^{\text{bat}}(t) \leq \bar{p}_i^{\text{bat}} \quad (4c)$$

Another operational constraint is on the voltage that is regulated within a range. Suppose that the voltage lower bound and upper bound are defined as $\underline{v}_i, \bar{v}_i$ respectively, the voltage at bus i is constrained by

$$\underline{v}_i \leq v_i \leq \bar{v}_i \quad (5)$$

For example, if the allowable voltage deviation is 5% from the nominal, then the voltage at bus $i \in \mathcal{N}_+$ is regulated in a per unit range $0.95^2 \leq v_i \leq 1.05^2$.

The optimal power flow (OPF) problem typically has an objective function related to the power injections at buses and the currents on the lines, such as minimizing the real power generation or total power losses. Moreover, the objective function is usually an aggregate of the local cost at each agent. In this work, we focus on the following objective function:

$$f(p) = \sum_{t=1}^T \sum_{i \in \mathcal{N}} \mathbb{C}_{i2}(t)(p_i^{\text{gen}}(t))^2 + \mathbb{C}_{i1}(t)p_i^{\text{gen}}(t) \quad (6)$$

where $\mathbb{C}_{i2}(t)$ and $\mathbb{C}_{i1}(t)$ are the terms associated with generation price at bus i at time t . For a network equipped with smart devices such as storage elements or controllable loads, the quadratic cost function tends to shave the peak load and flatten the generation profile, which are favorable for the efficiency of conventional power plants [10].

To summarize, the OPF problem for radial networks is formulated as below:

$$\begin{aligned} \text{OPF: } & \min && (6) \\ & \text{over} && v, L, P, Q, p, q \\ & \text{s.t.} && (1), (2), (3), (4), (5) \end{aligned}$$

The equality constraint (1d) is nonconvex. By relaxing this constraint, one can obtain a second-order cone program (SOCP) and gain higher numerical efficiency. However, the solution of the relaxed problem can be infeasible for the grid operation when violating (1d). A detailed analysis regarding the possible inexactness can be found in [15].

III. REAL-TIME DISTRIBUTED OPF ALGORITHMS

In this section, we first review a proximal regularized block-coordinate descent method, and then apply it to the nonconvex OPF problem. Subsequently, practical acceleration strategies are proposed for the real-time grid operation.

A. Preliminary

The proximal regularized block-coordinate descent method is a powerful framework for solving the minimization problem of a nonconvex function satisfying the Kurdyka-Lojasiewicz property [2]. It tackles the problem of the following structure:

$$\min_x g(x_1, \dots, x_p) + \sum_{i=1}^p h_i(x_i)$$

where vector $x = (x_1, \dots, x_p)$ belongs to the space $\mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_p}$, g is a C^1 continuous function with locally Lipschitz continuous gradient, and $h_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R} \cup \{+\infty\}$,

$i = 1, 2, \dots, p$ is a proper lower semicontinuous function. It is proven in [2] that a bounded sequence $\{x_1^k, \dots, x_p^k\}_{k \in \mathbb{N}}$ generated by an inexact block-coordinate descent algorithm converges to a critical point of $g(x_1, \dots, x_p) + \sum_{i=1}^p h_i(x_i)$ under the following two assumptions:

Assumption 1: (sufficient decrease) For $a_i^k > 0$, the following inequality condition holds:

$$\begin{aligned} h_i(x_i^{k+1}) + g(x_1^{k+1}, \dots, x_{i-1}^{k+1}, x_i^{k+1}, \dots, x_p^k) + \frac{a_i^k}{2} \|x_i^{k+1} - x_i^k\|^2 \\ \leq h_i(x_i^k) + g(x_1^{k+1}, \dots, x_{i-1}^{k+1}, x_i^k, \dots, x_p^k). \end{aligned} \quad (8)$$

Assumption 2: (relative error tolerance) For $b_i^k > 0$ and the subgradient $v_i^{k+1} \in \partial h_i(x_i^{k+1})$, the following inequality condition holds:

$$\|v_i^{k+1} + \nabla_{x_i} g(x_1^{k+1}, \dots, x_i^{k+1}, x_{i+1}^k, \dots, x_p^k)\| \leq b_i^k \|x_i^{k+1} - x_i^k\|. \quad (9)$$

Let us also define the proximal operator and indicator function for the remainder of the paper. Let $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function which is bounded from below, and $\alpha > 0$ a positive parameter. The proximal operator of f with coefficient α is defined as

$$\text{prox}_\alpha^f(x) := \arg \min_y f(y) + \frac{\alpha}{2} \|y - x\|^2$$

An indicator function of a closed subset Ω of \mathbb{R} is defined as

$$\delta_\Omega(x) := \begin{cases} 0 & \text{if } x \in \Omega \\ +\infty & \text{if } x \notin \Omega \end{cases}$$

B. OPF Decomposition

The OPF problem formulated in Section II has the following properties that facilitate decomposition:

- The objective function (6) is fully decomposable to each bus in the network.
- The equality constraints (1a - 1c) are linearly coupled between neighboring buses.
- The rest of equality and inequality constraints (1d - 4) are private at each bus.

Inspired by [20], we first decouple the coupling constraints (1a - 1c) by introducing a set of slack variables $x_{A,i}$ and $x_{C,i} := (x_{C1,i}, x_{C2,i}, \dots, x_{Cr,i}, \dots)$ at bus i as local estimates of the coupling variables associated with its neighbors, i.e. an ancestor A_i and a group of children C_i . Thus, each bus i manages a set of local variables $x_i := (x_{B,i}, x_{A,i}, x_{C,i})$ as follows:

$$x_{B,i} := (v_i^{(x)}, L_i^{(x)}, P_i^{(x)}, Q_i^{(x)}, p_i^{(x)}, q_i^{(x)})$$

$$x_{A,i} := v_{A_i}^{(x)} \equiv v_{A_i}^{(x)}$$

$$x_{Cr,i} := (L_{Cr,i}^{(x)}, P_{Cr,i}^{(x)}, Q_{Cr,i}^{(x)}) \equiv (L_{Cr,i}^{(x)}, P_{Cr,i}^{(x)}, Q_{Cr,i}^{(x)})$$

where $\{\cdot\}^{(x)}$ represents local variables.

As the local variables $x_{A,i}, x_{C,i}$ estimated at bus i should be in consensus with the corresponding variables x_{A_i}, x_{C_i} estimated at the adjacent bus A_i and C_i , a set of slack

variables in charge of communication are defined at the neighborhood around each bus i :

$$z_i := (v_i^{(z)}, L_i^{(z)}, P_i^{(z)}, Q_i^{(z)})$$

where $\{\cdot\}^{(z)}$ represents consensus variables.

With these additional variables, the OPF problems can be decomposed and reformulated into the following form:

$$\min \sum_{i=1}^N f_i(x_i) \quad (11a)$$

$$\text{s.t. } A_i x_i = b_i \quad i \in \mathcal{N} \quad (11b)$$

$$x_i^T E_i x_i = 0 \quad i \in \mathcal{N} \quad (11c)$$

$$x_{B,i} = z_i, x_{A,i} = z_{A_i}, x_{C,i} = z_{C_i} \quad i \in \mathcal{N} \quad (11d)$$

$$x_i \in \Omega_i \quad i \in \mathcal{N} \quad (11e)$$

where $A_i x_i = b_i$ corresponds to the local linear equality constraints (1a-1c), (2), (4a), $x_i^T E_i x_i = 0$ corresponds to the nonlinear equality constraint (1d), and Ω_i represents the local inequality bounds (3), (4b-4c), (5) associated with x_i .

C. Distributed OPF Algorithm

To tackle the nonconvex OPF problem, we first use the classical augmented Lagrangian method as the outer frame of our algorithms. The augmented Lagrangian associated with the OPF problem (11) is as follows:

$$\begin{aligned} \mathcal{L}_\rho(x, z; \mu, \gamma, \lambda) = & f(x) + \mu(Ax - b) + \gamma(x^T E x) + \lambda(x - z) \\ & + \frac{\rho}{2} \|Ax - b\|^2 + \frac{\rho}{2} \|x^T E x\|^2 + \frac{\rho}{2} \|x - z\|^2 \end{aligned} \quad (12)$$

subject to $x \in \Omega$, where $\Omega := \Omega_1 \times \dots \times \Omega_n$, $\rho > 0$ is a penalty parameter, $x := (x_1, \dots, x_n)$ is the group of local variables, $z := (z_1, \dots, z_n)$ is the group of consensus variables, $\mu := (\mu_1, \dots, \mu_n)$, $\gamma := (\gamma_1, \dots, \gamma_n)$, $\lambda := (\lambda_1, \dots, \lambda_n)$ are Lagrange multipliers associated with the constraints (11b-11d). Specifically, $\lambda_i := (\lambda_{B,i}, \lambda_{A,i}, \lambda_{C,i})$, where $\lambda_{B,i}, \lambda_{A,i}, \lambda_{C,i}$ correspond to the three different consensus constraints associated with (11d).

The standard augmented Lagrangian algorithm solves (11) by the primal-dual sequences:

- Primal update:

$$\{x^{k+1}, z^{k+1}\} := \arg \min_{x \in \Omega, z} \mathcal{L}_{\rho^k}(x, z; \mu^k, \gamma^k, \lambda^k) \quad (13)$$

- Dual update:

$$\mu^{k+1} := \mu^k + \rho^k (Ax^{k+1} - b) \quad (14a)$$

$$\gamma^{k+1} := \gamma^k + \rho^k (x^{k+1 T} E x^{k+1}) \quad (14b)$$

$$\lambda^{k+1} := \lambda^k + \rho^k (x^{k+1} - z^{k+1}) \quad (14c)$$

where k is the iteration indicator.

The dual update is a simple algebraic calculation, which can be split down to each bus for parallel computing. On the contrary, the primal update is a joint minimization of a nonconvex function over two coupling blocks of variables x and z , which is not easy to solve. We will next apply

the proximal block-coordinate descent method to tackle the primal update in a distributed and efficient way.

As described in Section III-B, $\mathcal{L}(x, z)$ is fully decomposable within the variable block x and z :

$$\mathcal{L}(x, z) = \sum_{i=1}^N \mathcal{L}_{x,i}(x_i, z) = \sum_{i=1}^N \mathcal{L}_{z,i}(x, z_i)$$

where $\mathcal{L}_{x,i}$ and $\mathcal{L}_{z,i}$ are the local augmented Lagrangian functions for the local variables x_i and consensus variables z_i at bus i respectively:

$$\begin{aligned} \mathcal{L}_{x,i}(x_i, z) &= f_i(x_i) + \mu_i(A_i x_i - b_i) + \frac{\rho}{2} \|A_i x_i - b_i\|^2 \\ &\quad + \gamma_i(x_i^\top E_i x_i) + \frac{\rho}{2} \|x_i^\top E_i x_i\|^2 \\ &\quad + \lambda_i(x_{B,i} - z_i) + \frac{\rho}{2} \|x_{B,i} - z_i\|^2 \\ &\quad + \lambda_{A,i}(x_{A,i} - z_{A_i}) + \frac{\rho}{2} \|x_{A,i} - z_{A_i}\|^2 \\ &\quad + \sum_{C r_i \in C_i} (\lambda_{C r_i, i}(x_{C r_i, i} - z_{C r_i}) + \frac{\rho}{2} \|x_{C r_i, i} - z_{C r_i}\|^2) \end{aligned}$$

$$\begin{aligned} \mathcal{L}_{z,i}(x, z_i) &= -\lambda_i(z_i - x_{B,i}) + \frac{\rho}{2} \|z_i - x_{B,i}\|^2 \\ &\quad - \lambda_{C r^*, A_i}(z_i - x_{C r^*, A_i}) + \frac{\rho}{2} \|z_i - x_{C r^*, A_i}\|^2 \\ &\quad - \sum_{C r_i \in C_i} (\lambda_{A, C r_i}(z_i - x_{A, C r_i}) + \frac{\rho}{2} \|z_i - x_{A, C r_i}\|^2) \\ &\quad + f_i(x_i) + \mu_i(A_i x_i - b_i) + \frac{\rho}{2} \|A_i x_i - b_i\|^2 \\ &\quad + \gamma_i(x_i^\top E_i x_i) + \frac{\rho}{2} \|x_i^\top E_i x_i\|^2 \end{aligned}$$

where $\{\cdot\}_{C r^*, A_i}$ refers to a specific child variable estimated at bus A_i that corresponds to the variable at bus i .

It can be noticed that $\mathcal{L}_{x,i}$ is a nonconvex function of x_i and $\mathcal{L}_{z,i}$ is a convex function of z_i . Inspired by [6], we tackle the primal update (13) by the following alternating minimization scheme:

$$\begin{aligned} x_i^{m+1} &\in \arg \min_{x_i \in \Omega_i} \langle x_i - x_i^m, \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m, z^m) \rangle \\ &\quad + \frac{c_i^m}{2} \|x_i - x_i^m\|^2, i \in \mathcal{N} \end{aligned} \quad (15)$$

$$z_i^{m+1} \in \arg \min_{z_i} \mathcal{L}_{z,i}(x^{m+1}, z_i) + \frac{d_i^m}{2} \|z_i - z_i^m\|^2, i \in \mathcal{N} \quad (16)$$

where m is the iteration indicator, $c_i^m > 0$ and $d_i^m > 0$ are the regularization coefficients. In practice, the coefficient d_i^m can be taken as a small constant, and c_i^m can be determined by a backtracking procedure to satisfy the following condition:

$$\begin{aligned} &\mathcal{L}_{x,i}(x_i^{m+1}, z^m) + \alpha \|x_i^{m+1} - x_i^m\|^2 \\ &\leq \mathcal{L}_{x,i}(x_i^m, z^m) + \langle x_i^{m+1} - x_i^m, \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m) \rangle + \frac{c_i^m}{2} \|x_i^{m+1} - x_i^m\|^2 \end{aligned} \quad (17)$$

where $\alpha > 0$ is a small positive constant.

Proposition 1: Suppose that (17) holds. Then the minimization steps (15) and (16) satisfy the inequality (8) and (9).

Proof: The x -update (15) implies

$$\langle x_i^{m+1} - x_i^m, \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m) \rangle + \frac{c_i^m}{2} \|x_i^{m+1} - x_i^m\|^2 + \delta_{\Omega_i}(x_i^{m+1}) \leq \delta_{\Omega_i}(x_i^m) \quad (18)$$

As the coefficient c_i^m satisfies (17), we have

$$\begin{aligned} &\delta_{\Omega_i}(x_i^{m+1}) + \mathcal{L}_{x,i}(x_i^{m+1}) + \alpha \|x_i^{m+1} - x_i^m\|^2 \\ &\leq \delta_{\Omega_i}(x_i^{m+1}) + \mathcal{L}_{x,i}(x_i^m) + \langle x_i^{m+1} - x_i^m, \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m) \rangle + \frac{c_i^m}{2} \|x_i^{m+1} - x_i^m\|^2 \end{aligned} \quad (19)$$

Substituting (18) into the right side of (19), one obtains

$$\delta_{\Omega_i}(x_i^{m+1}) + \mathcal{L}_{x,i}(x_i^{m+1}) + \alpha \|x_i^{m+1} - x_i^m\|^2 \leq \delta_{\Omega_i}(x_i^m) + \mathcal{L}_{x,i}(x_i^m) \quad (20)$$

By summing up (20) for all the buses, the *sufficient decrease* assumption (8) is satisfied.

The optimality condition of (15) is

$$\begin{aligned} &\exists v_i^{m+1} \in \partial \delta_{\Omega_i}(x_i^{m+1}), \\ &\nabla_{x_i} \mathcal{L}_{x,i}(x_i^m) + c_i^m (x_i^{m+1} - x_i^m) + v_i^{m+1} = 0 \end{aligned}$$

Then we have

$$\begin{aligned} &\|v_i^{m+1} + \nabla_{x_i} \mathcal{L}_{x,i}(x_i^{m+1})\| \\ &\leq \|v_i^{m+1} + \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m)\| + \|\nabla_{x_i} \mathcal{L}_{x,i}(x_i^{m+1}) - \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m)\| \\ &= c_i^m \|x_i^{m+1} - x_i^m\| + \|\nabla_{x_i} \mathcal{L}_{x,i}(x_i^{m+1}) - \nabla_{x_i} \mathcal{L}_{x,i}(x_i^m)\| \\ &\leq (c_i^m + L_i) \|x_i^{m+1} - x_i^m\| \end{aligned}$$

where L_i is a Lipschitz constant of $\mathcal{L}_{x,i}(x_i)$. The x -update (15), therefore, satisfies the *relative error* assumption (9).

For the z -update, (16) directly implies

$$\mathcal{L}_{z,i}(z_i^{m+1}) + \frac{d_i^m}{2} \|z_i^{m+1} - z_i^m\|^2 \leq \mathcal{L}_{z,i}(z_i^m)$$

Additionally, the optimality condition of (16) suggests that

$$\nabla_{z_i} \mathcal{L}_{z,i}(z_i^{m+1}) + d_i^m (z_i^{m+1} - z_i^m) = 0,$$

which implies $\|\nabla_{z_i} \mathcal{L}_{z,i}(z_i^{m+1})\| \leq d_i^m \|z_i^{m+1} - z_i^m\|$. Thus, the z -update satisfies (8) and (9) as well. ■

Algorithm 1 Distributed OPF Algorithm

- 1: **Inputs:**
grid infomation, inner loop termination tolerance ϵ ,
outer loop termination tolerance η
 - 2: **Initialize:**
 z, μ, γ, λ , choose $\rho > 0, \beta > 1$
 - 3: **repeat**
 - 4: **repeat**
 - 5: $x_{\text{prev}}, z_{\text{prev}} \leftarrow x, z$
 - 6: $x \leftarrow$ update x_i in parallel (15)
 - 7: $z \leftarrow$ update z_i in parallel (16)
 - 8: **until** $\|(x - x_{\text{prev}}, z - z_{\text{prev}})\| \leq \epsilon / \rho$
 - 9: $\mu, \gamma, \lambda \leftarrow$ update $\mu_i, \gamma_i, \lambda_i$ in parallel (14)
 - 10: $\rho \leftarrow \beta \rho$
 - 11: **until** $\|(Ax - b, x^\top E x, x - z)\|_2 \leq \eta$
-

Combining (12-17), our distributed algorithm for the non-convex OPF problem is presented in Algorithm 1. The outer

loop employs the classical augmented Lagrangian method with a growing penalty parameter ρ and a tightening primal update tolerance ϵ , while the inner loop follows the proximal regularized alternating minimization framework. As the augmented Lagrangian with a box constraint (12) is a semi-algebraic polynomial function satisfying the Kurdyka-Lojasiewicz (KL) property [5], Algorithm 1 has the following convergence properties.

Theorem 1: The sequence $\{x^m, z^m\}$ generated by the inner loop of Algorithm 1 converges to a critical point of $\mathcal{L}_\rho(x, z; \mu, \gamma, \lambda) + \delta_\Omega(x)$, and it has a finite length.

Proof: The sequence $\{x^m, z^m\}$ is constrained by (1-5), and thus $\mathcal{L}_\rho(x, z; \mu, \gamma, \lambda) + \delta_\Omega(x)$ is bounded from below. Based on Proposition 1, the inner loop of Algorithm 1 is a two-block case of the inexact regularized Gauss-Seidel method, which is proven to converge by Theorem 6.2 in [2]. ■

Theorem 2: The sequence $\{x^k, z^k, \mu^k, \gamma^k, \lambda^k\}$ generated by the outer loop of Algorithm 1 converges to a KKT point of $\mathcal{L}_\rho(x, z, \mu, \gamma, \lambda) + \delta_\Omega(x)$.

Proof: Based on Theorem 1, a proof for the outer loop convergence can be found in [11]. ■

From the computational perspective, Algorithm 1 enables efficient computation at all the agents simultaneously. The x -update (15) minimizes the proximal regularization of the nonconvex $\mathcal{L}_\rho(x)$ linearized at the point x^m , and it can be further split into a forward-backward scheme at each bus i :

$$x_i^{m+1} = \text{prox}_{c_i^m}^{\delta_{\Omega_i}}(x_i^k - \frac{1}{c_i^m} \nabla_{x_i} \mathcal{L}_{x_i}(x_i^m, z^m))$$

The backward proximal mapping to the indicator function $\delta_{\Omega_i}(x_i)$ is a projection to a box constraint that has a closed form solution. The z -update (16) can be viewed as a communication step of local variable estimates among the neighborhood around each bus i . It minimizes a regularized quadratic function, the solution of which is a system of linear equations and thus can be recovered in closed form as well. As these computations are distributed to each agent, Algorithm 1 is expected to be scalable for large power networks.

D. Real-time Distributed Parametric Tracking

As the OPF problem needs to be solved more and more frequently in a power network with the growing penetration of renewable energy resources, in this section we further investigate practical acceleration strategies with the purpose of providing a highly feasible and suboptimal solution within a limited number of iterations for real-time closed-loop grid operation.

A short-term demand forecast provided by the state-of-art techniques has around 2% - 4% mean absolute percentage error (MAPE) in 24 hours [22], which indicates that the updated demand information of a new OPF problem generally does not differ significantly from the demand predicted at the previous time step. This motivates us to initialize the

distributed computation for a new OPF problem from the solution of the previous one:

$$z_t^0 \leftarrow z_{t-1}^\dagger, (\mu_t^0, \gamma_t^0, \lambda_t^0) \leftarrow (\mu_{t-1}^\dagger, \gamma_{t-1}^\dagger, \lambda_{t-1}^\dagger)$$

where z_t and λ_t are the consensus variable and multiplier of a new OPF problem solved at time t , and $\{\cdot\}_{t-1}^\dagger$ represents the solution of the previous problem. If the prediction horizon of a multi-stage OPF covers a daily period of time such as 24 hours, we can even warm-start the new problem by the shifted solution of the previous problem to fully take advantage of the solution match. This online warm initialization is expected to provide a good initial guess for solving the new OPF problem, and hence accelerate the online algorithmic performance.

In spite of warm-initialization, Algorithm 1 may still require longer computational time than allowed in the real-time scenario, as the fact that the distributed algorithm based on the first order method generally needs a considerable number of iterations and that the network communication operates at a limited rate much slower than the on-chip exchange. When truncating the iterations of Algorithm 1 is necessary, we would reduce the number of outer loop iterations but maintain the number of inner loop iterations, meanwhile enlarging the penalty parameter ρ in order to provide a highly feasible and suboptimal OPF solution within a limited number of iterations.

Algorithm 2 Distributed OPF Parametric Tracking Scheme

1: **Inputs:**

grid information, allowed number of iterations M, K
new parameters, old solution $z_{t-1}^K, \mu_{t-1}^K, \gamma_{t-1}^K, \lambda_{t-1}^K$

2: **Initialize:**

$z_t^0 \leftarrow z_{t-1}^K, (\mu_t^0, \gamma_t^0, \lambda_t^0) \leftarrow (\mu_{t-1}^K, \gamma_{t-1}^K, \lambda_{t-1}^K)$
choose $\rho > 0$

3: **for** $k = 1, 2, \dots, K$ **do**

4: **for** $m = 1, \dots, M$ **do**

5: $x^m \leftarrow$ update x_i in parallel (15)

6: $z^m \leftarrow$ update z_i in parallel (16)

7: **end for**

8: $x^k \leftarrow x^M, z^k \leftarrow z^M$

9: $\mu^k, \gamma^k, \lambda^k \leftarrow$ update $\mu_i, \gamma_i, \lambda_i$ in parallel (14)

10: **end for**

By integrating the initialization and termination strategies into our scheme, a distributed parametric tracking algorithm is presented in Algorithm 2 for the real-time OPF problem. The outer loop of the tracking scheme consists of a fixed number of primal-dual sequences, while the inner loop solves the nonconvex primal update approximately by a fixed number of proximal alternating minimization steps. Provided that the change of OPF parameters, i.e. the load profile and the battery state, over time is relatively slow, it is proven in [13] that the tracking error of Algorithm 2 has the following property:

$$\begin{aligned} \|\hat{w}_{t+1} - w_{t+1}^*\|_2 &\leq \beta_w(\rho, M) \|\hat{w}_t - w_t^*\|_2 \\ &\quad + \beta_s(\rho, M) \|s_{t+1} - s_t\|_2 \end{aligned} \quad (21)$$

where

$$\beta_w(\rho, M) = C_1(1 + \rho C_2)(1 + \frac{C_3}{\rho})M^{-\frac{1}{2(3^{n_p-1}-1)}} + C_3/\rho$$

$$\beta_s(\rho, M) = C_1(1 + \rho C_2)C_3M^{-\frac{1}{2(3^{n_p-1}-1)}} + C_3C_4/\rho$$

Here, \hat{w}_t is the solution at time t , w_t^* is the KKT point closest to \hat{w}_t , s_t is the parameter of the OPF problem, $n_p \geq 2$ is the number of primal variables, M is number of inner loop iterations, and C_1, C_2, C_3, C_4 are constants.

It can be noticed in (21) that in order to ensure the stability of tracking error, $\beta_w(\rho, M)$ and $\beta_s(\rho, M)$ need to be sufficiently small. This motivates us to enlarge the penalty parameter ρ and the number of inner loop iterations M . By doing so, the contraction of the tracking error is theoretically guaranteed, and thus a series of safety-critical tracking solutions can be provided to the real-time OPF problems.

Remark 1: The number of inner loop iterations M that bounds β_w and β_s to a sufficiently small magnitude needs to grow nonlinearly when the problem dimension n_p increases. However, the choice of M is restricted by the given computational time and network communication rates in practice.

IV. NUMERICAL EXAMPLES

Numerical simulation results are presented in this section to demonstrate that

- Algorithm 1 is capable of solving general nonconvex OPF problems, for which the convex relaxation approach in [8] fails to provide a feasible solution.
- Algorithm 2 is capable of providing a highly feasible and suboptimal closed-loop control policy for the OPF problems in real-time grid operation.

Algorithm 1 and 2 are tested on a 9-bus network with 24-hour prediction horizon involving around 4600 variables and 1600 constraints. The network instance is obtained from the test case in [15]. The demand profiles are generated as the product of the static network demand and the normalized hourly demand statistics of European countries [1]. A battery with 0.6 p.u. capacity is artificially inserted into the network, and its stored energy is set to be 0.2 p.u. at 0h every day.

A. Offline algorithm convergence

Given plenty of offline computational time, we first solve the nonconvex OPF problem to optimality with Algorithm 1. The initial penalty parameter is set as $\rho = 1$ with a growing factor $\beta = 1.1$. The stopping criterion are set as $\eta = 10^{-4}$ and $\epsilon = 10^{-4}$. Figure 2 shows the feasibility residual and objective value against the outer loop iterations. The feasibility error is defined as $r = \|(Ax - b, x^T Ex, x - z)\|_2$ and the relative objective difference is defined as $f(x)/f(\hat{x}) - 1$. Here, x is the OPF solution from Algorithm 1 and \hat{x} is the solution of the corresponding convex relaxed problem [8]. It can be seen that after 60 outer loop iterations Algorithm 1 converges to a solution that satisfies the feasibility tolerance $r = 10^{-4}$. In addition, it ends up with an objective value about 2.8% higher than that of a convex relaxed solution,

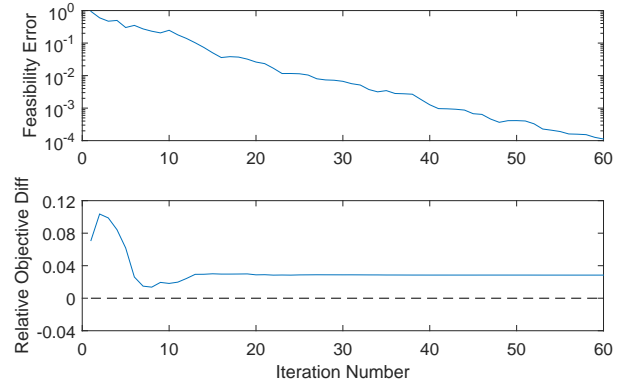


Fig. 2. Feasibility and objective during the iterations of Algorithm 1

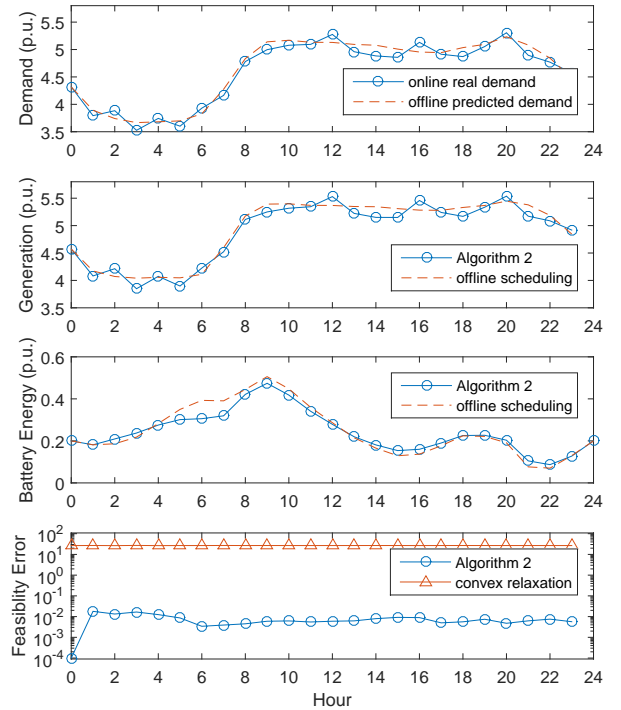


Fig. 3. Closed-loop grid operation controlled by Algorithm 2

which suffers large feasibility error $r = 26.6$ due to its violation of the nonconvex equality constraint (1d).

B. Online closed loop operation

For the online grid operation, we assume a 24h demand forecast service updated on an hourly basis, and solve the OPF problems with the updated grid information at each hour repeatedly. The demand forecast error is artificially created by shifting the demand periodically with a random perturbation up to 4%. Figure 3 shows the performance of Algorithm 2 with 10 outer loop iterations, maximum 500 inner loop iterations and a penalty parameter $\rho = 1$. It can be seen that under the closed loop control of Algorithm 2 the

operations of the generators and batteries are different from the day-ahead scheduling and adapt to the online demand changes. We can also observe that the feasibility errors of the truncated real-time solutions remain at a relatively low level, significantly better than the solution of the convex relaxation approach. This demonstrates the potential of Algorithm 2 in providing a highly feasible and sub-optimal control policy for the real-time grid operation.

V. CONCLUSION

Distributed algorithms based on the augmented Lagrangian method and proximal alternating minimization techniques were proposed to address the nonconvex OPF problem whenever the convex relaxation fails to be exact, either in a network with mesh topology or in a radial network where the convexified OPF solution violates the nonconvex constraint. Compared with centralized approaches, our distributed algorithms split the computations to each agent, hence are more scalable and privacy-sensitive. Furthermore, the proposed real-time OPF tracking scheme can potentially provide a highly feasible solution within a limited number of agent coordinations when feasibility prioritizes, meanwhile sub-optimally controlling power flow and smart devices for closed-loop grid operation in real-time.

REFERENCES

- [1] European network of transmission system operators for electricity, 2014. www.entsoe.eu.
- [2] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forwardbackward splitting, and regularized Gauss-Seidel methods. *Mathematical Programming*, 137(1-2):91–129, August 2011.
- [3] M.E. Baran and F.F. Wu. Optimal capacitor placement on radial distribution systems. *IEEE Transactions on Power Delivery*, 4(1):725–734, January 1989.
- [4] M.E. Baran and F.F. Wu. Optimal sizing of capacitors placed on a radial distribution system. *IEEE Transactions on Power Delivery*, 4(1):735–743, January 1989.
- [5] J. Bolte, A. Daniilidis, and A. Lewis. The ojasiewicz Inequality for Nonsmooth Subanalytic Functions with Applications to Subgradient Dynamical Systems. *SIAM Journal on Optimization*, 17(4):1205–1223, January 2007.
- [6] J. Bolte, S. Sabach, and S. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, 146(1-2):459–494, July 2013.
- [7] E. Dall’Anese, Hao Zhu, and G.B. Giannakis. Distributed Optimal Power Flow for Smart Microgrids. *IEEE Transactions on Smart Grid*, 4(3):1464–1475, September 2013.
- [8] L. Gan, N. Li, U. Topcu, and S.H. Low. Exact Convex Relaxation of Optimal Power Flow in Radial Networks. *IEEE Transactions on Automatic Control*, 60(1):72–87, January 2015.
- [9] L. Gan, A. Wierman, U. Topcu, N. Chen, and S. H. Low. Real-time Deferrable Load Control: Handling the Uncertainties of Renewable Generation. In *Proceedings of the Fourth International Conference on Future Energy Systems, e-Energy ’13*, pages 113–124, New York, NY, USA, 2013. ACM.
- [10] D. Gayme and U. Topcu. Optimal power flow with large-scale storage integration. *IEEE Transactions on Power Systems*, 28(2):709–717, May 2013.
- [11] J.-H. Hours and C.N. Jones. An augmented Lagrangian coordination-decomposition algorithm for solving distributed non-convex programs. In *American Control Conference (ACC), 2014*, pages 4312–4317, June 2014.
- [12] J.-H. Hours and C.N. Jones. An Alternating Trust Region Algorithm for Distributed Linearly Constrained Nonlinear Programs, Application to the AC Optimal Power Flow. *arXiv:1502.03777 [math]*, February 2015. [arXiv: 1502.03777](https://arxiv.org/abs/1502.03777).
- [13] J.-H. Hours and C.N. Jones. A Parametric Non-Convex Decomposition Algorithm for Real-Time and Distributed NMPC. *IEEE Transactions on Automatic Control*, PP(99):1–1, 2015.
- [14] M.R. Irving and M.J.H. Sterling. Efficient Newton-Raphson algorithm for load-flow calculation in transmission and distribution networks. *Generation, Transmission and Distribution, IEE Proceedings C*, 134(5):325–330, September 1987.
- [15] B. Kocuk, S.S. Dey, and X.A. Sun. Inexactness of SDP Relaxation and Valid Inequalities for Optimal Power Flow. *IEEE Transactions on Power Systems*, PP(99):1–10, 2015.
- [16] A.Y.S. Lam, Baosen Zhang, and D.N. Tse. Distributed algorithms for optimal power flow problem. In *2012 IEEE 51st Annual Conference on Decision and Control (CDC)*, pages 430–437, December 2012.
- [17] S.H. Low. Convex Relaxation of Optimal Power Flow - Part I: Formulations and Equivalence. *IEEE Transactions on Control of Network Systems*, 1(1):15–27, March 2014.
- [18] S.H. Low. Convex Relaxation of Optimal Power Flow - Part II: Exactness. *IEEE Transactions on Control of Network Systems*, 1(2):177–189, June 2014.
- [19] S. Magnusson, P. C. Weeraddana, and C. Fischione. A Distributed Approach for the Optimal Power Flow Problem Based on ADMM and Sequential Convex Approximations. *arXiv:1401.4621 [math]*, January 2014. [arXiv: 1401.4621](https://arxiv.org/abs/1401.4621).
- [20] Q. Peng and S. H. Low. Distributed Optimal Power Flow Algorithm for Balanced Radial Distribution Networks. *arXiv:1404.0700 [math]*, April 2014. [arXiv: 1404.0700](https://arxiv.org/abs/1404.0700).
- [21] P. Sulc, S. Backhaus, and M. Chertkov. Optimal Distributed Control of Reactive Power Via the Alternating Direction Method of Multipliers. *IEEE Transactions on Energy Conversion*, 29(4):968–977, December 2014.
- [22] J. W. Taylor. Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204(1):139–152, July 2010.