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 $\mathcal{N} := \{0, ..., n\}$ and a set of lines $\varepsilon := \{1, ..., n\}$. 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 $C_i := \{C_1, C_2, ..., C_R\}$, 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 $s_i = p_i + q_i$ 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 $Z_i = R_i + jX_i$ denote the complex impedance and $S_i = P_i + jQ_i$ 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 (FP7/2007-2013)/ ERC Grant Agreement n. 307608 (BuildNet).

Yuejiang Liu, yuejiang.liu@outlook.com
Jean-Hubert Hours, jhours@gmail.com
Giorgos Stathopoulos, georgios.stathopoulos@epfl.ch
Colin N. Jones, colin.jones@epfl.ch
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 $v_i$, $\bar{v}_i$ respectively, the voltage at bus $i$ is constrained by

$$v_i \leq v_i \leq \bar{v}_i$$

(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:

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

(6)

where $\mathcal{C}_{i2}(t)$ and $\mathcal{C}_{i1}(t)$ 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 shape 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:

**OPF:** \[ \min_{v, L, P, Q, p, q} f(p) \]

over \[ v, L, P, Q, p, q \]

s.t. \[ (1), (2), (3), (4), (5) \]

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-Łojasiewicz property [2]. It tackles the problem of the following structure:

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

where vector $x \in (x_1, \ldots, x_p)$ belongs to the space $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p}$, $g$ is a $C^1$ continuous function with locally Lipschitz continuous gradient, and $h_i : \mathbb{R}^{n_i} \to \mathbb{R} \cup \{+\infty\}$,
In consensus with the corresponding variables \( x^k \) estimated at the adjacent bus \( A \), as follows:

\[
\begin{align*}
\min & \quad \sum_{i=1}^{N} f_i(x_i) \\
\text{s.t.} & \quad A_i x_i = b_i, \quad i \in \mathcal{N} \\
& \quad x_i^T E_i x_i = 0, \quad i \in \mathcal{N} \\
& \quad x_{B,i} = z_i, x_{A,i} = z_{A,i}, x_{C,i} = z_{C,i}, \quad i \in \mathcal{N} \\
& \quad x_i \in \Omega_i, \quad i \in \mathcal{N}
\end{align*}
\]

where \( A_i x_i = b_i \) corresponds to the local linear equality constraints (1a-1c), \( x_i^T E_i x_i = 0 \) corresponds to the nonlinear equality constraint (1d), and \( \Omega_i \) 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:

\[
\begin{align*}
\mathcal{L}_p(x, z; \mu, \gamma, \lambda) &= f(x) + \mu(Ax - b) + \gamma(x^T b_0) + \gamma(x^T E x) + \lambda(x - z) \\
&\quad + \rho \|Ax - b\|^2 + \rho \|x^T b_0\|^2 + \rho \|x - z\|^2 
\end{align*}
\]

subject to \( x \in \Omega \), where \( \Omega := \Omega_1 \times \cdots \times \Omega_n \), \( \rho > 0 \) is a penalty parameter, \( x := (x_1, \ldots, x_n) \) is the group of local variables, \( z := (z_1, \ldots, z_n) \) is the group of consensus variables, \( \mu := (\mu_1, \ldots, \mu_n) \), \( \gamma := (\gamma_1, \ldots, \gamma_n) \), \( \lambda := (\lambda_1, \ldots, \lambda_n) \) 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:**

  \[
  \{x^{k+1}, z^{k+1}\} := \arg \min_{x \in \Omega, z} \mathcal{L}_p(x, z; \mu^k, \gamma^k, \lambda^k) 
  \]

- **Dual update:**

  \[
  \begin{align*}
  \mu^{k+1} &= \mu^k + \rho \left( A x^{k+1} - b \right) \\
  \gamma^{k+1} &= \gamma^k + \rho \left( (x^{k+1})^T E x^{k+1} \right) \\
  \lambda^{k+1} &= \lambda^k + \rho \left( x^{k+1} - z^{k+1} \right)
  \end{align*}
  \]

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
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
\[
\langle x_{i}^{m+1} - x_i^m, \nabla_x L_{z,i}(x_i^m) \rangle + \frac{c_i^m}{2} \| x_{i}^{m+1} - x_i^m \|^2 + \delta_{i} (x_{i}^{m+1}) \leq \delta_{i} (x_i^m)
\]
(18)

As the coefficient \( c_i^m \) satisfies (17), we have
\[
\delta_{i} (x_{i}^{m+1}) + L_{z,i}(x_i^m) + \alpha \| x_{i}^{m+1} - x_i^m \|^2 \\
\leq \delta_{i} (x_i^m) + L_{z,i}(x_i^m) + (x_{i}^{m+1} - x_i^m, \nabla_x L_{z,i}(x_i^m)) + \frac{c_i^m}{2} \| x_{i}^{m+1} - x_i^m \|^2
\]
(19)

Substituting (18) into the right side of (19), one obtains
\[
\delta_{i} (x_{i}^{m+1}) + L_{z,i}(x_i^m) + \alpha \| x_{i}^{m+1} - x_i^m \|^2 \leq \delta_{i} (x_i^m) + L_{z,i}(x_i^m)
\]
(20)

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

The optimality condition of (15) is
\[
\exists i_{m+1} \in \partial \delta_i (x_i^m), \\
\nabla_{x_i} L_{z,i}(x_i^m) + c_i^m (x_{i}^{m+1} - x_i^m) + e_i^m = 0
\]

Then we have
\[
\|e_i^m + \nabla_{x_i} L_{z,i}(x_i^m)\| \\
\leq \|e_i^m + \nabla_{x_i} L_{z,i}(x_i^m)\| + \|\nabla_{x_i} L_{z,i}(x_i^m) - \nabla_{x_i} L_{z,i}(x_i^m)\|
\]
\[
= c_i^m \|x_{i}^{m+1} - x_i^m\|^2 + \|\nabla_{x_i} L_{z,i}(x_i^m) - \nabla_{x_i} L_{z,i}(x_i^m)\|
\]
\[
\leq (c_i^m + L_i) \|x_{i}^{m+1} - x_i^m\|^2
\]

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

For the \( z \)-update, (16) directly implies
\[
L_{z,i}(z_{i}^{m+1}) + \frac{d_i^m}{2} \| z_{i}^{m+1} - z_i^m \|^2 \leq L_{z,i}(z_i^m)
\]

Additionally, the optimality condition of (16) suggests that
\[
\nabla_{z_i} L_{z,i}(z_{i}^{m+1}) + d_i^m (z_{i}^{m+1} - z_i^m) = 0
\]

which implies \( \|\nabla_{z_i} L_{z,i}(z_{i}^{m+1})\| \leq d_i^m \|z_{i}^{m+1} - z_i^m\| \).

Thus, the \( z \)-update satisfies (8) and (9) as well.

**Algorithm 1** Distributed OPF Algorithm

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

Combining (12-17), our distributed algorithm for the nonconvex OPF problem is presented in Algorithm 1. The outer
The previous time step. This motivates us to initialize the dated demand information of a new OPF problem generally operation. A short-term demand forecast provided by the state-of-art 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 \( \lambda_t \) 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 \( \rho \) 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: \( z_t^0 \leftarrow z_{t-1}^\dagger, (\mu_t^0, \gamma_t^0, \lambda_t^0) \leftarrow (\mu_{t-1}^\dagger, \gamma_{t-1}^\dagger, \lambda_{t-1}^\dagger) \)
3: choose \( \rho > 0 \)
4: for \( k = 1, 2, \ldots, K \) do
5: for \( m = 1, \ldots, M \) do
6: \( x^m \leftarrow \text{update } x_i \) in parallel (15)
7: \( z^m \leftarrow \text{update } z_i \) in parallel (16)
8: end for
9: \( x^K \leftarrow x^M, z^K \leftarrow z^M \)
10: \( \mu^K, \gamma^K, \lambda^K \leftarrow \text{update } \mu_i, \gamma_i, \lambda_i \) in parallel (14)

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:

\[
\| \bar{w}_{t+1} - w_{t+1}^\dagger \|_2 \leq \beta_w(\rho, M) \| \bar{w}_t - w_t^\dagger \|_2 \\
+ \beta_s(\rho, M) \| s_{t+1} - s_t \|_2
\]

(21)
where
\[
\beta_w(\rho, M) = C_1(1 + \rho C_2)(1 + \frac{C_3}{\rho^{\frac{1}{2(3^{n_p - 1})}}}) + C_3/\rho
\]
\[
\beta_s(\rho, M) = C_1(1 + \rho C_2)C_3M^{\frac{1}{2(3^{n_p - 1})}} + C_3 C_4/\rho
\]
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 \(\rho\) 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 \(\beta_w\) and \(\beta_s\) 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 \(r = \|(Ax - b, x^\top E x, x - z)\|_2\) 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, 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


